EPJ Data Science
a SpringerOpen Journal

RESEARCH   **Open Access**

Detecting coordinated and bot-like behavior in Twitter: the Jürgen Conings case

Bart De Clerck[1,2]* , Juan Carlos Fernandez Toledano[2], Filip Van Utterbeeck[2] and Luis E.C. Rocha[1,3]

*Correspondence:
bart.declerck@ugent.be
[1] Department of Economics, Ghent University, Sint-Pietersplein 5, Ghent, Belgium
[2] Department of Mathematics, Royal Military Academy, Renaissancelaan 30, Brussels, Belgium
Full list of author information is available at the end of the article

Abstract

Social media platforms can play a pivotal role in shaping public opinion during times of crisis and controversy. The COVID-19 pandemic resulted in a large amount of dubious information being shared online. In Belgium, a crisis emerged during the pandemic when a soldier (Jürgen Conings) went missing with stolen weaponry after threatening politicians and virologists. This case created further division and polarization in online discussions. In this paper, we develop a methodology to study the potential of coordinated spread of incorrect information online. We combine network science and content analysis to infer and study the social network of users discussing the case, the news websites shared by those users, and their narratives. Additionally, we examined indications of bots or coordinated behavior among the users. Our findings reveal the presence of distinct communities within the discourse. Major news outlets, conspiracy theory websites, and anti-vax platforms were identified as the primary sources of (dis)information sharing. We also detected potential coordinated behavior and bot activity, indicating possible attempts to manipulate the discourse. We used the rapid semantic similarity network for the analysis of text, but our approach can be extended to the analysis of images, videos, and other types of content. These results provide insights into the role of social media in shaping public opinion during times of crisis and underscore the need for improved strategies to detect and mitigate disinformation campaigns and online discourse manipulation. Our research can aid intelligence community members in identifying and disrupting networks that spread extremist ideologies and false information, thereby promoting a more informed and resilient society.

Keywords: Social network analysis; Polarization; Coordinated behavior
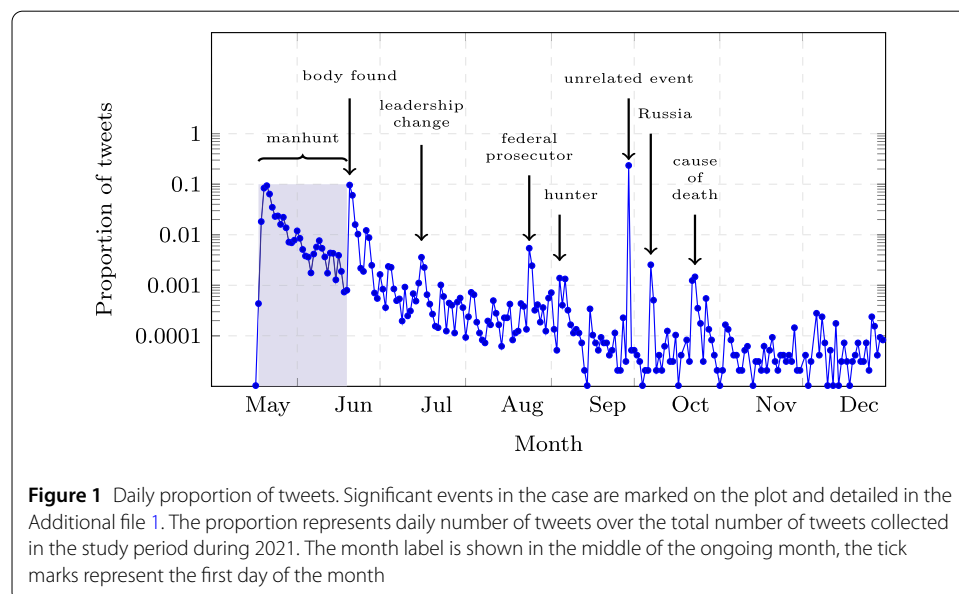
1 Introduction

During the COVID-19 pandemic, people faced an unprecedented situation in which a new virus spread rapidly across the globe. In addition to the virus, people were subjected to a relentless barrage of information. Despite efforts to accurately inform the public, misinformation, hoaxes, and conspiracy theories remain a concern, spreading within the general population and calling the attention of media outlets [1]. In addition to this "innocent" reporting, attempts by foreign actors to influence domestic debate (disinformation campaigns) have been observed within Europe [2, 3]. To keep up with the volume of information, researchers and fact-checkers had to put in overtime. Meanwhile, much research on

the spread of (false or misleading) information and other aspects of this period has been published [4–8].

This paper focuses on Twitter (now X) discourse related to the Jürgen Conings case that occurred in Belgium in 2021. Conings was a Belgian soldier and alleged right-wing extremist who threatened politicians and virologists during the COVID-19 crisis. Conings stole weapons from his barracks and went missing, sparking an international manhunt. The case drew a lot of attention on social and traditional media, with tens of thousands of people rallying behind him online, and also led to real-life protests. His body was eventually discovered, and the cause of death was ruled to be suicide, sparking additional controversy and conspiracy theories. Belgium's multilingual context (three official languages: Dutch, French and German), implies that communication (including news outlets) is segregated across geographical locations. Six major media outlets (including two state-sponsored) control most of the media in each region and no coordination exists between those outlets. In addition to the major official languages (Dutch and French), the discussion also unfolded in other languages, particularly English, indicating the presence of international actors, thereby adding another level of complexity. In Belgium, approximately 13% of the population uses Twitter, with the highest usage among young, higher-educated adults aged 20 to 35 [9, 10]. Figure 1 shows important events and the associated Twitter activity related to the case (see more details in the SI). According to Belgian intelligence services, foreign actors used the situation to try to shape the narrative [11].

We propose a methodology that incorporates both network science and content analysis to understand the dissemination of information on social media. Studying the users' interactions with network methods helps to understand (dis)information dynamics while content analysis helps to uncover trends in narratives, language, and themes and can contribute to revealing the goals and techniques of disinformation campaigns. It can also assist in identifying disinformation sources such as malevolent actors, bots, or state-sponsored organizations, allowing for a targeted response. Our methodology aims to unveil whether there is coordinated user activity and bot-like behavior around this particular case, taking into account real-time communication, language barriers, and not directly related

Figure 1 Daily proportion of tweets. Significant events in the case are marked on the plot and detailed in the Additional file 1. The proportion represents daily number of tweets over the total number of tweets collected in the study period during 2021. The month label is shown in the middle of the ongoing month, the tick marks represent the first day of the month

communication (e.g. other sources of disinformation or external events) that may affect the quality of the detection algorithms. We aim to identify conspiracy theories, potential sources of disinformation, and how (extreme) narratives spread on social networks, using tweets as a study case.

We first exploit the content of the messages, including the websites shared by the users. By analyzing the emerging bipartite user-website network, we investigate the discernible patterns of website sharing among users. Certain websites may be shared on a regular basis by users, implying the presence of specific narratives, ideologies, or interests within the user base. This dynamics may generate network structures, particularly network communities, reflecting that users with similar interests of beliefs are clustered. We identified language barriers related to website sharing and observed the presence of controversial websites observed in other work focusing on misinformation spreading.

We then study the potential presence of coordinated behavior by focussing on the time proximity between (re)tweets to generate rapid interaction networks. The rapid-retweet network and the rapid-semantic similarity network are two activity-based techniques that we propose for identifying coordinated behavior. In the rapid-retweet network, we found indications of coordinated behavior, and observed that conspiracy content was shared more rapidly than other content. The rapid-semantic similarity network allowed us to identify coordinated behavior that would have been missed by the rapid-retweet network and bot detection algorithms. This is a novel approach to detect coordinated behavior, which can be extended to other types of content such as images.

Finally, we look for evidence of bots (i.e. automated user) activity. Bots may be used to amplify specific messages, spread false information, or steer a specific narrative in an online debate. The presence of bot activity could indicate that external actors are attempting to exploit the case for their own purposes, such as propaganda dissemination or influencing public opinion. We found evidence of bot-like behavior for accounts from news outlets, users hijacking the topic, and users expressing conspiracy theories or distrust in the government.

1.1 Related work

Network analysis has seen a wide range of applications related to COVID-19, both in the real and virtual worlds. In social media, network science can be applied to the study of user interactions and discourse. When analyzing the public discourse on Twitter during the COVID-19 pandemic in Italy, researchers discovered that misinformation had a significant impact on the discourse and that the majority of unreliable news articles were shared by right-wing political groups [12]. Another study analyzed the semantic network during the initial Italian lockdown and discovered that discursive communities and traditional political parties overlapped substantially. Dis- or misinformation-related themes were uncommon and unevenly distributed across communities [13].

The relevance of narrative analysis [14], its shaping, and evolution in understanding the social aspects of a pandemic cannot be overlooked. In this context, narratives refer to the stories and viewpoints individuals share about the pandemic, its impacts, and the responses to it. The term evolution here refers to the study of how these narratives are formed and how they change over time, either naturally or due to external influences. In March 2020, the World Health Organization (WHO) declared an *infodemic* (a phenomenon resulting from many interacting and overlapping processes such as the production, consumption, and amplification of - potentially harmful - information online [15])

related to COVID-19. In addition to the virus outbreak, the profusion of incorrect information was regarded as a major issue. An Infodemic Risk Index was devised to quantify the extent to which nations are exposed to unreliable news [16]. The authors discovered that quantifiable surges of possibly unreliable information preceded the rise of COVID-19 infections, whereas reliable information tends to become more prevalent as infections increase. In May 2020, during the early months of the COVID-19 pandemic, the documentary 'Plandemic' was published with the intention of spreading misinformation and conspiracy theories on Twitter. It was discovered that the publication had an effect on the community structure and communication patterns of Twitter users discussing the term *plandemic* [17]. During the second wave of COVID-19 in Australia in 2020, two inter-related hashtag campaigns arose in response to the government's handling of the situation. The authors demonstrated, using a mixed-methods approach, how a small number of hyper-partisan pro- and anti-government campaigners were able to mobilize ad hoc Twitter communities and co-opt journalists and politicians to amplify their message [18]. Using a survey experiment in Dutch-speaking communities of Belgium, it was determined that exposure to disinformation about COVID-19 (e.g. virus origin, vaccine efficacy) alters the perception of vaccine efficacy [19]. A Latent Dirichlet Allocation (LDA) [20] analysis of the evolution of COVID-19 vaccine-related discussions on Twitter in Japan revealed a transition in public interest over time [21]. Different phases have been identified in an analysis of English-language news media narratives on COVID-19 from October 2019 to May 2020 worldwide. The Early and Peak stages showed worldwide convergence, whereas the pre-pandemic and recovery stages indicated regional heterogeneity [8]. An analysis of the strategies used by the Disinformation Dozen (Disinfo12), a group responsible for spreading a large volume of COVID-19-related misinformation on Twitter identified by the Center for Countering Digital Hate [22], revealed that they were able to spread misinformation by using both original content and content from other accounts [23]. This content was distributed via private websites and YouTube videos. The study also identified two categories of accounts re-sharing Disinfo12's content: those connected to low-credibility sources and those politically active and linked to the United States Republican Party.

In the analysis of discursive communities on Twitter during political and societal debates, it was discovered that these communities predominantly exhibit bow-tie structures during such debates, with low-quality content being produced and shared primarily within the strongly connected component sector, which aligns with the *infodemic* concept mentioned previously [24].

Although sometimes overlapping, bot-like activity and coordinated activity are distinct phenomena in the realm of online behavior. Bot-like activity refers to automated or semi-automated actions carried out by software applications, or "bots", which frequently mimic human behavior. While these bots are becoming increasingly sophisticated, they may not fully capture the nuances and complexities of genuine human interactions in all contexts. These bots can perform tasks such as posting repetitive content, disseminating misinformation, and amplifying specific messages. Coordinated activity, on the other hand, consists of the collaborative efforts of multiple actors, who can be either human or automated, working toward a common goal. These actors collaborate to influence public opinion, manipulate online discourse, or advance a specific agenda. Multiple datasets have been made available to researchers studying the dissemination of misinformation about COVID-19 [25–28]. The Twitter disinformation operations dataset [29] is another source of labeled

data linked to disinformation campaigns. These datasets are highly curated, containing only accounts associated with a disinformation campaign. The authors of [30] used an unsupervised entropy-based method to identify bot-like behavior in the Twitter disinformation operations dataset. In [31] a language and platform agnostic framework for identifying trolling activities in online social networks is proposed. The framework focuses on the activity sequences (trajectories) of accounts. The Twitter environment is modeled as a Markov decision process by considering both the actions taken by accounts and the feedback received. The behavior of accounts is represented by states and actions, and identification of trolls is approached as a binary classification problem. AMDN-HAGE is an unsupervised generative model for detecting disinformation campaigns on social media that uses Neural Temporal Point Processes and Gaussian Mixture Model to jointly model account activities and concealed groups [32]. The method does not require prior knowledge of synchronized accounts or predefined characteristics. The authors discovered coordinated organizations promoting anti-vaccine, anti-mask, and conspiracy theories based on COVID-19 data.

Our investigation seeks to understand the following: How do linguistic divides, bot-like behaviors, and coordinated user activities influence the dissemination of narratives, particularly misinformation and conspiracy theories, in the Twitter discourse surrounding the Jürgen Conings case? We propose the following hypotheses:

H1: Belgium's linguistic divides create distinct information exchange clusters within the multilingual online discourse, potentially leading to differing levels of susceptibility to misinformation, challenges in the coordination and effectiveness of countermeasures, and variations in the resilience and spread of different narratives.

H2: Bot-like behavior and coordinated user activities are prevalent and instrumental in amplifying narratives, especially those involving misinformation and conspiracy theories.

H3: The rapid dissemination of narratives, facilitated by bot-like behavior and coordinated activities, suggests targeted efforts to influence public opinion or manipulate the discourse.

2 Methods

2.1 Data collection

The study period spans from May until December 2021 and the data were downloaded from the Twitter platform between January 12 and January 21, 2022. We combined Twarc [33] with the search API for Twitter (with academic access). We started with a set of case-specific keywords, consisting of variations of Jürgen Conings' name and the initial expression of support (which translates for French and Dutch to "I am Jürgen"). Without imposing any language restriction, we used snowball sampling to find hashtags that were not on our initial list of keywords. Using expert knowledge of the case and a cutoff frequency (1% of the most popular hashtag), we selected the most relevant hashtags. These were added to the query and we then collected the tweets that contained them (Table 1). To capture the broader discussion, we used Twarc to gather the entire conversation thread (if applicable) from all of the recovered tweets.

Our query also captured a large number of unrelated messages, this was due to a UEFA Champions League match between Porto and Liverpool played on September 29, 2021. Although the soccer match has nothing to do with the case, *Jürgen* (one of the keywords) Klopp was Liverpool's head coach at the time, which lead to tweets about this event being

Table 1 Twitter API Search terms (combination of Dutch, French and English)

Phase	Keywords
Initial	conings, ikbenjurgen, jconings, jesuisjurgen, jurgen, jurgen_c, jurgenc, jurgenconings, ouestjurgen, waarisjurgen
Snowball	1pourjurgen, 1voorjurgen, ABL, ADIV, als1achterjurgen, armee, armée, belgianrambo, belgischleger, belgiumfailedstate, commissiondefense, dedonder, defense, défense, defensie, defensieminister, doofpot, extreemrechts, extremedroite, extrêmedroite, freejurgen, iksteunjurgen, JC, jurgen, jurgenconingsdead, jurgenslifematters, justiceforjurgen, klopjacht, korporaal, leger, marsfurjurgen, militair, militaire, ministredeladefense, NationaalParkHogeKempen, ocad, peperkoekenhartje, peperkoekmyass, rambo, rambobelge, ripjurgenconings, safehouse, SGRS, staatsveiligheid, teamjurgen, veiligheidsmachtiging, vlaamsfront, vlaamslegioen, wearejurgen, whereisjurgen

captured. The shared first name and the high-visibility of this event can account for this. Because there was no overlap with any significant events related to the Jürgen Conings case, we dropped the data covering the time span from the day leading up to the match until the day after the match (September 28th until September 30th) for the remainder of the analysis.

2.2 Content analysis

Transformer-based models such as BERT [34] or GPT-3 [35] learn latent representations (embeddings) of words and sentences in context. Sentence-BERT [36] (specifically `paraphrase-multilingual-mpnet-base-v2` [37]) was used to encode the content of the messages. We also used OpenAI's `text-embedding-ada-002` [38], as an additional embedding model to obtain a vector representation of the messages. Each of the embedding models are trained on a large corpus of text, can deal with multiple languages, can capture unique aspects of the language data (e.g. semantic meaning, syntactic structures, polysemy, homonymy, sentiment), ensuring no key features are overlooked. These two models differ in their training data, underlying architecture (BERT vs GPT-3), maximum context size (512 vs 8192 tokens) and embedding dimension (768 vs 1536). This dual-model strategy increases result validity, i.e. similar findings from both models reinforce our confidence in the insights. Before computing the embeddings, the mentions (`@user`) and urls were removed from the messages. Given two messages *A* and *B* and their embedding vectors $\mathbf{a}$ and $\mathbf{b}$, we can evaluate their semantic similarity by computing the cosine similarity of their embedding vectors: $S_c = \cos(\theta) = \frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{a}\|\|\mathbf{b}\|}$, where $\mathbf{a} \cdot \mathbf{b}$ denotes the inner product and $\| \cdot \|$ denotes the Euclidean norm of a vector. Because the embeddings are generated in a shared semantic space, regardless of the language of the input text, the embeddings of texts in different languages can be compared directly [39]. In the multilingual context of Belgium, this allows us to detect patterns cross languages, narratives, and strategies in the dissemination of information and disinformation.

2.3 Website sharing

The bipartite user-website network is composed of two types of nodes: users and websites. An edge between two nodes indicates that a user has at least once shared a URL to a website. The URLs are extracted from the tweets of the users. The websites are identified from the URLs either by extracting the domain name or by retrieving the original URL and then extracting the domain name (when link shortening services are employed). For instance, a shortened URL https://bit.ly/3zZ5Z0N that refers to the URL https://www.nieuwsblad.be/cnt/dmf20210730_094 would become the website node

nieuwsblad.be. We removed the website *twitter.com* from the network to avoid excessively increasing the density of the network. Interactions are studied by projecting the bipartite network onto a network of websites. This projection will result in a network where websites are nodes, and two websites are connected if they were shared by the same user. This helps to understand website categories, content preferences, and the dynamics of information flow among websites. Such an approach has been used to study the spread of disinformation on Facebook pages [40] and to detect echo chambers on Twitter [41].

Given an observed network G^*, the maximum-entropy method [42–45] consists of constructing an ensemble of networks $\mathcal{G}$ whose topology is random, apart from a controlled set of structural constraints, **C**, measured on G^*. The ensemble is found by maximizing the Shannon entropy S
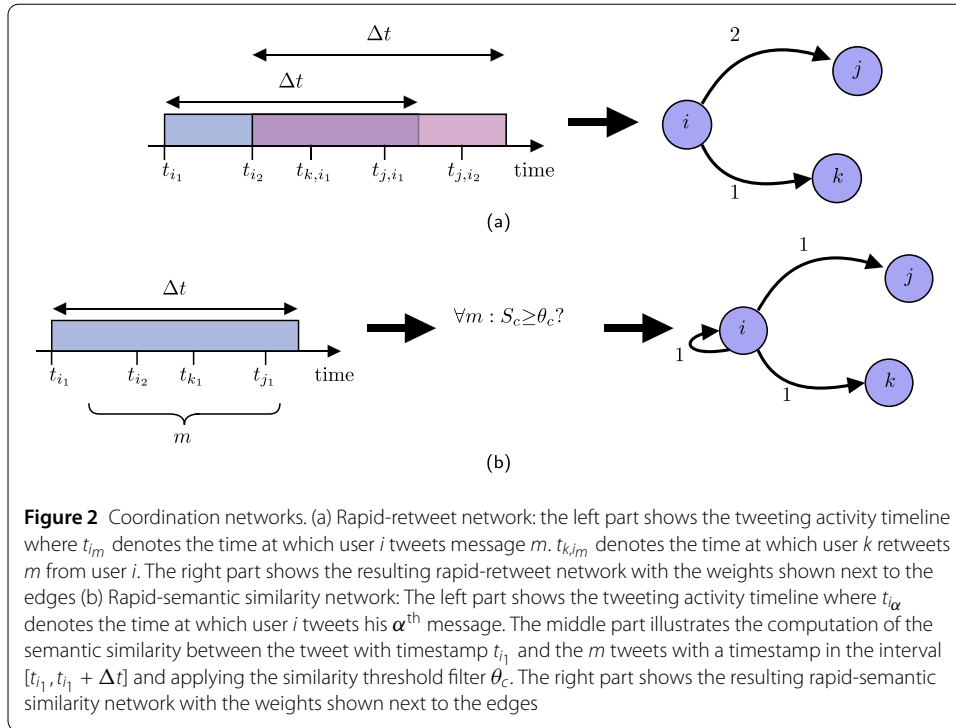
$$S = -\sum_{G \in \mathcal{G}} P(G) \ln P(G)$$

to obtain the least-biased ensemble. These random network models can be used to find statistically significant structures on real-world networks. The user-website interaction network is a bipartite network for which an entropy-based null-model has been defined. The bipartite configuration model [46, 47] can be used to analyze statistical significance of a network pattern. When considering the bipartite user-website network, we want to find patterns of website sharing that cannot be explained by the random behavior of users or by the popularity of the websites alone. This can reveal groups of websites that tend to be shared together. Clusters of websites can indicate the various interests, specific (political) narratives or ideologies of the user base. We used the Python package NEMTROPY [48], which has implemented multiple null models that allow to run the analysis.

2.4 Coordinated behavior

Coordinated behavior is studied by constructing rapid interaction networks. These networks are constructed by considering the time proximity between (re)tweets. Analyzing time proximity in user behavior to help identify potential cases of misinformation, manipulation, and polarization is motivated by a variety of factors. For starters, the real-time nature of human interactions on these platforms results in dynamic and engaging conversations, highlighting the importance of temporal proximity in understanding user behavior. Second, social media platforms' algorithms tend to prioritize recent and popular content, incentivizing coordinated actions to take place in close temporal proximity in order to increase visibility and reach. Furthermore, the online attention economy suggests that coordinated activities that occur in close temporal proximity to one another are more likely to capture users' attention and maximize the impact of shared content. We use Δt to denote the maximum difference in seconds between the timestamps of two tweets for them to be considered as rapid.

The rapid-retweet network [49, 50] is then defined as a weighted, directed and static network in which nodes represent users and edge weights represent the number of rapid retweets sent between the respective users. We use $\mathcal{G}_{RR}$ to denote a rapid retweet network. The edge's direction corresponds to the direction of information flow [51]. Figure 2a depicts the principle behind the creation of the rapid retweet network. The in-degree of a node of $\mathcal{G}_{RR}$ (k_{in}) can be interpreted as a metric for rapid retweet behavior with respect to

Figure 2 Coordination networks. (a) Rapid-retweet network: the left part shows the tweeting activity timeline where t_{i_m} denotes the time at which user i tweets message m. t_{k,i_m} denotes the time at which user k retweets m from user i. The right part shows the resulting rapid-retweet network with the weights shown next to the edges (b) Rapid-semantic similarity network: The left part shows the tweeting activity timeline where t_{i_α} denotes the time at which user i tweets his α^{th} message. The middle part illustrates the computation of the semantic similarity between the tweet with timestamp t_{i_1} and the m tweets with a timestamp in the interval $[t_{i_1}, t_{i_1} + \Delta t]$ and applying the similarity threshold filter θ_c. The right part shows the resulting rapid-semantic similarity network with the weights shown next to the edges

multiple users and the edge weight can be interpreted as the intensity of the rapid retweet behavior between two specific users.

Two messages are semantically similar if the cosine similarity of their associated embedding vectors is greater than a threshold value θ_c. If a message A is semantically similar to and is posted within Δt seconds of a message B, it is considered to have a high rapid-semantic similarity. The rapid-semantic similarity network is then defined as a weighted, directed and static network, with nodes representing users and weighted edges representing the number of rapid semantic similar messages. We use $\mathcal{G}_{RSS}$ to denote a rapid semantic similarity network. The direction of the edges matches the flow of information (Fig. 2b). The in-degree of a node (k_{in}) can be interpreted as a metric for rapid semantic similarity behavior with respect to multiple users and the edge weight (w) can be interpreted as the intensity of the rapid semantic similarity behavior between two specific users. The out-degree of a node (k_{out}) can be interpreted as a metric for the propensity of a user's messages to be rapidly similar to others. We will refer to a $\mathcal{G}_{RSS}$ obtained by using `text-embedding-ada-002` as an "ADA network" and refer to a $\mathcal{G}_{RSS}$ obtained using `paraphrase-multilingual-mpnet-base-v2` as a "MPNet network".

In both $\mathcal{G}_{RR}$ and $\mathcal{G}_{RSS}$ a low edge weight might represent an occasional rapid interaction, not necessarily implying underlying coordinated behavior. The higher the edge weight, the more likely it is that the rapid semantic similarity behavior is coordinated.

2.5 Bot-like behavior

To detect bot-like behavior, we used three existing techniques. The first technique examines the accounts that the platform has suspended. This is done by querying the Twitter API to ascertain the status of each account in the dataset. This a posteriori evaluation it is not a piece of evidence that the account was a bot, but has been used to demonstrate that

suspended users have a tendency to promote divisive issues [52]. As a second method, we used Botometer [53] to identify potential bots in our dataset. We use $S_{botometer}$ to denote the bot-like behavior according to the botometer score of a user. For the third approach, we used digital DNA to evaluate the predictability of user behavior and categorize accounts as either bots or real people [54, 55]. Compression statistics of a user' behavior (encoded in the digital DNA) is combined with other features such as retweets, replies, hashtag usage and URL sharing to create the digital DNA score. We use S_{DNA} to denote the bot-like behavior according to the digital DNA score of a user. The users' bot score (S_{bot}) is defined as the average of the botometer and digital DNA scores:

$$S_{bot} = \frac{S_{botometer} + S_{DNA}}{2} \tag{1}$$

As both scores are normalized to the interval $[0, 1]$, the resulting S_{bot} is also normalized to the interval $[0, 1]$. The requirement to gather a consecutive sequence of tweets is a known limitation of the last two approaches for the user under examination. Collecting all messages is subject to rate limiting, and can be a disadvantage of the botometer and the digital DNA approach.

2.6 Website controversy
Under the assumption that users who are banned from Twitter are more likely to share controversial content, we use the status of the users to examine the relationship between the content of a website and the users who link to it. The ratio of banned users to the total number of users linking to the website is used as a measure of website controversy ($S_{controversy}$):

$$S_{controversy} = \frac{number\ of\ banned\ users}{number\ of\ users} \tag{2}$$

This metric in itself can provide a distorted view of the controversy of a website, as it does not take into account the size of the user base. When evaluating the controversy of a website, we will also need to consider the number of users linking to the website to account for its popularity. A similar approach has been used to identify controversial YouTube videos, where the removal of a video was used as a proxy to label low-credibility content [56], which in turn can be used to identify important spreaders of controversial content.

3 Results
3.1 Descriptive statistics and tweet narratives
The dataset contains 96,651 tweets, of which 21,978 (22.74%) are original messages, 60,527 (62.62%) are retweets, 11,502 (11.90%) are replies and 2644 (2.74%) are quotes. In total, 41 different languages are observed, with the most common ones being Dutch (55.86%), English (20.17%) and French (11.72%). There are 51,841 unique users, of which 3278 (6.32%) are verified and 48,563 (93.68%) are not. The distribution of the users' status and their verification at the moment of the analysis is as follows: 45,959 active, 1341 closed and 1263 suspended non-verified users and 3276 active and 2 closed verified users. The dataset contains 17,114 urls from 681 unique domains (682 when including *twitter.com*).

By clustering the tweet messages based on their semantic similarity, we can identify the most popular themes in the discourse in addition to factual news reporting: (i) Speculation and commentary surrounding the manhunt (ii) Marc Van Ranst being in a safe house; (iii) Skepticism, disbelief and conspiracy theories related to the discovery of the body by the mayor of Maaseik (town where the body of Conings was found) during a bike ride; (iv) Criticism of media coverage and portrayal of Jürgen Conings; (v) Criticism of the way Conings was treated compared to other individuals; (vi) Expressions of support, frustration with the government, and concerns about the political climate in Belgium; (vii) Sarcasm, satire, and jokes about the situation and Belgian politics in general.

We also found unrelated topics in the data set, such as cryptocurrency market updates, international conflicts, and other non-related messages from users who managed to get some attention by hijacking one of the - at the time - trending hashtags.

3.2 Bot-like behavior and controversial websites

Both S_{DNA} and $S_{\text{botometer}}$ exhibit a bimodal distribution, indicating that both methods are capable of discerning two distinct groups of Twitter accounts. The digital DNA approach identifies peaks within the intervals of $[0, 0.15]$ and $[0.6, 0.7]$, while the botometer score shows peaks in the intervals $[0, 0.4]$ and $[0.6, 1.0]$. The Spearman rank-order correlation coefficient between both measures is 0.39 ($p < 0.00001$). Given these disparities, we adopted an averaging strategy to reconcile the cautiousness exhibited by the digital DNA method with the leniency demonstrated by the botometer score into a single metric (Equation (1)). Because most reliable news outlets are verified accounts and have an elevated bot score, we analyze the verified and unverified accounts separately. Based on the distribution of the bot score, we define accounts with $S_{\text{bot}} > 0.6$ as accounts with a high bot score. The verified accounts with an elevated bot score are mainly posting original messages (98.97%), with limited other activities (retweets: 0.85%, replies: 0.18%). For the non-verified accounts with an elevated bot score, we observe a different distribution, dominated by original messages (56.68%) and retweets (41.89%), and a limited number of quotes (0.17%) and replies (1.26%). Because we could only evaluate the scores of accounts that were active at the time of the analysis, we were unable to obtain the botometer and digital DNA analysis for the closed or suspended accounts (cf. methods section). These made up 5% of the total number of accounts and generated 6.2% of the total volume of tweets in the dataset.

When ranking the verified accounts based on their bot score, we observe that the majority of the accounts are news outlets. One account that stands out in this ranking is *Sputnik France*, this is a French version of the Russian state-sponsored news outlet *Sputnik*, which is known for influence operations [57, 58] and has been banned in Europe since March 2022 [59]. Other verified accounts with a high bot score are mostly politicians, journalists, and public figures (notably Marc Van Ranst, one of the threatened virologists). For the unverified accounts, the following types of accounts are identified as potential bots: unverified local news outlets, automated accounts sharing trending stories, spambots hijacking trending topics to push their own agenda (e.g. cryptocurrencies, free Palestine . . .), normal users actively sharing news, and users who are very active by either retweeting or reacting to specific tweets. This latter category of accounts is characterized by expressions of sympathy or support for Conings, skepticism toward the official narrative surrounding his death, criticism of media coverage, and distrust in the government and authorities.

The bipartite user-website network contains a total of 18,973 users, and 681 websites. When projected onto a network of websites we obtain a network of 681 websites, of which 294 (43.17 %) are isolated, and 3916 edges. The website *HLN.be* shows the highest degree and is a popular newspaper among Dutch language readers in Belgium. The giant connected component contains 354 websites (51.98% of all websites) and 3898 edges (99.54 % of the edges in the projected network). The second largest component is composed of only five domains and contains Italian news outlets some of which are associated with unreliable information (*voxnews.info* and *ilprimatonazionale.it*) [40]. The remaining components are all composed of just two nodes and contain websites from Italian mainstream news outlets, French right wing websites, Dutch alternative news outlets, and unrelated topics such as cryptocurrency, investment advice and blogs. The giant connected component is composed mainly of websites of major news outlets, but also includes some smaller websites. In the giant connected component we identified six communities, found by modularity maximization using the Leiden algorithm [60], who mainly differ in language (three Dutch, one French and two English[1]).

Within the network, we find domains that have been previously identified as untrustworthy or contentious [40, 56, 61, 62] that concentrate on the COVID-19 pandemic, or that have propagated conspiracy theories and false information, frequently with a U.S. basis (Table 2). Additionally, multiple "local" anti-vax websites or alternative news websites are present in the network, e.g. *stichtingvaccinvrij.nl* is known by the Dutch authorities for spreading anti-vax narratives from the U.S. [63] (Table 3).

Table 2 List of unreliable or controversial websites observed in the bipartite user-website network common with other works

bitchute.com	bongino.com
breitbart.com	childrenshealthdefense.org
dailystar.co.uk	friatider.se
hannity.com	ilprimatonazionale.it
oann.com	pjmedia.com
rt.com	rumble.com
theepochtimes.com	thefederalist.com
thegatewaypundit.com	voxnews.info
washingtontimes.com	washingtonexaminer.com
wnd.com	zerohedge.com

Table 3 Examples of "local" conspiracy websites present in the website network

Website	Content
ninefornews.nl	A conspiracy website that is part of the Dutch speaking community. In the past the website published articles about chemtrails, COVID-19 vaccine and the world being run by reptilian overlords. Regarding the Jurgen Conings case, the website doubted his remains actually being buried.
reactnieuws.net	A self-proclaimed independent news website that is part of the Dutch speaking community. It has strong ties with (extreme) right wing politics. It called upon the people to participate in a commemoration of the death of Jurgen Conings and pushed Telegram as an alternative communication channel to avoid government censorship.

[1]First Dutch community: mostly Belgian news outlets, centered around *HLN*; Second Dutch community: mostly Belgian news outlets, centered around *De Morgen*; Third Dutch Community: mostly news outlets form the Netherlands, centered around *NOS*; French community: mostly Belgian and French news outlets, centered around *Le Soir*; First English community: international news outlets; Second English community: topic hijacking and IT related websites (unrelated to the case).

One of the network's most prominent and contentious domains is *Russia Today (RT)*, an international news television network that receives funding from the Russian government and is frequently utilized for propaganda [64, 65]. This domain is found in the French community. It presents Jürgen Conings as a representative of Flemish nationalism and anti-establishment sentiment, appealing to disgruntled people who feel disregarded by the political elite. They draw attention to Conings' endorsements on social media and his affiliation with right-wing nationalist organizations. In one article [66], RT highlights the dissatisfaction of right-leaning voters who feel excluded from the political system and the lack of opportunities for right-wing parties for participation in governmental positions. On March 2nd 2022, following the Russian invasion of Ukraine, RT was sanctioned by the EU for spreading misinformation [59]. Figure 6d shows the relationship between the number of users linking to a website and the number of non-active users linking to the website.
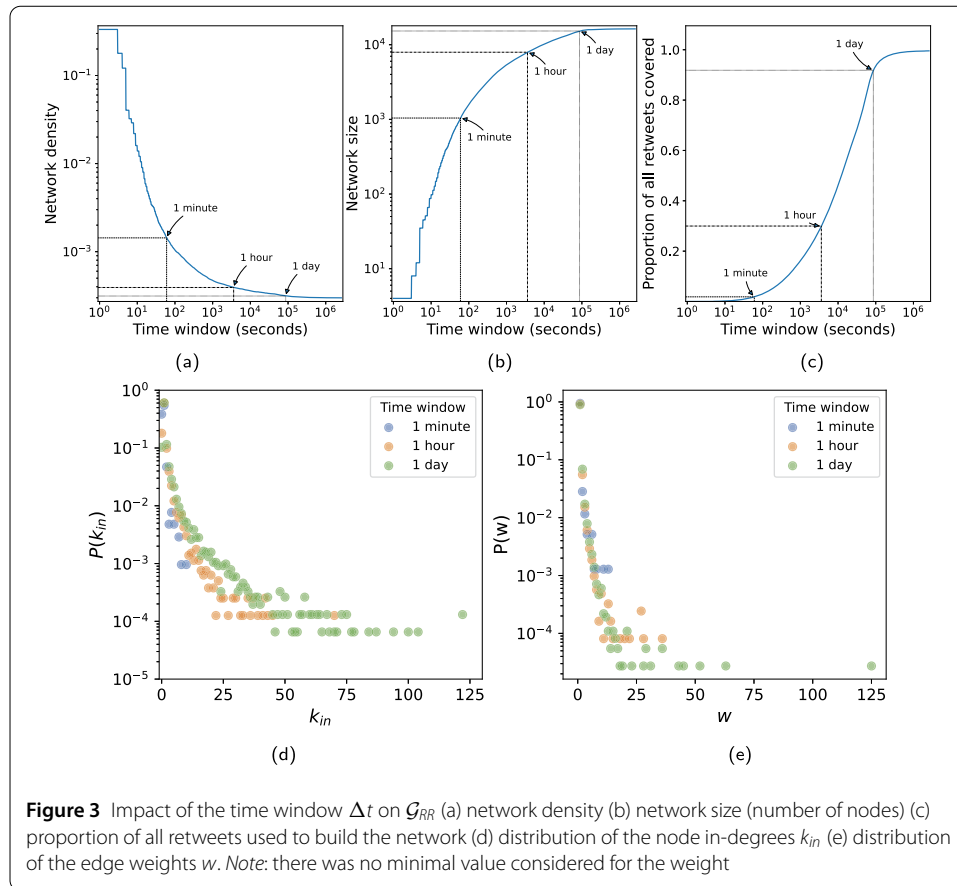
When considering the filtered domain network, we observe one component of almost exclusively Dutch speaking websites (both from Belgium and the Netherlands) and one of almost exclusively French speaking (only from Belgium) websites showing a higher number of interactions than what one would expect under the null model, computed by comparing the observed website sharing with its expected distribution. Within the Dutch speaking component, we observe the presence of the German newspaper *Bild*, which is explained by the fact that the person who discovered the body sold the images to this newspaper. We also observe the presence of the webpage *petities.com*, which is a platform for creating and signing online petitions. The anonymous author of the petition asked to spare Conings' life, claiming that "The state that has used him [Conings] for years to do its dirty work, now wants to kill him". The petition gathered thousand of signatures and was taken offline soon after it gained mainstream media coverage.

There are no validated links crossing the language barrier between the Dutch and French speaking components. Even news outlets that are part of the same media conglomerate such as *Le Vif* and *Knack*, which serve as reciprocal counterparts in the French and Dutch languages respectively, are not connected across the language barriers.

3.3 Rapid retweet network

To evaluate the impact of the parameter Δt on $\mathcal{G}_{RR}$, we considered network density, network size, and the proportion of all retweets in the dataset that contributed to $\mathcal{G}_{RR}$ (Fig. 3). There is an exponential decrease in density as the time window lengthened. In contrast, network size increased substantially beyond minute-long time windows before plateauing for longer time windows. The proportion of all retweets covered showed a sigmoid-shaped curve, with values close to zero for a limited time window and close to one for a time window exceeding one day. The distributions of the users' in-degrees and edge weights for different values of Δt are also shown on Fig. 3. The in-degree captures the number of accounts a user is propagating whereas the weights capture the intensity of the coordination between two specific users. Most of the users and edges are concentrated in the left side (low values) of both the in-degree and weight distributions.
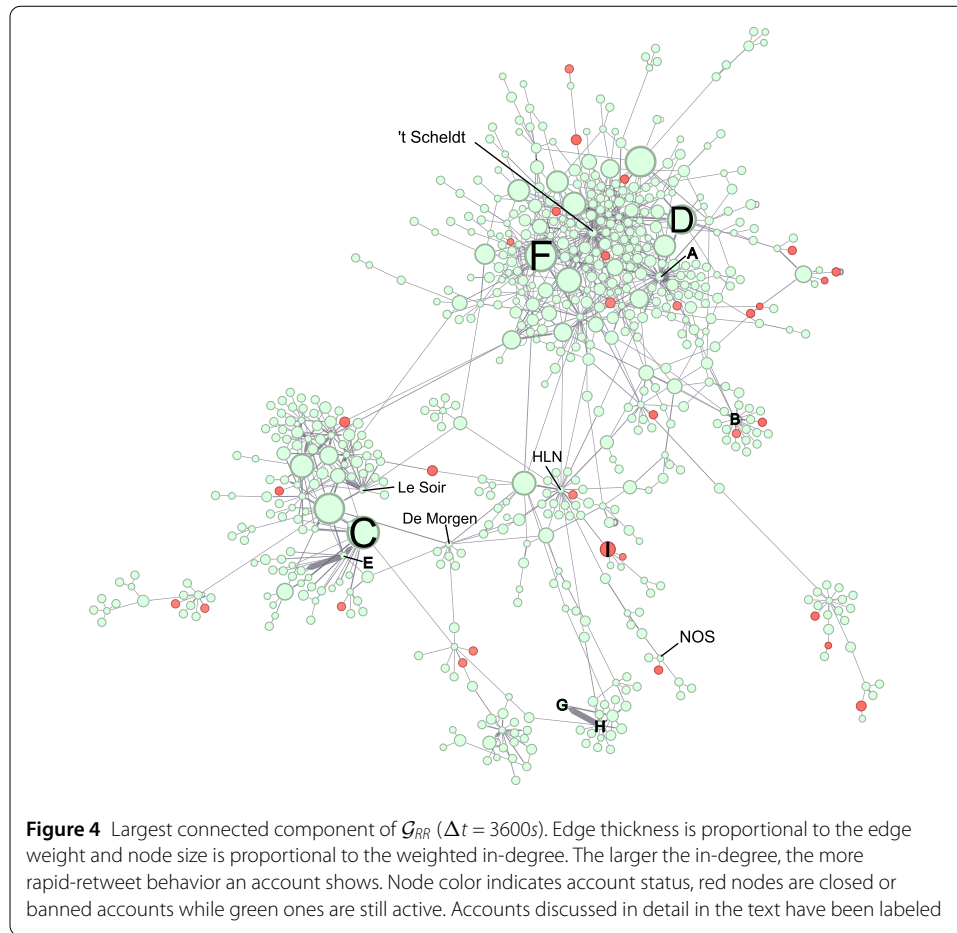
In earlier work, star structures have been found in platform manipulation campaigns [67]. To strike a balance between the needs of capturing coordinated behavior and keeping network density reasonable, we considered time windows of 60 seconds and one hour. Picking a small time windows could prove too restrictive, thus excluding coordinated activities that may be relevant, whereas a large time window could include too many users.

Figure 3 Impact of the time window Δt on $\mathcal{G}_{RR}$ (a) network density (b) network size (number of nodes) (c) proportion of all retweets used to build the network (d) distribution of the node in-degrees k_{in} (e) distribution of the edge weights w. *Note*: there was no minimal value considered for the weight

The time window should allow automated accounts enough time to retweet messages while also allowing them to spread out their activity. With a 60 seconds time window, we observe some star-like-interaction structures, spread across many small components. We mostly observe fully automated accounts, with no indication of suspicious activity. For example, Belgian level news blog `@nbbelgie` is systematically retweeted by its Flemish equivalent `@nbvln`. Other examples include accounts related to sports or news outlets.

When considering a one hour time window, we find additional star-like interaction structures, spread across many small components. Most of these star-like structured are not related to the case at all and are discussing other topics such as a missing person, bit-coins, trending topics and so on. They are picked up because their authors capitalize on the trending hashtags to spread their own message. The network also has a giant weakly connected component of users discussing the Jürgen Conings case (Fig. 4). The rapid retweet network $\mathcal{G}_{RR}$ is composed of 860 nodes and 1047 edges, and the giant component accounts for 618 nodes (71.86%) and 885 edges (84.53%).
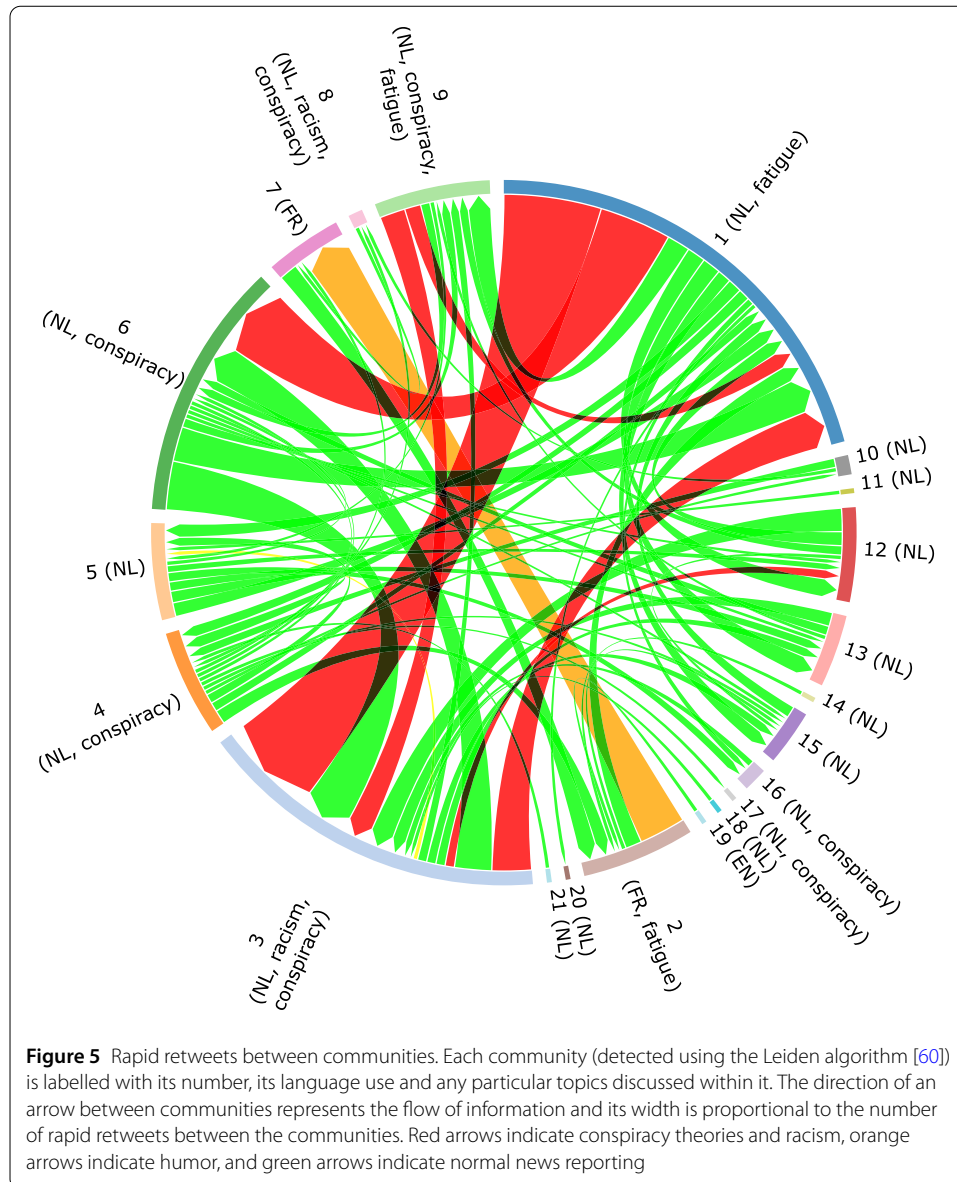
Within the giant component, we can observe several communities, which are characterized first of all by their language, and secondly by the nature of their content. Some communities are merely reporting facts, while others are discussing the case in combination with other topics such as fatigue of COVID-19 restrictions, forms of racism (e.g. toward muslims) when comparing the treatment of Jürgen Conings with other individuals, mistrust in the government and conspiracy theories, and some contain mainly satire and comedy. Figure 5 shows the exchange of information between these communities. Com-

Figure 4 Largest connected component of $\mathcal{G}_{RR}$ ($\Delta t = 3600s$). Edge thickness is proportional to the edge weight and node size is proportional to the weighted in-degree. The larger the in-degree, the more rapid-retweet behavior an account shows. Node color indicates account status, red nodes are closed or banned accounts while green ones are still active. Accounts discussed in detail in the text have been labeled

munity 1 appears to be a seed community for conspiracies and racism. We also observe that the exchanges between different language communities are limited.

The most rapidly retweeted account in the $\mathcal{G}_{RR}$ is *'t Scheldt*, a controversial Flemish satirical website associated with the far-right. We also find Belgian mainstream news outlets to be part of the most shared accounts. The accounts that have the highest out-degree who are not news outlets, criticize the government's handling of the case. Table 4 describes the content of some of these accounts. The accounts with the highest in-degree, indicating rapid retweet behavior, are all unverified users. Table 5 describes the content of some of these accounts. It is worth noting that `User  C` has a node betweenness centrality score of the same order of magnitude as the news outlets. Incidentally, `User C` and `User D` also have the most elevated S_{bot} rating from all nodes in the $\mathcal{G}_{RR}$ (0.75 and 0.78 respectively). We also identify accounts that are consistently retweeting each other. `User  G` and `User H` are two such accounts. They hold unconventional or extreme views, particularly related to conspiracy theories or alternative medicine (called "*wappies*" in Dutch).

We can also observe a limited number of banned users in the largest component of the rapid retweet network: 31 out of 618 (5%). The banned account for which we observed the largest amount of messages has a description in Cyrillic expressing explicit support for Russia (`User  I`). Several of the banned accounts are active in multiple languages. A common observation is skepticism about the official story, sympathy and support for

Figure 5 Rapid retweets between communities. Each community (detected using the Leiden algorithm [60]) is labelled with its number, its language use and any particular topics discussed within it. The direction of an arrow between communities represents the flow of information and its width is proportional to the number of rapid retweets between the communities. Red arrows indicate conspiracy theories and racism, orange arrows indicate humor, and green arrows indicate normal news reporting

Jürgen Conings, criticisms of public figures involved in the case and frustration with the portrayal of Conings in the media.

Overall, we observe that users who are critical of the government's handling of the Jürgen Conings case and of the COVID-19 measures tend to generate more engagement, as measured by the number of rapid retweets they receive or generate, which might indicate a potential area of coordinated behavior. For instance, `User A`, who systematically criticizes the government's handling of the case, posted the most rapid retweets. The results align with the findings of a previous study [68], indicating that people sharing conspiracy theory content had greater impulsive behavior in their posting.

3.4 Rapid semantic similarity network

The previous analysis considered only the activity patterns of tweeting and retweeting. We now introduce semantic information and analyze the rapid-semantic similarity net-

Table 4 Examples of users with the highest out-degrees, i.e. users who's content is shared the most, in the rapid retweet network. The description was generated based on the user's messages using GPT-4 with an 8k context length

	User Description
A	This user has strong far-right beliefs, sympathizes with Jürgen Conings, and criticizes the Belgian government, media, and left-leaning figures. Influenced by anti-establishment sentiment and emotions like frustration, distrust, and resentment, they view Conings as a symbol against perceived hypocrisy. The user questions the manhunt's efficiency and suggests conspiracies. Comparing Conings to Che Guevara, the person feels alienated from mainstream politics and seeks a symbol to challenge the status quo. Their perspective is based on right-wing ideology, anti-establishment attitude, and emotions like anger and disillusionment, reinforcing their support for Conings and distancing from mainstream views.
E	This user consistently expresses concern over Jurgen Conings and radicalization, stressing vigilance against extremist ideas and their online spread. The user discusses predicting potential radicalization, sympathize with terrorism victims, and criticize downplaying dangers. With emotions centered on empathy and concern, the user critiques far-right ideologies, the handling of the Conings case, and authorities' ineffective response to online extremism.
B	This user's stance on Jürgen Conings is driven by a combination of factors including distrust of the government and mainstream media, criticism over how the situation has been managed, sympathy for Conings' personal challenges, and an apparent disdain for misinformation and biased narratives. These elements appear to be interconnected and support a broader anti-establishment, conspiracy-minded viewpoint.

Table 5 Examples of users with the highest in-degrees, i.e. users who show the most rapid retweet behavior in the rapid retweet networks. The description was generated based on the user's messages using GPT-4 with an 8k context length

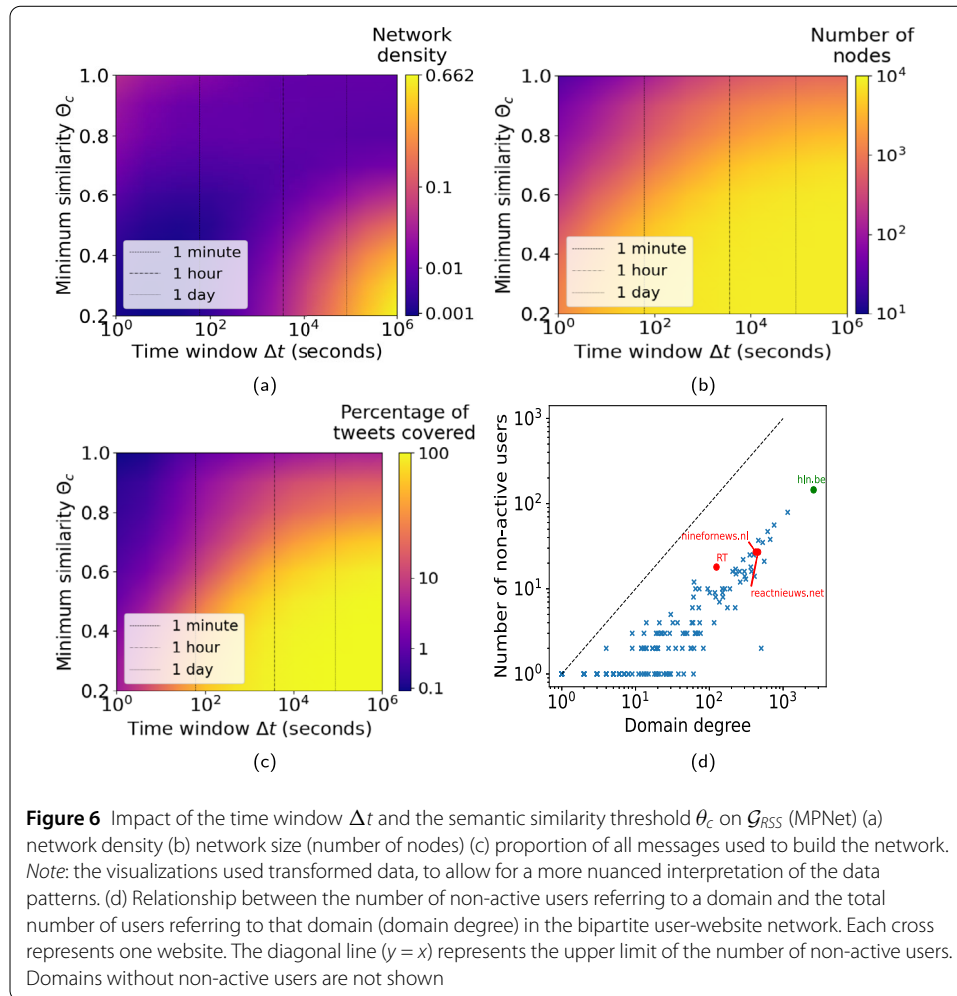|	User Description
C	The broad themes in the messages include: the urgency for a parliamentary investigation into the statements made by the president of the Standing Intelligence Agencies Review Committee, concerns about the state of the military intelligence and security service, demands for the resignation of the director of military intelligence, questions about the role and responsibility of the Minister of Defense, criticism of the handling of the Jürgen Conings case, discussions about possible radicalization and extremism within the army, and updates on the search for Jürgen Conings.
F	The broad themes from the messages include: suspicion of a conspiracy or cover-up surrounding Jurgen Conings' death, criticism of the government and media for their handling of the situation, support for Jurgen Conings and his actions, distrust of Marc Van Ranst and accusations of his role in the situation, concern about the use of excessive force in the search for Jurgen Conings, and a general frustration and discontent with the current political situation in Belgium.
D	The broad themes that emerge from the list of messages include the following: the search for Jurgen Conings, criticism of the government and media handling of the case, comparisons to other criminal cases, accusations of a political agenda, suspicions and conspiracy theories, sympathy for Conings or support for his actions, and questions about the level of response and resources used in the search.
G	The broad theme that emerges from these messages is the discussion and support for the fugitive Belgian soldier Jürgen Conings. The messages indicate a range of opinions, with some expressing sympathy and admiration for Conings, while others criticize his actions or express concern about the support he is receiving.

work $\mathcal{G}_{RSS}$. Retweets were excluded to generate this network, because they are semantically identical and are already covered by the $\mathcal{G}_{RR}$. The value of the parameter θ_c should be adjusted based on the choice of embedding model, as each model can produce a unique distribution of similarity scores. To illustrate the impact of the underlying embedding model on similarity scores between messages (and thus the choice of parameter θ_c), we provide an example of semantic similarity for the different embedding models in Table 6. Starting from a reference message, we provide other messages (both the original one as well as the translation to English if applicable) for decreasing values of semantic similarity. From these examples a similarity of 0.65 for one model still makes sense, while for another model, the content of the message is completely unrelated.

Table 6 Examples of semantic similarity between tweets. The first entry is the reference message, for the subsequent entries the similarity score is shown. The examples span multiple languages, a translation is provided when applicable, the original text is given in italics. The *ADA* and *MPNet* columns indicate the similarity score computed with `text-embedding-ada-002` and `paraphrase-multilingual-mpnet-base-v2` embedding models respectively. Note: The examples in this table are illustrative and do not come from the analyzed dataset

Message	ADA	MPNet
"The manhunt for Jurgen Conings is still ongoing. Authorities are searching for the missing soldier in the Hoge Kempen National Park. #absurdistan" (*La chasse à l'homme pour retrouver Jurgen Conings est toujours en cours. Les autorités recherchent le soldat disparu dans le parc national de la Haute Campine. #absurdistan*)	1.00	1.00
"The body of Jurgen Conings has been found. The heavily armed soldier had been on the run for more than a month. According to the police his dead is ruled a suicide"	0.85	0.68
"More and more people are convinced that Jurgen Conings is a hero and that he is being treated unfairly by the authorities." (*Steeds meer mensen zijn ervan overtuigd dat Jurgen Conings een held is en dat hij oneerlijk wordt behandeld door de autoriteiten.*)	0.82	0.41
"How can we prevent another Jurgen Conings? In an interview with VRT NWS, the head of the Belgian State Security Service said that the case of Jurgen Conings is a wake-up call." (*Hoe kunnen we een nieuwe Jurgen Conings voorkomen? In een interview met VRT NWS zei het hoofd van de Belgische Staatsveiligheid dat het geval van Jurgen Conings een wake-up call is.*)	0.82	0.44
"I don't believe a word of what the government says, his death was not a suicide, they are lying to us! #JurgenisaHero #JusticeForJurgen" (*Ik geloof geen woord van wat de regering zegt, zijn dood was geen zelfmoord, ze liegen tegen ons! #JurgenisaHero #JusticeForJurgen*)	0.78	0.35
"A Facebook group in support of the heavily armed soldier has more than 45,000 members and support with the general public is growing." (*Een Facebookgroep ter ondersteuning van de zwaarbewapende soldaat heeft meer dan 45.000 leden en de steun onder het grote publiek groeit.*)	0.77	0.27
"The Minister of Defence is getting worried that the actions of Jurgen Conings could inspire others…Well they should be afraid, they are the cause of all of this #failedstate" (*Le ministre de la Défense s'inquiète du fait que les actions de Jurgen Conings pourraient inspirer d'autres personnes…Et bien ils devraient avoir peur, ils sont la cause de tout cela #failedstate*)	0.83	0.35
"Mayonnaise is the best sauce in the world, especially on fries. #belgium"	0.73	0.14

Considering the values of Δt used for $\mathcal{G}_{RR}$ and the semantic similarity results, we considered time windows of 60 seconds and one hour in combination with similarity threshold values of 0.7 (MPNet network) and 0.85 (ADA network). These threshold values reflect the difference between the embedding models and a sensible threshold value for semantic similarity based on samples of the data. Figure 6 depicts the evolution of the network density, size, and the proportion of all messages in the datasets (excluding retweets) that contributed to $\mathcal{G}_{RSS}$.

The MPNet network has 933 nodes and 3612 edges with 57 non-active nodes (6.11%) and 135 bot nodes (14.47%) while the ADA network has 904 nodes and 8333 edges with 54 non-active nodes (5.97%) and 127 bot nodes (14.05%). For each $\mathcal{G}_{RSS}$, we observed a weak correlation between k_{in} and k_{out}. Despite the fact that all correlations were statistically significant ($p < 0.001$), their absolute values were less than 0.3. This implies that those who frequently disseminate rapid semantically similar tweets (high out-degree) may not receive a comparable number of such tweets (high in-degree), and vice versa. At the same time, when comparing the ranks of k_{in} and k_{out} for nodes common between the rapid-semantic similarity networks obtained when using different embedding models, we observed a stronger correlation (all above 0.7, $p < 0.001$). While a strong correlation does not necessarily indicate the absence of embedding model sensitivity, we found that all nodes from the ADA network were also present in the MPNet. We also observe the presence of a giant component discussing the Jürgen Conings case, which accounts for 66.04% of nodes and 95.18% of edges (MPNet) and 67.50% of nodes and 98.96% of edges (ADA network).

Figure 6 Impact of the time window Δt and the semantic similarity threshold θ_c on $\mathcal{G}_{RSS}$ (MPNet) (a) network density (b) network size (number of nodes) (c) proportion of all messages used to build the network. *Note*: the visualizations used transformed data, to allow for a more nuanced interpretation of the data patterns. (d) Relationship between the number of non-active users referring to a domain and the total number of users referring to that domain (domain degree) in the bipartite user-website network. Each cross represents one website. The diagonal line ($y = x$) represents the upper limit of the number of non-active users. Domains without non-active users are not shown

The accounts of news websites have both a high in-degree and a high out-degree in the rapid-semantic similarity network. This implies that these accounts are both important senders and receivers of rapid semantically similar messages. This is expected, because news websites act as primary information disseminators. The unverified users on the network with a high in-degree or out-degree appear to be actively involved in discussions, which can range from regular topics to conspiracy theories. These can act as a bridge, allowing misinformation to cross from isolated or extreme groups into larger, mainstream conversations. This highlights the potential for misinformation to have far-reaching consequences, and emphasizes the importance of counter-strategies that take into account the full range of user behavior within the network.

What can be more revealing in term of coordinated behavior are edges with high weights. We identify three categories of accounts that have edges with elevated weights: local news outlets, spambots and presumably coordinated accounts. We observe multiple account pairs posting almost identical messages, typically following a new post from a news outlet or a prominent user. These messages often follow a specific template, with only minor variations, indicating coordinated and possibly automated behavior (Table 7). The only accounts of this group that have been removed from the platform are the spam-

Table 7 Examples of rapid semantically similar posting templates. Square brackets indicate variable content

Number of occurences	Content template
270	[@user] 1 / WAW…50% of tests are false-positive [URL] 2 / A different King in Belgium ???? Massive support for Jürgen Conings on social media [URL]
250	[@user] [series of trending hashtags] [URL]
149	[@multiple_users] Bad day for the fascists, election defeat, suicide of jurgen conings [variable series of emojis]
98	Think for yourself people, Marc Van Ranst and the WHO are criminals! Here's the link [URL]! [series of hashtags] [URL]

bots. The accounts involved in this type of behavior all have a low S_{bot} rating (except for the spambots), so these would not have been flagged as bots.

4 Discussion

4.1 User-website network

The user-website network demonstrates the language divide between the Dutch and French speaking communities in Belgium. Some users are active in both communities, but this is not the norm. Back in the 1970s, the terminology "cultural apartheid" was used to characterize the situation in Belgium. Over 50 years ago, it was predicted that inter-regional communication would become increasingly difficult [69]. The Dutch speaking part of the network is more connected to Anglophone websites than the French speaking part. This is also reflected by the different unreliable websites with which the Dutch and French speaking communities interact. Understanding how different linguistic communities engage with various media platforms can reveal whether certain narratives are more prevalent or spread differently within the linguistic divide. Such distinctions were also found in the Russia-Ukraine conflict, where Russian and Ukrainian military bloggers had significantly different Telegram message content [70]. We can also draw a parallel between the linguistic communities of Belgium and the existence of different subreddits on Reddit. Each subreddit has its own user base, and information posted in one subreddit may not reach another subreddit. Previous research has demonstrated that community membership might hinder the dissemination of information or viewpoints among (linguistic) communities [7]. Malicious actors could take advantage of this division.

4.2 Coordination networks

For both the rapid retweet network and the rapid semantic similarity network, we observe a giant connected component discussing the Jürgen Conings case. In both networks we find indications of coordinated behavior, but the nature of the coordination is different. The mechanisms that underpin these two networks are fundamentally different. Retweeting is a direct, intentional act of information dissemination by a user [71], whereas semantic similarity is an indirect measure of content overlap that can occur even when users do not interact directly. The rapid-semantic similarity model can find more subtle forms of coordinated behavior (i.e. similar statements across users), indicating possible synchronization that is not obvious in standard retweet networks. By showing influential nodes inside these coordinated clusters, it can contribute to a fuller understanding of information

transmission dynamics and help enhance counter-strategies. While there is some overlap in the users involved in rapid retweeting and rapid semantic similarity behavior, each network captures a distinct set of users and interactions, indicating that the specific interactions captured by rapid retweets and rapid semantic similarity are largely distinct. This underscores the value of considering both types of behavior in our analysis.

Our findings indicate that banned accounts in the rapid retweet network do not behave as bridging nodes, implying that their removal has little effect on information diffusion. This emphasizes the need for more effective intervention measures, which might include targeting important nodes to reduce information dissemination. Furthermore, while individual account scores may be valuable, they may not be effective in preventing societal-wide knowledge transmission, an issue that vaccination tactics face too. Future studies could concentrate on developing a more sophisticated knowledge of network dynamics and intervention options.

The threshold used to determine semantic similarity may have an impact on network structure [72, 73]. With a high threshold, many semantically similar messages may be excluded, reducing the number of edges in the semantic similarity network. Another critical factor to consider is user behavior. Users' proclivity to retweet versus create original content, even if it mirrors existing messages, can vary greatly. Because of this variation in user behavior, a user may be represented in one network but not the other. The nature of the tweet content may also play a role in the observed differences. Certain tweets, such as those from popular users or those about trending topics, may be more likely to be retweeted [74]. The parameters that were chosen for our study's rapid retweet and rapid semantic similarity networks are still subject to some arbitrariness. Despite this limitation, our approach provides a useful system for comprehending and identifying coordinated social media behavior. Furthermore, because Twitter messages are (usually) very short, it is possible that a text embedding model does not perfectly capture the specific intent of the message.

4.3 Social media platform

Twitter does not have the same influence on the Belgium's social media landscape as other social media platforms do, but it offers a reasonable proxy for online communication in real-time. According to the most recent surveys, Facebook (66.1%), YouTube (57.5%), Instagram (39.2%), TikTok (19%) and Snapchat (18.9%) have a higher penetration rate than Twitter (13.1%) among Belgians over 12 years old [9]. The platform is popular for political discourse and news sharing and is often directly referenced by news outlets. Specific adoption rates of the platform by Belgian journalists and for every political level are not available, but 87% of members of federal parliament and 100% of Belgian members of European parliament are active on Twitter [75, 76]. Some elements of the composition of the Twitter user base in Belgium are known: the platform has more adoption among students, males and higher educated users in urban areas [10]. Given Twitter's global use and influence, our study's emphasis on it may still be able to shed light on how people behave on social media. It's also worth noting that the data collection occurred prior to Elon Musk's acquisition of the platform (April-October 2022) and before some of Twitter's internal workings were made open source (March 2023). A malicious actor focusing solely on this platform could now analyze how it works and attempt to amplify its own agenda further [77, 78]. Twitter withdrew from a voluntary European Union agreement to combat online

disinformation in May 2023, a code to which the platform had adhered since 2018. Furthermore, the policy change regarding API access and the associated costs may have an impact on the ability to collect data from the platform in the future. The methods presented in this paper are not limited to Twitter and can be applied to other (social media) platforms.

While images, videos, and linked media are frequently included in social media posts, the textual content of tweets was the focus of our research. These factors were not considered in our analysis, which may have limited our understanding of coordinated social media behavior. Image segmentation, object recognition, OCR, and speech-to-text techniques generate textual information that can be used to supplement message analysis and improve the ability to decipher the intended meaning of a user's messages [79].

Our current methodology, which is based on retrospective queries through Twitter's search endpoint, has limitations. For example, we could overlook messages or accounts that were removed from the platform between the time of creation and collection. These difficulties, however, can be overcome by modifying our approach to include a number of improvements. The transition to Twitter's Stream API is an important step in this evolution. This API allows us to collect data in real-time by providing a continuous feed of tweets based on specified parameters or keywords. This change will allow us to access live tweets as they are posted, enabling us to detect and analyze trends as they emerge. Incorporating dynamic query modification increases the adaptability of our system. This entails real-time adjustments to the parameters or keywords used for data collection in response to emerging trends or new developments. As the discourse evolves, this responsiveness ensures that our data collection remains relevant to our research objectives. Another critical aspect of this transformation is dynamically updating our social network analysis as new data arrives from the Stream API. This ongoing updating process enables us to gain timely insights into the evolution of discourse, the emergence of new narratives or actors, and the detection of abrupt shifts that may indicate coordinated activity or new information campaigns. Finally, human curation can improve our system even further: human judgment and specific subject matter expertise can provide invaluable insights at certain stages of the data collection and analysis process. A human analyst can supervise dynamic query modification, deciding on new parameters based on their understanding of the evolving situation. Humans can also review the results of social network analysis updates, spot anomalies or interpret complex patterns that algorithms alone may not be able to capture.

4.4 Computational aspect

The computation of embeddings for each model can be time-consuming, with the method used largely determining the duration. Local computational power will be the bottleneck for on-premises inference, whereas rate limiting will be the bottleneck when using OpenAI's API. However, it is important to note that these embeddings must only be computed once, which reduces the impact of this computational overhead. Our computations were carried out on either a standard workstation or a GPU server. The GPU server was primarily used for its speed and ability to execute multiple tasks concurrently, which was especially useful for parameter analysis. The dataset under consideration here could be analyzed on a workstation. The methodology can be applied in (near) real-time, which is fundamental to blocking content or users swiftly.

5 Conclusions

Politically driven statements or world-changing events can spark passionate debate online, especially if they align with polarized ideas and goals. The Jürgen Conings case in Belgium is an example of a major topic of discussion and disagreement on social media platforms. To gain a deeper understanding of the potential for opinion manipulation in such divisive debates in social media, we conducted an in-depth investigation of the online discourse of this case on Twitter.

The main narratives and themes in the discourse cover factual information about the case, different opinions on the government's handling of the situation, conspiracy theories surrounding the case itself and the COVID-19 pandemic in general. We found that both the sharing of websites and the discussion of the case were largely divided across language communities. Major news outlets, conspiracy theory websites, and anti-vax platforms were identified as the primary sources of website sharing. The emergence of new conspiracy theories and the alignment with pre-existing ones (in this case related to the COVID-19 pandemic), is a phenomenon that has been previously observed [6].

The analysis of the rapid retweet network revealed the presence of a giant and well-connected component discussing the Jürgen Conings case, dominated by users who were critical of the government's handling of the case and, by extension, the COVID-19 measures. Our analysis of the bot-like and coordinated activity revealed a wide range of user behaviors such as automated retweeting and hashtag hijacking, but also systematically posting of similar messages, suggesting potential disinformation campaigns. The rapid semantic similarity network allowed us to identify coordinated behavior that would have gone unnoticed in the rapid retweet network. Therefore, the combination of both methods is necessary to increase the power of detecting potential coordinated behavior.

There are some limitations to this study, such as the focus on textual content and the exclusion of other social media platforms. However, our methodology can be applied to other social media. For example, for Reddit, one could connect users if they publish semantically similar comments on a post within a set time window. Similarly, for Facebook, one could connect users if they link to the same post or group within a set time window. Although the semantic similarity analysis was restricted to textual content within the confines of Belgium, given the number and geographical distribution of languages in Belgium, we could demonstrate the detection quality of the method across different languages, indicating its potential to be deployed in other contexts. For example, this approach can also be used to construct a rapid semantic similarity network based on shared images, audio or videos, by using an appropriate model to quantify semantic similarity for desired content type.

Combining network science with content analysis (as done with the rapid retweet and rapid semantic similarity networks) can assist in identifying individuals and groups spreading extremist ideologies, and uncovering coordinated activity and flagging potential disinformation campaigns. There is also some level of false positive detection, however, this is inherent in real-world social network data. Members of the intelligence community can use the proposed scores for automated flagging of users for further in-depth evaluation, early discovery and prevention of misinformation escalation, including disruption of such networks. This would help to reduce the ability of certain actors to manipulate the public opinion and spread false information, potentially assisting in the early discovery and prevention of disinformation escalation, potential acts of violence or terrorism.

As time goes on, the development of advanced AI technologies will pose additional challenges in combating the spread of false or misleading information. To effectively address these challenges and promote a more informed and resilient society, we must adapt our defensive measures and strengthen collaboration among researchers, policymakers, and society at large. By identifying patterns of information dissemination and coordinated behavior, our research can help inform the development of strategies to protect against the influence of false information and promote the integrity of online discourse.

Abbreviations
WHO, World Health Organization; LDA, Latent Dirichlet Allocation; Disinfo12, Disinformation Dozen; BERT, Bidirectional Encoder Representations from Transformers; GPT, Generative Pre-trained Transformer; RT, Russia Today.

Supplementary information
Supplementary information accompanies this paper at https://doi.org/10.1140/epjds/s13688-024-00477-y.

Additional file 1. Details of the Jürgen Conings case (PDF 96 kB)

Declarations

Competing interests
The authors declare that they have no competing interests.

Author details
[1]Department of Economics, Ghent University, Sint-Pietersplein 5, Ghent, Belgium. [2]Department of Mathematics, Royal Military Academy, Renaissancelaan 30, Brussels, Belgium. [3]Department of Physics and Astronomy, Ghent University, Proeftuinstraat 86, Ghent, Belgium.

References
1. Lwin MO, Panchapakesan C, Lee SY, Tandoc E (2023) Mainstream news medias role in public health communication during crises: assessment of coverage and correction of covid-19 misinformation. Health Commun 38(1):160–168. https://doi.org/10.1080/10410236.2021.1937842
2. Nestoras A, Cirju R (2021) The rise of China in the information domain? Measuring Chinese influence in Europe during the covid-19 pandemic. EU Policy Rev 1:199–201
3. Bentzen N (2020) Covid-19 foreign influence campaigns: Europe and the global battle of narratives
4. Roozenbeek J, Schneider CR, Dryhurst S, Kerr J, Freeman ALJ, Recchia G, van der Bles AM, van der Linden S (2020) Susceptibility to misinformation about covid-19 around the world. R Soc Open Sci 7(10):201199. https://doi.org/10.1098/rsos.201199
5. Imhoff R, Lamberty P (2020) A bioweapon or a hoax? The link between distinct conspiracy beliefs about the coronavirus disease (covid-19) outbreak and pandemic behavior. Soc Psychol Pers Sci 11(8):1110–1118. https://doi.org/10.1177/1948550620934692
6. Shahsavari S, Holur P, Wang T, Tangherlini TR, Roychowdhury V (2020) Conspiracy in the time of corona: automatic detection of emerging covid-19 conspiracy theories in social media and the news. J Comput Soc Sci 3(2):279–317. https://doi.org/10.1007/s42001-020-00086-5

7.   Wu W, Lyu H, Luo J (2021) Characterizing discourse about covid-19 vaccines: a reddit version of the pandemic story. Health Data Sci 2021:9837856. https://doi.org/10.34133/2021/9837856

8.   Ng R, Chow TYJ, Yang W (2021) News media narratives of covid-19 across 20 countries: early global convergence and later regional divergence. PLoS ONE 16(9):1–12. https://doi.org/10.1371/journal.pone.0256358

9.   Kemp S (2022) Digital 2022: Belgium - DataReportal – global digital insights. DataReportal – global digital insights. https://datareportal.com/reports/digital-2022-belgium. Accessed 13 April 2022

10. Hellemans J, Willems K, Brengman M (2021) The new adult on the block: daily active users of tiktok compared to Facebook, Twitter, and Instagram during the covid-19 crisis in Belgium. In: Martínez-López FJ, López López D (eds) Advances in digital marketing and eCommerce. Springer, Cham, pp 95–103

11. VRT NWS (2021) Defensie beschuldigt "Russische internettrollen" van inmenging in zaak-Jürgen Conings: waar gaat dit over? https://www.vrt.be/vrtnws/nl/2021/10/07/russische-trollen/. Accessed 20 May 2022

12. Caldarelli G, De Nicola R, Petrocchi M, Pratelli M, Saracco F (2021) Flow of online misinformation during the peak of the covid-19 pandemic in Italy. EPJ Data Sci 10(1):34. https://doi.org/10.1140/epjds/s13688-021-00289-4

13. Mattei M, Caldarelli G, Squartini T, Saracco F (2021) Italian Twitter semantic network during the covid-19 epidemic. EPJ Data Sci 10:47

14. Bamberg M (2012) Narrative analysis. In: APA handbook of research methods in psychology, vol 2. Research designs: quantitative, qualitative, neuropsychological, and biological. American Psychological Association. Washington, pp 85–102. https://doi.org/10.1037/13620-006

15. Briand SC, Cinelli M, Nguyen T, Lewis R, Prybylski D, Valensise CM, Colizza V, Tozzi AE, Perra N, Baronchelli A, Tizzoni M, Zollo F, Scala A, Purnat T, Czerniak C, Kucharski AJ, Tshangela A, Zhou L, Quattrociocchi W (2021) Infodemics: a new challenge for public health. Cell 184(25):6010–6014

16. Gallotti R, Valle F, Castaldo N, Sacco P, De Domenico M (2020) Assessing the risks of 'infodemics'in response to covid-19 epidemics. Nat Hum Behav 4(12):1285–1293. https://doi.org/10.1038/s41562-020-00994-6

17. Nazar S, Pieters T (2021) Plandemic revisited: a product of planned disinformation amplifying the covid-19 "infodemic". Front Public Health 9:649930. https://doi.org/10.3389/fpubh.2021.649930

18. Graham T, Bruns A, Angus D, Hurcombe E, Hames S (2021) #istandwithdan versus #dictatordan: the polarised dynamics of Twitter discussions about Victoria's covid-19 restrictions. Media Int Aust 179(1):127–148. https://doi.org/10.1177/1329878X20981780

19. Balcaen P, Buts C, Bois CD, Tkacheva O (2023) The effect of disinformation about covid-19 on consumer confidence: insights from a survey experiment. J Behav Exp Econ 102:101968. https://doi.org/10.1016/j.socec.2022.101968

20. Blei DM, Ng AY, Jordan MI (2003) Latent Dirichlet allocation. J Mach Learn Res 3:993–1022

21. Kobayashi R, Takedomi Y, Nakayama Y, Suda T, Uno T, Hashimoto T, Toyoda M, Yoshinaga N, Kitsuregawa M, Rocha LEC (2022) Evolution of public opinion on covid-19 vaccination in Japan: large-scale Twitter data analysis. J Med Internet Res 24(12):41928. https://doi.org/10.2196/41928

22. CCDH (2021) The disinformation dozen: why platforms must act on twelve leading anti-vaxxers. Technical report, The Center for Countering Digital Hate

23. Nogara G, Vishnuprasad PS, Cardoso F, Ayoub O, Giordano S, Luceri L (2022) The disinformation dozen: an exploratory analysis of covid-19 disinformation proliferation on Twitter. In: 14th ACM web science conference 2022. WebSci '22. Association for Computing Machinery, New York, pp 348–358. https://doi.org/10.1145/3501247.3531573

24. Mattei M, Pratelli M, Caldarelli G, Petrocchi M, Saracco F (2022) Bow-tie structures of Twitter discursive communities. Sci Rep 12(1):12944. https://doi.org/10.1038/s41598-022-16603-7

25. Memon SA, Carley KM (2020) CMU-MisCov19: a novel Twitter dataset for characterizing COVID-19 misinformation. Zenodo. https://doi.org/10.5281/zenodo.4024154

26. Cui L, Lee D (2020) CoAID: COVID-19 healthcare misinformation dataset. arXiv:2006.00885

27. Elhadad MK, Li KF, Gebali F (2021) Covid-19-fakes: a Twitter (Arabic/English) dataset for detecting misleading information on covid-19. In: Barolli L, Li KF, Miwa H (eds) Advances in intelligent networking and collaborative systems. Springer, Cham, pp 256–268

28. Lamsal R (2020) Design and analysis of a large-scale covid-19 tweets dataset. Appl Intell. https://doi.org/10.1007/s10489-020-02029-z

29. Twitter (2021) Twitter transparency reports. Technical report. https://transparency.twitter.com/en/reports/information-operations.html. Accessed 2021-11-10

30. De Clerck B, Rocha LEC, Van Utterbeeck F (2022) Maximum entropy networks for large scale social network node analysis. Appl Netw Sci 7(1):69

31. Ezzeddine F, Luceri L, Ayoub O, Sbeity I, Nogara G, Ferrara E, Giordano S (2022) Characterizing and detecting state-sponsored troll activity on social media. arXiv:2210.08786

32. Sharma K, Zhang Y, Ferrara E, Liu Y (2021) Identifying coordinated accounts on social media through hidden influence and group behaviours. In: Proceedings of the 27th ACM SIGKDD conference on knowledge discovery & data mining. KDD '22. Association for Computing Machinery, New York, pp 1441–1451. https://doi.org/10.1145/3447548.3467391

33. Summers E, Brigadir I, Hames S, van Kemenade H, Binkley P, Tinafigueroa, Ruest N, Walmir, Chudnov D, Betsy, Chartier R, Celeste, Lin H, Alice, Chosak A, Lenz M, McCain RM, Milligan I, Segerberg A, Shahrokhian D, Walsh M, Lausen L, Woodward N, Eggplants, Ramaswami A, Nguyen B, Hereñú D, Milajevs D, Elwert F, Westerling K DocNow/twarc: V2.13.0. https://doi.org/10.5281/zenodo.7484102

34. Devlin J, Chang M-W, Lee K, Toutanova K (2019) BERT: pre-training of deep bidirectional transformers for language understanding. In: Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, vol 1. Association for Computational Linguistics, Minneapolis, pp 4171–4186. (Long and Short Papers). https://doi.org/10.18653/v1/N19-1423

35. Brown TB, Mann B, Ryder N, Subbiah M, Kaplan J, Dhariwal P, Neelakantan A, Shyam P, Sastry G, Askell A, Agarwal S, Herbert-Voss A, Krueger G, Henighan T, Child R, Ramesh A, Ziegler DM, Wu J, Winter C, Hesse C, Chen M, Sigler E, Litwin M, Gray S, Chess B, Clark J, Berner C, McCandlish S, Radford A, Sutskever I, Amodei D (2020) Language models are few-shot learners. CoRR. arXiv:2005.14165

36. Reimers N, Gurevych I (2019) Sentence-bert: sentence embeddings using Siamese bert-networks. In: Proceedings of the 2019 conference on empirical methods in natural language processing. Association for Computational Linguistics, Minneapolis. arXiv:1908.10084

37. Reimers N, Gurevych I (2021). Sentence-transformers/paraphrase-multilingual-mpnet-base-v2. https://huggingface.co/sentence-transformers/paraphrase-multilingual-mpnet-base-v2. Accessed 10 June 2022

38. OpenAI (2022) New and improved embedding model. https://openai.com/blog/new-and-improved-embedding-model. Accessed 5 Jan 2023

39. Reimers N, Gurevych I (2020) Making monolingual sentence embeddings multilingual using knowledge distillation. CoRR. arXiv:2004.09813

40. Guarino S, Pierri F, Di Giovanni M, Celestini A (2021) Information disorders during the covid-19 infodemic: the case of Italian Facebook. Online Soc Netw Media 22:100124. https://doi.org/10.1016/j.osnem.2021.100124

41. Pratelli M, Saracco F, Petrocchi M (2023) Entropy-based detection of Twitter echo chambers. arXiv:2308.01750

42. Squartini T, Garlaschelli D (2011) Analytical maximum-likelihood method to detect patterns in real networks. New J Phys 13(8):083001. https://doi.org/10.1088/1367-2630/13/8/083001

43. Squartini T, Mastrandrea R, Garlaschelli D (2015) Unbiased sampling of network ensembles. New J Phys 17(2):023052. https://doi.org/10.1088/1367-2630/17/2/023052

44. Squartini T (2017) Maximum-entropy networks: pattern detection, network reconstruction and graph combinatorics. Springer, Cham

45. Garlaschelli D, Loffredo MI (2008) Maximum likelihood: extracting unbiased information from complex networks. Phys Rev B **78**(1). https://doi.org/10.1103/physreve.78.015101

46. Saracco F, Di Clemente R, Gabrielli A, Squartini T (2015) Randomizing bipartite networks: the case of the world trade web. Sci Rep 5(1):10595. https://doi.org/10.1038/srep10595

47. Saracco F, Straka MJ, Clemente RD, Gabrielli A, Caldarelli G, Squartini T (2017) Inferring monopartite projections of bipartite networks: an entropy-based approach. New J Phys 19(5):053022. https://doi.org/10.1088/1367-2630/aa6b38

48. Vallarano N, Bruno M, Marchese E, Trapani G, Saracco F, Cimini G, Zanon M, Squartini T (2021) Fast and scalable likelihood maximization for exponential random graph models with local constraints. Sci Rep 11(1):15227. https://doi.org/10.1038/s41598-021-93830-4

49. Pacheco D, Flammini A, Menczer F (2020) Unveiling coordinated groups behind white helmets disinformation. CoRR. arXiv:2003.01313

50. Pacheco D, Hui P, Torres-Lugo C, Truong BT, Flammini A, Menczer F (2020) Uncovering coordinated networks on social media. CoRR. arXiv:2001.05658

51. Suresh VP, Nogara G, Cardoso F, Cresci S, Giordano S, Luceri L (2023) Tracking fringe and coordinated activity on Twitter leading up to the US Capitol attack. arXiv:2302.04450

52. Majó-Vázquez S, Congosto M, Nicholls T, Nielsen RK (2021) The role of suspended accounts in political discussion on social media: analysis of the 2017 French, UK and German elections. Soc Media Soc **7**(3). https://doi.org/10.1177/20563051211027202

53. Yang K-C, Varol O, Hui P-M, Menczer F (2020) Scalable and generalizable social bot detection through data selection. Proc AAAI Conf Artif Intell 34(01):1096–1103. https://doi.org/10.1609/aaai.v34i01.5460

54. Cresci S, Di Pietro R, Petrocchi M, Spognardi A, Tesconi M (2017) The paradigm-shift of social spambots: evidence, theories, and tools for the arms race. In: Proceedings of the 26th international conference on world wide web companion, pp 963–972

55. Pasricha N, Hayes C (2019) Detecting bot behaviour in social media using digital dna compression. In: Irish conference on artificial intelligence and cognitive science

56. Yang K-C, Pierri F, Hui P-M, Axelrod D, Torres-Lugo C, Bryden J, Menczer F (2021) The covid-19 infodemic: Twitter versus Facebook. Big Data Soc 8(1):20539517211013861. https://doi.org/10.1177/20539517211013861

57. Wagnsson C (2023) The paperboys of Russian messaging: rt/sputnik audiences as vehicles for malign information influence. Inf Commun Soc 26(9):1849–1867. https://doi.org/10.1080/1369118X.2022.2041700

58. Furman IO, Gürel KB, Sivaslıoğlu FB (2023) "As reliable as a kalashnikov rifle": how sputnik news promotes Russian vaccine technologies in the Turkish twittersphere. Soc Media Soc 9(1):20563051221150418. https://doi.org/10.1177/20563051221150418

59. Council of the EU (2022) EU imposes sanctions on state-owned outlets RT/Russia Today and Sputnik's broadcasting in the EU. https://www.consilium.europa.eu/en/press/press-releases/2022/03/02/eu-imposes-sanctions-on-state-owned-outlets-rt-russia-today-and-sputnik-s-broadcasting-in-the-eu/. Accessed 05 May 2022

60. Traag VA, Waltman L, van Eck NJ (2018) From Louvain to Leiden: guaranteeing well-connected communities. CoRR. arXiv:1810.08473

61. Muric G, Wu Y, Ferrara E (2021) Covid-19 vaccine hesitancy on social media: building a public Twitter data set of antivaccine content, vaccine misinformation, and conspiracies. JMIR Public Health Surveill 7(11):30642. https://doi.org/10.2196/30642

62. Rawnsley A (2020) Right-wing media outlets duped by a Middle East propaganda campaign. https://www.thedailybeast.com/right-wing-media-outlets-duped-by-a-middle-east-propaganda-campaign. Accessed 20 Aug 2023

63. Land Information Manoeuvre Centre (2020) Weekly update des-/misinformatie. Technical report, LIMC. https://open.overheid.nl/documenten/ronl-d6520cb2-8c35-42b6-85f6-823123c88a65/pdf. Accessed Jan 20, 2024

64. Yablokov I (2015) Conspiracy theories as a Russian public diplomacy tool: the case of Russia Today (RT). Polit 35(3–4):301–315. https://doi.org/10.1111/1467-9256.12097

65. Elswah M, Howard PN (2020) "Anything that causes chaos": the organizational behavior of Russia Today (RT). J Commun 70(5):623–645. https://doi.org/10.1093/joc/jqaa027

66. Wilson D (2021) Why fed-up Belgians are backing heavily armed 'Flemish Rambo' fugitive as a symbol of anti-establishment frustration. https://www.rt.com/op-ed/526763-flemish-rambo-jurgen-conings-belgium/. Accessed 11 Feb 2022

67. Nizzoli L, Tardelli S, Avvenuti M, Cresci S, Tesconi M, Ferrara E (2020) Charting the landscape of online cryptocurrency manipulation. CoRR. arXiv:2001.10289

68. Batzdorfer V, Steinmetz H, Biella M, Alizadeh M (2022) Conspiracy theories on Twitter: emerging motifs and temporal dynamics during the covid-19 pandemic. Int J Data Sci Anal 13(4):315–333. https://doi.org/10.1007/s41060-021-00298-6

69. Stephenson GV (1972) Cultural regionalism and the unitary state idea in Belgium. Geogr Rev 62(4):501–523. Accessed 2024-01-29
70. Oleinik A (2024) Telegram channels covering Russia's invasion of Ukraine: a comparative analysis of large multilingual corpora. J Comput Soc Sci. https://doi.org/10.1007/s42001-023-00240-9
71. Metaxas PT, Mustafaraj E, Wong K, Zeng L, O'Keefe M, Finn S (2015) What do retweets indicate? Results from user survey and meta-review of research. In: International conference on web and social media
72. Putra JWG, Tokunaga T (2017) Evaluating text coherence based on semantic similarity graph. In: Proceedings of TextGraphs-11: the workshop on graph-based methods for natural language processing. Association for Computational Linguistics, Vancouver, pp 76–85. https://doi.org/10.18653/v1/W17-2410
73. Antognini D, Faltings B (2019) Learning to create sentence semantic relation graphs for multi-document summarization. arXiv:1909.12231
74. Park CS, Kaye BK (2019) Expanding visibility on Twitter: author and message characteristics and retweeting. Soc Media Soc **5**
75. van Vliet L, Törnberg P, Uitermark J (2020) The Twitter parliamentarian database: analyzing Twitter politics across 26 countries. PLoS ONE 15(9):1–24. https://doi.org/10.1371/journal.pone.0237073
76. van Vliet L, Chueri J, Törnberg P, Uitermark J (2023) Political groups over national parties: measuring the europeanization of the political arena through meps Twitter interactions. Party Polit 30(3):479–492. https://doi.org/10.1177/13540688231158486.
77. Perra N, Rocha LEC (2019) Modelling opinion dynamics in the age of algorithmic personalisation. Sci Rep 9(1):7261. https://doi.org/10.1038/s41598-019-43830-2
78. Chen L, Chen J, Xia C (2022) Social network behavior and public opinion manipulation. J Inf Secur Appl 64:103060. https://doi.org/10.1016/j.jisa.2021.103060
79. Varshney D, Vishwakarma DK (2022) A unified approach of detecting misleading images via tracing its instances on web and analyzing its past context for the verification of multimedia content. Int J Multimed Inf Retr 11(3):445–459. https://doi.org/10.1007/s13735-022-00235-8

Publisher's Note