

# Flying Under the Radar : How Smaller Social Media Groups Are More Susceptible to Misinformation <sup>\*</sup>

Christine Sowa Lepird<sup>1</sup>[0000-0002-9350-5627], Dr. Kathleen M. Carley<sup>1</sup>[2222--3333-4444-5555], and Anna Wu<sup>2</sup>[0009-0006-3840-8329]

<sup>1</sup> Carnegie Mellon University, Pittsburgh PA 15213, USA [csowa@andrew.cmu.com](mailto:csowa@andrew.cmu.com)  
<http://www.casos.cs.cmu.edu>

<sup>2</sup> Troy High School, Fullerton CA 92831, USA [anna.aq.wu@gmail.com](mailto:anna.aq.wu@gmail.com)  
<https://www.fjuhsd.org/Troy>

**Abstract.** This paper analyzes posts sharing news articles pertaining to the 2022 U.S. midterm elections originating from Twitter, Reddit, and Facebook pages. After classifying the news articles shared, this analysis found that pink slime and misinformation receive more engagement per group size (on platforms with community features) than authentic local and national news sites. Furthermore, it finds that the distribution of types of news sources that a Facebook Page shares changes with its size. Finally, the network features of groups sharing authentic local news are found to have a higher betweenness centrality than those of groups sharing pink slime.

**Keywords:** cross-platform · misinformation · elections.

## 1 Introduction

Social media platforms play a critical role in facilitating the dissemination of information and the propagation of campaign narratives during events of political significance, such as the U. S. midterm elections. The ability of social media to reach large audiences in minimal time has established it as an important source of news for voters. It is important, therefore, to determine the quality and credibility of news shared over social media and assess differences in user response and engagement by type of news media.

Previous research in this area correlated demographics [8], ideological bias [7] and ideological extremity [10] with news type, typically “real” or “misinformation,” finding a positive relationship between age, conservatism, and polarization

---

<sup>\*</sup> This work was supported in part by the Office of Naval Research (ONR) Award N00014182106, the Knight Foundation, the Center for Computational Analysis of Social and Organizational Systems (CASOS), and the Center for Informed Democracy and Social-cybersecurity (IDeaS). The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the ONR or the U.S. government.

and sharing misinformation. Susceptibility to misinformation as a metric of the average user without any qualifiers or restrictions is mostly overlooked, with the exception of [6]. That research found that, in a general multi-platform dataset, real news and misinformation exhibited similar spread patterns, potentially indicating that users are unable or unwilling to distinguish between credible and non-credible news.

Furthermore, while the existing literature predominantly compares real and misinformation dissemination, more research is necessary in the comparison of local news and pink slime, a type of low-credibility, often partisan news disguised as a local publication to garner trust. This paper analyzes the relationship between type of news (real, misinformation, local, pink) and user engagement, measured in like count and corrected for the size (implied popularity) of its page of origin.

## 2 Data and Methods

Data was collected from Twitter, Reddit, and Facebook pertaining to the United States 2022 Midterm Elections in regions with the most contention elections. The posts pulled from each of the platforms contain URLs to external sites for further analysis. The elections took place on November 8, 2022, and the data was collected from October 1, 2022 to December 1, 2022. Elections selected for this analysis included the most competitive districts and regions in Arizona, Georgia, Pennsylvania, Nevada, and Wisconsin. For the full set of races and keywords, contact the authors.

The Twitter researcher, Reddit’s Pushshift API [13], and Facebook’s Crowd-Tangle API [1] were all used to pull the data for this research.

Once the data was collected, the URLs that the posts linked to were cleaned, and their Media/Bias Fact Check ratings were included [2]. Additionally, throughout this paper, “known sources of real news, misinformation, or pink slime” will be defined as news articles originating from domains listed as “real, misinformation, or pink slime” in the media thesaurus compiled by the CASOS University Center at Carnegie Mellon University. The media thesaurus has been compiled from multiple publicly available lists of news media URLs and media organizations’ Twitter accounts: Media Bias/Fact Check [2] lists many news sites and rates how factual and credible the reporting is for many; the George Washington University Dataverse [12] has a list of over 9600 Twitter accounts for media organizations, derived from over 160 million tweets between 2016 and 2020; the Columbia Journalism Review site has been a source for hundreds of “pink slime” news outlet domains [11] that often publish biased, algorithmically produced stories; there is also a Github repository [3] of unreliable, misleading, and/or “misinformation” news sources that includes lists from Snopes Field Guide, Melissa Zimdars’ OpenSources, Wikipedia, and others. There is often overlap between these sources, particularly for the less factual news outlets; to resolve any conflicts that emerge between the sources, the thesaurus errs on the side of not labeling a news source in question as misinformation. Finally, the

labels of “local news” were through list of authentic local news sites owned by companies [4] as well as this Github repository [9].

### 3 Analysis and Results

The full dataset included 1,306,829, 25,221, and 15,867 posts from Twitter, Facebook Pages, and Reddit, respectively. Of those, 756,084 (Twitter) , 12,885 (Facebook Pages) , and 7,275 (Reddit posts) linked to URLs that had a designated CASOS rating. Per Table 1, Reddit leads with the highest percentage of shared news on the topic being to real news sites and the lowest percentage going to misinformation sites - possibly due to the work of subreddit moderators. While Facebook doesn’t have as high of a proportion of news going to misinformation sites as Twitter, the Facebook Pages lead in links going to local news sites as well as pink slime sites. Contrary to [5] there were instances of pink slime sites shared on Reddit. While their research did not find pink slime sites on Reddit (perhaps to only searching political subreddits), this dataset found 11 references to these sites, mostly on smaller subreddits geared towards a local community.

Table 1: Breakdown of news types shared on the three platforms.

	Twitter	Facebook	Reddit
Real News Sites	80.39%	71.67%	89.84%
Misinformation Sites	5.42%	4.50%	0.64%
Local News Sites	13.63%	22.5%	9.39%
Pink Slime Sites	0.56%	1.32%	0.13%

#### 3.1 Relative Engagement

Facebook posts were made to Facebook Pages, and through CrowdTangle we were able to find the number of Facebook users who like the page. Since Twitter doesn’t have an appropriate grouping feature, it was excluded from this portion of the analysis. Additionally, the Reddit data only had group size for individual posts to subreddits and not for comments made on the platform; since this removed a large portion of the pink slime posts, this platform was also excluded from the analysis. For Facebook Pages, the median group has 64,403 subscribers and the smallest 25% of groups have fewer than 5,895 subscribers. To understand how well posts made to these groups performed, a relative engagement metric was established. For Facebook Pages this is the number of likes a post received divided by the number of followers the page has, referred to in a platform-neutral way as Engagement Per Group Size.

After calculating the Engagement Per Group Size, the posts were broken into the four news types and compared. Per Figure 1, posts sharing pink slime received the highest engagement per group size followed by posts sharing misinformation. Real news received the least relative engagement.

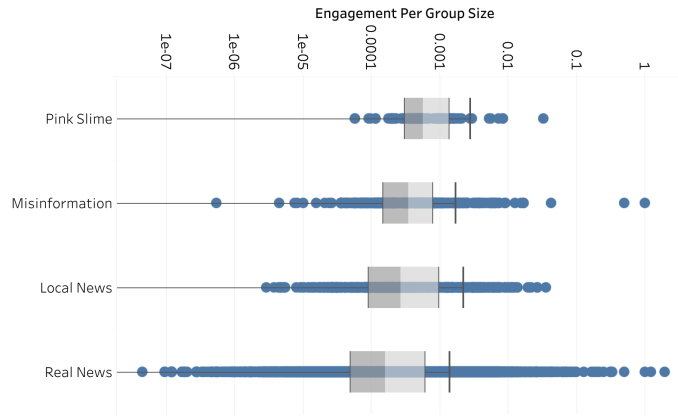


Fig.1: Logarithmic Distribution of the Boxplot Distribution of Engagement/Group Size by News Type for Facebook and Reddit

### 3.2 Network Properties

To get an understanding of the network properties of how these news sites spread, a network created exclusively of accounts sharing pink slime and another created exclusively of accounts sharing local news sites were created. In Fig 2, Facebook Pages and subreddits linking to sites shared by more than one group are shown. Interestingly, none of the groups from these platforms sharing pink slime link to other domains sharing pink slime in this dataset. For local news, 7.1% of the groups that shared a local news site shared more than one.

When these network are analyzed from Twitter users to news domains in Fig 3, there are some Twitter users sharing more than one pink slime domain (2.3%), but that is far lower than the Twitter users sharing more than one local news domain (11.2%). Upon further inspection, the individuals sharing multiple pink slime domains on Twitter are sharing domains that are targeting different regions as opposed to pink slime sites operating in the same region.

## 4 Discussion and Conclusion

The information gleaned from the relative engagement analysis highlights the importance of pink slime sites on Facebook Pages. While it makes up a small minority of posts during the Midterm election, posts made to Facebook containing these links received more likes (normalized by page size) than any other news type. These sites are designed to target hyper local regions with content related to their community; increased relative engagement indicates that the people seeing the content resonate with and appreciate it. While the posts fly under the radar of researchers looking at the top news story by absolute engagement metrics, it shows that promoting messaging at hyper local small groups are

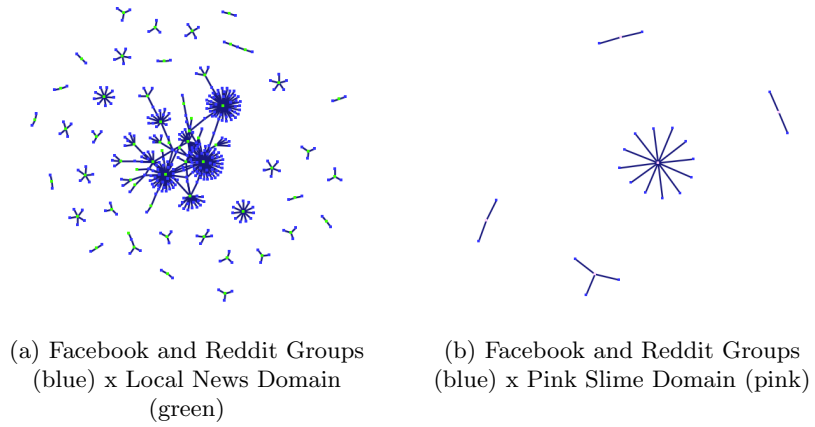


Fig. 2: Facebook and Reddit Networks of Local and Pink Slime Sites Shared More than Once

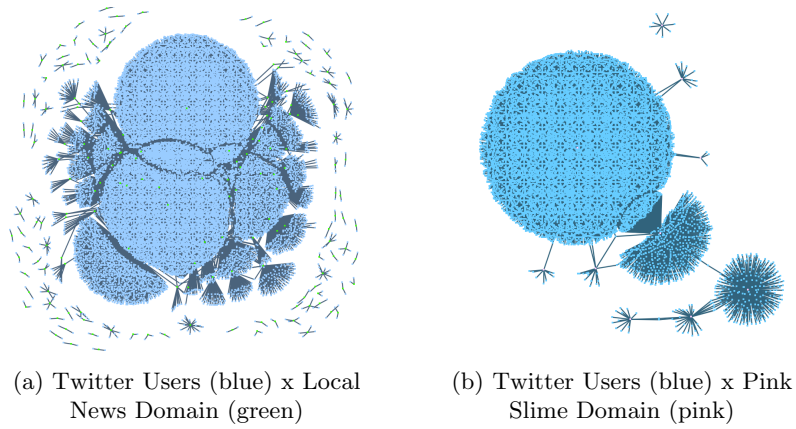


Fig. 3: Twitter Networks of Local and Pink Slime Sites

having an outsized impact. Real news sources receiving the lowest engagement per group size is another concerning finding.

With regards to the network findings, it indicates that individuals on Twitter and groups on Facebook and Reddit are consuming and sharing pink slime sites in a silo. Individuals on Twitter and groups on Facebook and Reddit are much more likely to share multiple sources of authentic local news sites than those sharing pink slime sites are to share more than one.

Future work will include an analysis of the individual Twitter accounts sharing pink slime and local news and will assess the distribution of bot account sharing the sites.

## References

1. CrowdTangle, <https://www.crowdtangle.com/>
2. Media bias/fact check news, <https://mediabiasfactcheck.com/>
3. unreliable-news/data at master · hearvox/unreliable-news, <https://github.com/hearvox/unreliable-news>
4. Who owns the media?, <https://www.freepress.net/issues/media-control/media-consolidation/who-owns-media>
5. Burton, A.G., Koehorst, D.: Research note: The spread of political misinformation on online subcultural platforms. *HKS Misinfo Rev* **1**(6), 10–37016 (2020)
6. Cinelli, M., Quattrocioni, W., Galeazzi, A., Valensise, C.M., Brugnoti, E., Schmidt, A.L., Zola, P., Zollo, F., Scala, A.: The covid-19 social media infodemic. *Scientific Reports* **10**(1) (2020). <https://doi.org/10.1038/s41598-020-73510-5>
7. Edelson, L., Nguyen, M.K., Goldstein, I., Goga, O., McCoy, D., Lauinger, T.: Understanding engagement with u.s. (mis)information news sources on facebook. *Proceedings of the 21st ACM Internet Measurement Conference* (2021). <https://doi.org/10.1145/3487552.3487859>
8. Guess, A., Nagler, J., Tucker, J.: Less than you think: Prevalence and predictors of fake news dissemination on facebook. *Science Advances* **5**(1) (2019). <https://doi.org/10.1126/sciadv.aau4586>
9. Clemm von Hohenberg, B., M.T.E.C.A.W.M.: A list of over 5000 us news domains and their social media accounts., <https://doi.org/10.5281/zenodo.7651047>
10. Hopp, T., Ferrucci, P., Vargo, C.J.: Why do people share ideologically extreme, false, and misleading content on social media? a self-report and trace data-based analysis of countermedia content dissemination on facebook and twitter. *Human Communication Research* **46**(4), 357–384 (2020). <https://doi.org/10.1093/hcr/hqz022>
11. JOURNALISM, T.C.F.D.: Domains as of august 3, 2020, <https://datawrapper.dwcdn.net/TqILa/2/>
12. Littman, J., Wrubel, L., Kerchner, D., Bromberg Gaber, Y.: News outlet tweet ids. <https://doi.org/10.7910/DVN/2FIFLH>, <https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/2FIFLH>
13. Pushshift: Pushshift/api: Pushshift api, <https://github.com/pushshift/api>