

# Gab: The Alt-Right Social Media Platform

Yuchen Zhou<sup>1</sup>, Mark Dredze<sup>1[0000–0002–0422–2474]</sup>, David A. Broniatowski<sup>2</sup>,  
William D. Adler<sup>3</sup>

<sup>1</sup> Center for Language and Speech Processing  
Johns Hopkins University, Baltimore, MD 21218 yzhou111@jhu.edu,  
mdredze@cs.jhu.edu

<sup>2</sup> Department of Engineering Management and Systems Engineering  
The George Washington University, Washington, DC 20052 broniatowski@gwu.edu

<sup>3</sup> Department of Political Science, Northeastern Illinois University, Chicago, IL 60625  
w-adler@neiu.edu

**Abstract.** This study proposes the use of Gab as a vehicle for political science research regarding modern American politics and the Alt-Right population. We collect several million Gab messages posted on Gab website from August 2016 to February 2018. We conduct a preliminary analysis of Gab platform related to site use, growth and topics, which shows that Gab is a reasonable resource for Alt-Right study.

**Keywords:** Alt-Right · Gab · politics · topic model

## 1 Introduction

The 2016 election and its aftermath has provided an unusual glimpse into the so-called alt-right, a new movement of the far-right in the United States with extreme positions on issues such as race and religion. Once a fringe movement, it became elevated as a result of its support for President Donald Trump. The movement continued to gain prominence with its ties to Steve Bannon, Trump’s campaign manager and later a top White House official, culminating in the Charlottesville Unite the Right rally in August 2017.

Scholars have conducted some work on the rise of alt-right on social media, specifically on Facebook and Twitter [8] or by using survey data. However, it is difficult to identify such a narrowly specific population on those large sites. We propose here to investigate the alt-right by examining Gab (<http://gab.ai>), a relatively new social media outlet that was founded in August 2016 with an explicit agenda of allowing unencumbered free speech to all views, no matter how extreme. Although the sites founders disavow any affiliation with the alt-right, provocateurs such as Milo Yiannopolous and Richard Spencer are some of the most popular users of the site. As a site that caters to this subpopulation, a close analysis can tell us much about the views of this group. This paper represents a preliminary study of the sites users, posts, primary topics of discussion, and its growth. This work also can provide scholars a way to test theories that some are concerned may not generalize to politically conservative populations [5], even

though this clearly only represents a partial subset of the larger American right wing.

## 2 Data Collection

We obtained a large collection of Gab messages by crawling the Gab platform. Gab uses a sequential message id system, in which the first messages on the platform began with an id of 1 and each subsequent message incremented the id by 1. Therefore, we query the website for the message with each message id starting with 1 until our data collection ended (February 23, 2018). This process resulted in 16,771,455 messages posted from August 10, 2016 to February 23, 2018. Every message contains several metadata fields, including message id, created time, repost status (i.e. retweet), the message text, and some author information. This dataset contains 146,188 unique author accounts.

Many ids did not have a message, which may be because the messages were deleted, or some other unknown reason. We found a large number of missing messages, with peaks in the number of missing messages around December 2016, August 2017 and September 2017. This may suggest that Gab may have removed certain accounts or specific messages, an interesting observation given that Gab purports to be an open, free speech platform (although they do have standards for when messages will be deleted.) We do not currently have sufficient information to study this behavior, but plan to examine it in future work.

Recently published work used a similar procedure to crawl the platform [9]. Since our data collection methods differ, we refer the reader to their publication for a comparison of site statistics.

## 3 Analysis

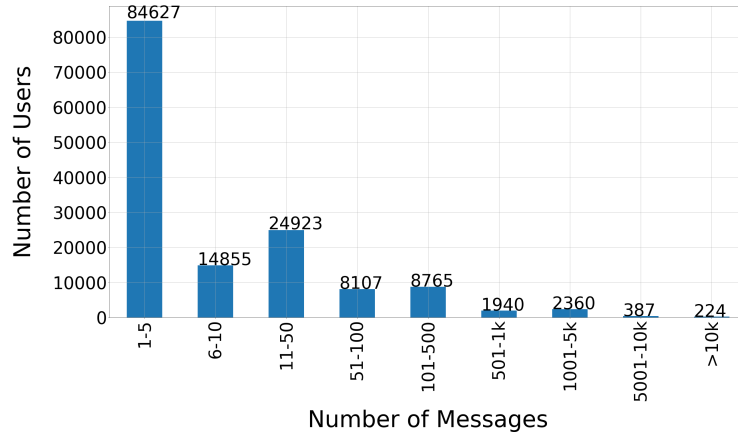
### 3.1 User Statistics

We begin by examining user statistics of the platform. We compute the number of messages authored by each user. Since users are derived from authored messages, this excludes users with an account on the platform but who have not authored a message. Therefore, every user in this analysis has authored at least one message that was available when we downloaded the data.

Figure 1 shows a histogram of the number of users who authored each range of messages. Most users (56.67%) had only a few messages (less than 5), and only a tiny number of users had a very large number of messages (2% of users with more than 1000 messages.) Most gab users do not post regularly, but there are very prolific users. It may be that a few users lead the discussion in Gab, with most users consuming or reacting to this content.

### 3.2 User Interactions

We next consider how users interact on the platform. We count the number of messages with @ mentions (references to other users), the number of messages



**Fig. 1.** The number of users who have posted different numbers of messages.

that repost/share content (i.e. retweet) and the number of messages sharing links. For comparison, we compute these same statistics for a sample of Twitter data. We used all tweets from the 1% sample Twitter streaming API for 17 days in December 2017.

Table 1 shows the results of this comparison. Interestingly, while Twitter and Gab users post links and report content with similar frequency, they are much less likely to mention other users in their post. This may suggest fewer user interactions on the platform as compared to Twitter.

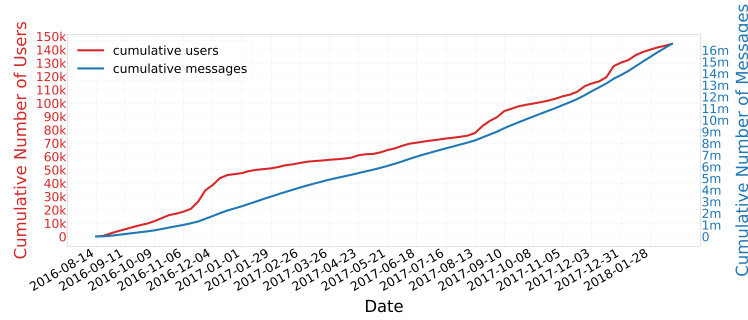
**Table 1.** Percentage of Gab and Twitter messages which contain @ mentions, shared links and which are reposts.

Platform	@ mentions	Shares	Links
Gab	15.67%	42.16%	36.15%
Twitter	68.59%	50.03%	43.74%

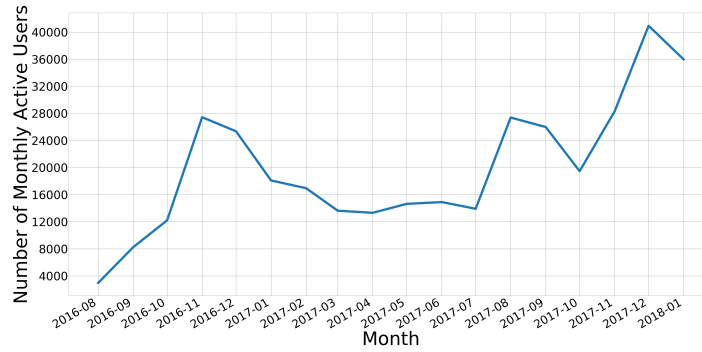
### 3.3 Site Growth

How has the platform grown from its founding? We measure user growth by recording the time at which a user first posts to the platform, and counting how many users have joined the platform before a given date. We also count the number of observed messages over time. Figure 2 plots both of these trends over time.

Most weeks have a consistent number of new users, around 1000. However, we observe two time periods when the number of new users showed a large increase: November 2016 (average 5065 users per week) and August to September



**Fig. 2.** The total number of users and messages over time.



**Fig. 3.** Number of active users each month

2017 (average 3346 users per week). These time periods coincide with major events occurring in the United States. The US presidential election happened in November 2016, and the election of the Republican candidate Donald Trump likely fueled interest in the site. A terrorist attack in Barcelona as well as the “Unite the Right” rally in Charlottesville, Virginia both occurred in August 2017. Hurricane Irma in Florida was in September 2017. These events may have attracted users to Gab and increased its exposure, as these events were important to the Alt-Right community. The analysis of [9] showed a similar trend in the number of new accounts.

We next consider the number of active users over time. Do users sign up en masse for the site, and then stop using it? Or does the number of actively engaged users continue to grow? We measure this by recording the number of monthly active users: the number of users who post in a given month. Figure 3 shows the number of monthly active users, and the trend continues to grow upward trend, although it has spikes around the time of new sign-up surges.

### 3.4 Topics

**Topic Model** To study what is discussed on the Gab platform, we use a topic model to extract prominent topics among Gab messages. A topic model is a type of unsupervised machine learning algorithm, which clusters words into likely topics based upon word and document co-occurrence statistics. This enables us to efficiently summarize tens of millions of Gab messages.

Each Gab message is treated as one document for the purposes of training the LDA model. We keep the 20,000 most frequent types (words) excluding those that occur in more than 90% of the documents. We exclude documents with less than 3 tokens.

We use the Sprite topic model [7, 2] and train with 50 topics, yielding a probability distribution over the vocabulary words for each topic. We assigned each document to its modal topic. We reviewed the topic word distributions and assigned a label to each topic.

**Topic Analysis** Since themes repeat across the 50 topics, we group together similar topics to produce 33 meaningful topic groups. Table 2 lists the topics and their prevalence. Topics related to politics comprised 56.4% of all the Gab messages. It is clear the platform is heavily focused on political content.

Finally, we identify topics that were prevalent during the two periods when sign ups surged on the platform. In November 2016, the top topics are “2016 Election and contemporary political debates”, “MAGA”, “sex scandals” and “Christmas and New Year”. This supports the observation that sign ups were driven by the US presidential election results. During the August and September 2017 surge, top topics included “Right and left”, “terrorism”, “state related issues” (related to the storms in Texas and Florida) and “Ideology, religion, and race”. These correlate with major events happening at that time, including the protests in Charlottesville and the hurricanes.

## 4 Future Work

Gab represents a unique platform that scholars can use to study issues relevant to a variety of literatures. Here, we suggest several directions of inquiry, but these should not be seen as limiting other potential avenues.

First, Gab is a source of information about a specific sub-population within the American right wing movement that is generally difficult to study directly. This could allow us to test academic theories that some have been concerned will not generalize to politically conservative populations [4].

There has also been substantial research on the spread of conspiracy theories in recent years [6], particularly on the problem of so-called fake news and how it spreads [1]. Gab’s ties to the alt-right, a population that may be prone to sharing such theories, make it a fruitful area of investigation to see how these stories multiply along with a potential venue through which they originate.

**Table 2.** Topics discussed on Gab and their prevalence, sorted into two groups: politics related and other.

Politics Related		Other	
Ideology, religion and race	10.23%	Men and women	6.26%
Trump, Clinton and conspiracies	5.10%	Social media	4.83%
Right and left	5.08%	Profanity	4.63%
Miscellaneous politics	4.62%	Gab	3.29%
2016 Election and contemporary debates	3.42%	Pop culture	2.52%
Muslims and Europe	2.79%	Food	2.24%
MAGA	2.76%	Christmas and New Year	2.24%
Wars and international politics	2.39%	Literature and photos	2.02%
Las Vegas shooting and terrorism	2.29%	Economics	1.93%
Taxes and government spending	2.21%	Education and children	1.89%
Sex scandals	2.18%	Sports	1.65%
Guns	2.13%	Language	1.25%
Immigration	2.11%	German conversation	1.23%
State level issues	2.10%	Technology	1.13%
Globalism	1.95%	Health and nutrition	1.11%
Climate change	1.81%		
Legal	1.94%		
The media	1.25%		

There is also a substantial literature on the question of whether using social media mobilizes people to be more politically engaged [3]. Gab could be used to explore the connections between users of this platform and alt-right political activities.

## References

1. Allcott, H., Gentzkow, M.: Social media and fake news in the 2016 election. *Journal of Economic Perspectives* **31**(2), 211–36 (2017)
2. Benton, A., Paul, M.J., Hancock, B., Dredze, M.: Collective supervision of topic models for predicting surveys with social media. In: AAAI. pp. 2892–2898 (2016)
3. Bond, R.M., Fariss, C.J., Jones, J.J., Kramer, A.D., Marlow, C., Settle, J.E., Fowler, J.H.: A 61-million-person experiment in social influence and political mobilization. *Nature* **489**(7415), 295 (2012)
4. Duarte, J.L., Crawford, J.T., Stern, C., Haidt, J., Jussim, L., Tetlock, P.E.: Political diversity will improve social psychological science 1. *Behavioral and Brain Sciences* **38** (2015)
5. Haidt, J.: *The righteous mind: Why good people are divided by politics and religion*. Vintage (2012)
6. Oliver, J.E., Wood, T.J.: Conspiracy theories and the paranoid style (s) of mass opinion. *American Journal of Political Science* **58**(4), 952–966 (2014)
7. Paul, M.J., Dredze, M.: Sprite: Generalizing topic models with structured priors. *Transactions of the Association for Computational Linguistics* **3**, 43–57 (2015)
8. Tucker, J., Guess, A., Barberá, P., Vaccari, C., Siegel, A., Sanovich, S., Stukal, D., Nyhan, B.: Social media, political polarization, and political disinformation: A review of the scientific literature. March 2018 report for the Hewlett Foundation (2018)
9. Zannettou, S., Bradlyn, B., De Cristofaro, E., Sirivianos, M., Stringhini, G., Kwak, H., Blackburn, J.: What is gab? a bastion of free speech or an alt-right echo chamber? arXiv preprint arXiv:1802.05287 (2018)