# Investigating Political Misinformation on Reddit

GitHub Repo: https://github.com/TreadyUCSD/DSC180B_Group5


**Aditi Shrivastava**
UC San Diego Address
Email adshriva@ucsd.edu

**Tyler Ready**
UC San Diego Address
Email tready@ucsd.edu



## Abstract

In this paper, we collect posts from 28 political subreddits over a two year period and analyze the spread of misinformation in outgoing links. We categorize misinformation using a compiled list of untrustworthy domains, and use the resulting data to create networks of users and subreddits showing the spread of misinformation. Additionally, we utilize natural language processing techniques to analyze the sentiment of comments on posts with respect to whether the post contained misinformation. We compare our findings to a paper analyzing the spread of misinformation through Twitter users.


## 1 Introduction

We investigate the dissemination and characterization of political misinformation spread on the online discussion site Reddit, looking specifically at a large subset of randomly selected American political subreddits (using the Reddit API). In the political sphere, the problem of online misinformation is one that has not only majorly dictated the public sentiment of certain political figures and events, but has also sowed a general distrust of political authority in most of the American public, leading to changes in outcomes of major elections and more. Reddit, which is one of the largest American social platforms, has had a number of known incidents regarding misinformation, including the banning of numerous controversial subreddits. While suspected instances of misinformation in these cases are typically identified manually by thorough investigation into not only the contents of the posts, but also the source, author, popularity, and grammatical context, we are limited in our ability to do so given the breadth of information present on Reddit. As such, we explore a more automated approach to distinguishing true information from false information using outgoing URLs, building off of methods used in the Twitter-based exploration of misinformation in 'Anatomy of an online misinformation network' by Shao et al [1]. By taking advantage of Reddit's hierarchical structure of subreddits, moderators, users, posts, and comments, we replicate the investigation on the spread of misinformation on Twitter with a similar investigation on the spread of political misinformation on Reddit in the years 2020 and 2021.

We use a sample of approximately 30 political subreddits to perform our analysis on how misinformation spreads both within and between subreddits. This includes identifying misinformation by looking at the domain of shared links, as well as identifying the main spreaders of misinformation by creating social network graphs and tracking influential instances of misinformation through the use of k-core decomposition. Also, we use the top-level comments from each of the gathered posts to further characterize and attempt to isolate the textual metadata present in posts that contain misinformation, as opposed to those that don't. Given the structure of Reddit, we will compare how misinformation spreads among different subreddits, as well as between subreddits and across demographics, which we hope will illuminate not only the extent to which American political boards/forums/subreddits on Reddit share political misinformation, but also how that misinformation commonly presents itself.

## 2 Data

Through the use of the Reddit API, which is well-documented and commonly used in similar studies, we collect a robust dataset with which to do our analyses. Our datasets are primarily gathered through repeated calls to the Pushshift API along with the python Reddit API wrapper (PRAW), which allow us to store metadata on approximately 2,000,000 Reddit posts from our list of political subreddits. We curated this list of about 30 subreddits in hopes of creating a sample that accurately represents all the differing communities of political information shared on Reddit, ranging from 'r/LibertarianSocialism' to 'r/Anarchism' to 'r/Communist'. We then extracted all posts containing URLs over a two year span from January 2020 to December 2021 in JSONL format, separated by subreddit. These posts contain specific information on the post title, ID, upvote ratio, score, author, and several other metrics that once again highlight the difference in structure between Twitter and Reddit. In order to determine if a given Reddit post is an instance of misinformation, we created a database of 800+ unreliable domains sourced from Iffy.com [2] that are known to regularly publish false information, as verified by low MBFC (Media Bias/Fact Check) and FR (Factual Reporting) scores. We then use this list of highly non-credible sources to cross-check and determine the presence of a matching URL that commonly spreads misinformation in a given Reddit post, using this as our final metric of information categorization.

After doing so, we use the gathered Reddit posts to extract all associated comments using the Python Reddit API wrapper, which produces a CommentForrest object allowing us to perform a breadth-first search to access all top-level comments and their corresponding replies. As we are only interested in the most popular top-level (can not be a reply) comments, we then use the JSONL post files to cross check the Reddit post link and download up to 1000 top-level comments in CSV format. We note that the comments are parsed as messy and unformatted strings, with several non-English and non-unicode characters that will require processing before use. The entirety of our data is stored on the DSMLP server, and is under 1GB in total. With our dataset now complete, we observe some notable discrepancies between subreddits-- namely that the number of posts ranges from tens to hundreds of thousands (r/republicanism vs r/politics), that some subs appear to have very few unique users (indicating the possibility of an echo chamber effect, i.e. r/uspolitics), and that certain subs generally have much longer comments than others (r/AnarchoPacifism vs r/alltheleft). Below are the complete summary statistics on the extracted political Reddit posts and comments across all gathered subreddits.

| Subreddit | Number of Posts | % of Unique Users | Number of Users w/ >10 Posts | Average Post Sore |
|---|---|---|---|---|
| r/alltheleft | 12580 | 16.4% | 152 | 8.6 |
| r/AmericanPolitics | 13659 | 9.9% | 118 | 1.3 |
| r/Anarchism | 19216 | 34.3% | 273 | 10 |
| r/anarchist | 139 | 42.4% | 2 | 1 |
| r/AnarchoPacifism | 101 | 38.6% | 3 | 1.4 |
| r/blackflag | 360 | 21.7% | 2 | 1.1 |
| r/capitalism | 4166 | 40.2% | 53 | 3.6 |
| r/communist | 342 | 26.0% | 9 | 1.1 |
| r/Conservative | 281397 | 19.0% | 2169 | 9.9 |
| r/conservatives | 58978 | 10.8% | 365 | 3.5 |
| r/conspiracy | 218914 | 24.6% | 3009 | 7.2 |
| r/democracy | 1474 | 35.0% | 25 | 1.6 |
| r/democrats | 33285 | 31.1% | 368 | 5.8 |
| r/greenparty | 3885 | 17.5% | 45 | 3.9 |
| r/Liberal | 7410 | 28.0% | 79 | 4.5 |
| r/Libertarian | 46884 | 20.8% | 590 | 5.4 |
| r/LibertarianSocialism | 1355 | 27.1% | 19 | 2 |
| r/Liberty | 925 | 22.4% | 9 | 1.2 |
| r/moderatepolitics | 7988 | 29.9% | 140 | 5.9 |
| r/neoprogs | 57 | 87.7% | 0 | 1 |
| r/politics | 579206 | 17.5% | 5854 | 65.9 |
| r/progressive | 5798 | 25.1% | 59 | 4.4 |
| r/republicanism | 46 | 67.4% | 0 | 1.2 |
| r/Republican | 54901 | 17.6% | 449 | 3.7 |
| r/republicans | 13179 | 14.0% | 132 | 1.5 |
| r/socialdemocracy | 3296 | 25.7% | 63 | 2.8 |
| r/socialism | 26380 | 35.6% | 264 | 8.8 |
| r/uspolitics | 27175 | 8.1% | 162 | 1.8 |

Figure 1, above: Number of Posts, percentage of unique users, number of users with over ten posts, and average post score, by subreddit.

| Subreddit | Number of Top-Level Comments | % of Posts with Comments | Avg Comment Length (Words) | Avg # Comments per Post |
|---|---|---|---|---|
| alltheleft | 10221 | 81.2% | 27.5 | 4.9 |
| AmericanPolitics | 10295 | 75.4% | 41.6 | 2 |
| Anarchism | 26609 | 38.5% | 27.7 | 4.6 |
| anarchist | 61 | 43.9% | 50.1 | 3 |
| AnarchoPacifism | 36 | 35.6% | 98.6 | 1.4 |
| blackflag | 86 | 23.9% | 49 | 1.2 |
| capitalism | 11664 | 28.0% | 38.2 | 3.8 |
| communist | 69 | 20.2% | 30.1 | 1.4 |
| Conservative | 43358 | 15.4% | 25.8 | 7 |
| conservatives | 76572 | 39.8% | 24.7 | 13.3 |
| conspiracy | 12692 | 5.8% | 38.9 | 7.4 |
| democracy | 891 | 60.4% | 35.7 | 1.7 |
| democrats | 9202 | 27.6% | 26.9 | 4.2 |
| greenparty | 4457 | 14.7% | 35.4 | 2.9 |
| Liberal | 34829 | 47.0% | 28.4 | 11.7 |
| Libertarian | 38942 | 83.1% | 29.2 | 6.9 |
| LibertarianSocialism | 1134 | 83.7% | 34.7 | 2.5 |
| Liberty | 567 | 61.3% | 33 | 2.9 |
| moderatepolitics | 18400 | 23.3% | 66.5 | 8.2 |
| neoprogs | 5 | 8.8% | 28.6 | 1.7 |
| politics | 94699 | 16.3% | 35.6 | 15.9 |
| progressive | 4425 | 76.3% | 35.1 | 3.7 |
| republicanism | 34 | 73.9% | 20.8 | 1.4 |
| Republican | 26750 | 48.7% | 37.7 | 3.7 |
| republicans | 5046 | 38.3% | 70.9 | 1.4 |
| socialdemocracy | 10552 | 32.1% | 37.9 | 5 |
| socialism | 6969 | 26.4% | 33.8 | 3.6 |
| uspolitics | 11890 | 43.8% | 32.3 | 2.5 |

Figure 2, above: Number of top-level comments, percentage of posts with comments, average comment length in words, and average number of comments per post, by subreddit.

## 3 Methods and Analysis

### 3.1 Social Network Graphs

Utilizing the list of non-credible domains, we cross reference every outgoing link in our dataset one subreddit at a time to gather data on the proportion of links likely containing misinformation to the total number of links in the subreddit. We count the total number of posts for each subreddit over the two year span that contains urls to an outgoing site, along with the number of those urls that match a domain on our list. These data points are compiled on a CSV file with each row containing the subreddit name, total posts, and total posts with misinformation.

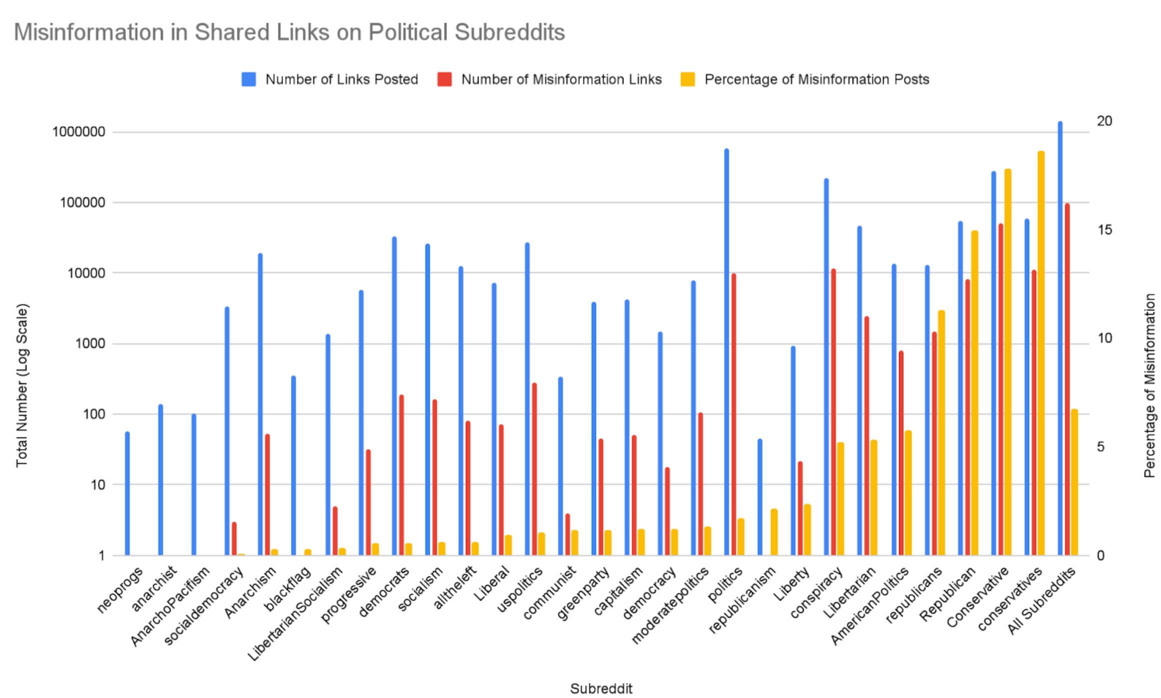


Figure 3, above: Number of Total Posts, Number of Misinformation Posts, and Percentage of Misinformation in Total Posts for each subreddit, sorted by Percentage of Misinformation, with aggregate data for all subreddits to the far left.

As can be seen from the graph, the average rate of misinformation in posts among all of the subreddits is around 7%. There is no correlation between total number of posts and percentage of misinformation among posts, but there is a correlation between percentage of misinformation and political leaning, with the four subreddits with above average misinformation all being right-leaning political subreddits. The subreddit with the most posts, r/politics, has a rate of misinformation just under 2%. Interestingly, all but one of the subreddits, r/capitalism, with a lower rate of misinformation than r/politics is either left-leaning or neutral, while 7 of the 9 subreddits with a higher rate of misinformation than r/politics are right-leaning, with the other two being neutral. This suggests that misinformation is much more likely to be spread in right-leaning subreddits, despite the fact that there are many more left-leaning subreddits on the site.

We also construct a graph between users and subreddits to visualize how users posted across subreddits and the amount of misinformation spread. Users and subreddits are both represented by a node, with an edge from a user to a subreddit if the user has posted in that subreddit. The size of the nodes for users is proportional to their total number of posts, and the color of the user node is blue if they have not posted misinformation, and orange if they have. The edges are also colored according to misinformation, with blue edges indicating a user has not posted misinformation to that subreddit and orange edges indicating that they have. For readability, the nodes of subreddits are a constant size and are colored a deep red color, with labels over each subreddit in black. Due to the large number of users, only users with more than 100 posts are included on the graph to avoid unnecessary clutter. Additionally, we found that 42,362 of the posts had their poster delete their account some time before we collected the data, so we removed the deleted users from the graph so that they would not be misrepresented as one user with many posts in the visualization. Of these posts, 1,686 contain misinformation, or 3.97% of all the posts with deleted users. After construction, we conduct a k-core decomposition on the graph and extract the k-core from the result, finding the k-core at k=15. This number is representative of the maximum number of subreddits that any user posted to in the two year span.

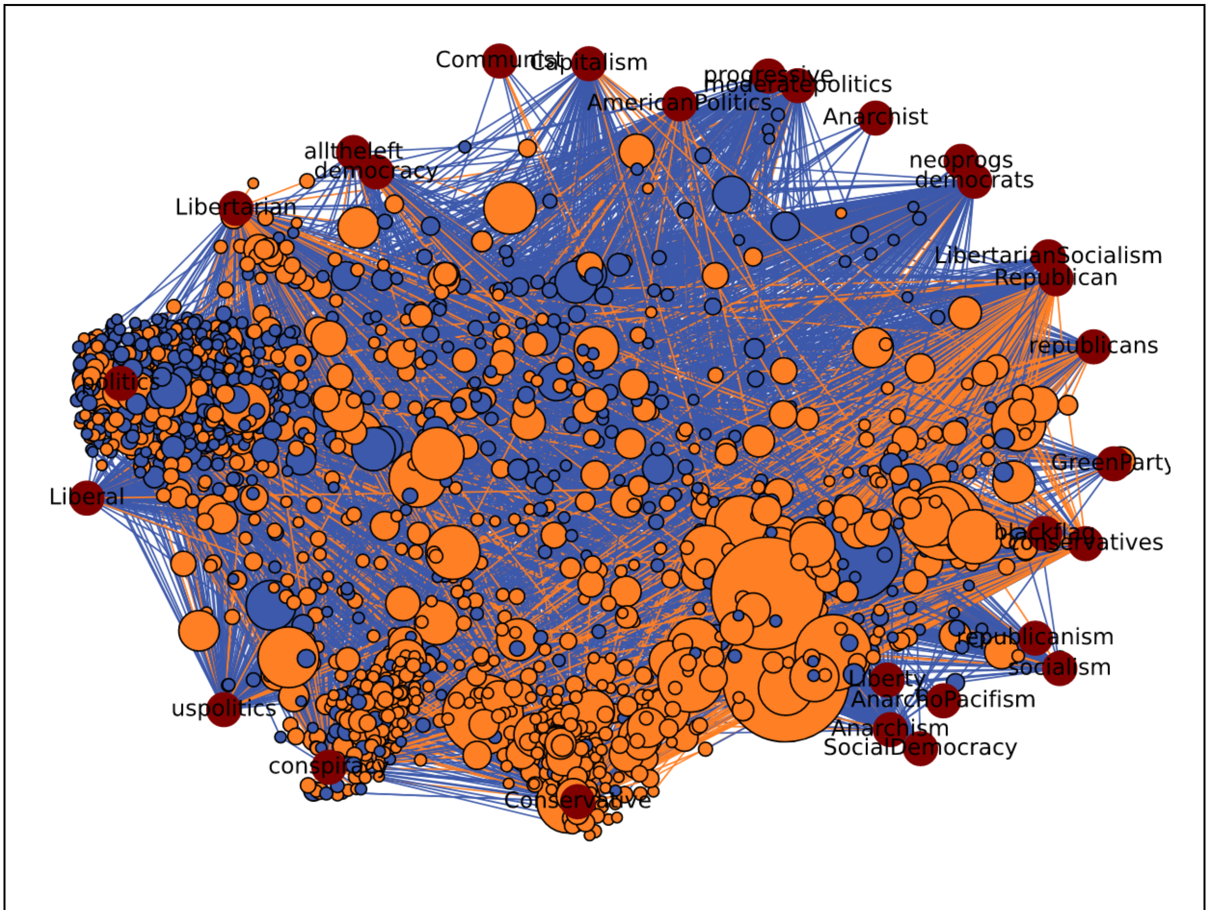


Figure 4, above: Graph of users and subreddits

This visualization correlates with our findings from the examination of misinformation in subreddits, with users posting misinformation clustered around right leaning subreddits. As a user is shown as orange if they have posted any misinformation, it is important to note that the size of the node representing a user does not correlate with the amount of misinformation spread, but instead with the total number of posts. Even if a user has only posted one link containing misinformation, they still appear orange in the visualization. Only 32% of users in the dataset shared misinformation, with an average percentage of misinformation shared by users who posted any misinformation of 13.4%. On average, users posted to 2.3 subreddits, with 24 being the most subreddits posted to by any one user. The graph shows that most of the users who posted a large amount had also posted misinformation at some point. This is backed up by an analysis of the users, as users who did not share misinformation posted an average of 45 times while users who did share misinformation posted an average of 136 times, with the maximum posts by a user not sharing misinformation at 8,535 and the maximum posts by a user sharing misinformation at 15,443. This suggests that users who post more are more likely to post misinformation, however it is unclear if this is simply a result of posting more or if posters who post more are more likely to read sources of misinformation.

Additionally, we create a graph of only users based on what subreddits users have in common with each other to investigate how users might interact with each other. First we remove any users from the data who had only posted in a single subreddit, and then we remove any users who posted fewer than ten times to reduce the size of the resulting graph so that it is easier to conduct analysis. This leaves a set of users who frequently posted on more than one subreddit over the two year span. Since the comment data gathered does not include who posted the comments, we can not construct this graph based on direct interactions between users, so instead we connect two users with an edge if they have posted to at least two of the same subreddits. Since users who post on a subreddit are often following that subreddit, it is more likely that users would have interacted with each other if they had posted to the same subreddits and could have seen each other's posts. The resulting graph has 7,599 users with 10,223,225 connections between them, and represents an approximation of the network of information spread between all users in the dataset. We utilize k-core analysis to decompose the graph into its core, and from that we are able to extract the core at k=2,911. This number is representative of the number of users who posted in the same subreddits as the users who posted in the most subreddits over the two year period. Although having posted in the same subreddit does not guarantee that a given user has seen another user's post, having subreddits in common leads to an increased likelihood that any two users will interact. Therefore, the k-core value can also be seen as an upper approximation of the influence of the most frequent posters by the number of fellow posters reached.

### 3.2 Comment Exploration

To accompany the aforementioned multi-level network analysis, we perform additional analyses on the associated top-level comments in order to further investigate the spread of political misinformation by characterizing a difference in the textual metadata of posts that contain true information and posts that contain misinformation. As such, we construct two descriptive models that allow us to compare across two categories (true and false information) the polarity and subjectivity, identify any highly associated words or phrases, and ultimately describe any significant discrepancies in the comments not only between the two target categories, but also across all subreddits. Before proceeding with the analysis, we clean and process the textual comment data to allow for uniform and unbiased NLP-- specifically, we use the NLTK and regex tokenize and replace functions (respectively) to split each comment into clean, lower-case, punctuation-free, and tokenized sentences. We then create a custom list of stop words building off of the vocabulary suggestions provided in the paper 'Raising the Flag: Monitoring User Perceived Disinformation on Reddit' by Achimescu et al. [3], which allows us to filter out superfluous or unneeded words/phrases such as 'is', 'he's', or 'bot'. We then loop through each subreddit and merge the CSV comment data with the JSONL post data to keep only the posts with at least one top-level comment (keeping each subreddit separate), and apply the same aforementioned technique of identifying misinformation using a predetermined list of unreliable domains. Finally, given that the size of our data has more than quadrupled after storing all textual comment data, we convert the data into lazy out-of-core dataframes using Vaex, which allows us to perform very large computational and visualization tasks in an on-the-fly, memory efficient manner.

The first visual below, which compares comment sentiment polarity and subjectivity instances of real vs. fake information, is constructed using the TextBlob module for NLP, which uses the NLTK package to achieve sentiment analysis without the use of a lexicon or dictionary based approach. As such, we use semantic labels beyond just unicode characters to accurately extract the polarity and subjectivity values of each comment, given that the comments data contains many non-lexical instances of tone indicators such as exclamation marks, emoticons, emojis, and more. Also, we use intensity as a potential indicator of subjectivity, noting that the presence of certain words or phrases (such as 'very'), expletives, and strong punctuation may indicate highly subjective comments. Here, polarity is scored on a [-1, 1] scale (with -1 indicating a highly negative sentiment and 1 indicating a highly positive sentiment), and subjectivity on a [0,1] scale (with 0 indicating a highly objective/neutral sentiment and 1 indicating a highly subjective sentiment). Using the Holoviews package to implement the visualization, we create a scatterplot depicting the polarity and subjectivity values for each comment, categorized by type of information. We also append two histograms on either axis to more accurately describe the numerical distribution of both metrics across each information type. The visual also holds interactive components, allowing the user to toggle between the descriptive tooltips, which contains the polarity value, subjectivity value, source subreddit, and type of misinformation, on each point (comment). Here, given that we generally consider the comments of posts that instances of misinformation to be more volatile and negative than those that don't, we expect to observe a right skew in polarity and a left skew in subjectivity.

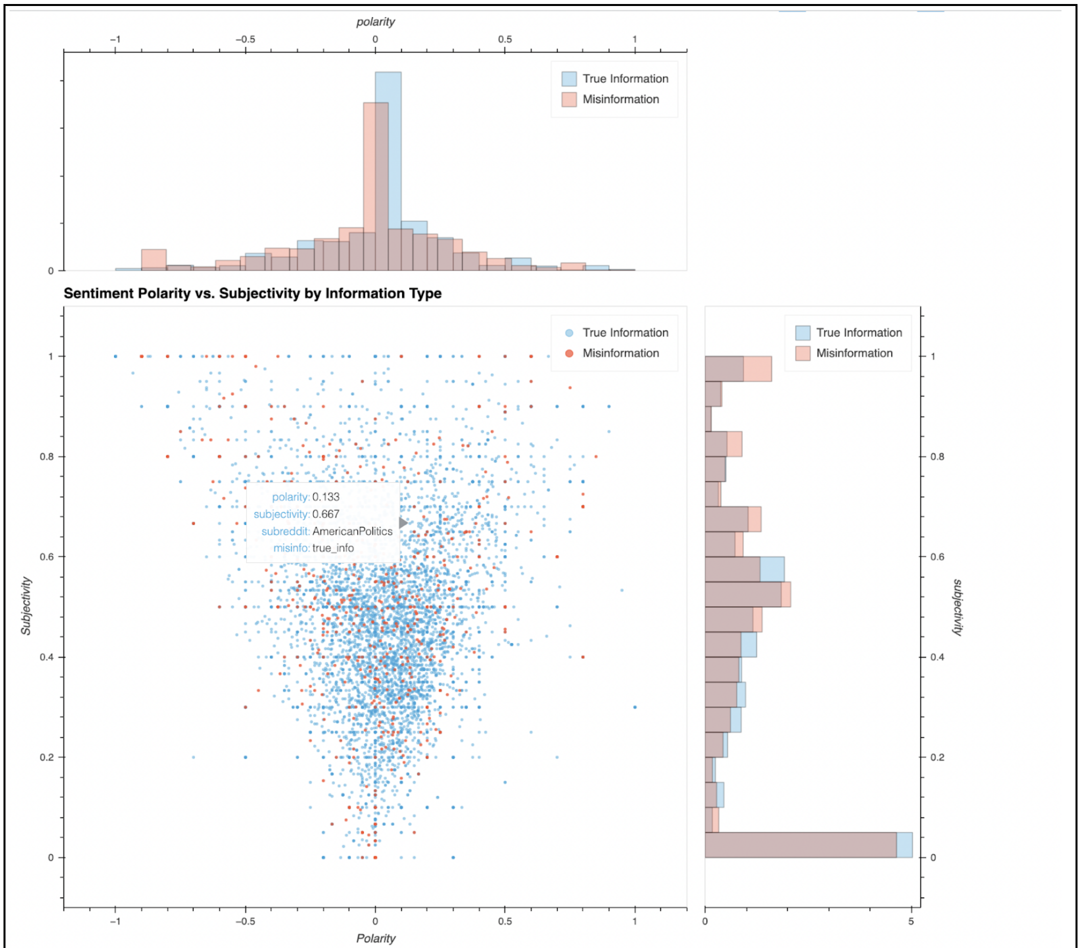


Figure 5, above: Screenshot depicting the relationship and distribution of top-level comment polarity and subjectivity by type of misinformation, across all gathered political subreddits.

In the plot above, we see that both the comments for posts that contain misinformation and those that don't follow nearly the same spread and shape in regards to polarity and subjectivity. We see that both categories (blue and red) tend to remain close to 0 in polarity, but disperse in value as subjectivity rises. This is expected behavior, given that we would intuitively not expect a textual comment to be either highly positive or negative in sentiment without also being highly subjective. Also, we find (as shown in both histograms) that the two corpora follow nearly the same distribution in both polarity (unimodal, centered around approximately 0, with some prominent outliers around -1) and subjectivity (peaks around the extrema and the center, with a tendency towards 1). These findings are not completely consistent with what we expected to find, but we do note that polarity in misinformation comments is slightly right skewed in comparison to true information comments, indicating that comments on posts with false information tend to be more negative in sentiment. Also, we see that misinformation posts' comments appear to be slightly more subjective than those of posts. In other words, users who comment on posts that include links to fake news tend to do so in a more negative and subjective tone-- this may mean not only that the comments are opposed to or resistant to the content of the post, but also that they're expressing a strong opinionated sentiment towards the post.

The second plot we generate allows us to visually compare the corpora of comments from posts that contain misinformation and those that don't using a scatter plot made up of textual words and phrases. We use the ScatterText package to create a visualization of highest-ranking words/phrases from each category, allowing us to display hundreds of category-representative points. The tool takes in as input the same merged post and comment dataset, and uses a PyTextRank algorithm to rank words by their score and plot according to the category they most align with. Here, the word score is made up of several measures, namely precision, recall, non-redundancy, and characteristicness. By implementing all of these metrics in the descriptive model, we ensure that the statistically most meaningful words and phrases are extracted for each category, and that discriminative power (regardless of frequency), word frequency, co-occurrence, and association (respectively) are used to determine category. The resulting visual then takes advantage of relative position, color, and interactive tooltips to display in full the most characteristic words present in comments associated with both true and false information. Also, we include a search bar that allows the user to type in any word or phrases and click to view not only which category it belongs to, but also how often it appears in either category, what subreddit it is sourced from, and the original comment itself. [4] Below, figure 6, is a table detailing the top 15 most highly scored words for each category, and figure 7, a screenshot of the resulting visual.

| True Information | Misinformation |
|---|---|
| thank you | witness information |
| images | fuck china |
| fuck trump | new law |
| domain require | private gun |
| patient review | blacklist conservatives |
| review thank | sexes |
| require moderation | serious problem |
| moderation please | law changed |
| please message | austin |
| make mistakes | witness identity |
| war iran | given defense |
| sometimes make | savaged |
| im sometimes | sex gender |
| mistakes questions | voter id |
| however removed | inspector general |

Figure 6, above: Top 10 most highly associated words/phrases for true information and misinformation

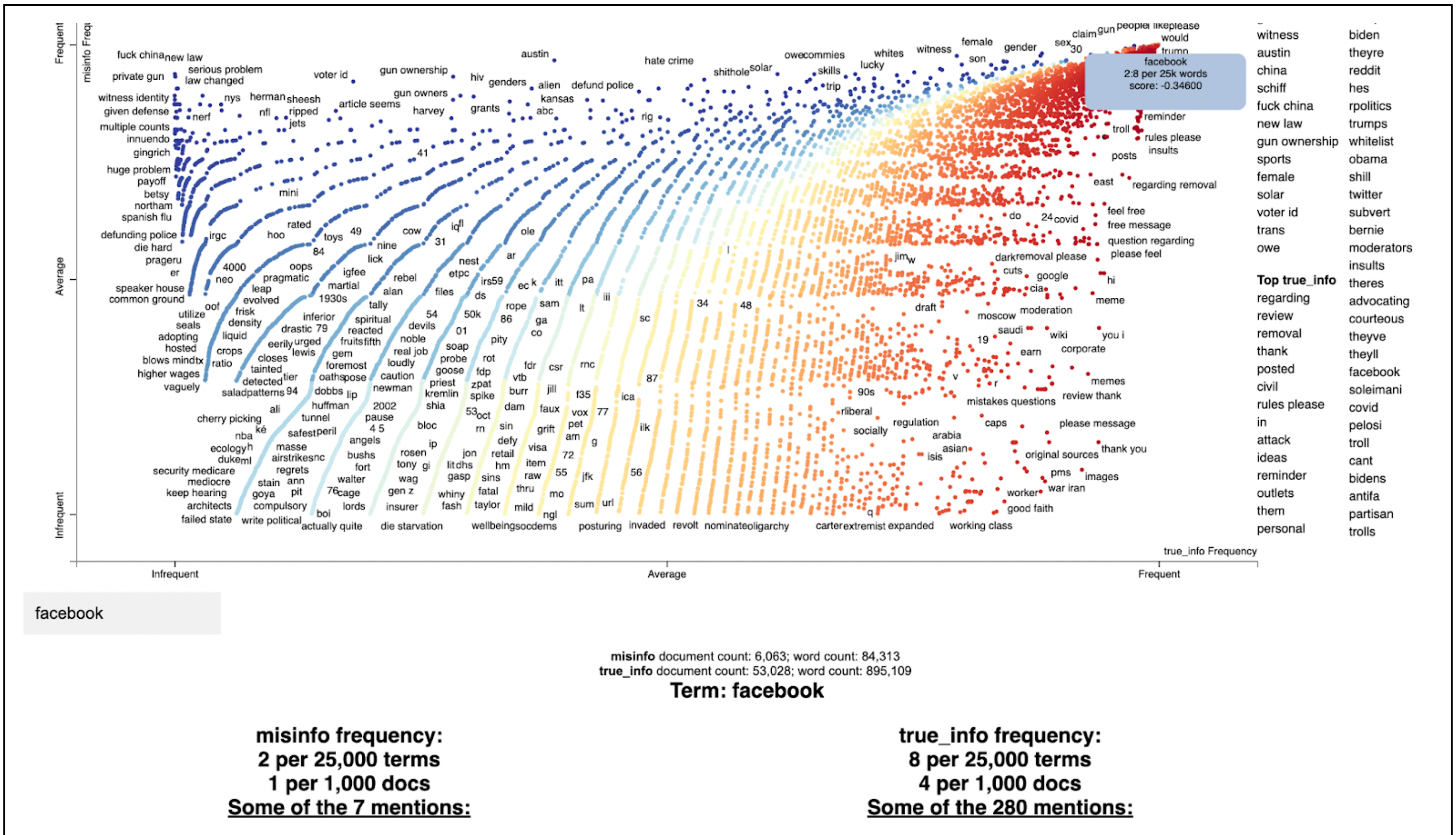


Figure 7, above: Resulting scattertext plot, depicting top comment words for both misinformation and true information, highly characteristic words, term frequency, and search term ‘Facebook’ highlighted.

The visual displayed above is a still capture of the generated scattertext plot, and shows the difference in word usage in comments from posts from both categories, misinformation (y-axis) and true information (x-axis). Using the table above to facilitate legibility, we see that the most associated words for each category are ‘thank you’, ‘images’, and ‘fuck trump’ for true information, and ‘witness information’, ‘fuck china’, and ‘new law’ for misinformation. Notably, while the extracted words and phrases for the former are rather general and non-correlated (with the exception of ‘trump’), the extracted words for misinformation tend to relate to rather divided and often-debated topics in political communities. This indicates that when misinformation is being shared on the gathered political forums, the content likely usually pertains to topics such as foreign countries, sexuality and gender, policy law, and other similarly divisive topics (usually in the right-leaning communities, as noted earlier). Upon further exploration, we find that there are some terms and phrases that tend to be characteristic of only one category and have nearly no presence in the other category, such as ‘Soleimani’ (misinformation) and ‘Bernie’ (true information). Unsurprisingly, there appears to be a stark difference in the type of content commonly spread in both types of political Reddit posts, and this visual highlights this by providing both document and word counts for comparison.

## 4 Results

Our social network analysis of both user and subreddit interactions reveals to us that approximately 7% of all information posted is misinformation, while some subreddits (typically right-leaning) and users are more prone to containing misinformation than others. We also construct a graph of only user to user interactions to track the spread of misinformation among users, noting that users who post frequently in the same forums are more likely to interact with each other. After performing k-core decomposition on both graphs, we generate k values of 15 and 2,911, respectively. While these values are not directly comparable to those found in the k-core decomposition of Twitter tweets, we do see that our k values are considerably larger than those of Twitter, which is to be expected given the structure of Reddit. The k-value of 15 on the user subreddit interaction graph indicates that the network has a core of size 15 subreddits-- or that the most densely populated core of political misinformation being spread exists within a core of 15 subreddits. This is a much larger community than the ~7-core found in Twitter (given that each subreddit can have thousands of users), and is highly indicative of an echo chamber effect among users of those 15 subreddits. And when comparing the user-user Reddit interactions to the user-user Twitter interactions, we find that misinformation exists mainly within tightly knit communities of nearly 3,000 users in the gathered subreddits. By finding these cores, we've shown that the same users and subreddits appear to spread misinformation, and through textual analysis we report similar findings for the comment' content.

To further explore the difference in comment data for posts that do and don't contain misinformation, we gain a more thorough understanding of what may textually characterize a given post as misinformation. By creating a scatterplot of the relationship sentiment polarity and subjectivity among gathered comments, we see that those of posts with misinformation tend to be slightly more negative in sentiment and potentially more subjective in nature (indicating possible dissent)-- in other words, without yet analyzing individual features in the comments, we find that a given post is more likely to contain misinformation if the comments are low in polarity and high in subjectivity. Upon employing NLP textrank and f-score techniques to gain a better understanding of topics discussed, we create a scattertext plot to visualize the most highly associated words in both categories (misinformation and true information). Here, we find that while the content of comments of posts with true information is rather random, the content of posts with misinformation tends to focus on a subset of often divisive political topics, especially those pertaining to right-leaning communities. In conjunction with the network analysis, these findings confirm our belief that information shared in the cores of misinformed posts tend to follow an echo chamber effect, wherein a nearly homogenous set of users, subreddits, and topics are being discussed.

In conclusion, we present an investigation into both the structure and content of misinformation being shared in political forums on the social media site Reddit in the years 2020 and 2021. By gathering a large sample of data from nearly 30 political subreddits, we utilize network graph and NLP analysis to not only track the spread of misinformation between and among subreddits and users, but also visually compare the sentiment and content of the comments in posts that contain misinformation and those that don't. Our findings reveal potential indicators of misinformation through both user and comment statistics, and we hope to continue this analysis in the future by potentially altering the way we categorize misinformation, and expanding our search to other media and social media formats.

## 5 Sources

[1] Shao C, Hui P-M, Wang L, Jiang X, Flammini A, Menczer F, et al. (2018) Anatomy of an online misinformation network. PLoS ONE 13(4): e0196087. https://doi.org/10.1371/journal.pone.0196087

[2] https://iffy.news/iffy-plus/

[3] Achimescu, V.; Chachev, P.D. Raising the Flag: Monitoring User Perceived Disinformation on Reddit. Information 2021, 12, 4. https://dx.doi.org/10.3390/info12010004

[4] arXiv:1703.00565 **[cs.CL]**

## 6 Appendix

Project Proposal:

https://docs.google.com/document/d/1vn_KJA59UAi-uJyJ1waaTc4aNAvrsPYiA4JGssosMzU/edit?usp=sharing