Preprints.org

Article

Not peer-reviewed version

Identifying Key Energy Influencers on Twitter: A Multiplex Network Analysis Using Graph Traversal Techniques

Vincenzo De Leo * , Michelangelo Puliga , Martina Erba , Cesare Scalia , Andrea Filetti , Alessandro Chessa

Posted Date: 12 March 2026

doi: 10.20944/preprints202603.0942.v1

Keywords: social network analysis; multiplex networks; Twitter; energy sector; machine learning; community detection; node classification; climate change

Article

Identifying Key Energy Influencers on Twitter: A Multiplex Network Analysis Using Graph Traversal Techniques

Vincenzo De Leo [1,*], Michelangelo Puliga [1], Martina Erba [2], Cesare Scalia [2], Andrea Filetti [2] and Alessandro Chessa [1,3]

[1] Linkalab, CoMPLeX SySTeMS CoMPuTaTioNaL LaBoRaToRy, Viale Elmas, 142, 09122, Cagliari (Italy)
[2] Eni S.p.A, Piazzale Enrico Mattei, 1, 00144, Rome (Italy)
[3] Data Lab, Luiss University, Viale Romania, 32, 00197, Rome (Italy)
* Correspondence: vincenzo.deleo@linkalab.it

Abstract

In this work, we inspected the friendship network on Twitter (recently rebranded as X), concentrating on individuals and organizations intertwined with the energy field. We particularly focus on seasoned professionals, corporate entities, and domain specialists, all connected through 'following' relationships. By meticulously examining these ties, we uncover several distinct groupings within the network, each defined by the unique roles its members occupy. Our analysis demonstrates that the natural emergence of such clusters on social platforms exerts a profound influence on public discourse regarding energy and other critical matters, including climate change. Furthermore, we reveal that the ever-changing interplay of misleading information catalyzes the formation of ideologically divided factions, which often leads to reduced engagement in online conversations. These emergent clusters, characterized by their shared communication styles, form relatively compact communities where the exchange of information is infrequent compared to larger networks and is usually confined to accounts created for specific commercial objectives. Additionally, by leveraging a machine learning approach, we are able to pinpoint pivotal actors within these niche segments and elucidate the mechanisms that sustain their connectivity. This method provides novel insights into how corporate communication unfolds on social media, offering a refreshed perspective on professional networking. Ultimately, our findings highlight the ways in which companies within the energy sector take advantage of Twitter to coordinate their initiatives, with key institutions serving as central nodes in maintaining the organization of these networks.

Keywords: social network analysis; multiplex networks; Twitter; energy sector; machine learning; community detection; node classification; climate change

---

1. Introduction

In recent times, digital social platforms have evolved into strategic assets across a wide array of fields including advertising, marketing, retail, e-commerce, as well as education, scholarly research, policymaking, and beyond. These context offer a shared space where corporations, public institutions, and individuals are free to discuss about their opinions in an effort to affect public debate. Nonetheless, the incessant torrent of content, often originating from sources whose credibility is questionable, is neither curated nor prioritized, resulting in a conflation of accurate data with deceptive assertions. In the absence of a systematic mechanism for content verification, this environment becomes fertile ground for the rapid dissemination of misinformation, thereby leaving news vulnerable to misinterpretation and potential exploitation.

This study introduces an innovative methodology aimed at deciphering online discussions centered on energy and climate change, topics that have critical global significance. By delving into

 doi:10.20944/preprints202603.0942.v1

Twitter networks through the lens of Machine Learning (ML) techniques, our research illuminates the pathways through which both valid and erroneous information propagate, the influence exerted by clusters of experts on shaping public opinion, and the dual role that digital platforms play in either obstructing or enhancing public comprehension of multifaceted issues.

Machine Learning, a branch within the expansive field of artificial intelligence, comprises algorithms and statistical models that empower computers to execute tasks without explicit programming by leveraging pattern recognition and inferential logic. This capability is essential for searching through vast datasets to unearth latent patterns or forecast upcoming trends. In social networks such as Twitter, ML is applied to categorize nodes (whether individual accounts or organizations) and their interconnections, thereby revealing the formation of clusters or subgraphs (i.e. segments of a larger graph that include a subset of nodes along with all the connections that interlink them) based on common interests or interactions.

Twitter, by design, encourages the exploration of media-generated information via a connectivity mechanism that allows users to follow others and access their posts without necessitating reciprocal relationships. This distinctive configuration enables one to cultivate a broad audience by selectively subscribing to updates from chosen contacts, rather than engaging in mutual following. We exploited this feature to investigate how professionals within the energy sector establish connections, utilizing a technique that traces and follows these ties to construct a network of industry-centric dialogues, termed the Friendship Network. In a manner analogous to studies on global health organization networks [1], this method allows for the identification of pivotal individuals and clusters in the energy debate, using technology to select relevant participants for a deeper analysis.

Once the relationships among Twitter users with an interest in energy topics were charted and the data refined using an ML classifier, we turned our attention to a detailed subnetwork comprising thousands of nodes. This refined subgraph was subsequently scrutinized for further connections, with a particular focus on user mentions. Essentially, our investigation considered two distinct forms of connectivity that together form a layered, complex network, where the same users might interact via different modalities. This approach enabled us to discern discrete clusters within the overall network, each characterized by its own unique communication patterns, such as the frequency of tweets, mentions, or retweets.

The principal aim of our work is to unravel the internal dynamics of Twitter communities that engage in discussions about energy and climate change. By leveraging machine learning to classify nodes and to comprehend how these nodes coalesce into specialized subgraphs or communities around these critical topics, we address issues that are both urgent and significant. Climate change, one of the most important challenges of our era, affects global ecosystems, economies, and societies, while the transition in energy practices plays an equally vital role in mitigating its impacts, thus calling for widespread public engagement and judicious policymaking.

Our analysis further uncovers a spectrum of communication styles across these clusters. For instance, energy institutions frequently emerge as central connectors among various smaller groups, whereas companies tend to exhibit more reserved, indirect interactions with one another. This methodology offers a nuanced perspective on the dynamics of discussions in social networks, casting light on the intricate web of relationships among institutions, corporations, experts, and journalists invested in energy topics. Each sub-group within the larger network manifests its own distinct manner of engaging and disseminating information, underscoring the rich diversity of dialogue in the digital sphere of energy discussions.

2. Background

Social media has been shown to amplify misinformation, creating echo chambers where like-minded users exchange biased views [2–4]. Although original content is available, its interpretation often becomes distorted [5]. Combined with natural cognitive biases, this information deluge fosters polarization and a competitive struggle for attention, as noted by Bessi et al. [6] and Weng et al. [7].

The overwhelming flow of notifications can lead to "information overload", causing users to disengage from discussions, a phenomenon documented by Gunaratne et al. [8] and linked to significant economic losses, such as the 650 billion USD reported in 2007 [9]. Specifically, empirical observations suggest that when users receive more than 30 notifications per hour, they tend to disengage from discussions. This pattern of 'information overload' is a critical factor that compromises both communication quality and marketing effectiveness within the energy discourse. Algorithmic filtering further intensifies these echo chambers, making it challenging to access high-quality content and underscoring the need to amplify authoritative voices.

Advanced computational methods such as machine learning for node classification and graph analysis have been used effectively to decode the structure of social networks [10,11]. These techniques help identify key influencers and organizations within Twitter communities, a theme reinforced by studies on community detection [12,13]. Researchers such as Fortunato [14] and Rani et al. [15] have demonstrated how grouping users by shared interests or interactions via mentions, hashtags, or mutual follows, can reveal "communities of practice"[1].

Modern analyses use multilayer graphs to capture the complexity of these interactions, providing a nuanced picture of how communities form and communicate. Recent research has also highlighted the increase in polarization in climate change discussions, with studies showing that events like the COP summits intensify ideological divides [16]. Multidimensional sentiment analysis further illustrates the varied global perceptions of climate change [17], while the strategic framework of fossil fuel companies adds another layer to online discourse [18].

Our work builds on these insights by moving beyond traditional community detection. We incorporate node classification and multiplex network analysis [19], which examines how users interact across different contexts such as public mentions versus direct connections. This approach, which also finds application in fields ranging from biology [20] to societal studies [21] and transportation [22], allows us to assess user importance through measures like multidegree and to identify cohesive groups that persist across multiple interaction layers [23].

3. Materials and Methods

A key novelty of our approach lies in constructing the expert user network not by extracting interactions (such as mentions, retweets, or replies) from tweets filtered through a topic-specific Bag-of-Words (BOW) method, but rather by anchoring the network on an initial set of verified users, recognized as primary influencers in the subject area, as shown in Figure 1.

[1] https://scaledagileframework.com/communities-of-practice/

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 March 2026 doi:10.20944/preprints202603.0942.v1

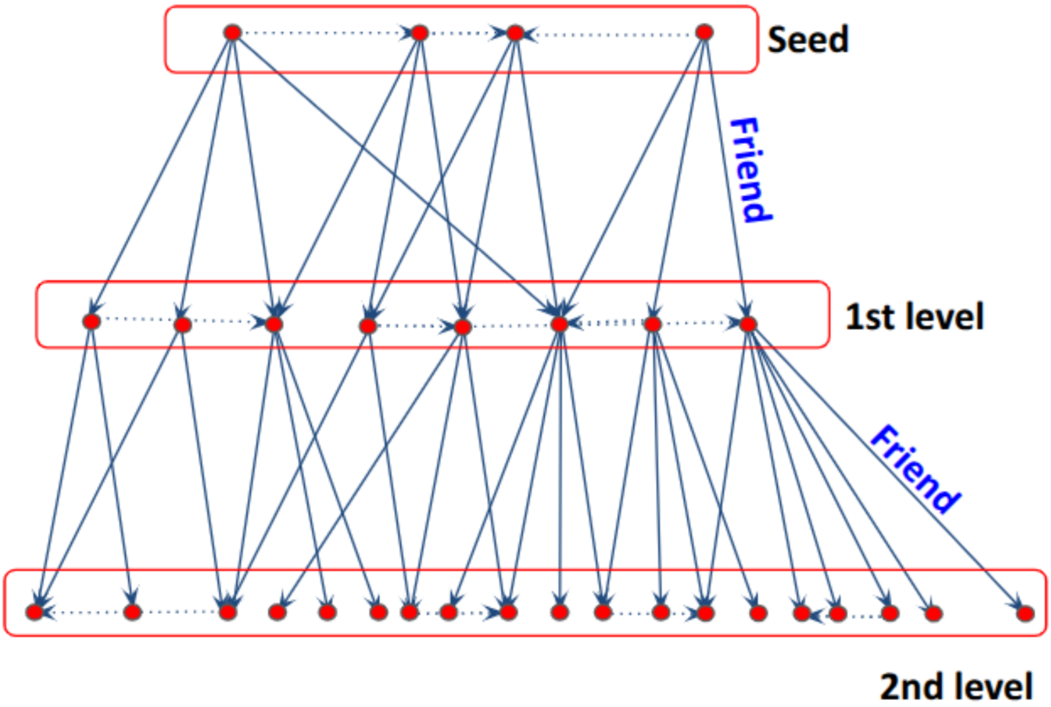

Figure 1. Friendship Network generation method.

From this seed of verified accounts, we used the Twitter API to download the profile descriptions of their immediate "Friends." These profiles were then manually evaluated and sorted based on their relevance to the discussion topics at hand. The manually labeled data subsequently served as the training set for a supervised Machine Learning model, which automatically classified the profile descriptions of the second-tier Friends, that is, those connected to the already identified first-level users engaged in the pertinent discussions.

The final network is thus organized into three tiers: the seed (comprising tens of users), a manually curated first-level group (numbering in the hundreds), and a second-level group (automatically classified by our ML model, consisting of thousands of users).

After assembling this layered network, we retrieved the timelines of all users across the three tiers. This dataset allowed us to analyze discussion themes, investigate community structures based on the graph's topological features, examine tweet content, identify peaks in communication volumes over time, and more.

Further methodological details are available in the supplementary material[2].

3.1. Data Preparation

To construct a network of high-quality users focused on a specific subject, we began by carefully selecting an initial seed list of accounts, detailed in the supplementary material[3] using the most rigorous criteria available.

Considering the discussion's context, the seed was derived from a roster of Italian energy sector companies. This list, generated via Bloomberg terminal queries, provided 32 companies. We then used the Twitter API to retrieve the corresponding user profile descriptions.

A manual filtering process followed, where we screened these profiles for keywords associated with energy and the green economy. This preliminary refinement narrowed the seed to 14 users, as outlined in the supplementary material.

[2] https://tinyurl.com/twcomm-supplementary-material
[3] https://tinyurl.com/twcomm-supplementary-material

 doi:10.20944/preprints202603.0942.v1

With the seed accounts established, we executed Twitter API calls using get_friends_list() to gather the complete set of friends for these users, resulting in 740 unique accounts (see supplementary material for further details).

Next, we manually classified these 740 users by referencing the seed users' descriptions to extract relevant keywords (e.g., energy, gas, methane, fuel). This process yielded a subset of 210 English-language users deemed pertinent to energy discussions (details in the supplementary material).

Building on this, we employed additional get_friends_list() API calls on the positively classified first-level Friends, collecting over 60,000 unique second-level accounts. With these manually curated datasets in hand, we then developed a Machine Learning model to automatically classify the second-level Friends based on their Twitter profile descriptions.

To determine the most accurate classifier for our needs, we conducted a GridSearch analysis over several ML algorithms, including:

- KNeighbors: An instance-based learner that predicts based on the majority vote among its k nearest neighbors.
- SGD: An SGDClassifier optimized via stochastic gradient descent, well-suited for large-scale datasets.
- DecisionTree: A model that segments the feature space into a hierarchy of decision nodes for interpretable results.
- GradientBoosting: An iterative technique where successive models correct the errors of their predecessors.
- RandomForest: An ensemble method that aggregates multiple decision trees to boost accuracy and robustness.
- Multi-Layer Perceptron: A deep neural network capable of capturing complex non-linear relationships with its multiple hidden layers.

The DecisionTree classifier not only achieved the highest accuracy (0.80) but was also selected for its ability to handle the hierarchical structure of decision nodes, which is particularly suitable for interpreting user profile descriptions. A subsequent manual quality check on 100 random samples confirmed that 95% of the accounts were perfectly aligned with the expected energy-related profiles, with only 1% being clearly irrelevant.

This optimized model was applied to the second-level Friends dataset, automatically classifying these accounts based on their profile content. The outcome identified 4,620 users as positive.

A manual quality check was conducted on a random sample of 100 users from the ML-classified set by reviewing their profile descriptions and their five most recent tweets. The analysis showed that 95% of the accounts were perfectly aligned with the expected profiles, 4% were ambiguous, and only 1% did not match the criteria.

In total, the final positive dataset comprises 4,844 non-unique users: 14 from the seed, 210 from the manually curated first-level Friends, and 4,620 from the automatically classified second-level Friends.

More details about data preparation can be found in the supplementary material[4].

3.2. Networks Construction

Leveraging the datasets described earlier, we constructed an interaction network where users are linked by their "Friend" status, that is, one user appearing in another's friend list. This criterion yielded a network of 4,844 unique users interconnected by 13,806 links.

In addition to this Friendship Network, we retrieved the timelines of all these users to develop a second network based on mentions and retweets. This Mentions Network's largest connected component comprises 2,094 nodes and 32,824 links. Although the starting nodes were the same as those in the Friendship Network, this approach also produced several minor isolated clusters, which were excluded from our analysis due to their limited analytical value.

[4] https://tinyurl.com/twcomm-supplementary-material

This dual-network strategy allowed us to compare centrality metrics and community structures across both networks. For example, as shown in Figure 2 and Figure 3, the Friendship Network's degree distribution aligns more closely with a power-law trend, leading to the formation of hubs, a point discussed further in the Results section.

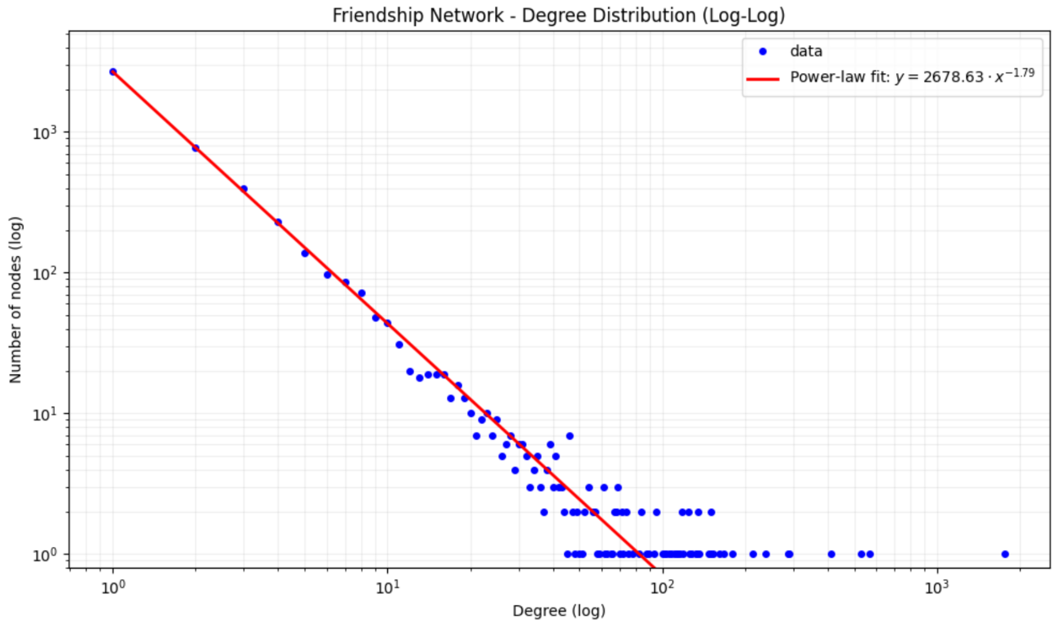

Figure 2. Degree distribution of Friendship Network.

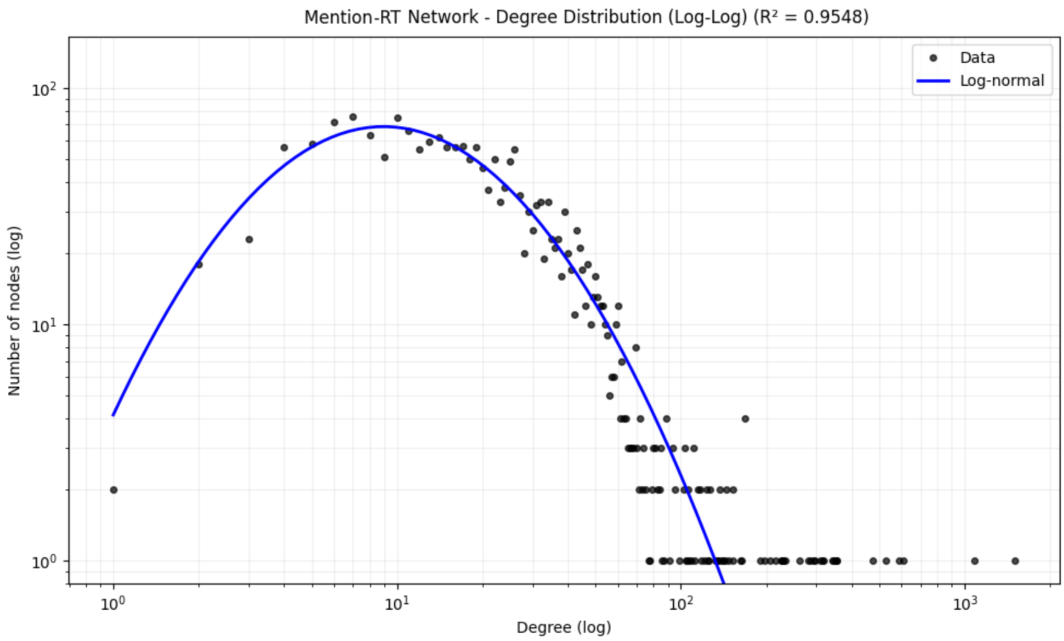

Figure 3. Degree distribution of Mentions Network.

Additional information regarding network construction is available in the supplementary material[5].

3.3. Users Classification Based on Company, Institution, and Expert User Labels

To better understand the role of these hubs, we developed an additional classifier that distinguishes among four node categories: institutions, companies, experts, and others. Using manual

[5] https://tinyurl.com/twcomm-supplementary-material

 doi:10.20944/preprints202603.0942.v1

training, our ML system learned to recognize these classes and assign labels accordingly. The data reveal that most hubs fall into the "institution" category, as evidenced by the higher average degree reported in Table 1.

Table 1. The users classified by role with their average measures.

	Twitter(X) Metadata			Friendship Network Analysis
	avg followers	avg friends	avg statuses	avg degree
Company	13k	1.0k	4k	7
Expert User	10k	1.5k	8k	4
Institution	17k	1.5k	7k	12
Other	21k	1.6k	10k	6

We prepared the data using a Stratified K-Folds cross-validator to generate train/test splits, a necessary step due to the imbalance between classes. The classifier was built on a Naive Bayes model, particularly effective for imbalanced datasets, and used user descriptions as the primary feature. A subsequent community detection analysis showed that in the Friendship Network, distinct communities emerged for each category (Institutions, Companies, etc.), confirming well-differentiated roles and communication styles.

Furthermore, the Mentions Network is heavily influenced by institutions (see supplementary material[6] for details), reinforcing the observation that institutions serve as the primary connectors among other users and drive the overall discussion.

4. Results

Both the Friendship and Mentions networks were examined using traditional complex network techniques alongside modern multiplex approaches. In Figure 4, the Friendship Network is depicted with hubs and community clusters highlighted by distinct node colors, while Figure 5 presents the Mentions Network with similar visual markers. The fundamental concept is that user connections can be captured through both direct friendship links (unweighted) and mentions (weighted by frequency). The weighted Mentions Network, by filtering out weaker ties, helps identify users who are highly engaged in conversation, so-called active nodes[7].

[6] https://tinyurl.com/twcomm-supplementary-material
[7] An active node is a user that engages more frequently with others through actions such as sending messages, posting updates, or contributing to discussions. Active nodes often play crucial roles in disseminating information, influencing others, and keeping the network dynamic. Their activity can help sustain the network's vibrancy and facilitate the spread of ideas, knowledge, and practices within communities, including CoPs

doi:10.20944/preprints202603.0942.v1

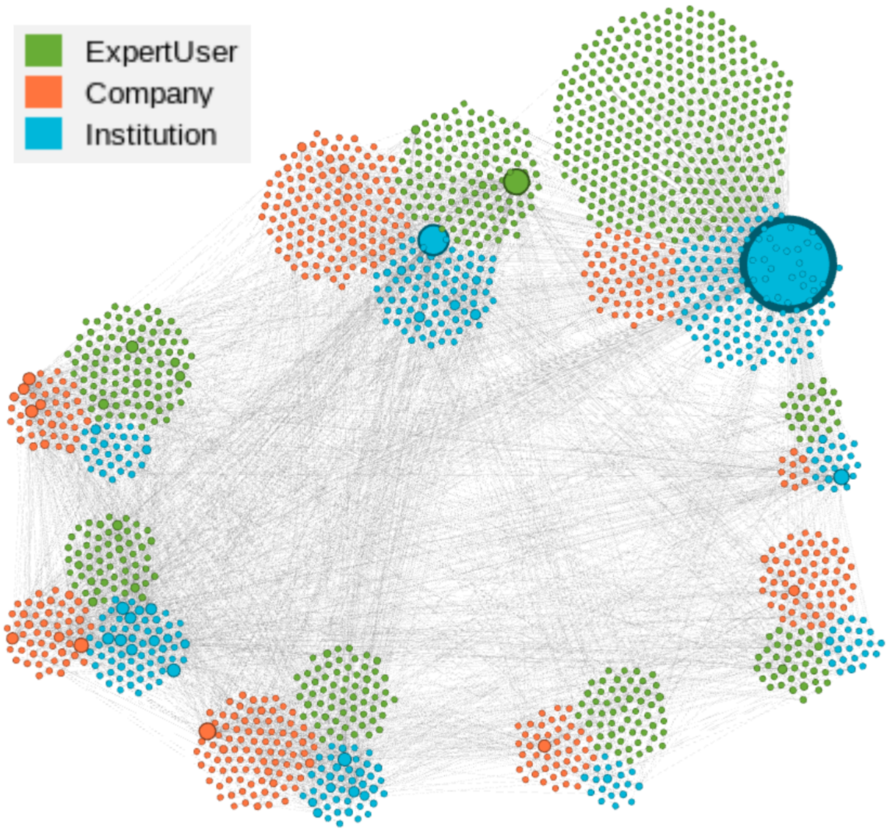

Figure 4. The Friendship Network.

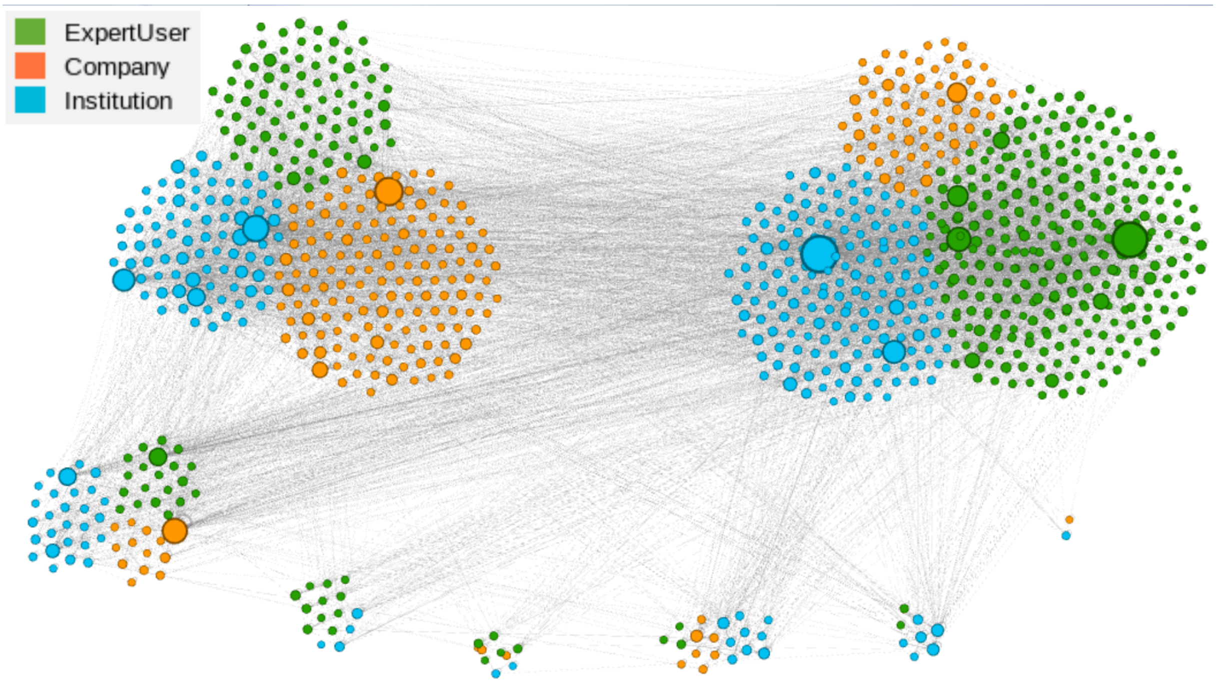

Figure 5. The mentions and retweets network.

In brief, our analysis yielded a Friendship Network comprising roughly 5k nodes and 14k links, compared to a Mentions Network with approximately 2k nodes and 32k links. Notably, only about 5% of the links are shared between the two networks, suggesting that merely following another user does not necessarily imply active engagement in discussions, at least within this Energy community of practice. One key analytical tool is the degree distribution; by fitting its tail to a power law of the form $y = x^{-\alpha}$, we can assess whether the network follows a power-law pattern, a signature of hubs. As

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 March 2026 doi:10.20944/preprints202603.0942.v1

detailed in [24], while strongly scale-free networks are rare, most networks are better characterized by log-normal distributions, making our Friendship Network a strong benchmark for energy discussions.

While the Friendship Network's degree distribution closely aligns with a power-law (exponent of 1.8, $R^2 = 0.99$), indicating a scale-free structure, the Mentions and Retweet Network does not exhibit this behavior. Instead, it fits a log-normal distribution ($R^2 = 0.95$), suggesting that conversational engagement follows different growth dynamics compared to the establishment of formal follow relationships.

These hubs, with their vast numbers of connections, serve as primary broadcasters of information, ensuring messages reach wide audiences. To further elucidate their role, we developed a classifier that segments nodes into four categories: "institutions", "companies", "experts", and "others". After training the model manually, Table 1 shows that the majority of hubs fall under the "institution" label and exhibit high average degrees. This dominance of institutions implies that they act as central connectors, linking companies and expert users, a list of the most connected nodes is provided in Table 2.

Table 2. The most important nodes by degree and their role.

twitter account	degree	reference
EU_ENV	1771	Institution
GreenCogEU	530	Institution
arikring	411	Expert User
EnergyDemand	292	Institution
EUEnvironment	288	Institution
euenergyweek	239	Institution
DrJoeNyangon	214	Expert User
Albanian_Energy	181	Company
Eurogas_Eu	168	Company
Climate_Rescue	152	Other
EU_ManagEnergy	152	Institution
Ed_Crooks	151	Expert User

We also conducted community detection on a composite multiplex network integrating both friendship and mention links (see Figure 6). The detailed community detection methodology is provided in the supplementary material[8]. Summary statistics (Table 3) and link density data (Table 4) reveal, for instance, that community "3" (comprising 1,062 nodes) has a high average degree, indicative of many institutions (making up 21% of total nodes) and a balanced mix of experts and companies. This community forms the dense core of business and institutional interactions, while other communities show lower connectivity. Interestingly, community "2" has the highest proportion of expert users yet a lower average degree, suggesting that experts tend to be less interconnected and instead vie for user attention. The cross-density analysis reveals a significant directional asymmetry in communication. For instance, the interaction density from 'Institution' to 'Other' ($36 \cdot 10^{-4}$) is higher than the reverse 'Other' to 'Institution' ($27 \cdot 10^{-4}$). This reflects the proactive nature of institutional outreach, where official entities act as primary information broadcasters rather than mere participants in bidirectional conversations.

[8] https://tinyurl.com/twcomm-supplementary-material

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 March 2026 doi:10.20944/preprints202603.0942.v1

Table 3. Community detection on multiplex network, with the social impact of each community with at least 10 nodes each.

	Twitter (X) Metadata			Friendship Network Analysis	Roles Distribution Inside Community				
community	avg foll.	avg friends	avg tweets	avg deg.	Expert %	Comp. %	Instit. %	Oth. %	Nodes
1	15k	1.7k	10k	7	28.9	25.2	12.2	33.7	5273
2	14k	900	3.6k	4	38.5	5.6	15.1	40.9	1475
3	15k	1.2k	4.3k	12	27.9	26.6	20.9	24.6	1062
4	18k	1.5k	7.4k	4	34.5	3.7	9.1	52.7	804
5	26k	800	3.8k	4	21.5	9.1	23.8	45.7	265
6	5k	1.2k	4.2k	2	11.4	6.3	13.9	68.4	158
7	1k	1.8k	1.8k	2	44.4	11.1	11.1	33.3	18
8	8k	800	4.7k	3	16.7	0.0	0.0	83.3	12

Table 4. The density of links between the classified nodes, allows repetition as we have two equivalent layers of the multiplex network.

source	target	cross density ($\times 10^{-4}$)
Institution	Institution	54
Institution	Other	36
Other	Other	33
Institution	ExpertUser	32
Institution	Company	28
Other	Institution	27
Other	ExpertUser	27
Company	Institution	17
Company	Company	17
Other	Company	14
ExpertUser	ExpertUser	13
Company	Other	12
ExpertUser	Other	12
ExpertUser	Institution	11
Company	ExpertUser	10
ExpertUser	Company	7

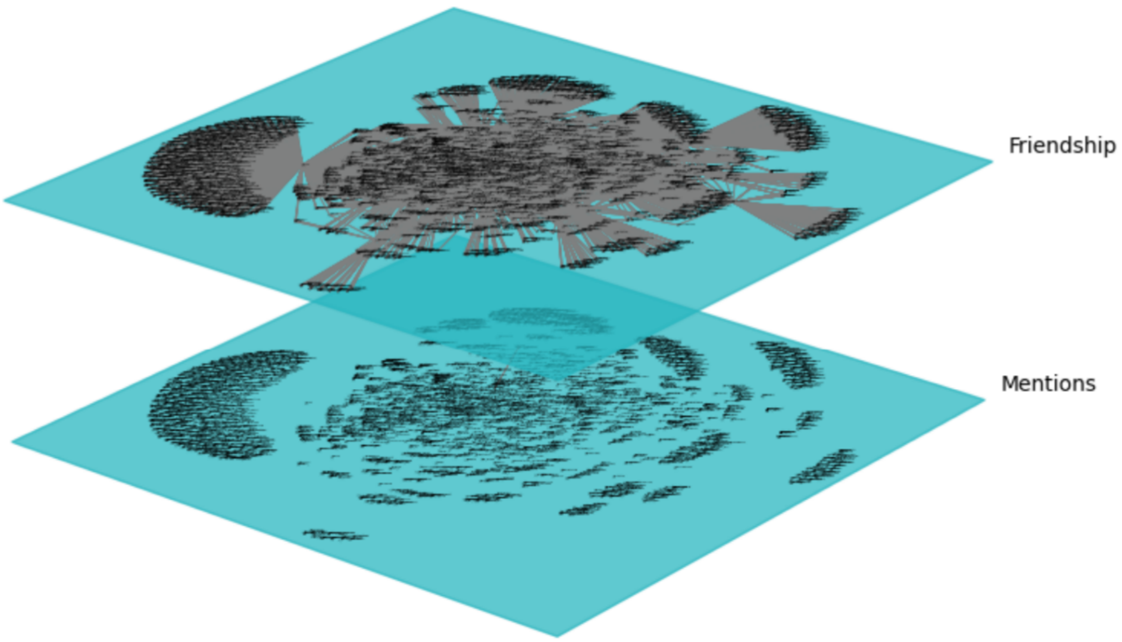

Figure 6. The Multiplex network.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 12 March 2026 doi:10.20944/preprints202603.0942.v1

These findings are consistent with existing literature on energy and climate change discussions. For example, Cortés P.A. and Quiroga R. [25] (2023) compared platforms and underscored Twitter's unique role in real-time engagement, while Pearce W. et al. [26] (2014) demonstrated how major climate events can catalyze community formation. Dellmuth L. and Shyrokykh K. [27] (2023) expanded on the policy implications of Twitter data, and Fownes J. R. et al. [28] (2018) earlier highlighted Twitter's influence on public opinion. Additional insights from Erokhin and Komendantova [29] (2023), Kolic et al. [30] (2022), and Torricelli et al. [31] (2023) further enrich our understanding of how external events and active node classifications shape online discourse.

Moreover, our approach extends prior work on active node classification and multiplex network analysis. Compared to Laishram et al. [32] (2019), our study distinguishes itself by applying ML to pinpoint active nodes, thereby enhancing insights into community formation and interaction patterns. Similarly, while Himelboim et al. [33] (2017) explored Twitter's structural dynamics, our emphasis on active participation offers a deeper look at how individuals influence information spread. The socioeconomic and political perspectives provided by Abitbol et al. [34] (2021) and Orhan et al. [35] (2023) are complemented by our focus on communities of practice, an approach also echoed by Li and Chang [36] (2023) in identifying influential nodes in misinformation campaigns. Hanteer and Rossi [37] (2019) also share thematic similarities, yet our multiplex framework provides a more dynamic perspective on user engagement and information flow.

The application of multiplex network analysis, as advocated by Kivelä et al. [38] (2014), is crucial for capturing the layered nature of interactions on platforms like Twitter. Our innovative method, integrating machine learning for node classification with graph traversal in a multiplex setting, extends standard approaches such as those described by Zhai X. et al. [39] (2018) and aligns with Fortunato's emphasis on detecting sub-communities [14] (2010). Furthermore, our observation that institutions serve as key nodes is in line with Cha et al. [40] (2010), who demonstrated that follower count alone does not capture influence, which is better measured by active engagement such as retweets and mentions. Finally, our results concerning the impact of network structure on information spread resonate with findings by Gonçalves et al. [41] (2011), emphasizing the role of network saturation in shaping user engagement.

5. Conclusions

This manuscript makes several important contributions. First, it presents an advanced technique for dissecting social networks, specifically Twitter, by pinpointing communities of practice. This approach deepens our insight into the interactions of individuals and groups around targeted topics. Second, by concentrating on the classification of active nodes and applying graph traversal within a multiplex framework, the study sheds light on the evolution and dissemination of information, ideas, and even misinformation across various communities. Third, the method enables the identification of influential actors within these communities, a feature that holds significant value for marketers, policymakers, and social scientists seeking to understand and steer online conversations. Finally, recognizing communities of practice on social media can guide organizations in crafting tailored communication strategies that resonate with specific audiences, while also aiding policymakers and activists in formulating effective messages.

Our research demonstrates that graph traversal benefits substantially from prior active node classification via ML. By preemptively determining the role of a node, computational resources are conserved and subgraph reconstruction is made more efficient. In our study, a modest seed of energy-related nodes expanded into a vast Friendship Network of energy experts, institutions, and companies on Twitter. This analysis was further enriched by constructing a second, multilayer Mentions Network, tracking how these same users reference one another over several years. By combining community detection with node popularity measures (including centrality, degree, and assortativity) and leveraging specialized ML classifiers, we could differentiate whether a node represents an institution, a company, or an energy expert. This fine-grained classification reveals that institutions serve as the cohesive force

 doi:10.20944/preprints202603.0942.v1

in the network, broadcasting decisions, new regulations, and news, while experts comment on these updates and companies, largely using Twitter for marketing purposes, remain relatively detached from direct public discussion. The numerical indicators we obtained confirm that the energy domain is particularly responsive to official institutional communications.

The analysis identifies institutions as the 'glue' of the multiplex network. They maintain the network's function by providing the core content that experts and companies then react to. Without this institutional core, the flow of specialized information across the different layers (friendship and mentions) would be significantly more fragmented.

Methodologically, we introduced a hybrid framework that combines ML with network analysis. On one hand, centrality measures across all layers of the multiplex network clarify node popularity; on the other, role-based node classification illuminates distinct communication styles. Together, these methods empower researchers to uncover communities of practice and analyze the communication strategies of each user group. Moreover, our multilayer network framework, which accounts for even the ephemeral links created by mentions, proves to be an exceptionally effective tool for community detection. Interestingly, our analysis shows that, despite the presence of friendship links, not all nodes actively engage in dialogue: many (notably companies) instead function as consumers and disseminators of information rather than as participants in discussions.

Looking ahead, future research should aim to streamline this technique into a more universally applicable, "ready-to-use" system. Currently, the process involves multiple stages of classification, graph traversal, and filtering. Ideally, an intelligent system would automatically classify users by analyzing their content and profiles with minimal human oversight. While this prospect is promising, it also presents challenges in developing a predominantly unsupervised method for group detection.

The practical implications of this study are far-reaching. By mapping the dynamics of Twitter discussions on energy and climate change using ML techniques, stakeholders can identify key influencers and the formation of communities around these topics. This insight can facilitate the tailoring of messages to specific audiences and enable direct engagement with influential users to amplify critical messages on energy transition and climate action. For policymakers, understanding social media discourse offers valuable perspectives on public opinion and concerns, which can lead to policies that better reflect societal expectations. Additionally, by identifying where misinformation is concentrated, efforts can be better directed to counteract false narratives with accurate information. Researchers and advocacy groups, in turn, can leverage these insights to spot gaps in public understanding and emerging debate areas, guiding both future research and targeted advocacy campaigns. Finally, companies in the energy sector can utilize these findings to gauge public sentiment on energy sources, sustainability, and climate policies, informing corporate strategies, communication plans, and product development. Overall, this study provides an important tool for navigating the complex landscape of digital public discourse, ultimately aiding a diverse range of stakeholders in addressing the global challenges of energy and climate change.

Author Contributions: Conceptualization, V.D.L. and A.C.; methodology, V.D.L. and M.P.; software, M.P.; formal analysis, V.D.L. and M.P.; investigation, M.E., C.S. and A.F.; writing—original draft preparation, V.D.L.; writing—review and editing, M.P., A.C., M.E., C.S. and A.F.; supervision, A.C. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable. This article does not contain any studies with human participants performed by any of the authors.

Informed Consent Statement: Not applicable. This article does not contain any studies with human participants performed by any of the authors.

Data Availability Statement: The datasets generated and/or analyzed during the current study are available from the corresponding author on reasonable request.

 doi:10.20944/preprints202603.0942.v1

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ML	Machine Learning
BOW	Bag of Words
API	Application Programming Interface
COP	Conference of the Parties
CoP	Community of Practice

References

1. Lopreite, M.; Puliga, M.; Riccaboni, M.; De Rosis, S. A social network analysis of the organizations focusing on tuberculosis, malaria and pneumonia. *Social Science & Medicine* **2021**, *278*, 113940. https://doi.org/https://doi.org/10.1016/j.socscimed.2021.113940.
2. Barreto, A.M. The Word-of-Mouth Phenomenon in the Social Media Era. *International Journal of Market Research* **2014**, *56*, 631–654. https://doi.org/10.2501/IJMR-2014-043.
3. Alzanin, S.M.; Azmi, A.M. Detecting rumors in social media: A survey. *Procedia Computer Science* **2018**, *142*, 294–300. Arabic Computational Linguistics, https://doi.org/https://doi.org/10.1016/j.procs.2018.10.495.
4. Del Vicario, M.; Bessi, A.; Zollo, F.; Petroni, F.; Scala, A.; Caldarelli, G.; Stanley, H.E.; Quattrociocchi, W. The spreading of misinformation online. *Proceedings of the National Academy of Sciences* **2016**, *113*, 554–559. https://doi.org/10.1073/pnas.1517441113.
5. Eismann, K. Diffusion and persistence of false rumors in social media networks: implications of searchability on rumor self-correction on Twitter. *Journal of Business Economics* **2021**. https://doi.org/10.1007/s11573-020-01022-9.
6. Bessi, A.; Zollo, F.; Del Vicario, M.; Puliga, M.; Scala, A.; Caldarelli, G.; Uzzi, B.; Quattrociocchi, W. Users Polarization on Facebook and Youtube. *PLOS ONE* **2016**, *11*, 1–24. https://doi.org/10.1371/journal.pone.0159641.
7. Weng, L.; Flammini, A.; Vespignani, A.; Menczer, F. Competition among memes in a world with limited attention. *Scientific Reports* **2012**, *2*. https://doi.org/10.1038/srep00335.
8. Gunaratne, C.; Baral, N.; Rand, W.; Garibay, I.; Jayalath, C.; Senevirathna, C. The effects of information overload on online conversation dynamics. *Computational and Mathematical Organization Theory* **2020**, *26*, 255–276. https://doi.org/10.1007/s10588-020-09314-9.
9. Roetzel, P.G. Information overload in the information age: a review of the literature from business administration, business psychology, and related disciplines with a bibliometric approach and framework development. *Business Research* **2019**, *12*, 479–522.
10. Shirzadian, P.; Antony, B.; Gattani, A.G.; Tasnina, N.; Heath, L.S. A time evolving online social network generation algorithm. *Scientific Reports* **2023**, *13*, 2395. https://doi.org/10.1038/s41598-023-29443-w.
11. Davies, A.; Ajmeri, N. Realistic Synthetic Social Networks with Graph Neural Networks, 2022, [arXiv:cs.SI/2212.07843].
12. Zhu, Y.; Du, Y.; Wang, Y.; Xu, Y.; Zhang, J.; Liu, Q.; Wu, S. A Survey on Deep Graph Generation: Methods and Applications, 2022, [arXiv:cs.LG/2203.06714].
13. Liu, Z.; Barahona, M. Graph-based data clustering via multiscale community detection. *Applied Network Science* **2020**, *5*. https://doi.org/10.1007/s41109-019-0248-7.
14. Fortunato, S. Community detection in graphs. *Physics Reports* **2010**, *486*, 75–174. https://doi.org/https://doi.org/10.1016/j.physrep.2009.11.002.
15. Rani, S.; Mehrotra, M. Community Detection in Social Networks: Literature Review. *Journal of Information & Knowledge Management* **2019**, *18*, 1950019. https://doi.org/10.1142/S0219649219500199.
16. Falkenberg, M.; Galeazzi, A.; Torricelli, M.; Di Marco, N.; Larosa, F.; Sas, M.; Mekacher, A.; Pearce, W.; Zollo, F.; Quattrociocchi, W.; et al. Growing polarization around climate change on social media. *Nature Climate Change* **2022**, *12*, 1114–1121. https://doi.org/10.1038/s41558-022-01527-x.
17. Effrosynidis, D.; Sylaios, G.; Arampatzis, A. Correction: Exploring climate change on Twitter using seven aspects: Stance, sentiment, aggressiveness, temperature, gender, topics, and disasters. *PLoS ONE* **2023**, *18*, e0293151. https://doi.org/10.1371/journal.pone.0293151.

18. Debnath, R.; Ebanks, D.; Mohaddes, K.; Roulet, T.; Alvarez, R.M. Do fossil fuel firms reframe online climate and sustainability communication? A data-driven analysis. *npj Climate Action* **2023**, *2*, 47. https://doi.org/10.1038/s44168-023-00086-x.
19. Horvat, E.A.; Zweig, K.A., Multiplex Networks. In *Encyclopedia of Social Network Analysis and Mining*; Alhajj, R.; Rokne, J., Eds.; Springer New York: New York, NY, 2018; pp. 1430–1434. https://doi.org/10.1007/978-1-4939-7131-2_25.
20. Kinsley, A.C.; Rossi, G.; Silk, M.J.; VanderWaal, K. Multilayer and Multiplex Networks: An Introduction to Their Use in Veterinary Epidemiology. *Frontiers in Veterinary Science* **2020**, *7*. https://doi.org/10.3389/fvets.2020.00596.
21. Xiong, X.; Zeng, Z.; Feng, M.; Szolnoki, A. Coevolution of relationship and interaction in cooperative dynamical multiplex networks. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **2024**, *34*. https://doi.org/10.1063/5.0188168.
22. Shanmukhappa, T.; Ho, I.W.H.; Tse, C.K.; Wu, X.; Dong, H. Multi-layer Public Transport Network Analysis. In Proceedings of the 2018 IEEE International Symposium on Circuits and Systems (ISCAS), 2018, pp. 1–5. https://doi.org/10.1109/ISCAS.2018.8351818.
23. Magnani, M.; Hanteer, O.; Interdonato, R.; Rossi, L.; Tagarelli, A. Community Detection in Multiplex Networks, 2021, [arXiv:cs.SI/1910.07646].
24. Broido, A.D.; Clauset, A. Scale-free networks are rare. *Nature Communications* **2019**, *10*. https://doi.org/10.1038/s41467-019-08746-5.
25. Cortés, P.A.; Quiroga, R. How academic research and news media cover climate change: a case study from Chile. *Frontiers in Communication* **2023**, *8*, 1226432. https://doi.org/10.3389/fcomm.2023.1226432.
26. Pearce, W.; Holmberg, K.; Hellsten, I.; Nerlich, B. Climate Change on Twitter: Topics, Communities and Conversations about the 2013 IPCC Working Group 1 Report. *PLOS ONE* **2014**, *9*, 1–11. https://doi.org/10.1371/journal.pone.0094785.
27. Dellmuth, L.; Shyrokykh, K. Climate change on Twitter: Implications for climate governance research. *WIREs Climate Change* **2023**, *14*, e848, [https://wires.onlinelibrary.wiley.com/doi/pdf/10.1002/wcc.848]. https://doi.org/https://doi.org/10.1002/wcc.848.
28. Fownes, J.R.; Yu, C.; Margolin, D.B. Twitter and climate change. *Sociology Compass* **2018**, *12*, e12587, [https://compass.onlinelibrary.wiley.com/doi/pdf/10.1111/soc4.12587]. e12587 SOCO-1346.R1, https://doi.org/https://doi.org/10.1111/soc4.12587.
29. Erokhin, D.; Komendantova, N. Climate Discussion on Twitter. In Proceedings of the Climate Mis/Disinformation in a Post-Factual Era, 2023.
30. Kolic, B.; Aguirre-López, F.; Hernández-Williams, S.; Garduño-Hernández, G. Quantifying the structure of controversial discussions with unsupervised methods: a look into the Twitter climate change conversation, 2022, [arXiv:cs.SI/2206.14501].
31. Torricelli, M.; Falkenberg, M.; Galeazzi, A.; Zollo, F.; Quattrociocchi, W.; Baronchelli, A. How does extreme weather impact the climate change discourse? Insights from the Twitter discussion on hurricanes. *PLOS Climate* **2023**, *2*, e0000277. https://doi.org/10.1371/journal.pclm.0000277.
32. Laishram, R.; Wendt, J.D.; Soundarajan, S. Crawling the community structure of multiplex networks. In Proceedings of the Proceedings of the Thirty-Third AAAI Conference on Artificial Intelligence and Thirty-First Innovative Applications of Artificial Intelligence Conference and Ninth AAAI Symposium on Educational Advances in Artificial Intelligence. AAAI Press, 2019, AAAI'19/IAAI'19/EAAI'19. https://doi.org/10.1609/aaai.v33i01.3301168.
33. Himelboim, I.; Smith, M.A.; Rainie, L.; Shneiderman, B.; Espina, C. Classifying Twitter Topic-Networks Using Social Network Analysis. *Social Media + Society* **2017**, *3*, 2056305117691545, [https://doi.org/10.1177/2056305117691545]. https://doi.org/10.1177/2056305117691545.
34. Abitbol, J.L.; Morales, A.J. Socioeconomic Patterns of Twitter User Activity. *Entropy* **2021**, *23*. https://doi.org/10.3390/e23060780.
35. Orhan, Y.E.; Pirim, H.; Akbulut, Y. Building Political Hashtag Communities: A Multiplex Network Analysis of U.S. Senators on Twitter during the 2022 Midterm Elections. *Computation* **2023**, *11*. https://doi.org/10.3390/computation11120238.
36. Li, J.; Chang, X. Combating Misinformation by Sharing the Truth: a Study on the Spread of Fact-Checks on Social Media. *Information Systems Frontiers* **2023**, *25*, 1479–1493. https://doi.org/10.1007/s10796-022-10296-z.

 doi:10.20944/preprints202603.0942.v1

37. Hanteer, O.; Rossi, L. An Innovative Way to Model Twitter Topic-Driven Interactions Using Multiplex Networks. *Frontiers in Big Data* **2019**, *2*, 9. PubMed-not-MEDLINE, https://doi.org/10.3389/fdata.2019.00009.
38. Kivelä, M.; Arenas, A.; Barthelemy, M.; Gleeson, J.P.; Moreno, Y.; Porter, M.A. Multilayer networks. *Journal of Complex Networks* **2014**, *2*, 203–271, [https://academic.oup.com/comnet/article-pdf/2/3/203/9130906/cnu016.pdf]. https://doi.org/10.1093/comnet/cnu016.
39. Zhai, X.; Zhou, W.; Fei, G.; Liu, W.; Xu, Z.; Jiao, C.; Lu, C.; Hu, G. Null Model and Community Structure in Multiplex Networks. *Scientific Reports* **2018**, *8*, 3245. https://doi.org/10.1038/s41598-018-21286-0.
40. Cha, M.; Haddadi, H.; Benevenuto, F.; Gummadi, P.K. Measuring User Influence in Twitter: The Million Follower Fallacy. In Proceedings of the Proceedings of the Fourth International Conference on Weblogs and Social Media, ICWSM 2010, Washington, DC, USA, May 23-26, 2010; Cohen, W.W.; Gosling, S., Eds. The AAAI Press, 2010.
41. Gonçalves, B.; Perra, N.; Vespignani, A. Modeling Users' Activity on Twitter Networks: Validation of Dunbar's Number. *PLOS ONE* **2011**, *6*, 1–5. https://doi.org/10.1371/journal.pone.0022656.