# Social Convos: Capturing Agendas and Emotions on Social Media

**Ankita Bhaumik, Ning Sa, Gregorios Katsios, Tomek Strzalkowski**
Rensselaer Polytechnic Institute
Troy, New York
{bhauma, san2, katsig, tomek}@rpi.edu

**Abstract**

Social media platforms are popular tools for disseminating targeted information during major public events like elections or pandemics. Systematic analysis of the message traffic can provide valuable insights into prevailing opinions and social dynamics among different segments of the population. We are specifically interested in influence spread, and in particular whether more deliberate *influence operations* can be detected. However, filtering out the essential messages with telltale influence indicators from the extensive and often chaotic social media traffic is a major challenge. In this paper we present a novel approach to extract influence indicators from messages circulating among groups of users discussing particular topics. We build upon the concept of a *convo* to identify influential authors who are actively promoting some particular agenda around that topic within the group. We focus on two influence indicators: the (control of) agenda and the use of emotional language.

**Keywords:** agenda, emotions, social networks, campaigns

## 1. Introduction

Popular social media platforms, such as Twitter (now X), Facebook, Instagram, and others, have become global, wide reaching communication channels that remain largely unregulated for access and content. They are an ideal vehicle for placing precisely targeted commercial advertising and other types of outreach activities. Social media have also become means of choice for spreading propaganda, disinformation, and for running influence operations. In this paper we focus on this last phenomenon; more specifically: how the presence of deliberate influence operations can be detected among the overall message traffic.

While at the first glance the social media message traffic may appear haphazard, it actually mirrors various social phenomena found in human interactions, albeit on a much larger scale. Social media is thus an extensive repository of information to study group dynamics at scale, including sociolinguistic behaviors such as agenda control and influence, as well as the methods and techniques deployed to achieve them (Del Vicario et al., 2017; Bovet and Makse, 2019; Caldarelli et al., 2020).

At the same time, a crucial challenge in effectively analyzing message traffic lies in distinguishing specific interactions from the multitude that may be happening concurrently, amid a background of largely unrelated "noise". To effectively analyze a potential influence operation, we need to focus on a specific set of messages exchanged among potential influencers, their targets, and many onlookers. It is within this potentially large group where the key social behaviors can be observed with sufficient clarity. Past research has explored representing these behaviors on social media as a landscape of *convos* (Katsios et al., 2019), and this paper builds upon that foundation.

Katsios et al., 2019 define a *convo* as an online social phenomenon where people are engaged around a topic or an activity. A *convo* can be formed among users who read, edit, comment or forward an information artifact, which could be a document, a project, a topic, or an idea. Some examples can be a repository on GitHub, a subreddit on Reddit or a hashtag group on Twitter. We hypothesize that *convos* can be used to detect focused groups of messages around a particular topic, which can be further used for detecting influence operations.

Social media plays a pivotal role in shaping public opinion, especially during national elections or emergencies, like the COVID-19 pandemic (Karlsen and Enjolras, 2016; Jungherr, 2016; Badawy et al., 2019). These operations are driven by groups of users and characterized by the specific agendas they aim to promote. We deploy the concept of *convos* to extract the subsets of social media traffic devoted to popular topics arising within the specified time periods, e.g., during French national elections in 2022.

By examining a *convo*'s traffic, we can isolate the messages related to its main topic and identify a network of influential authors promoting a particular viewpoint, or agenda. We characterize the *convo* representation using established influence indicators, namely agenda control (Broadwell et al., 2013) and the use of emotional language (Bhaumik et al., 2023). This analysis helps pinpoint the *convo*'s most impactful members and reveals the network of relationships between them. Our hypothesis is that a dense network of mutual connections among the key influencers indicates a deliberate influence operation. This may be contrasted with a more or-



*LREC-COLING 2024*, pages 14984–14994

ganic message traffic where individual influencers are acting largely independently of one another.

Existing works on extracting influence indicators from social media has been performed on a message level (Mather et al., 2022; Bhaumik et al., 2023). After tagging individual messages with agenda and emotion labels, the aggregate agendas and emotions in the dataset are typically obtained by calculating a relative distribution or density of each label. However, this bottom-up approach has several drawbacks. Firstly, by working on individual messages, this approach underestimates the semantic and social links between the messages, thus treating them as independent events. Secondly, by ignoring duplicate messages and reposts (retweets on Twitter), this approach weighs each unique message equally and ignores the differences in their popularity. Thirdly, although there are predefined emotion categories, the set of agendas germane in each type of event are typically unknown in advance. As a result, some amount of human annotation may be needed in order to obtain the relevant agenda labels.

To mitigate the above drawbacks, we propose a novel top-down approach to detect agendas (and who promotes them) and emotions (and who uses them) simultaneously from *convo* messages. We use instruction tuned pre-trained large language models (LLMs) that have the ability to process collections of messages as opposed to fine-tuned models that work at a message level.

We perform our experiments on the 2022 French Election Twitter dataset to study popular influence operations, such as #Frexit (Wikipedia, 2023), that promoted France's exit from the European Union. We outline the complete methodology of extracting *convos* from a social media corpus, identification of influential authors in a convo and characterizing them using agendas and emotions in Section 2. Though our experiments have been carried out on a Twitter dataset, our approach has been generalized for multiple social media platforms like Reddit or public forums which do not use hashtags.

## 2. Methodology

Our approach to identify *convo* networks comprises of three major steps (Fig. 1):

1. Identification of a convo using hashtag or topic groups around the keywords of interest.

2. Identification of authors who act as top influential users in a convo.

3. Detection of entities, agendas and emotions from the convo.

In the following subsections, we outline every step in detail.

### 2.1. Identification of Convos

As described by Katsios et al., the definition of a convo depends on the type of social media under examination. For example, on Twitter, convos typically form around hashtag groups and on Reddit they arise in subreddits discussing similar topics. Therefore, given a set of terms of interest, the first step in our approach is to identify these hashtag or topic groups in the input corpus.

Employing a keyword search mechanism to identify the hashtag or topic groups on social media is a less effective approach due to the inherent complexities of the medium. Social media messages are often short and may not mention certain keywords directly. For example, when examining tweets from the 2022 French Elections presidential campaign, its not guaranteed that the terms `France`, `president`, or `election` will be explicitly mentioned. To address this challenge, we first identify hashtag groups on Twitter, or topic groups in other social media platforms, then we use these groups to gather messages related to keywords of interest, allowing for a more comprehensive and accurate analysis.

To identify hashtag convos, we create a co-occurrence based distance matrix of the top 6000 tokens in the hashtag vocabulary of the corpus. Using this as our feature representation, we perform dimensionality reduction via UMAP (McInnes et al., 2018) and then use HDBSCAN (Campello et al., 2013) to cluster them into hashtag groups. Some examples of hashtag groups are listed in Table 1. For datasets from other social media like Reddit or public forums, we use the LDA topic modeling algorithm to identify topic distributions in the input corpus (Blei et al., 2003). The convos of interest are then identified by searching for a set of terms of interest in the hashtag or topic groups. Finally, for each convos of interest, we extract messages to perform further analysis.

| **Hashtag Groups** |
|---|
| #frexit, #ue, #asselineau, #reprenonslecontrôle, #schengen |
| #stopputin, #stoprussia, #standwithukraine, #closetheskyoverukraine, #noflyzoneoverukraine |
| #covid19, #covid, #covid_19, #omicron, #vaccine |
| #passvaccinal, #passevaccinal, #passsanitaire, #passdelahonte, #passevaccinaldelahonte |
| #mckinseygate, #macronmckinsey, #macronassassin, #macronmckinseygate, #macroncnon |

Table 1: Examples of hashtag groups extracted from the French Elections 2022 dataset, showing top 5 hashtags in each group.

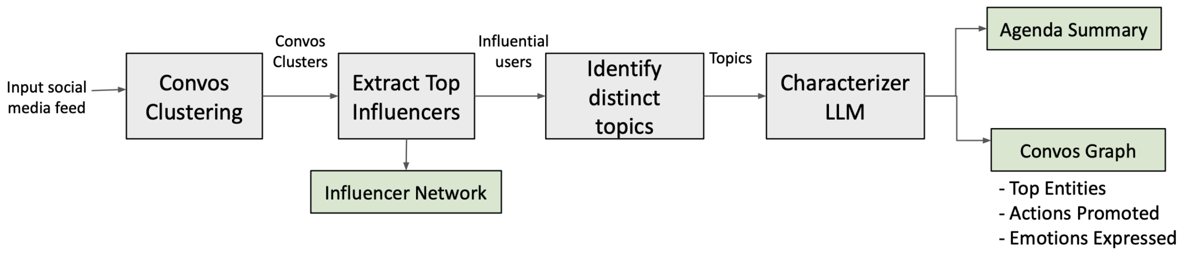



Figure 1: Overview of our approach to extract and characterize *convos*.

### 2.2. Identification of the Top Influencers

As previously discussed, not all messages within a hashtag convo carry equal significance. The weight of a message is determined by the number of retweets, reflecting broader agreement among users. Conversely, a collection of messages from a single user can provide a more comprehensive expression of their ideas and positions compared to a single message. Therefore, instead of analyzing all tweets in a convo, we prioritize the examination of tweets from influential users, who are often retweeted.

To characterize these influential tweets we categorize them by topic, allowing us to extract the agendas and emotions associated with each topic group. We use a two-level hierarchical clustering method to identify distinct topics within bigger groups of messages (Grootendorst, 2022). The agendas and emotions expressed by these influential users serve as a representative sample of the entire convo.

Furthermore, we explore the underlying efforts driving these agendas by scrutinizing the tweeting behavior of influential users and exploring the connections within the cross-influencer network. We hypothesize that in a targeted influence operation these authors would be actively retweeting each other to promote similar ideas.

### 2.3. Agenda and Emotion Detection

In order to effectively perform the tasks of agenda and emotion detection across a collection of messages, we utilize large language models (LLMs) as they have a large context window. Existing fine-tuned models are not suitable for these tasks, primarily due to their limitations in processing single messages within specific domains. Recent advancements in pre-trained LLMs have demonstrated remarkable capabilities in zero-shot summarization across diverse domains (Zhang et al., 2023; Yang et al., 2023). Leveraging these breakthroughs, we have developed an instruction-tuned model for agenda and emotion detection.

In our approach, we provide the LLM with a global context that acts as an instructional guide for its intended task. We adopt a prompt-based template to extract both the agendas and emotions from a collection of messages. Additionally, we employ an output template designed to steer the LLM in generating outputs in a format conducive to visualizing all components within the convo network. This output template replicates a JSON file structure, enhancing its utility for seamless downstream processing. We design prompts to direct the model to answer the questions and accomplish the following tasks:

- Prompt 1: *What are the top distinct entities (maximum 5) mentioned in several messages?*
- Prompt 2: *What are the authors promoting about each entity? Give me 1 phrase for each.*
- Prompt 3: *What are the emotions expressed towards each entity?*
- Prompt 4: *Combine the entities, promoted actions and emotions in the output template.*

The prompts are used to fill out an output JSON template:

```
output = [
{
"entity": {entity},
"promoted_actions": {action},
"emotions": {emotion}
},
...
]
```

This yields a file that contains information regarding the entities, agendas, and emotions within the conversation.

## 3. Experiments

### 3.1. Dataset

We run our experiments on a publicly available dataset from the 2022 French presidential elections (Daignan, 2022). It contains 45 million tweets and retweets from Nov. 12, 2021 to Apr. 03, 2022, focused around the main actors in the election. French is the dominant language of the dataset, followed by English. This dataset contains all retweet information about the tweets which makes it an ideal choice to identify influential message groups in this corpus. We choose two keywords of interest that were prominent topics during the French election campaign, identified through analysis of news articles and online forums: #frexit and #covid_19 (Wikipedia, 2023).

### 3.2. Implementation

We pre-process the original dataset to create a traceable network of tweets and linked retweets. Emojis, hyperlinks and tweets having less than 3 textual tokens are removed. To select the most influential messages in a *convo* we examine the messages by the top 10 most retweeted authors in that *convo*. Alternatively, this value could be determined based on the proportion of messages contributed by each author in a *convo*. Next, we use multilingual BERTopic to separate the most influential messages into smaller topic groups (Grootendorst, 2022).

To implement the agenda and emotion detection system we use a LLM with a sufficiently large context window. We perform our experiments using Llama-2-13b-chat model with a context length of 4096 tokens (Touvron et al., 2023). We provide the global context in the system prompt and use chat prompts to further probe the LLM. The complete prompt template is listed in the Appendix A.1. We use nucleus sampling with $p$ as 0.9 for generating responses and limit the number of new tokens to 500. We also tested our approach using ChatGPT to achieve comparable performance (Ouyang et al., 2022). The experiments are performed on 2 NVIDIA A100 GPUs.

### 3.3. Results

The original corpus is pre-processed to produce a linked dataset comprising of 16.7 million tweets and retweets. On further removal of emojis, hyperlinks, and tweets having less than 3 textual tokens, we arrive at 2.8m tweets and their associated retweet information.

Our first step to identify all hashtag groups in the corpus results in more than 40 hashtag groups (Table 1). We select the two that contain our terms of interest, *frexit* and *covid_19*, to extract the hashtag *convos* around them. Table 2 lists the total number of authors, tweets and retweets in each of these convos. We then select the messages by the top 10 retweeted authors. In the #frexit convo, the top influencers' messages account for 5% of total messages in the convo and comprise around 44% of the total retweets (20,135 out of 44,859). Whereas for the larger #covid_19 convo, that has a larger number of tweets and authors, the top influencers contribute 1% of all tweets and their retweet number comprise 28% of the total retweets.

To understand the connections among the top 10 influencers, we construct a network based on their retweet behavior in the whole corpus (Fig. 2). The nodes in Fig. 2 represent the authors and the weight of the edges represent the number of retweets between them, balloons on top of the node indicate instances of self-retweeting by that author. We construct a graph for each convo and notice that there is no overlap in top influencers between the two, signifying that these authors are predominantly dedicated to promoting specific agendas. While the top 10 influencers in the #Frexit convo (Fig.2a) are connected by a dense network of retweet or self-retweet, the network for #covid_19 (Fig. 2b) contains only 8 authors.

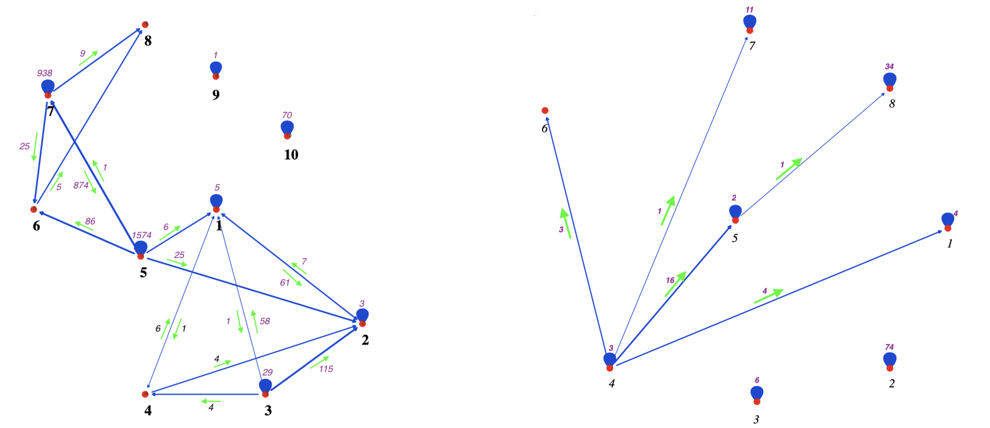
| | #Frexit | | #Covid_19 | |
|---|---|---|---|---|
| | Inf. | Convo | Inf. | Convo |
| No. authors | 10 | 3,572 | 10 | 10,239 |
| No. tweets | 902 | 16,406 | 403 | 39,719 |
| No. RT count | 20,135 | 44,859 | 38,144 | 133,457 |

Table 2: Descriptive results of influencers (Inf.) in the #Frexit and the #Covid_19 convo

We cluster the influencers' messages in each convo into smaller topic groups to facilitate more effective content analysis. We focus on characterizing only the top 5 largest topic groups from each convo. Next, we apply our emotion and agenda detection model to the message clusters to characterize the influence indicators within each influential convo. It is to be noted here that any established emotion/agenda detection model that can process groups of messages can be used for this step. Fig 3 visualizes the json outputs we obtain to depict the top entities, promoted actions and emotions of two selected message clusters from each convo. We remove any json outputs that have more than 50% overlap of entities. The generated agenda summaries, listed in Table 3, provide us with a concise and informative snapshot of the convos around the topic of interest. Some additional convos clusters have been illustrated in the Appendix A.2.

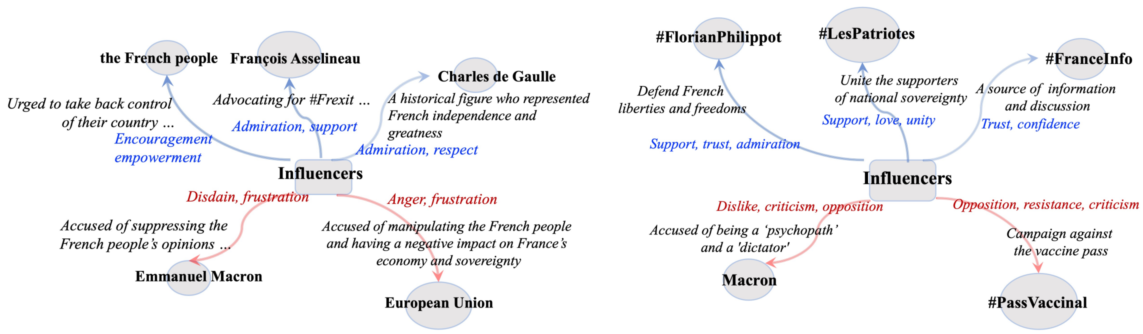
**#frexit convo (Fig 3a):** In the first message cluster of this convo, the influencers promote the candidacy of François Asselineau and advocate for France leaving the EU. The first snapshot also illustrates a similar idea: the influencers support *François Asselineau* who advocates for #Frexit, encourage *the French people* to take back control of their country, and admire *Charles de Gaulle* who represented French independence and greatness. They also accuse *Emmanuel Macron* for suppressing the French people's opinion and are angry about *EU* for manipulating the French people and damaging France's economy and sovereignty.

The second influential message group in the same convo represents another view of the influencers where they advocate for another candidate

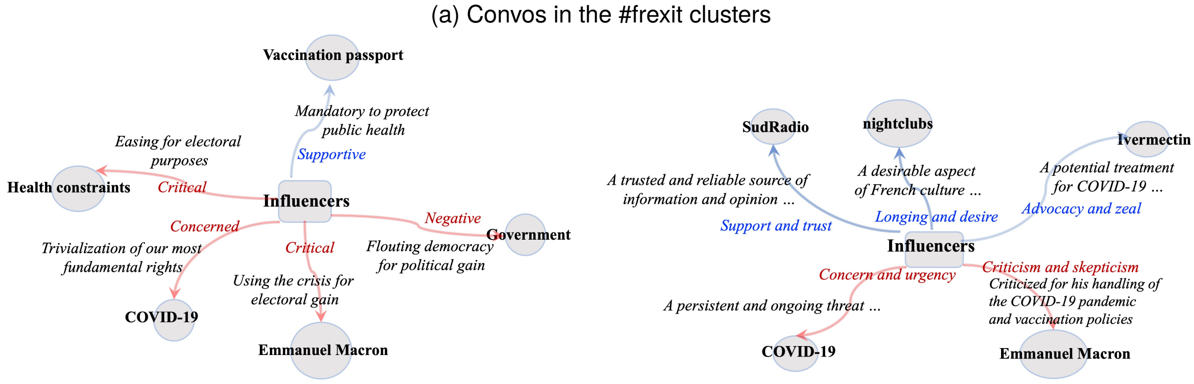


(a) #Frexit convo (b) #covid_19 convo

Figure 2: Influencer Networks among top influencers in each convos. Each node (red point) represents an influencer. The edges are labeled by the weight of the connection between the users. The blue circles represent self-retweet links.




(a) Convos in the #frexit clusters




(b) Convos in the #covid_19 clusters

Figure 3: Different clusters in the *convos* about #Frexit and #covid_19. They are characterized by the indicators of top entities, promoted actions, and emotions. The *influencers* are the top 10 retweeted authors in that convo. The expanded network of the influencers are illustrated in Fig. 2a and 2b. Entities are grey blobs and promoted actions are the grey edge labels. Emotion labels over the edges are colored as negative/positive ones. The model also generates an agenda summary, given in Table 3.

in the election: *Florian Philippot* and support the *LesPatriotes* party he leads. They also criticize president *Macron* and call for a campaign against his policy on *PassVaccinal*.

**#covid_19 convo (Fig. 3b):** In both the influential message clusters of this convo, the users feel concerned about *COVID19* and criticize *Emmanuel Macron* for using the crisis for electoral gain. In the first message cluster, *Vaccination passport* is supported and thought to be able to protect public health, while the *Government* is accused for easing the *Health constraints* for political purposes. Whereas in the second message cluster, people

| Message Cluster | Agenda Summary |
|---|---|
| Fig. 3a - left | The agenda behind this set of tweets is to promote the candidacy of François Asselineau for the French presidency in 2022 and to advocate for France's withdrawal from the European Union (#Frexit) |
| Fig. 3a - right | The agenda behind this set of tweets is to promote the political movement #LesPatriotes and its leader, Florian Philippot, and to campaign against the current French President Emmanuel Macron and his policies, particularly the controversial vaccine pass |
| Fig. 3b - left | The agenda behind this set of tweets is primarily focused on the French government's handling of the COVID-19 pandemic, specifically the use of vaccination passports, the easing of health constraints, and the perceived political motivations behind these actions |
| Fig. 3b - right | The agenda behind this set of tweets is to criticize and scrutinize the actions and words of French President Emmanuel Macron regarding the COVID-19 pandemic and vaccination |

Table 3: Summarized agendas of the messages clusters shown in Fig.3

show desire for *nightclubs*, advocate for *Ivermectin* as a possible treatment for Covid-19, and think *SudRadio* is a trustful resource of information and opinion.

### 3.4. Analysis and Discussion

To validate our hypotheses, we manually analyze the generated convo visualizations and influencer networks. In addition, to ascertain that the generates agendas are correct representations of the broader corpus of data, we compare them against news articles and online forums.

The numbers in Table 2 show that in both the convos, the top 10 influencers generate a very small fraction (5% and 1%) of the tweets in the convo, but receive a large proportion of retweets (44% and 28%). Therefore, we can reasonably extract the tweets by the top influencers to represent the influential or popular messages in the convo. A noticeable difference between the two convos is that, the number of the influential tweets in the #Frexit convo is more than twice of the number in the #covid_19 convo. This signifies that the influencers in the #Frexit convo have been much more active in generating content to drive their agendas.

We closely examine the network of influencers to analyze their retweet behavior. The#frexit influencer network (Fig. 2a) is strongly connected by a large number of retweets. Bi-directional retweeting is also found between 4 pairs of authors. On the contrary, in the #covid_19 influencer network, the connections are notably weaker. All the connections are one-directional and majority of them originate from one author. We hypothesize that a combination of the influencers' efforts in generating the tweets and the density of their retweet network may serve as an early sign of a potential influence operation within the convo. Consequently, though the influencers' receive a substantial number of retweets in both convos, it is plausible that an influence operation may be more prominent in the #Frexit convo compared to the #covid_19 convo.

Next, we analyze the agendas and convos snapshots generated by our model. In the #Frexit convo, both the message clusters in Fig. 2a criticize president Macron but they advocate for different candidates [1]. The first cluster focuses more on France's withdrawal from the EU, and support François Asselineau who is a lead character in the #Frexit movement [2]. However, the second cluster advocates for Florian Philippot and his party, and shows strong opposition against the vaccine pass [3]. Influencer *I4* and *I1* are the top contributors in the first cluster, while more than 90% of the tweets in the second cluster come from *I6*. Fig 2a also depicts that *I4* and *I1* retweet each other, and *I6* is not directly connected to either of them. Hence, we can deduce that even within the same convo, distinct subgroups of influencers and their respective retweeting users tend to agree on specific topics while disagreeing on others. Similarly, the agendas in the #covid_19 convo discuss several topics around the vaccine pass in France and criticize Macron for his handling of the pandemic [4].

[1] https://www.thenationalnews.com/world/europe/2022/04/13/marine-le-pen-denies-frexit-agenda-as-presidential-rivals-clash-over-europe/

[2] https://international.la-croix.com/news/politics/frances-frexit-presidential-candidate/4821

[3] https://www.thelocal.fr/20180219/les-patriotes-florian-philippot-what-you-need-to-know-france-newest-far-right-party

[4] https://www.france24.com/en/europe/20211014-macron-s-covid-health-pass-a-success-in-overcoming-france-s-vaccine-scepticism

Therefore, these visualizations of the focused message groups in the convos provide a concise representation of the ongoing discussions among popular authors around that topic. Extracting such detailed analysis would require significant manual effort and time from experienced analysts. Together with the influencer networks, they provide valuable insights about targeted operations in a corpus.

# 4. Related Work

## 4.1. Influence Operations on Social Media

The use of social media for disseminating influence operations has been extensively studied across various fields such as social, political science and marketing (Watts and Dodds, 2007; Katz et al., 2017). Different methods have been explored to identify influential activity and users across significant events like the Covid-19 pandemic and Presidential Elections (Sameh, 2013). One popular approach involves the representation of social media as a network of users interlinked by various influence measures. Riquelme and González-Cantergiani (2016) present a comprehensive study of various influence measures proposed for social network analysis using Twitter. Cha et al. (2010) investigate how measures like in-degree, retweets, and mentions between users indicate distinct influential behaviors. They also study how the user's influence varies across different topics.

While numerous studies have focused on understanding influential behavior and metrics, effectively filtering out influential messages remains a major challenge. Our work takes a different approach by initially filtering and summarizing essential messages from the huge corpus of social media traffic. We also take a novel approach of using LLMs for this task due to their competitive performance in summarization tasks (Bang et al., 2023; Pu et al., 2023). We have performed some initial experiments using models like Pegasus, *bart-large-cnn-samsum* (Zhang et al., 2020). The biggest drawbacks of these models were the inability to process large pieces of text from diverse social media platforms.

## 4.2. Agenda and Emotion Detection

In previous studies, agenda and emotion detection are usually presented as individual tasks. Most of the work on agenda theory focus on the agenda setting process between the media and the audience, between different types of media, or between different groups of people (Ceron et al., 2016; Guo, 2019; McCombs et al., 1997; Vargo et al., 2014).

In (McCombs et al., 1997), the authors investigate the images of the election candidates presented by mass media and shaped among the voters. Document level content analysis is performed on TV news and newspaper stories, and the voters' opinions towards the candidates are collected by surveys. The agendas in any political campaign are usually candidate names, their personalities, ideologies, and qualifications, and whether their images are built to be positive, neutral, or negative. The authors compare the proportion of each agenda between the media and voters to understand the effects of media on the general population. With the emergence of social media, similar research has been performed to explore the agenda-setting dynamics between mass media and social media, to identify the roles of the agenda setter between them (Su and Borah, 2019; Conway et al., 2015).

The use of emotional language has also been an important indicator during the political campaigns (Grüning and Schubert, 2022). With the availability of deep learning models and annotated datasets, fine-tuning models for emotion detection in particular domains have become a popular approach for emotion detection (Cai and Hao, 2018; Huang et al., 2019; Chiorrini et al., 2021). Prompting techniques using pre-trained large language models have also been deployed to perform zero-shot emotion detection in unknown domains (Plaza-del Arco et al., 2022). Although there has been significant research on emotion detection over individual messages on social media, our work is dedicated to detecting emotions expressed in groups of messages towards specific entities or topics.

All the works mentioned above aim to clarify the phenomena of information flow from one source to another, and ignore the possible influence operations behind the information flow. In contrast, our approach addresses this missing piece by identifying the influencers in the dynamic process and analyzing their behaviors and their inter-connections. Additionally, our break-down and summary method works efficiently as it requires lesser human annotation, which is a necessary step in most of the previous works (Guo, 2019; Su, 2022; Subbiah et al., 2023).

# 5. Conclusion

In this work we introduce a novel systematic approach to identify influential message subsets from social media datasets. We build upon the idea of a *convo* to extract concise snapshots of these online discussions around particular events or topics. We characterize these *convos* using agendas and emotions to enable us to detect targeted influence operations by a group of influencers. We also create a network of these influencers to analyze the retweet behavior and group dynamics among them. A case study around prominent campaigns during the 2022 French Elections helps us to evaluate our proposed approach. Future work would include application of our approach on other types of social media and definition of concrete metrics to measure degrees of influence and precisely identify influence operations.

## 6. Acknowledgements

This paper is based upon work supported by the Defense Advanced Research Projects Agency (DARPA) under Contract No. HR001121C0186. Any opinions, findings and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of DARPA or the U.S. Government.

## 7. Limitations and Ethical Considerations

Given the nature of the proposed task and the role of LLM in our approach, this work does have several limitations. Firstly, we analyze our approach on a publicly available Twitter dataset. The dataset contains retweet information but does not differentiate between tweets, quotes, and replies. In future, we plan to extend to other platforms and other datasets with more diverse user interactions in order to better understand the group dynamics of online communities. Secondly, it's widely acknowledged that current LLMs exhibit a range of biases (Rozado, 2023). In some instances, we've observed that LLMs accurately summarize complementary opinions from diverse user groups, which leads us to believe that biases of LLMs may have minimal impact on this task. However, we acknowledge the possibility of biases reflecting in certain outputs, and we plan to conduct human evaluations to assess biases and report on them in future work.

We do not use or analyze any personally identifiable data in our experiments. The outputs generated are solely based on summaries generated by filtering popular messages on Twitter and the use of popular LLMs. They may reflect various political agendas or opinions of particular candidates. They must be verified by expert analysts before use in any downstream application.

## 8. Bibliographical References

Meysam Alizadeh, Jacob N Shapiro, Cody Buntain, and Joshua A Tucker. 2020. Content-based features predict social media influence operations. *Science advances*, 6(30):eabb5824.

Adam Badawy, Aseel Addawood, Kristina Lerman, and Emilio Ferrara. 2019. Characterizing the 2016 russian ira influence campaign. *Social Network Analysis and Mining*, 9:1–11.

Yejin Bang, Samuel Cahyawijaya, Nayeon Lee, Wenliang Dai, Dan Su, Bryan Wilie, Holy Lovenia, Ziwei Ji, Tiezheng Yu, Willy Chung, et al. 2023. A multitask, multilingual, multimodal evaluation of ChatGPT on reasoning, hallucination, and interactivity. *arXiv preprint arXiv:2302.04023*.

Francesco Barbieri, Jose Camacho-Collados, Luis Espinosa Anke, and Leonardo Neves. 2020. TweetEval: Unified Benchmark and Comparative Evaluation for Tweet Classification. In *Findings of the Association for Computational Linguistics: EMNLP 2020*, pages 1644–1650.

Ankita Bhaumik, Andy Bernhardt, Gregorios Katsios, Ning Sa, and Tomek Strzalkowski. 2023. Adapting emotion detection to analyze influence campaigns on social media. In *Proceedings of the 13th Workshop on Computational Approaches to Subjectivity, Sentiment, & Social Media Analysis*, pages 441–451.

David M Blei, Andrew Y Ng, and Michael I Jordan. 2003. Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan):993–1022.

Alexandre Bovet and Hernán A Makse. 2019. Influence of fake news in twitter during the 2016 us presidential election. *Nature communications*, 10(1):7.

George Aaron Broadwell, Jennifer Stromer-Galley, Tomek Strzalkowski, Samira Shaikh, Sarah Taylor, Ting Liu, Umit Boz, Alana Elia, Laura Jiao, and Nick Webb. 2013. Modeling sociocultural phenomena in discourse. *Natural Language Engineering*, 19(2):213–257.

Xiaofeng Cai and Zhifeng Hao. 2018. Multi-view and attention-based bi-LSTM for Weibo emotion recognition. In *2018 International Conference on Network, Communication, Computer Engineering*, pages 772–779.

Guido Caldarelli, Rocco De Nicola, Fabio Del Vigna, Marinella Petrocchi, and Fabio Saracco. 2020. The role of bot squads in the political propaganda on twitter. *Communications Physics*, 3(1):81.

Ricardo JGB Campello, Davoud Moulavi, and Jörg Sander. 2013. Density-based clustering based on hierarchical density estimates. In *Pacific-Asia conference on knowledge discovery and data mining*, pages 160–172. Springer.

Andrea Ceron, Luigi Curini, and Stefano M Iacus. 2016. First-and second-level agenda setting in the twittersphere: An application to the italian political debate. *Journal of Information Technology & Politics*, 13(2):159–174.

Meeyoung Cha, Hamed Haddadi, Fabricio Benevenuto, and Krishna Gummadi. 2010. Measuring user influence in twitter: The million follower fallacy. In *Proceedings of the international AAAI conference on web and social media*, volume 4, pages 10–17.

Zhuo Chen, Chris Chao Su, and Anfan Chen. 2019. Top-down or bottom-up? a network agenda-setting study of chinese nationalism on social media. *Journal of Broadcasting & Electronic Media*, 63(3):512–533.

Andrea Chiorrini, Claudia Diamantini, Alex Mircoli, and Domenico Potena. 2021. Emotion and sentiment analysis of tweets using BERT. In *EDBT/ICDT Workshops*.

B.A. Conway, K. Kenski, and D Wang. 2015. The rise of twitter in the political campaign: Searching for intermedia agenda-setting effects in the presidential primary. *Journal of Computer-Mediated Communication*, 20(4):363–380.

Jean-Michel Daignan. 2022. French presidential election 2022.

Munmun De Choudhury, Hari Sundaram, Ajita John, Doree Duncan Seligmann, and Aisling Kelliher. 2010. " birds of a feather": Does user homophily impact information diffusion in social media? *arXiv preprint arXiv:1006.1702*.

Michela Del Vicario, Fabiana Zollo, Guido Caldarelli, Antonio Scala, and Walter Quattrociocchi. 2017. Mapping social dynamics on facebook: The brexit debate. *Social Networks*, 50:6–16.

Maarten Grootendorst. 2022. Bertopic: Neural topic modeling with a class-based tf-idf procedure. *arXiv preprint arXiv:2203.05794*.

David J Grüning and Thomas W Schubert. 2022. Emotional Campaigning in Politics: Being Moved and Anger in Political Ads Motivate to Support Candidate and Party. *Frontiers in Psychology*, 12:6337.

Lei Guo. 2019. Media agenda diversity and intermedia agenda setting in a controlled media environment: A computational analysis of china's online news. *Journalism Studies*, 20(16):2460–2477.

Yen-Hao Huang, Ssu-Rui Lee, Mau-Yun Ma, Yi-Hsin Chen, Ya-Wen Yu, and Yi-Shin Chen. 2019. EmotionX-IDEA: Emotion BERT- an Affectional Model for Conversation. *arXiv preprint arXiv:1908.06264*.

Andreas Jungherr. 2016. Twitter use in election campaigns: A systematic literature review. *Journal of information technology & politics*, 13(1):72–91.

Rune Karlsen and Bernard Enjolras. 2016. Styles of social media campaigning and influence in a hybrid political communication system: Linking candidate survey data with twitter data. *The International Journal of Press/Politics*, 21(3):338–357.

G. Katsios, N. Sa, and T. Strzalkowski. 2019. Social convos: A new approach to modeling information diffusion in social media. In *Proceedings of Advances in Artificial Intelligence, Software and Systems Engineering (AHFE 2019)*, pages 3746–3753, Cham. Springer International Publishing.

Elihu Katz, Paul F Lazarsfeld, and Elmo Roper. 2017. *Personal influence: The part played by people in the flow of mass communications*. Routledge.

Brodie Mather, Bonnie Dorr, Adam Dalton, William de Beaumont, Owen Rambow, and Sonja Schmer-Galunder. 2022. From stance to concern: Adaptation of propositional analysis to new tasks and domains. In *Findings of the Association for Computational Linguistics: ACL 2022*, pages 3354–3367.

Maxwell McCombs, Juan Pablo Llamas, Esteban Lopez-Escobar, and Federico Rey. 1997. Candidate images in spanish elections: Second-level agenda-setting effects. *Journalism & Mass Communication Quarterly*, 74(4):703–717.

L. McInnes, J. Healy, and J. Melville. 2018. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. *ArXiv e-prints*.

Long Ouyang, Jeffrey Wu, Xu Jiang, Diogo Almeida, Carroll Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, et al. 2022. Training language models to follow instructions with human feedback. *Advances in Neural Information Processing Systems*, 35:27730–27744.

Flor Miriam Plaza-del Arco, María-Teresa Martín-Valdivia, and Roman Klinger. 2022. Natural language inference prompts for zero-shot emotion classification in text across corpora. In *Proceedings of the 29th International Conference on Computational Linguistics*, pages 6805–6817.

Xiao Pu, Mingqi Gao, and Xiaojun Wan. 2023. Summarization is (almost) dead. *arXiv preprint arXiv:2309.09558*.

Fabián Riquelme and Pablo González-Cantergiani. 2016. Measuring user influence on twitter: A survey. *Information processing & management*, 52(5):949–975.

David Rozado. 2023. The political biases of ChatGPT. *Social Sciences*, 12(3):148.

Ahmed Sameh. 2013. A twitter analytic tool to measure opinion, influence and trust. *Journal of Industrial and Intelligent Information*, 1(1).

Yan Su. 2022. Networked agenda flow between elite us newspapers and twitter: A case study of the 2020 black lives matter movement. *Journalism*, page 14648849221092521.

Yan Su and Porismita Borah. 2019. Who is the agenda setter? Examining the intermedia agenda-setting effect between Twitter and newspapers. *Journal of Information Technology & Politics*, 16(3):236–249.

Melanie Subbiah, Amrita Bhattacharjee, Yilun Hua, Tharindu Kumarage, Huan Liu, and Kathleen McKeown. 2023. Towards detecting harmful agendas in news articles. In *Proceedings of the 13th Workshop on Computational Approaches to Subjectivity, Sentiment, & Social Media Analysis*, pages 110–128, Toronto, Canada. Association for Computational Linguistics.

Hugo Touvron, Louis Martin, Kevin Stone, Peter Albert, Amjad Almahairi, Yasmine Babaei, Nikolay Bashlykov, Soumya Batra, Prajjwal Bhargava, Shruti Bhosale, et al. 2023. Llama 2: Open foundation and fine-tuned chat models. *arXiv preprint arXiv:2307.09288*.

Chris J Vargo, Lei Guo, Maxwell McCombs, and Donald L Shaw. 2014. Network issue agendas on twitter during the 2012 us presidential election. *Journal of communication*, 64(2):296–316.

Duncan J Watts and Peter Sheridan Dodds. 2007. Influentials, networks, and public opinion formation. *Journal of consumer research*, 34(4):441–458.

Wikipedia. 2023. Frexit — Wikipedia, the free encyclopedia. [Online; accessed 17-October-2023].

Xianjun Yang, Yan Li, Xinlu Zhang, Haifeng Chen, and Wei Cheng. 2023. Exploring the limits of ChatGPT for query or aspect-based text summarization. *arXiv preprint arXiv:2302.08081*.

Jingqing Zhang, Yao Zhao, Mohammad Saleh, and Peter Liu. 2020. Pegasus: Pre-training with extracted gap-sentences for abstractive summarization. In *International conference on machine learning*, pages 11328–11339. PMLR.

Tianyi Zhang, Faisal Ladhak, Esin Durmus, Percy Liang, Kathleen McKeown, and Tatsunori B Hashimoto. 2023. Benchmarking large language models for news summarization. *arXiv preprint arXiv:2301.13848*.

# A. Appendices

## A.1. Appendix A: Agenda and Emotion Detection Prompts

We have used Llama-2-13b-chat model for summarization and characterization of the extracted convos. Tables 4 and 5 show the prompts used for the tasks.

```
You are a helpful, respectful and
honest agenda detection assistant.

Read the list of messages given and
understand the hidden agendas behind
them. Please do not share false
information.

If you are unable to understand the
agenda, simply say "No agenda". What is
the overall agenda behind this set of
messages? Give me a short summary.

Messages: {input_text}
```

Table 4: System Prompt use to characterize convos using Llama-2-13b-chat

```
What are the top distinct entities
(maximum 5) mentioned in several
messages? What are the authors
promoting about each entity? Give me 1
phrase for each. Give the emotions
expressed towards each entity.

Combine the entities, promoted actions
and emotions in the following format:

output = [
    {
        "entity": {entity},
        "promoted_actions": {action},
        "emotions": {emotion}
    },
    ...
]
```

Table 5: Prompts and output template to extract entities, promoted actions and emotions

## A.2. Appendix B: Additional Results

As mentioned earlier we only visualize the top 5 largest topic groups for each topic of interest. We have selected the most distinct ones for the paper. Figure 4 illustrates the other clusters in the #frexit convo.


Figure 4: Additional Convos in the #frexit clusters