Weaponizing Words: Analyzing Fake News Accusations Against Two Online News Channels

Joni Salminen
Qatar Computing Research Institute
Doha, Qatar
jsalminen@hbku.edu.qa

Milica Milenkovic
Union–Nikola Tesla University
Belgrade, Serbia
micagmekic@gmail.com

Sercan Şengün
Illinois State University
Normal, United States
ssengun@ilstu.edu

Soon-gyo Jung
Qatar Computing Research Institute
Doha, Qatar
sjung@hbku.edu.qa

Bernard J. Jansen
Qatar Computing Research Institute
Doha, Qatar
bjansen@hbku.edu.qa

2021 8th International Conference on Behavioral and Social Computing (BESC) | 978-1-6654-0023-7/21/$31.00 ©2021 IEEE | DOI: 10.1109/BESC53957.2021.9635377

***Abstract*— Attacks against media channels are increasing in social media. The concept of fake news has been weaponized to label and discredit content with which one does not agree. Using data collected from Facebook and YouTube, we analyze attacks against online news channels to understand the logic behind them. Based on programmatic data collection of 4,980,783 comments, we calculate the prevalence of attacks against the channels to range from 0.53% to 2.10% for the media organization owning the channels. The qualitative analysis reveals five major themes in the attacks: Political, Societal, Environmental, Economic & Industrial, and Opinionated. We find that the political theme forms the major entry point for the practice of labeling a news outlet as fake news. Users predominantly do this through the assertion that a news piece does not address the "other side" of an argument and are very vocal about their perceptions. Implications highlight how fake news labeling has the potential to become as detrimental phenomenon as the actual fabrication of content.**

Keywords— Fake news, user-generated attacks, online news

I. INTRODUCTION

There is a growing use of the term "fake news" to discredit legitimate news. It is important to understand the causes of such attacks against media, which can inform detection algorithms, policies, and interventions to address the problem of fake news in society [4,27,28]. It is particularly important to understand the root causes of why online users attack news media. For this research, we define an online attack as a user-generated comment that employs 'fake news' or a similar label against a legitimate news channel.

Interpretative approaches to the analysis of the relationship between disenfranchised audience members and the media are often left unattended, with research focusing on the somewhat more tangential problem of detecting fake news [13,15,20,21]. Our argument is that, while important, detection cannot inform society about the weaponization of the concept of fake news [29], that is, the fact that the term is used by online users to attack legitimate news. Understanding the reasons behind these attacks is crucial for the design of effective mitigation of hate against the media, a rapidly increasing phenomenon [26]. If the relationship between news channels and their audiences continues to deteriorate, this could have dire consequences to society, including political unrest and the lack of legitimate sources of information [29].

In this research, our goal is to classify and analyze online attacks against a media organization, i.e., dealing with the weaponization of the term ("[using] the term to tag real news organizations—whose coverage [the users] disagree with—as purveyors of fake news" [30]). Our analysis is based on the user comments collected from the social media pages of two news channels (identities anonymized due to the sensitivity of the topic) belonging to the same media organization. These channels cover a variety of news topics, including sensitive political issues. A previous investigation of the user comments in the social media channels of this organization [26] shows that, among other targets, there is a high prevalence of attacks against the organization itself. For this reason, the comments provide an interesting dataset to examine attacks against media in the online news context. The goal of this study is to come closer to understanding the reasons for these attacks. We particularly focus on two research questions (RQs):

RQ1: *What themes or topics drive the attacks, and why?*

RQ2: *Are comments attacking the media supported by other users more than non-attacking comments?*

To address RQ1, we analyze a collection of user-generated attacks targeting an online news organization and its news stories. For RQ2, we compare the number of likes that attacking and non-attacking comments garner. Highlights of our findings include:

- The occurrence of fake news label is low (about 1%-2%) compared to another fake news category (i.e., fake news genre).
- "Politics" seems to be the most prevalent topic for fake news label.
- Most fake news label attackers are very direct. They tend to state clearly that the news media is bias/fake/propaganda.
- The attackers generally not receive more support than the non-attackers.

II. RELATED WORK

Fake news is defined as the creation and propagation of factually incorrect information, often of political nature [23]. However, this conventional definition overlooks another, perhaps equally important conceptualization of fake news; that of it being "unpleasant" or "biased" information, which is not necessarily untrue but criticized due to other reasons, such as not reporting the news objectively. Hence, fake news is a loaded term, with different meanings attached to it [30].

The concept of fake news exists in two distinct dimensions: as a genre and as a label [9]. Fake news as a genre encapsulates the deliberate creation of false content and can be accepted as "weaponized propaganda for assaulting agencies reared on public trust" [29]. In this context, it is possible to look at fake news as a content genre through the lenses of journalism, political communication, and even history. In fact, fake news as a genre is not a modern phenomenon. Previous research traces its origins as early as the pre-printing era historical materials of the Byzantine Empire during the 5th century [5]. The invention of the Gutenberg press helped manufacture large-scale news hoaxes and disinformation that are affiliated with events such as "conflicts, regime change, and catastrophes" [24].

However, in the present day, the phenomenon also merged with the issues of online virulence, Internet memes, and social media, dramatically amplifying the effectiveness of the genre, especially for political discourse. For example, in their research regarding the 2016 US presidential election, Allcott and Gentzkow assert that social media is a prevalent source of political information, and social media platforms "can make it difficult to judge an article's veracity." [1]. In response, platforms such as Google, Twitter, and Facebook perform first- and third-party fact-checking, and flagging of content—an initiative that comes with its own criticism [2].

These efforts highlight the second dimension of the phenomenon: fake news as a label. Labeling news as "fake" can be performed both by political actors as a way of "attacking the media" [22], "delegitimiz[ing…] political opponents" [11], and "influencing online trajectories of public discourse" [29], and by other people as a way to "cover tendentious news coverage, partisan rhetoric, and false or outrageous statements by politicians" [22]. The purpose of labeling news as fake is to weaponize the fake news as a genre concept in order to discredit the source of information. Often, this behavior stems from partisanship and group polarization (e.g., "us vs. them") [6]; however, there are other cases of discourses with prevalent fake news labeling that might or might not be based in political discourse [14]. An example is the fake news around COVID-19, labeled as an "infodemic" [19], that merges together medical, public health, national, and political concerns.

Although both of these dimensions can be considered as a process of "weaponization," they need different approaches for detection, analysis, and eventually tackling the problems they cause that may vary from "diminish[ing] democratic processes" [29] to having "negative consequences for public debate and trust in media and political institutions" [11]. On the one hand, research suggests that there is a growing awareness about fake news as a genre—e.g., [22] asserts that the public is aware of fake news; [8] highlights that 57% of journalists consider fake news as disinformation. On the other hand, other research underlines that fake news, as a label, receives less scholarly and public interest—[9]; [8] suggests that only 22% of journalists see the fake news label as a prevalent problem.

To this end, our research addresses a gap of mixed-method analyses of the weaponization of fake news as a label (i.e., attacks against the media). First, we will analyze around 80,000 social media comments that were flagged as fake news discussions found among 5 million comments from the social media channels of two separate news outlets, and provide descriptive statistics of the results. Then, we will further analyze 3,011 randomly selected comments qualitatively to determine the themes and categories within the discussions.

III. DATA COLLECTION

To collect the data, we used two application programming interfaces (APIs), YouTube Analytics API and Facebook Insights API, returning a total of 4,980,783 comments from the YouTube (YT) and Facebook (FB) pages of the two channels (see Table 1) belonging to the news and media organization. The pages in question contain news stories (videos on YT and text/video posts on FB). The comments were collected from news stories on FB and YT. The data were collected with the channel owner's permission and in accordance with the online platforms' terms of service.

TABLE I. NUMBER OF COMMENTS FOR EACH CHANNEL. TWO CHANNELS (C01 AND C02, REAL NAMES HIDDEN DUE TO THE SENSITIVITY OF THE RESEARCH TOPIC) BELONG TO THE SAME ORGANIZATION.

Platform	Channel	Number of total comments	Number of flagged comments*
YT	C01	253,499	5,752 (2.27%)
FB	C01	1,888,782	47,734 (2.53%)
YT	C02	783,130	13,348 (1.70%)
FB	C02	2,055,372	13,167 (0.64%)
	Grand totals	4,980,783	80,001 (1.61%)

*Number of flagged comments indicates the combined number of comments corresponding with the search terms ('bias', 'fake news', 'propaganda'. Percentage indicates share from total comments retrieved for the channel.

C01 is a "social media-only" news channel, which means all its news stories are published only on social media platforms (primarily on YT and FB, which is why we chose these two platforms for the analysis). **C02** can be considered a more "traditional" news channel, with both a website as the main hub for its news stories and social media outlets that typically link to the news stories on the main website. Hence, we found it interesting to compare the attacks on these two slightly different types of media channels that both report news stories.

Three keywords ("**bias**," "**fake news**," and "**propaganda**") were used to search the comment sections of the channels' news stories for comments that potentially attack the media channel or a given news story: These terms were selected because they relate to attacks against media [6]; that is, media is claimed to be biased and spreading fake news and propaganda. As the media organization does no active moderation, this sample represents a predominantly "unfiltered" view into the user-generated attacks against the channel and its content. Figure 1 shows the distribution of the flagged comments by channel and platform.

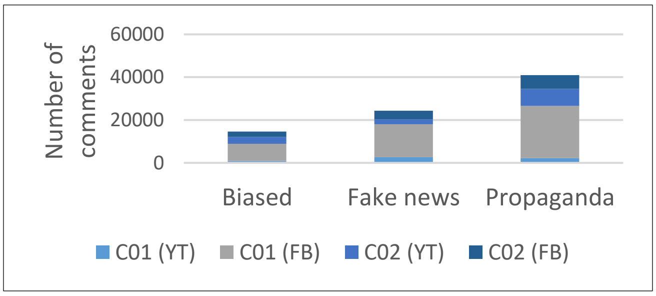

Fig. 1. Number of comments based on the keyword search. The keyword "propaganda" yielded the most comments, and the C01 FB channel accounted for most of the comments.

IV. DATA ANALYSIS

Because to the large number of comments that were identified as potential attacks, we decided to manually analyze 1,000 randomly sampled comments for each keyword. We applied content analysis of discussion forums [17], which involves code assignment (code = a conceptual label assigned to a piece of text, in this case, to the comments), cross-checking with another researcher, code interpretation, and analyses such as descriptive statistics (we calculated the frequency of each code). During the initial reading of the comments, we came across false positives (i.e., not actual attacks against the channel), and those were marked as such. Table 2 shows a categorization of results by keyword. One of the researchers coded the material, and another one verified the validity. Any disagreements were solved through discussion [17].

TABLE II. ANALYZED COMMENTS BY KEYWORD AND ATTACK CATEGORY.

Units per keyword	Attack category	Definition	Coded instances (N)s
Bias • 1003 comments • 807 true positives • 196 false positives	Biased Media/Channel	Accusation is made towards the media channel, their page, or the media community.	260 (32.22%)
	Biased Report/News	Accusation is made towards the report, video, article, or using the pronouns to point the particular coverage.	487 (60.35%)
	Biased Interviewer/Reporter	Accusation is made towards the person conducting the interview.	66 (8.18%)
Fake News • 1002 comments • 795 true positives • 207 false positives	Fake News Announcement	Pointing out that the report/news is actually fake news.	712 (89.56%)
	Fake News Possibility	A person is not entirely sure in their statement that the news is actually fake news.	13 (1.64%)
	Stop Reporting Fake News	Asking the media to stop reporting, spreading, using fake news or misinformation.	70 (8.81%)
Propaganda • 1006 comments • 865 true positives • 141 false positives	Propaganda Announcement	Pointing out that the report/news is a propaganda.	786 (98.87%)
	Stop the Propaganda	Asking the media to stop reporting, using or spreading propaganda.	63 (7.92%)
	Propaganda Possibility	A person is not entirely sure in their statement that the news represents propaganda.	16 (2.01%)

V. RESULTS

A. Data Exploration

Table 1 shows the different types of attacks. Accusations of bias are mostly targeting the specific news stories (60.4%). In turn, most labels of fake news and propaganda are more general (89.6% and 98.9%, respectively) in nature. A minority of commentators is asking the news channel to stop reporting fake news, but most comments only contain an accusation.

Using the true and false positive counts from Table 2, we can estimate the "real" prevalence of attacks (i.e., attack rate) for each channel, platform, and the organization as a whole. This is simply the number of flagged comments times the true positive rate (TPR). For example, for C01, FB has 26,535 flagged for the term "propaganda" (i.e., many comments contained that word). The TPR for "propaganda" is 86.0%, so the estimated true number of comments attacking C01 on FB based on it being "propaganda" is 26,535 × 0.86 = 22,816. Table 3 shows the estimated attack rates per channel and social media platform. The average attack rate of the whole organization (i.e., the number of TPR-corrected attacks in all channels and platforms divided by the number of total comments) is 1.33%.

TABLE III. ATTACK RATES BASED ON TPR-CORRECTED TOTALS. THE MACRO-AVERAGE IS THE AVERAGE OF YT AND FB ON THE BOTTOM ROW, AND C01 AND C02 IN THE THIRD COLUMN. C01 IS MORE TARGETED THAN C02, AND ACCUSATIONS ARE MORE COMMON IN YT THAN FB.

	C01	C02	Macro-average
YT	1.86 %	1.42 %	1.64%
FB	2.10 %	0.53 %	1.32%
Macro-average	1.98 %	0.98 %	

We also compared the number of likes in true positive and false positives to address if attacks against the media are "voted higher" than non-attacks by other users (RQ2). For this analysis, we divided the comments into two groups, true positives (attacks) and false positives (non-attacks), and tested the significance of either having a higher like count. A two-sample t-test assuming unequal variances showed that non-attacks ($M = 112.9$) have more likes than attacks ($M = 59.8$), $t(569) = -2.50$, $p = 0.013$. Percentage-wise, this difference is substantial (i.e., non-attacks have 88.8% more likes on average than attacks), which can be interpreted as "good news"—The attacks against media are generally not supported more strongly than "regular" (non-attack) comments, but the results indicate the opposite.

B. Qualitative Analysis

Overall, 63 initial topics were discovered manually, which we grouped into larger themes based on conceptual similarity. The main themes were Political, Societal, Environmental, Economic and Industrial, and Opinionated (see Figure 2). The themes are discussed in the following subsections.

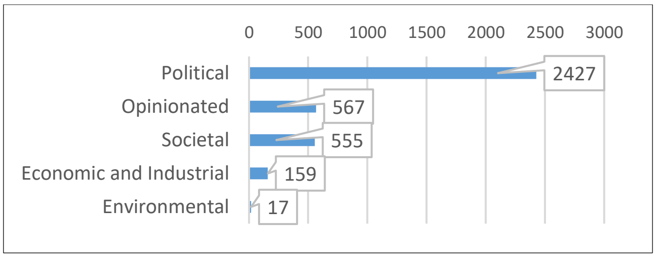

Fig. 2. Frequency of comments per topic.

a) Political: Political is the largest theme, with subtopics and examples shown in Table 4. The theme has five subtopics: (**S1**) regulations and policies, (**S2**) individual countries, (**S3**) public sector and politics in general, (**S4**) conflicts and disturbances, and (**S5**) other for comments that do not fit in the first four. In the **S1**, users discuss the "fakeness" of news pieces supporting or criticizing a public policy or a politician's view of that policy. The S1 also marks the assertions that some statements made by politicians did not happen or happened in a different context than the news piece reports. In the **S2**, users reject and label news reports about countries as fake or biased. In the **S3**, users discuss the validity of news pieces around general political events. In the **S4**, the discussions revolve around the idea of propaganda, especially the role of the news media in the conflict depending on the country from where the media originates. Finally, the **S5** includes discussions that cannot be categorized in the first four.

TABLE IV. TOPICS AND EXAMPLES OF THE POLITICAL THEME.

Subtopic	Example	
	News story	**Comment**
Regulations and Policies (Abortion policy, Vaccination policy, Firearm regulations)	"Donald Trump wants to end the right to abortion, but does he understand the science?"	"[C01] is biased and unprofessional. I'm with Trump on this topic. Anyone who supports abortion is a monster. You were all once defenseless foetuses. No one has the right to murder a defenseless human being."
Individual Countries (Mentions of various countries such as US, Russia, Syria, Israel, Yemen, Iran, France, The Philippines…)	"Officially, these kids don't even exist. Meet Malaysia's invisible children."	"The first time I admit [C01] is bullshit, but because of this all Malaysian together defending our country, because of one biased video"
Public Sector and Politics in general (Government Officials and Law Enforcement, Elections, and various political events)	"French President Emmanuel Macron walks towards President Donald J. Trump at the NATO summit, then turns to Angela Merkel at the last minute."	"only illiterates will buy this cheap propaganda,i see nothing wrong in what Macron did.Angela Merkel was the only female there at the upfront,so he chooses to hug her first before saluting the Men.#ofcourse ladies first. [C02] this is not an article,tell us what is happening in Manchester city,about the terror attack"
Conflicts and Disturbances (The Philippine Drug War, Syrian Civil War, The Ukraine – Russia Conflict, Israel – Palestine Conflict, Protests, Terrorism, Violence)	"'I hope you can remember us.' This teacher in Aleppo shares a harrowing "last call' video as government forces take control of the city."	"This is what your government Saudi Arabia has done. Your Isis will soon be eliminated. Stop your shitty propaganda. How about the genocide in Yemen. Still on that?"
Other topics (Left – Wing Politics, Refugees, The Catalan Independence, Terrorism, US Presidential Candidates)	"In Trump's America, Refugees Are Fleeing For Canada [Divided America, Pt. 2] | [C01] Docs"	"highly misleading propaganda. these are not refugees unless they come from the Syrian area. These are overwhelmingly economic migrants that are male between the ages of 16 and 40."

b) Societal: Societal includes topics such as race, religion, and social issues. Subtopics and examples of attacks are shown in Table 5. The societal theme has three subtopics: (**S6**) race, (**S7**) religion, and (**S8**) other social issues (such as gender, sexual orientation, violence, etc.). The **S6** encapsulates discussions around generalized racial language. Generalized racial language has been identified as a misleading linguistic tool to talk about membership in communities [25]. The **S7** includes discussions around how members of certain religious identities are portrayed. The **S8** includes other social issues, such as gender, sexual orientation, medicine, and violence.

TABLE V. TOPICS AND EXAMPLES OF THE SOCIETAL THEME.

Subtopic	Example	
	News story	**Comment**
Race	"Black Muslims in Britain are fed up with being erased."	"Who is erasing them? 👀 👀 Fake news to get black people hated up.. 😂 😂 😂"
Religion	"Canadian Prime Minister Justin Trudeau addresses the shootings at a Quebec City mosque."	"More fake news to make you sympathize with Muslims pathetic"
Social Issues (Gender (in)equality, along with the LGBT topic, and Drugs/Medicine topic, Violence)	"South Korea's Gender Wars: Trolls, Threats and Anger Online - 101 East"	"this documentary is so biased against men. the gender wars are not specific to korea, its turning into a worldwide phenomenon. explore the truth instead of just asking women for their opinions. in most cases, the life of a man revolves around his woman... she is the end; when women start replacing the man with the State and discard them and refuse to bear children or harbour desire for any maternal role, it's inevitable this war will break out. the destruction of the family and rise of consumerism has resulted in this mess. the world will never be the same after this war ends. the next age of machines and AI is coming. it's tragic to see this, but men and women will finally be free."

c) Environmental: This is both a topic and a theme (see Table 6). This topic/theme mostly includes discussions around the validity of climate change news and research.

TABLE VI. TOPICS AND EXAMPLES OF THE ENVIRONMENTAL THEME.

Subtopic	Example	
	News story	**Comment**
Environmental (mentions and news about climate change, environmental issues, and related events)	"It's so cold that iguanas are freezing in trees and seawater is turning to ice."	"climate change is not real, this is fake news."

d) Economic and Industrial: Economic and industry are closely connected matters, so we grouped attacks based on these matters under one theme (see Table 7). The theme has three subtopics: (**S9**) corporations, (**S10**) economic issues, and (**S11**) entertainment. The **S9** includes discussions around the "fakeness" of news about certain corporations. Often, these attacks overlap with other social or economic issues. The **S10** mostly focuses on creating propaganda for economic gains. Users highlight that the news outlet focuses on certain economic news to affect the market and business outcomes along local and national lines. Finally, in the **S11**, users discuss issues around sports, celebrities, etc. It is interesting to see that even such discussions can be marked as fake news or propaganda with economic or political outcomes.

TABLE VII. TOPICS AND EXAMPLES OF THE ECONOMIC AND INDUSTRIAL THEME.

Subtopic	Example	
	News story	Comment
Corporations (mentions of various companies, and their correlation to the news/report, or the media house)	"Is Victoria's big Secret that the company doesn't like trans models?"	"Well this backfired on [C01] typical biased media 😅"
Economic issues (this topic covers economic crisis, funding, and mentions of money on a global scale)	"India: The next global economic powerhouse? - Counting the Cost"	"keep doing your propaganda. ..we don't care. .."
Entertainment (World Cup, Olympic, Other sporting events, mentions of musicians, artists, actors, and other celebrities)	"When it's Barça vs. Real Madrid, it's not just soccer. It's politics."	"A biased report, obviously believing the pro-independence propaganda given out. Yes, the Barça fans are politicized, but it doesn't mean they represent the majority of Catalans, not even the majority of people who live in Barcelona. You need more information from Catalans who feel they are also Spanish."

e) Opinionated: Table 8 shows examples of the theme. The theme has four subtopics: (**S12**) accusations, (**S13**) the other side of the story, (**S14**) shame, and (**S15**) poor research. This theme encapsulates the opinions directly related to the news outlet. For example, accusations against the news outlet for being a propaganda tool, for not covering the other side of the story, shaming the news outlet for their general approach, or blaming them for poor background research.

TABLE VIII. TOPICS OF THE OPINIONATED THEME

Subtopic	Example	
	News story	Comment
Accusations (Accusations of media supporting terrorists, being manipulative or spreading hate)	"Inside Story - Shimon Peres: a man of peace or war criminal?"	"[C02]: a news channel or terrorist propaganda platform."
Other side of the story (the commentator gives additional opinion on why he things the media is biased, that news is fake, or propaganda)	"LIVE outside Al Aqsa Mosque in Jerusalem."	"Fake news! they are NOT letting them get in! They have metal detectors so their blood thirsty worthless bodies dont kill innocent israelis."
Shame (A commentator mentions that media should be ashamed of themselves for the particular news reported, or a history of reported news)	"Israel destroyed a Palestinian village in the Negev. It wants to build a Jewish village there instead."	"Liars!!! This are bedouin that are israeli citizens. They live in rahat and build houses illegally outside the city borders. They were asked by the government not to build there but they don't listen. Shame on you for telling fake news!"
Poor research (A commentator is pointing out that more research was necessary, that the report lacks evidence and facts)	"Muslims in China are being forced to abandon their traditions and live under constant surveillance. Now they're fleeing the country in huge numbers."	"thank you for being so biased without solid evidence or data."

VI. TOPIC PREVALENCE

Attacks on environmental topics are negligible, but politics makes for 70% of the attacks. Topics that were mentioned the most (see Table 9) are the Syrian civil war and the Israel-Palestine conflict, and various types of violence either presented in the news report or was mentioned in the comment, along with the opinions of the commentators. Unsurprisingly, the government official who received the most attention, both from the media itself and users, is President Donald Trump.

TABLE IX. MOST FREQUENT TOPICS

Topics	Frequency (% of total)
Syrian civil war	338 (20.5%)
Other side of the story	223 (13.5%)
Trump	206 (12.5%)
Violence	187 (11.3%)
Israel - Palestine conflict	163 (9.9%)
Politics - general	135 (8.2%)
India	105 (6.4%)
Economic issues	101 (6.1%)
Qatar sponsorship	98 (5.9%)
Religion - Islam	95 (5.8%)

When the political news involves conflict, disturbance, violence, or disruption, e.g., the Syrian civil war or Israel-Palestine conflict, the news or the comment itself mentioned a lot of violence; in our case, violence is defined as a whole variety of destructive events, from assaults, killings, bombings to genocide and mass destruction. At times, the audience is writing their own version of the news report. In their attacks, audience members may attempt to elaborate reasons for the events covered in the news, or to point out other similar events that the report did not cover, in order to prove a bias or propaganda, or to write the version that, according to them, is not "fake news".

A theme that contains the most emotional and personal attachment of the reader to the news report itself is the opinionated theme. The audience here accuses media of misconduct, a hidden agenda, shady businesses, or simply pointing out another version of the story (see Table 10).

TABLE X. RATIONALE OF THE OPINIONATED THEME

Rationale	Frequency
Other side of the story	223 (47.58%)
Poor research	75 (15.99%)
Accusation of supporting terrorists	67 (14.29%)
Shame	58 (12.37%)
Accusation of spreading hate	31 (6.61%)
Accusation of being manipulative	15 (3.20%)

VII. DISCUSSION, IMPLICATIONS, & LIMITATIONS

Out of the ~5M collected comments, only ~80K (1.6%) were flagged as fake news labeling (i.e., mentioning "bias," "fake news," and/or "propaganda"). Compared to previous research, it appears that fake news as a genre is more common than fake news as a label. For example, Moscadelli et al. [18]

assert that 23.1% of the articles shared in Italy about the COVID-19 pandemic between December 2019 and April 2020 can be considered fake news. Another study [12] suggests that 6.7% of political URLs shared on Twitter during the 2016 US presidential election came from fake news sources. Yet another study finds that 9% of the content shared about the 2019 European Parliament elections came from “disinformation news outlets” [7]. Thus, fake news labeling seems to be a less prevalent problem than fake news as a genre. An outlier is FB, asserting in 2017 that fewer than 0.1% of the user-generated content shared on their platform could be considered malicious; however, the methods of this finding are not clearly defined [16].

Our results also underline the difference between the two news outlets, as exemplified by the average percentages of fake news labeling comments (1.98% for **C01** vs. 0.98% for **C02**). This highlights a difference of perception between news outlets, with some of them garnering more fake news labels than others. Previous research suggests that “increased knowledge of media ownership may affect judgments of credibility in responding to print news” [3] as well as the news outlet being national versus international [10]. We believe that our results are the outcome of a combination of the outlet’s originating country and the subject matters it typically covers. In particular, **C01** is a ‘social media only’ news channel, which could possibly make it more susceptible to attacks than a more traditional news channel. There was also a slight difference between the social media channels (1.64% for YT vs. 1.32% for FB). This might be explained by the anonymity that YouTube offers for its users, as well as the comparison between the effects of moving image content versus textual or static image content.

Our qualitative analysis illustrates that the users predominantly label specific news pieces as biased (60.35%) while occasionally labeling the outlet as biased (32.22%). When users find a news piece that they perceive as biased, they announce that it is fake (89.56%) or propaganda (98.87%). The qualitative coding of a subsample of 3,011 comments revealed that the reasons for attacks are predominantly Political (65.2%), occasionally Opinionated (15.2%) and Societal (14.9%), and rarely Economic/Industrial (4.3%) or Environmental (0.50%). By further examining the Political, Societal, and Opinionated categories, it is possible to see that most loaded topics that cause fake news labeling are military conflicts (e.g., Syrian civil war, Israel-Palestine conflict) and policymaking (e.g., Donald Trump, India, etc.).

When faced with content that they perceived as fake news, users mostly try to highlight the “other side” of the story or complain that this other side is not represented (47.58%). This is followed by accusations of news reports being poorly researched (15.99%), news outlets (or the countries from which the outlet originates) supporting conflicts or terrorism (14.29%), as well as general shaming language (12.37%).

Our results have implications for future research. First, the phenomenon of using fake news as a label has the potential to become as prevalent as the fake news as a genre. It should be noted that fake news as a genre requires a more extensive effort; one needs to fabricate false information, publish it on channels that mimic news media, and disseminate it over social media channels. Fake news as a label, on the other hand, has a low barrier to mobilize as it piggybacks on existing news outlets and legitimately created content. Labeling a news piece or an outlet as fake news delegitimizes it without even a need for diametrically opposite content — even without fabricating fake news, it is possible to cause equal disinformation by labeling legitimate sources as fake news. This builds a public discourse wherein news can be contested without an antithesis or a disproof simply by labeling it as fake news.

Although identifying fake news content and overcoming it might be possible, it might be harder to overpower a group of voices that weaponize fake news as a label. Especially when mobilized by public figures or politicians with a wider reach, this practice has a strong potential of corroding political communication in the public sphere. To this end, it should be underlined that our number displays the tendency of the general public to use fake news as a label against a media outlet and does not include the behaviors of politicians and other public figures who might have a higher reach than regular social media users.

Regarding limitations, the keyword-based method we applied might not be effective enough to extract all types of fake news labels. Therefore, it is possible some forms of attacks labeling the news story or channel as fake news were missed. Future studies could inductively discover new keywords for fake news labeling detection to develop a more comprehensive lexicon. Future work could also apply automatic topic models for topic discovery among the attacks (e.g., LDA), as well as develop machine learning models for attack detection (e.g., text classification). Finally, while we analyze the fake news accusation attacks, defense methods against such attacks remain an important issue for future research.

VIII. CONCLUSION

Users actively point out what they perceive as mis- or disinformation in social media. Although other topics can also lead to labeling as fake news, the dominant source is politics. Because highlighting the “other side” of the argument is the main strategy mobilized by users in labeling content as fake news, news outlets labeled as fake news should focus on covering multiple angles to weaken or avoid such implications when reporting political news.

REFERENCES

[1] Hunt Allcott and Matthew Gentzkow. 2017. Social media and fake news in the 2016 election. *Journal of economic perspectives* 31, 2 (2017), 211–36.

[2] Jack Andersen and Sille Obelitz Søe. 2020. Communicative actions we live by: The problem with fact-checking, tagging or flagging fake news–the case of Facebook. *European Journal of Communication* 35, 2 (2020), 126–139.

[3] Seth Ashley, Mark Poepsel, and Erin Willis. 2010. Media Literacy and News Credibility: Does knowledge of media ownership increase skepticism in news consumers? *Journal of Media Literacy Education* 2, 1 (2010), 3.

[4] Laura Burbach, Patrick Halbach, Martina Ziefle, and André Calero Valdez. 2019. Who Shares Fake News in Online Social Networks? In *Proceedings of the 27th ACM Conference on User Modeling, Adaptation and Personalization* (UMAP ’19), Association for Computing Machinery, New York, NY, USA, 234–242. DOI:https://doi.org/10.1145/3320435.3320456

[5] Joanna M. Burkhardt. 2017. History of fake news. *Library Technology Reports* 53, 8 (2017), 5–9.

[6] Pere-Lluís Huguet Cabot, David Abadi, Agneta Fischer, and Ekaterina Shutova. 2021. Us vs. Them: A Dataset of Populist Attitudes, News Bias and Emotions. *arXiv:2101.11956 [cs]* (January 2021). Retrieved February 1, 2021 from http://arxiv.org/abs/2101.11956

[7] Matteo Cinelli, Stefano Cresci, Alessandro Galeazzi, Walter Quattrociocchi, and Maurizio Tesconi. 2020. The limited reach of fake news

on Twitter during 2019 European elections. *PloS one* 15, 6 (2020), e0234689.

[8] Jana Laura Egelhofer, Loes Aaldering, Jakob-Moritz Eberl, Sebastian Galyga, and Sophie Lecheler. 2020. From novelty to normalization? How journalists use the term "fake news" in their reporting. *Journalism Studies* 21, 10 (2020), 1323–1343.

[9] Jana Laura Egelhofer and Sophie Lecheler. 2019. Fake news as a two-dimensional phenomenon: a framework and research agenda. *Annals of the International Communication Association* 43, 2 (April 2019), 97–116. DOI:https://doi.org/10.1080/23808985.2019.1602782

[10] Mokhtar Elareshi and Barrie Gunter. 2012. Credibility of televised news in Libya: Are international news services trusted more than local news services? *Journal of Middle East Media* 8, 1 (2012).

[11] Kate Farhall, Andrea Carson, Scott Wright, Andrew Gibbons, and William Lukamto. 2019. Political Elites' Use of Fake News Discourse Across Communications Platforms. *International Journal of Communication* 13, (2019), 23.

[12] Nir Grinberg, Kenneth Joseph, Lisa Friedland, Briony Swire-Thompson, and David Lazer. 2019. Fake news on Twitter during the 2016 US presidential election. *Science* 363, 6425 (2019), 374–378.

[13] Momchil Hardalov, Ivan Koychev, and Preslav Nakov. 2016. In search of credible news. In *International conference on Artificial intelligence: methodology, systems, and applications*, Springer, 172–180.

[14] P. Sol Hart, Sedona Chinn, and Stuart Soroka. 2020. Politicization and Polarization in COVID-19 News Coverage. *Science Communication* 42, 5 (2020), 679–697.

[15] Georgi Karadzhov, Pepa Gencheva, Preslav Nakov, and Ivan Koychev. 2018. We built a fake news & click-bait filter: what happened next will blow your mind! *arXiv preprint arXiv:1803.03786* (2018).

[16] David MJ Lazer, Matthew A. Baum, Yochai Benkler, Adam J. Berinsky, Kelly M. Greenhill, Filippo Menczer, Miriam J. Metzger, Brendan Nyhan, Gordon Pennycook, and David Rothschild. 2018. The science of fake news. *Science* 359, 6380 (2018), 1094–1096.

[17] Rose M. Marra, Joi L. Moore, and Aimee K. Klimczak. 2004. Content analysis of online discussion forums: A comparative analysis of protocols. *Educational Technology Research and Development* 52, 2 (2004), 23.

[18] Andrea Moscadelli, Giuseppe Albora, Massimiliano Alberto Biamonte, Duccio Giorgetti, Michele Innocenzio, Sonia Paoli, Chiara Lorini, Paolo Bonanni, and Guglielmo Bonaccorsi. 2020. Fake news and covid-19 in Italy: Results of a quantitative observational study. *International journal of environmental research and public health* 17, 16 (2020), 5850.

[19] Salman Bin Naeem and Rubina Bhatti. 2020. The Covid-19 'infodemic': a new front for information professionals. *Health Information & Libraries Journal* 37, 3 (2020), 233–239.

[20] Preslav Nakov. 2020. Can We Spot the" Fake News" Before It Was Even Written? *arXiv preprint arXiv:2008.04374* (2020).

[21] Van-Hoang Nguyen, Kazunari Sugiyama, Preslav Nakov, and Min-Yen Kan. 2020. FANG: Leveraging social context for fake news detection using graph representation. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management*, 1165–1174.

[22] Rasmus Kleis Nielsen and Lucas Graves. 2017. " News you don't believe": Audience perspectives on fake news. (2017).

[23] Gordon Pennycook, Tyrone D. Cannon, and David G. Rand. 2018. Prior exposure increases perceived accuracy of fake news. *J Exp Psychol Gen* 147, 12 (December 2018), 1865–1880. DOI:https://doi.org/10.1037/xge0000465

[24] Julie Posetti and Alice Matthews. 2018. A short guide to the history of'fake news' and disinformation. *International Center for Journalists* 7, (2018), 2018–07.

[25] John R. Rickford. 2016. *Raciolinguistics: How language shapes our ideas about race*. Oxford University Press.

[26] Joni Salminen, Hind Almerekhi, Milica Milenković, Soon-gyo Jung, Jisun An, Haewoon Kwak, and Bernard J. Jansen. 2018. Anatomy of Online Hate: Developing a Taxonomy and Machine Learning Models for Identifying and Classifying Hate in Online News Media. In Proceedings of The International AAAI Conference on Web and Social Media (ICWSM 2018), San Francisco, California, USA.

[27] Suleiman Usman Santuraki. 2019. Trends in the Regulation of Hate Speech and Fake News: A Threat to Free Speech? Hasanuddin Law Review 5, 2 (2019), 140–158.

[28] Kai Shu, Amy Sliva, Suhang Wang, Jiliang Tang, and Huan Liu. 2017. Fake news detection on social media: A data mining perspective. ACM SIGKDD Explorations Newsletter 19, 1 (2017), 22–36.

[29] Christopher A. Smith. 2019. Weaponized iconoclasm in Internet memes featuring the expression 'Fake News.' Discourse & Communication 13, 3 (2019), 303–319.

[30] Edson C. Tandoc. 2019. The facts of fake news: A research review. Sociology Compass 13, 9 (2019), e12724. DOI:https://doi.org/10.1111/soc4.12724