A real-time platform for contextualized conspiracy theory analysis

David Chong, Erl Lee, Matthew Fan, Pavan Holur, Shadi Shahsavari, Timothy Tangherlini[†], Vwani Roychowdhury
{davidchong13807,elee,fanma,pholur,shadihpp,vwani}@ucla.edu, {tango}@berkeley.edu
Department of Electrical and Computer Engineering, University of California, Los Angeles
[†]*Department of Scandinavian and Culture Analytics, University of California, Berkeley*

***Abstract*—As the U.S. Capitol riots of January 6, 2021 attest, conspiracy theories can lead to civil unrest. Many ”netizens” of the radical online fora that host these conversations are not themselves radicalized. However, the limited in-line context necessary to evaluate their legitimacy, along with readers' predisposition to believe stories that reinforce their existing worldviews, can create interpretations of real-world events that conflict with fact. Recent work proposes automated computational methods to reconstruct the missing contexts of conspiracy theories. While such work has successfully exposed several recent conspiracy theories, a robust real-time system that builds on this success *at scale* is needed. This paper presents one such implementation with: (a) A plug-and-play module for adding and removing heterogeneous text sources, (b) A scalable infrastructure that caters to the constantly expanding knowledge space, (c) Automation to parse real-time stories, and (d) An intuitive interface with a feedback module for users to directly evaluate the coherence of the generated context. The system returns a dynamic infinite-vocabulary Knowledge Graph (KG) that represents the directed relationships (edges) between critical actors (nodes) that feature in conspiracy theories. Our current implementation features 6 heterogeneous data sources accessed for > 5000 posts/day with a low search-latency ($< 5ms$ for querying $> 100,000$ relationships). Evaluation (with Amazon Mechanical Turk (AMT)) suggests our visualization captures well the coherence of the resultant conspiracy theories, with a positive skew (> 21.0) toward positive ratings.**

***Index Terms*—Large-Scale Data Mining, Knowledge Representation and Reasoning, Natural Language Processing, Network Visualization**

I. Introduction

A. The emergence of conspiracy theories

Rumors and the closely allied genre of conspiracy theory have particular rhetorical power and are remarkably efficient vehicles for the communication of cultural ideology–the norms, beliefs, and values that help structure behavior in groups. Storytelling can be seen as a forum for the constant negotiation of these fundamental aspects of group and belonging and, consequently, can have considerable impact on an individual's world view and the real world actions that they are willing to take. Stories that suggest a profound existential threat to the persistence of the group are those most likely to trigger potentially violent reactions.

Prior to the rapid rise of the internet and the concomitant explosion in social media where the emergence of virtually conceived communities took hold, storytelling was largely constrained to low level, face-to-face interactions that carried with them the social control of repeated interaction. These constraints were loosened considerably with the enormous reach of social media, and the attendant speed of near-instantaneous internet communication. Consequently, stories circulating on these media have the potential to influence much larger swathes of a population without the constraints of prior social systems. Predictably, this aspect of social media has attracted malign actors whose intent is to disrupt or mislead, as well as other individuals whose stories would be otherwise limited in reach because the ideas did not align with the main cultural ideology of the group.

Once the expected social controls are relaxed, stories that were once considered to be ”fringe”, reflecting the ”paranoid style” that Hofstadter proposed as a characterization of both conspiracy theories and conspiracy theorists [1], move to a more central position in a group's discussions. Given the tendency of social media groups to fairly quickly self-select for members through processes such as preferential attachment and homophilous self-selection, conversations rapidly converge on an underlying narrative framework that validates the newly emerging group's cultural ideology. As the ideology forms, the group selects collectively both the types of threats they perceive to their group's cohesion and persistence, the possible strategies to deal with those threats, and the expected or reported outcomes of applying those strategies to those threats. In the case of conspiracy theories, the threats are considered to be constant, and to come from numerous and varied quarters. The inside group often perceives itself to be a last bastion against these existential threats.

During the Covid pandemic, the threat of the disease was compounded by the threat of the Chinese Communist Party, the threat of Bill Gates and junk science conspiring with “big telecom”, and the threat of Satanic child-eating Democrats [2]. Indeed, wherever one turned in these conversations, there was a threat to some imagined way of life, and an attendant vigorous discussion of “what we should do.” These stories contributed in turn to violent actions, include the burning of cellphone towers [3], an attack on a family pizzeria by an enraged gunman [4], efforts by armed militias to round up immigrants at the southern US border [5], deliberate interference with hospitals [6], a refusal to abide public health directives [7], and a concerted effort to encourage vaccine hesitancy [8].

2375-9259/21/$31.00 ©2021 IEEE
DOI 10.1109/ICDMW53433.2021.00021

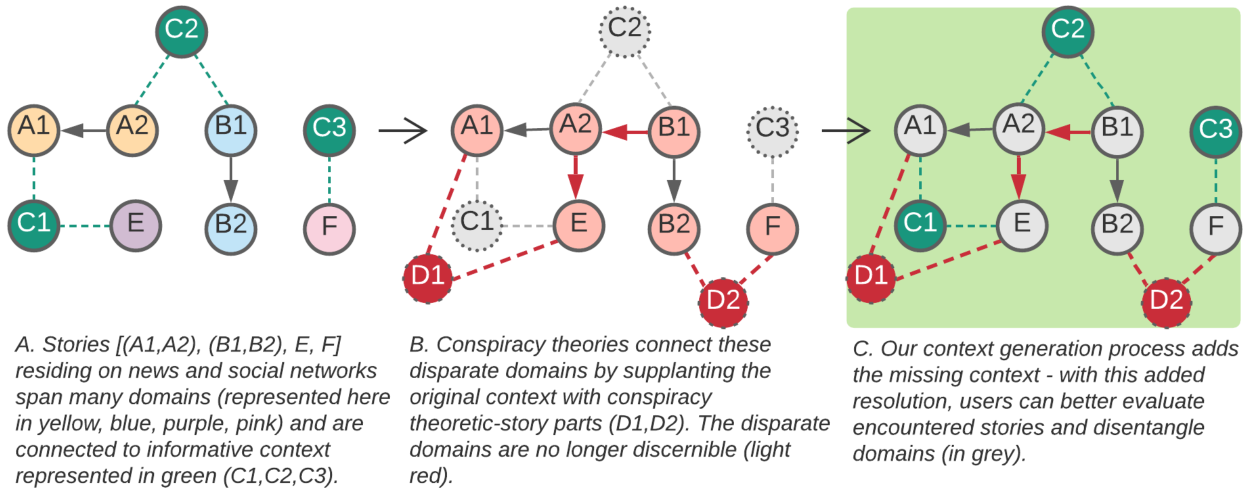

Fig. 1. An outline that describes the broad functionality of our tool. Stories that reside in different domains can be identified because of the contexts that distinguish them (A). Conspiracy theorists align these domains in an effort to create a unified conspiracy theory. Through this narrative coalescence of multiple domains and the addition of network components that resonate with the group's beliefs, perceived threats and potential strategies or the collective worldview of these communities (B), theories propagate quickly and haphazardly, rapidly mutating to adapt to varying belief systems. We attempt to augment these theories in real-time with the missing context to empower story witnesses to reach more informed judgement about the narrative(s) they encounter (C).

B. A generative model for conspiracy theories

The conspiracy theories considered here share certain important characteristics. First, each theory features a group of prominent and identifiable actors and other entities that interact with one another in specific ways (through text relationships). Second, each conspiracy theory incorporates multiple disparate domains, aligning them into a single narrative framework, redolent of the contours of the group members' world view.

For example, the Pizzagate conspiracy theory features politicians such as Hillary Clinton, her presidential campaign chair John Podesta, and the Washington DC pizzeria Comet Ping Pong. The conspiracy theory posits Hillary Clinton and John Podesta *ran a sex trafficking ring* through Comet Ping Pong. This theory connects the disparate domains of politics, sex trafficking, casual dining, and illegal behavior through tertiary and unverifiable hidden knowledge supposedly extracted from John Podesta's emails. In such a setting with multiple contexts and disparate relationships between actants (some of which are critical to the theories' validity but are hidden or false), an approach based on raw unstructured text parsing, especially for social media conversations, may not yield a holistic representation of the conspiracy theory.

These features suggest that a network representation of these actors (actants) and their inter-actant relationships along with their respective communities offers a clear way to intuitively capture the conspiracy theories' various narrative components. In such a model, the nodes correspond to actants (people, places, and objects) occurring in the narrative and edges correspond to interactions between actants. Communities represent the disparate domains that are linked together to constitute a narrative framework. Indeed, this KG can closely model a human's concept map that conspiracy theorists themselves sample as they attempt to concoct new stories or story parts (as social media posts). Our model of this structure of conspiracy theories is described in Fig. 1.

While the graph representation of the network of conspiracy theories offers a model of the mental pathways humans explore when creating a story, distilling the multilayered narratives behind the conspiracy theories poses various challenges.

C. Challenges of contextualizing conspiracy theories

Although unfounded rumors and complex conspiracy theories are widely acknowledged to be potentially detrimental to social order, they are difficult to identify in the noisy social media environments in which they take root and spread. Many of the seeds of these conversations emerge from obscure and deliberately anonymous data silos such as the transient image boards of 4chan and 8kun. The "recaps" found in the news media contain at best a partial context needed to understand the narrative frameworks driving the conspiracy theories. Without access to the narrative framework–the actants and interactant relationships that define the narrative domain–it is impossible to assess the legitimacy of an individual story or story part. This lack of context then places the burden of real-time judgement on *feel*–in short whether a person can align the story with their underlying beliefs and what they in that context consider to be plausible. Given the capacity for rumor and conspiracy theory to efficiently convey aspects of cultural ideology, this alignment is precisely the intended target of a rumor and its more complex sibling, the conspiracy theory. These stories can, accordingly, take advantage of the well-documented phenomenon that people are more likely to reinforce their preexisting beliefs than challenge them [9].

Because of the ubiquity of storytelling on social media, and the connection that online communication has to real-world action, there is a clear and pressing need for developing real-time methods that expose the broader structures surrounding a seemingly unstructured, context-sparse story (see Fig. 1). While recent solutions, such as Twitter's removal of tweets deemed misinformation, YouTube's addition of guidelines below topic-specific videos, or the augmentation of existing communication channels with more evaluative context, partially address the problem, they fail to address the broader story environment.

D. A framework to contextualize conspiracy theories

We attempt to present this broader environment using a general framework (Figs. 1,2) that empowers users to evaluate conversations in real-time through access to both the narrative framework driving a particular conversation and the broader contexts on which those conversations rest. Our framework achieves platform-independence through an abstract *Source Identification* layer. This allows users to identify the data sources most relevant to the discussions they wish to evaluate and easily refine their analysis by adding or removing sources from the pipeline. As the majority of discussions within social media and the news occur in text form, we aggregate and clean the text scraped from user-defined data sources. This aggregate text corpus is then fed into a Natural Language Processing (NLP) module in order to convert the natural language of media posts into a format suitable for knowledge extraction. We perform this knowledge extraction to distill the aggregate data into a concise and context-rich summary of current media discussions. Finally, a visualization layer allows users to explore the extracted summary through an intuitive interface.

Our work renders this framework in an implementation that a user can query in real-time to make immediately apparent the broader contexts of stories and story parts they encounter in their media traversals. This implementation is scalable, fast, and modular, and provides an interpretable overview of the underlying narrative framework powering conversations in these diverse media, helping the uninformed user to better judge stories and recognize the underlying cultural ideological space in which those stories circulate. We confirm this functionality by tasking human evaluators to rate the constitutive coherence of the generated, visualized context for queried theories and theory-relevant actors.

II. Related Works

There is a large body of prior work on extracting, analyzing, and visualizing data from social media and the news [10]–[13]. A significant portion of this work focuses on detecting the presence of misinformation [14], [15] and tracking its propagation across social networks [16], [17]. While this work is insightful, in many cases, there is a lack of explainability that results from the use of black-box Deep Learning (DL) models. In other cases, the systems lack a convenient and intuitive interface through which to explore the detected body of misinformation. Our pipeline addresses both of these limitations by introducing a novel visual interface supported by an explainable artificial intelligence (AI) engine. This interface allows one to explore semantically-coherent narratives aggregated directly from the raw text of conspiracy-focused discussions in social media and the news. In contrast to alternate misinformation discovery tools, such as FakeNewsTracker [18] or the CNN Facts First Database [19], the data processing section of our pipeline does not classify relationships as True or False and indeed does not require any class labels. Instead, our pipeline generates phrase-level narrative structures from user-selected data sources that *contextualize* the potentially conspiracy-theoretic data. By way of contrast, comparable unsupervised systems applied on this task of estimating the latent structure in unstructured text often rely either on graph generation over articles referenced by known users or on tracking interactions between metadata such as authors (of posts), topical tags (found as hashtags on Twitter) or content creators (on YouTube) [20], [21], which cannot be applied to popular, anonymous, and ill-maintained forums such as 4chan and 8kun where conspiracy theories are known to thrive [22].

Popular systems also motivated by the transparency and semantic richness of explainable graph architectures include SenticNet6, a high-performing model for sentiment analysis [23] [24]. World knowledge in that architecture is organized via symbolic AI to *logically* interpret the emotions within limited-domain posts. Our work is similar to this approach as we aggregate network *components* of narratives (a symbolic computational unit) in order to visualize the structure of conspiracy theories. The resulting computational graph is suitable for smart network traversal engines such as the model demonstrated in DeepPath [25], which applies Reinforcement Learning to answer logical questions given a KG.

To construct our platform, we host a real time, infinite-vocabulary KG generation pipeline with a unique visualization tool that affords intuitive and convenient exploration of the resultant KGs. Classical approaches to building such a KG often include aggregating relationships via OpenIE [26] or a comparable relationship extractor, which may then be condensed by fixed ontologies such as WikiData [27] and YAGO [28]. Our work extends this field and advances one state-of-the-art KG generation pipeline [29], which has previously been employed to discover the narratives undergirding the Bridgegate conspiracy, the Pizzagate conspiracy theory, Covid-19 conspiracies [30] and vaccine hesitancy. Such a constantly-evolving KG ensures that the conspiracy theory detection system maintains relevance as narrative frameworks evolve.

III. Methodology and Pipeline

A. Source Identification

While our proposed framework may handle any heterogeneous selection of data sources (plug-and-play), our current implementation employs 6 diverse corpora identified from news and social media, specifically biased (semi-supervised) to capture both formal and informal discourse surrounding

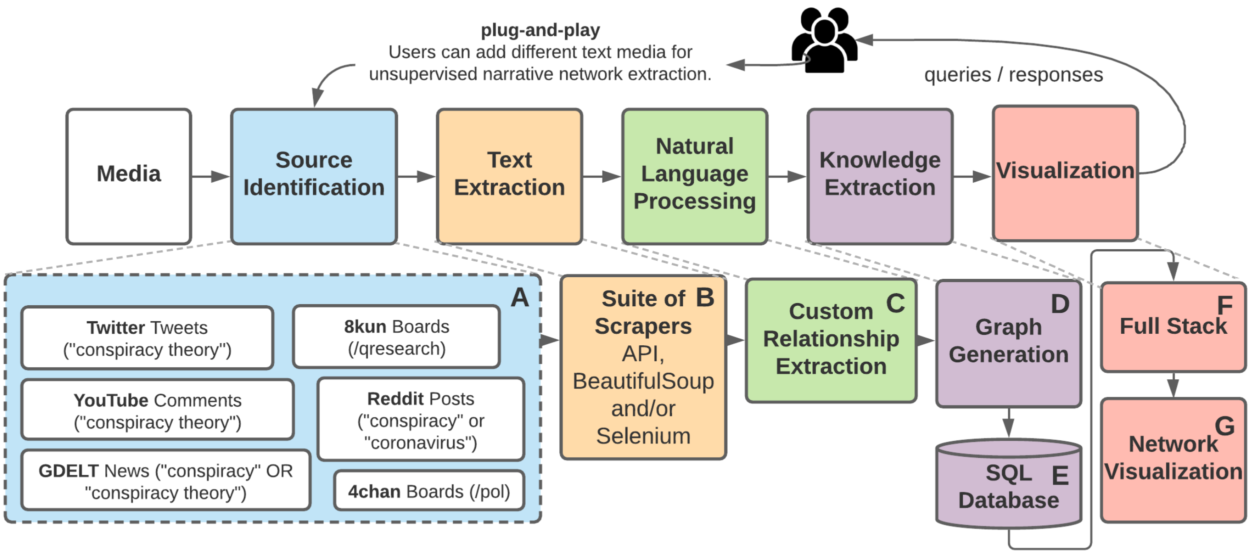

Fig. 2. An overview of the general framework (top) [4] is mapped to an implementation introduced in this work that is assembled and evaluated at scale (bottom). The system operates in conjunction with the existing media-consumer interactions. The submodules (B) - (G) are discussed in Section III. Users can query the system in module (G) to extract context relevant to encountered theories.

conspiracy theories (see Fig. 2, block A for a complete overview of the sources). By concurrently processing large-scale data from GDELT Doc [31], Twitter, YouTube, Reddit, 4chan, and 8kun, our selection represents a rich set of forums. For instance, Twitter posts are comparatively brief and limited by character count (280) in contrast to Reddit comments, which are longer. YouTube comments, on the other hand, elaborate on or extend an accompanying video or comment thread and represent a different user demographic. 4chan and 8kun posts are anonymous and less moderated. Indeed, 8kun is a popular forum for “Q drops”, a seemingly distributed and robust source of cryptic information spread by a group or an individual code-named “Q”. The different formats of our heterogeneous data sources influence the content extracted from them. Overall, our choice of sources forms a rich sample set in our attempt to capture the broader narrative of conspiracy theories.

B. Text Extraction

These user-identified sources form the backbone of a semi-supervised text extraction module that recursively parses, scrapes, and filters the heterogeneous and biased media to extract raw text, staged for NLP. For example, given the search term ”conspiracy theory” for YouTube, our text extraction system first finds the top 50 videos tagged to that search term and iteratively parses the comments section of each video. We find through qualitative analysis that YouTube comments constitute many partial threads surrounding conspiracy theories. Similarly 8kun hosts a dedicated page */qresearch* that contains posts inferring, *researching*, and investigating the beliefs of Q. Our BeautifulSoup scrapers parse each post and extract recursively the related text. Many conspiracy theories on Reddit are tagged with *conspiracy theory* and *coronavirus*, and so our scrapers (with the Reddit API) pull data from these seeds. In general, since our processing pipeline requires no supervision, data sources can be modified in a *plug-and-play* manner, with the resulting KGs a reflection of the data provided as input. As these sources, especially 4chan and 8kun, are highly volatile in the size of the raw data they feature daily, we limit the total data points aggregated across sources to ~ 5000 samples/day.

C. Network Generation from Data

The daily extractions of raw data and their metadata from these biased sources are compiled into a collective dataset. This dataset is parsed by a custom OpenIE-extended infinite-vocabulary relationship extractor [29] to extract relationship tuples that constitute a narrative (sub-)network. For example, the following relationship tuples were extracted from the data scraped on April 22, 2021:

$$\textit{bioweapon,lab,wuhan} \longrightarrow \text{created} \longrightarrow \textit{coronavirus,flu,vaccine,virus}$$

and

$$\textit{attack,drug,part,report} \longrightarrow \text{warns} \longrightarrow \textit{coronavirus,flu,vaccine,virus}.$$

These relationship tuples, alongside many others, form a narrative network describing the actants and relationships believed to characterize conspiracies. To summarize, this network generation process distills a narrative network from the data scraped daily that will, when aggregated over many days, form a larger time-indexed cumulative narrative network (see Fig. 2, blocks C,D, and Fig. 3).

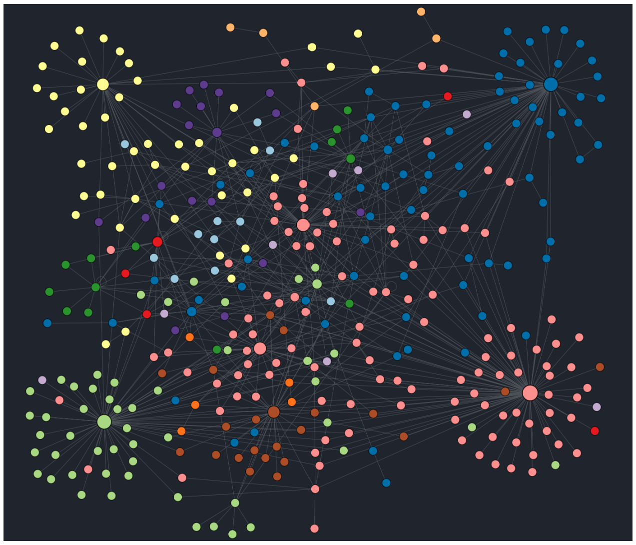

Fig. 3. A sample visualization of the narrative network generated on a given day. Each node represents an actor (actant) and each directed edge is a relationship between a pair of actants. The relative size of a node is proportional to the degree of the node. Color codes correspond to different communities of actors. These automatically differentiated communities (blue, yellow, green, red in Figure) interact with one another across different contexts to create or disband stories that potentially host conspiracy theories.

D. Daily Narrative Networks

The daily extracted networks can be represented as $G_t(V_t, E_t)$, where V_t is the set of nodes, E_t is the set of edges, and t is the date on which the network is extracted. Each node $v_{i,t} \in V_t$ corresponds to an actant present in the data on date t and each edge $e_{ij,t} \in E_t$ corresponds to a relationship from actant $v_{i,t}$ to actant $v_{j,t}$. Each node $v_{i,t}$ is labelled with a list of words $l_v(v_{i,t})$ that co-occur in the data and each edge is similarly labeled $l_e(e_{ij,t})$. This process attempts to map multi-gram actants to nodes that are semantically consistent and create interpretable relationships between them. After generating the resultant network for a given day, offensive words in any of the labels that feature in the CMU Corpus of Bad Words [32] are replaced with a wildcard token marking them as offensive. The output of this unsupervised data processing pipeline on a given day consists of the list of actants V_t and the list of relationships E_t along with their labels. To further expose any inherent structure in these networks, the actants are grouped into a set of semantically similar communities C_t, calculated with the Louvain method [33] applied to $G_t(V_t, E_t)$.

E. Cumulative Narrative Network

The daily generated networks combine to form a larger cumulative narrative network $G(\tilde{V}, \tilde{E})$ over the set of all days T (til present), where $\tilde{V}$ represents the union of the nodes mined to-date $\tilde{V} = \bigcup_{t \in T} V_t$ and $\tilde{E}$ similarly represents the union of edges: $\tilde{E} = \bigcup_{t \in T} E_t$.

Nodes and edges that comprise this cumulative graph G may be attributed to one or more of the time-indexed subnetworks $G_1, G_2, \ldots, G_{|T|}$. Since our cumulative graph captures the dynamics of constantly evolving narratives in media, we expect the networks extracted for a collection of days to have overlapping entities. Indeed, for a given actor $v \in \tilde{V}$, we can, in real-time, create a hierarchical network $G_v(\hat{V}, \hat{E})$ conditioned on the existence of node v such that $G_t \in \hat{V}$ for all $t \in T$where $v \in V_t$. In this way, $\hat{V}$ is a subset of daily-extracted narrative networks abstracted as nodes. $\hat{E}$ connects the node v to each node in $\hat{V}$. We provide one approach to extract such hierarchical networks in the ***Node2Hierarchy*** custom endpoint (see III-F).

F. Migrating the Networks to a Scalable Database

In order to augment the daily networks G_t to the cumulative network G in a scalable fashion required by our framework, we employ a MySQL [34] database (DB) that supports real-time queries (Fig. 2, blocks E,F present an overview). To this end, we introduce the following DB schema for our conspiracy theory network in Table I.

Each row in the *Nodes* Table represents a node $v_{i,t}$. The "node_id" and "Date" columns jointly form the primary key to index the rows of the *Nodes* Table. The column "node" contains a node's label $l_v(v_{i,t})$ identified by our unsupervised parsing and the column "community" contains an enumerated index from C_t that represents the community in which a node resides from the date-wise community detection process.

Each row in the *Relationships* Table represents a directed edge $e_{ij,t}$ between a pair of existing nodes $v_{i,t}$ and $v_{j,t}$, both of which are required to be present in the *Nodes* Table. "rel_id", "obj1", "obj2", and "Date" make up the primary key of the *Relationships* Table. In this design, "obj1" and "obj2" refer to the source and target of the edge, respectively. "obj1" and "obj2" reference the "node_id" in the *Nodes* Table. The "relation" field contains the human-interpretable label $l_e(e_{ij,t})$ of the relationship $e_{ij,t}$. The "rel_id" gives a way to index an edge $e_{ij,t}$ efficiently and to differentiate relationships between the same pair of nodes without looking at the "relation" field.

We construct a novel set of 3 primitive endpoints via a NodeJS server [35] that, when composed together, admit many interesting queries and query types. These endpoints mimic closely the approach employed by social media users as they pursue stories and storytelling:

- ***Node2Hierarchy***: Accessing filtered narrative networks indexed by time $t_i, t_j, \ldots, t_k$ that share a queried actant v. This empowers the user to trace an actant-centric context across time that provides information about the dynamics of theories headlining this actant.
- ***Node2Depth***: Accessing a subgraph (up to a certain depth) centered on an actant of interest on a particular day. This *node-hopping* simulates the connection of disparate domains that constitute a conspiracy theory along a path that includes multiple directed edges.
- ***Node2Vote***: Registering up-votes or down-votes for a community of actants and their relationships. This voting system provides real-time feedback about the quality of the results of our pipeline and enables our system to

Table Name	Field	Type	Key
nodes	Date	date	PRI
	node_id	int	PRI
	node	text	
	community	int	
relationships	Date	date	PRI
	rel_id	int	PRI
	obj1	int	PRI
	obj2	int	PRI
	relation	text	

TABLE I
SCHEMA FOR REPRESENTING OUR CONSPIRACY THEORY NETWORK

potentially aggregate and analyze the collective interest in parts of the cumulative narrative network G.

Both the schema and endpoints together form a foundation for creating informative queries. For example, a user can compose ***Node2Hierarchy*** and ***Node2Depth*** to query for the actor "Q" across the cumulative network to return a hierarchical network of Q-related narrative subnetworks where the maximum depth of each subnetwork is limited to some fixed depth d.

G. Front-End Design and User Interaction / Feedback

The endpoints developed in the previous section support an online exploration tool to efficiently query and analyze the cumulative narrative network G. A custom User Interface (UI) was developed using React [36] to support visualization. React updates the HTML Document Object Mode (DOM) in response to any change of *state*; i.e. it excels at building applications that focus on rendered dynamic visualizations with user interaction. Here, a *state* comprises (a) network data (of nodes and edges), generally a subgraph from G; (b) the search query that conditions the subgraph; and (c) any search filters to highlight specific nodes and edges. User interaction with the network is supported by a force-directed layout from d3 [37]. A user can interact with the graph by zooming, panning, and dragging nodes. Hovering over a node displays the node's label and hovering over an edge displays the entire *source node label* $\longrightarrow$ *verb phrase* $\longrightarrow$ *target node label* text relationship. Node size is scaled according to degree. Colors are assigned to nodes based on community membership (see Fig. 3,4). Directionality of edges is represented by an animation along each edge.

The details pane provides additional information about a selected node, including its name, community, community members, and associated relationships. Users can then select any community member or relationship, at which point the interface updates its state to highlight the new selected node. This exploration technique provides an interactive model of the generation and/or discovery process of conspiracy theories, as social media users explore and connect actants presented in the visualization tool. With this UI, users can explore how ideas and connections evolve through time and across contexts, discover trends, and track the evolution of conspiracy theories not otherwise immediately apparent (see Fig. 2, blocks F,G).

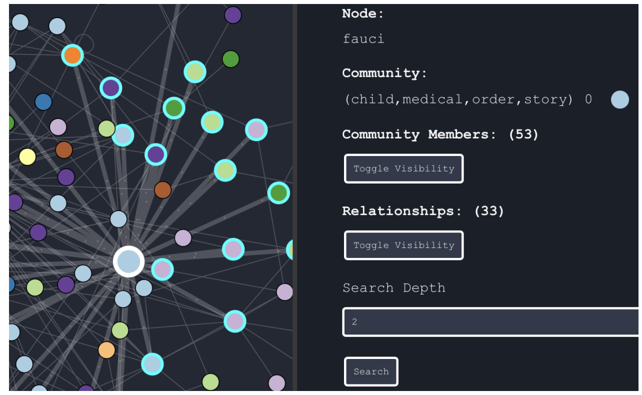

Fig. 4. Overview of the Visualization Tool: The Details pane on the right informs about the central actant in the narrative subgraph on the left, in this case the node labeled "fauci".

H. Quantitative Evaluation

We sought anonymous and media-informed users to interact with our interface so that we might understand (a) which graph components users find useful and (b) to collect feedback about our success in designing the general narrative extraction framework. We asked these evaluators to consider whether the nodes and the inter-node relationships made sense, irrespective of whether they were factual or even plausible. This evaluation task was one that interrogated the coherence of the cumulative narrative network, given what the evaluators already knew, and was not intended to be a measure of the veracity of the narrative described by the graph.

The evaluators were employed through AMT [38] over a 3 week period, and were requested to interact with the pipeline for at least 30 minutes or until they had voted 20 times (up-votes and down-votes combined), whichever came first. Evaluators were also vetted (Masters-only) and were required to reside in the United States (to be ideal consumers of the social media and conspiracy-theoretic stories that our pipeline is biased to distill).

To aggregate the interactions, we devised a consensus method to rate our cumulative narrative network. As users explore the dynamically visualized network, they are presented with the option of up-voting and down-voting selected actants. Every up-vote by user k to actant v_i adds a weight of $w_{ki} = 1$ and every down-vote adds a weight of $w_{ki} = -1$.

Since our cumulative graph is large and continues to expand (from January 12^{th}, 2021 to May 19^{th}, 2021, $|\tilde{V}| = 47271$ and $|\tilde{E}| = 139577$), a smoothing function is employed to boost the voting, such that if node $v_i \in \tilde{V}$ gets a vote, then emergent edges $e_{ij} \in \tilde{E}$, incident edges $e_{ji} \in \tilde{E}$ and neighboring nodes $v_j \in \tilde{V}$ for all feasible j get the *same* vote. Therefore, the cumulative vote for a node v_i and edge e_{ij}:

$$r_{\tilde{V}}(v_i) = \sum_{\forall k \in \text{users}} w_{ki} + \sum_{\forall k \in \text{users}} \{ \sum_{e_{ij} \in \tilde{E}} w_{kj} + \sum_{e_{ji} \in \tilde{E}} w_{kj} \}$$

,

$$r_{\tilde{E}}(e_{ij}) = \sum_{\forall k \in \text{users}} w_{ki} + w_{kj}$$

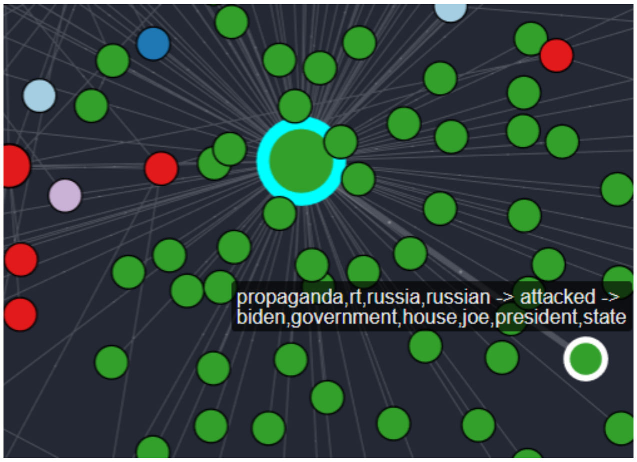

Fig. 5. A subgraph mined on February 10, 2021, depicting the high-degree actant corresponding to the Biden presidency (highlighted in blue) and its numerous associations. The aggregate-theorized connection to Russian propaganda (highlighted in white) is placed in the context of additional links to a wide range of actants, including George Soros and Bill Gates, as well as the Illuminati.

. This smoothing is semantically justified: a vote for a particular actant is likely to hold for its nearest neighbors, especially in a semantically-linked network like ours. In rare cases when the degree of a central node is too large, we subsample the connected edges and nodes at random. These weights are maintained in an auxiliary "weights" table in the MySQL server that connects node and edge IDs to the cumulative vote and the number of votes per node and per edge. Alongside plotting the cumulative vote distribution across the nodes and edges and mapping frequently high- or low- ranked votes to their communities, we also compute the correlation between node ratings and node degrees to identify whether cumulative ratings can provide information about network structure. Together, these metrics provide a fundamental quantitative template that contain information about the large-scale properties of our resultant graph.

IV. Results and Discussion

A. Qualitative Evaluation

Navigating these narrative networks uncovers sentiments consistent with popular conspiracy theories. For example, Fig. 6 displays the relationship tuple,

$$\textit{greene,marjorie,taylor} \longrightarrow \textit{overturn} \longrightarrow \textit{46th,barack,bill,clinton,obama,state}.$$

This relationship tuple was mined on January 21, 2021, two weeks after the January 6th storming of the United States Capitol. Many of the participants of this event supported the conspiracy theories spread by Q. Continued support for Q is captured in the relationship, which expresses the belief that QAnon-supporting Republican congresswoman Marjorie Taylor Greene will overturn her opposition, the Democratic leadership, represented by Democratic presidents, Joe Biden,

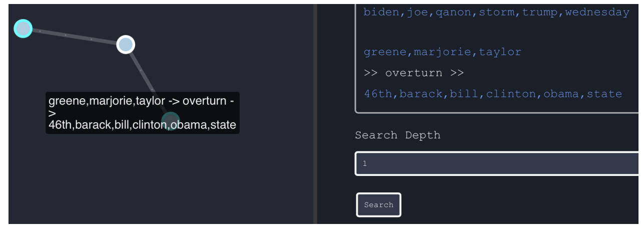

Fig. 6. A relationship mined on January 21, 2021, expressing the sentiment that QAnon-supporting congresswoman Marjorie Taylor Greene will overturn democratic leadership, represented by Democratic presidents Joe Biden, Barack Obama, and Bill Clinton. Similar sentiments were present during the Januray 6th Capitol Riots, which involved many supporters of QAnon.

Barack Obama, and Bill Clinton. Another relationship example, displayed in Fig. 5,

$$\textit{propaganda,rt,russia,russian} \longrightarrow \textit{attacked} \longrightarrow \textit{biden,government,house,joe,president,state},$$

reveals an aggregate discussion of Russian propaganda attacking the Biden presidency. This relationship appears in the context of the subgraph containing the high-degree actant, *biden,government,house,joe,president,state*, and demonstrates the significance of the Biden presidency within the aggregate discourse. We discover that, in addition to suffering attacks from Russian propaganda, the Biden presidency has supposedly been brought to power by George Soros and Bill Gates:

$$\textit{sorosngates} \longrightarrow \textit{brought} \longrightarrow \textit{biden,government,house,joe,president,state},$$

while also being brainwashed by the Illuminati:

$$\textit{illuminati} \longrightarrow \textit{brainwashed} \longrightarrow \textit{biden,government,house,joe,president,state}.$$

These relationships, among many others, reveal the breadth of circumstances within which the Biden presidency has been implicated, as well as the central role played by the Biden presidency within conspiracy theory-focused discussions.

We identify another significant actant, television host and political commentator Tucker Carlson, in Fig. 7, highlighted in the relationship tuple,

$$\textit{qanon} \longrightarrow \textit{brainwashed} \longrightarrow \textit{carlson,illuminati,tucker,youtube}.$$

This relationship tuple connects the group (or individual) identified by "QAnon" to the community of conservative media featuring Carlson, his supposed association to the Illuminati, and his presence on YouTube. While this relationship may not be a conspiracy theory on its own, it provides valuable context to the end user to assess other more conspiracy-theoretic tuples that may contain Tucker Carlson as a central figure. For example, another extracted relationship tuple,

$$\textit{anons,nunes,lukohova} \longrightarrow \textit{referenced} \longrightarrow \textit{carlson,illuminati,tucker,youtube},$$

contains the same object phrase "carlson,illuminati,tucker,youtube" as the previous tuple. In this case, however, the subject phrase "anons,nunes,lukohova" references a more tangible and well-defined conspiracy

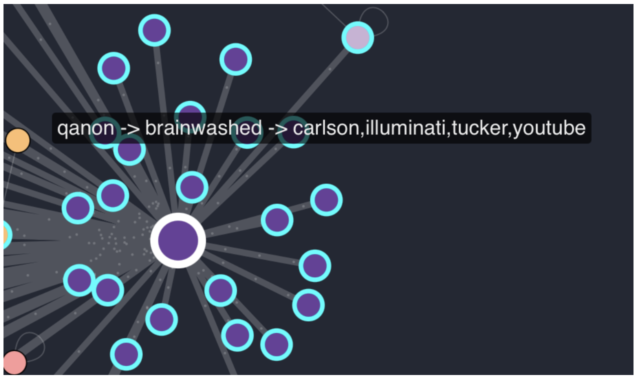

Fig. 7. A subgraph featuring a connection between conservative media host Tucker Carlson and the group (or individual) 'QAnon'. Such subgraphs provide users with valuable context for assessing the validity of conspiracy theories like those surrounding figures like Carlson.

theory that features the ex-Chair of the House Intelligence Committee, Devin Nunes. He famously claimed in 2019 (and continues to claim) that Democrats indirectly conspired with Svetlana Lukohova (Lokhova), an alleged Russian spy, to defame and ultimately indict Michael Flynn, the ex-National Security Advisor to President Donald Trump [39]. A social media user who encounters this theory, without our tool's context provider, may not have enough insight to assess the story's merit and its validity. If, instead, the user actively interacted with the visualization tool and discovered (a) the semantic link labeled *referenced* from this plot (captured in the node label "anons,nunes,lukohova") to the node referencing "Tucker Carlson" and (b) a secondary linkage labeled *brainwashed* from the node labeled "QAnon" to the same node referencing "Tucker Carlson", the user may be better equipped to appraise the real-time story. This example emphasizes the utility of a real-time network representation of narratives, a computational model in which disparate domains are connected in an explainable and intuitive way to create distilled and informative representations that not only summarize but also contextualize these stories' underlying narrative(s).

B. *Quantitative Evaluation*

Motivated by the sub-networks discovered in the Qualitative Evaluation, we now attempt to understand the structural and coherence properties of our cumulative network at scale. To this end, we first plot the degree distribution of our cumulative narrative network G on a log-log scale and empirically observe that a power-law convincingly models the spread (see Fig. 8). The estimated trend line in the distribution $y = -2.00x + 5.45$ implies that the degree distribution follows:

$$\text{Number of Nodes} = 10^{5.45}\text{degree}^{-2.00}.$$

The exponent -2.00 is weakly suggestive of a *Scale-Free Network* [40] and may imply a generative *Preferential Attachment Process* [41] underlying the network's construction. In such a model, new nodes preferentially attach to other nodes that

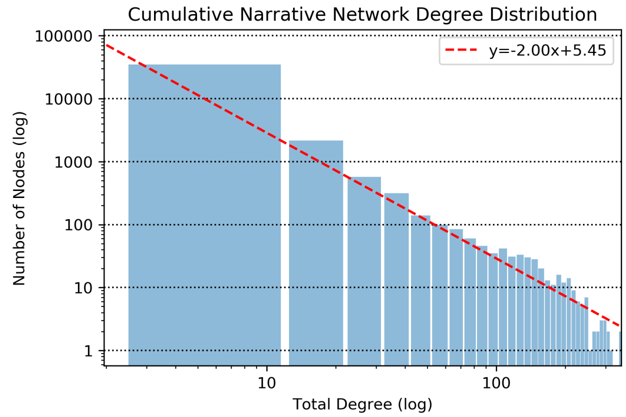

Fig. 8. The degree histogram plotted on a log-log scale. The trend line (of slope -2.00) describes the power-law, which convincingly captures the empirical distribution.

are already present in the graph and have a higher degree (or established trust). In the conspiracy theory setting, this process may be extended such that newly discovered narrative fragments are more likely to attach to popular and previously well-connected nodes or actants so as to appear more convincing, related, and/or relevant to the unsuspecting witness. This is an indication that the narratives that underlie conspiracy theories exhibit homophily; conspiracy theorists are largely driven by their pursuit of influence and popularity through their ideas and, as a result, try to entertain or include famous and consistent themes that resonate within communities of uninformed believers.

While such unsupervised distributions are indicative of the overarching structure evident in and the generative model powering our narrative network, the user feedback system provides additional targeted insights. Our voting feature built into the visualization engine provides a high-level overview of the coherency of our cumulative narrative network as well as reveals particular features of our network that users found most salient. A histogram with respect to the cumulative rating of the number of the nodes and edges is presented in Fig. 9.

The clear right-shifted bias of the distribution is indicative of the positive reception of our tool in exposing coherent narratives within the intertwined conspiracy theories that circulate in the media. Sweeping measures that quantify this trend are presented in Table II. These results indicate that between January 12^{th}, 2021 and May 19^{th}, 2021 and across nodes and relationships, there are far more upvotes than downvotes and this bias is further evidenced by the positive Pearson's moment coefficient of skewness [42],

$$g_{\tilde{V}} = \frac{\frac{1}{|\tilde{V}|}\sum_{i=1}^{|\tilde{V}|}(r_{\tilde{V}}(v_i) - \bar{r}_{\tilde{V}})^3}{\{\frac{1}{|\tilde{V}|}\sum_{i=1}^{|\tilde{V}|}(r_{\tilde{V}}(v_i) - \bar{r}_{\tilde{V}})^2\}^{\frac{3}{2}}},$$

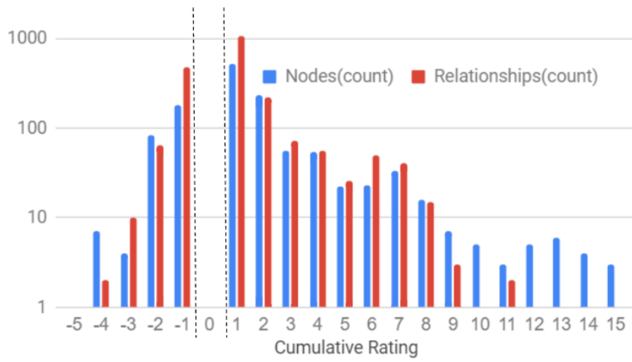

Fig. 9. A histogram of the cumulative rating on the nodes (in blue) and edges (in red). The skew of the vote distribution to the positive values indicates that users who query for specific actants more often than not find parts of our graph intelligible and rich in context. We expect the node / relationship distribution to be correlated as an artefact of the smoothing.

	Total Number		**Skewness**
	+ve vote	***-ve vote***	
Nodes	2660	409	21.0397
Rels.	2896	646	21.3189

TABLE II
METRICS SUMMARIZING THE PERFORMANCE OF OUR PIPELINE

$$g_{\tilde{E}} = \frac{\frac{1}{|\tilde{E}|}\sum_{i=1}^{|\tilde{E}|}(r_{\tilde{E}}(e_{ij}) - \bar{r}_{\tilde{E}})^3}{\{\frac{1}{|\tilde{E}|}\sum_{i=1}^{|\tilde{E}|}(r_{\tilde{E}}(e_{ij}) - \bar{r}_{\tilde{E}})^2\}^{\frac{3}{2}}}.$$

In these expressions, $\bar{r}_{|\tilde{V}|}$, $\bar{r}_{|\tilde{E}|}$ are the expected cumulative vote across nodes $\tilde{V}$ and edges $\tilde{E}$ respectively. Note that this value is computed without the large number of nodes and relationships that register a cumulative rating of 0.

Positive reception aside, user ratings also weakly predict consistent network structure demonstrated in Fig. 10, which presents a scatter plot of the cumulative rating per node against the node's degree. As before, this plot was computed without nodes registered with a cumulative rating of 0. We also omit nodes with a very high degree ($\geq$ 100, 99.37%ile), as such nodes contribute semantic value more often by connecting disparate domains than by possessing semantic meaning in isolation (for example, the node "people" on April 30th, 2021 has a degree of 107) – many of these nodes, in fact, have a rating of ± 1 as an artefact of smoothing. In other words, nodes with high degrees have multiple contexts into which they fit, resulting in the cumulative ratings for these nodes being close to zero while also being part of integral communities that feature more disparate ratings.

The positive slope of the estimated trend line in the scatter plot indicates that in addition to the semantic content of a node's label, its degree factors into users' perception of semantic relevance. This may be attributed to the rich context provided by the large number of inter-actant relationships connecting the *central* node to its neighboring nodes. As a node's degree increases, the first-neighbor context also expands and the node accumulates a larger number of relationships connecting to a wider variety of associated actants and contexts.

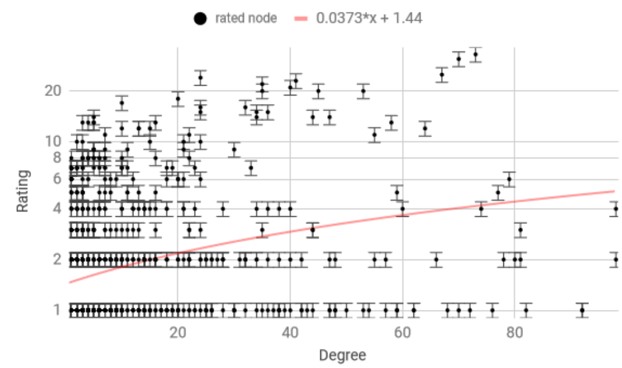

Fig. 10. The correlation between a node's rating (y-axis, log-scale) and its total degree (x-axis). We only consider nodes that have a degree $\leq$ 100 to avoid very high-degree nodes that are often not as rich in context by themselves and instead connect different domains.

V. CONCLUDING REMARKS

Conspiracy theories evolve rapidly and propagate in a far-reaching manner across many communities, inducing both fear and uncertainty and, in some cases, goading otherwise non-radicalized social media users to extreme action. Bolstered by frail contexts, these narratives ultimately sow deep distrust in social institutions among the groups caught in their shifting sands.

In this paper, we describe an implementation of the general framework for narrative extraction to aggregate the missing contexts of these conversations. Previous work left open the questions of whether (a) the narrative networks would be useful in a real-time setting with diverse data sources and a multitude of conspiracy-theoretic domains, and (b) one could evaluate the resulting networks at scale to quantify the performance of the proposed general framework for narrative extraction.

In our current work, we have addressed these questions. First, we have shown that a real-time implementation with (i) a modular Source Identification layer featuring plug-and-play datasets and (ii) a scalable infrastructure that supports querying different conspiracy theories across time and context does indeed create a powerful tool that renders meaningful and structured context concerning extant and emergent conspiracy theory narratives. Second, a consensus-based quantitative evaluation empirically confirms that the pipeline approach is effective at relaying the context of multiple conspiracy theories derived from their sources at large scale.

Platforms such as ours can help stitch together coexisting narratives that seem distinct as they circulate across the mediascape. It is common knowledge that many journalists, narrative enthusiasts, conspiracy theorists, and law enforcement agents have searched widely in vain for the origins of "Q", as well as for the motives, agendas, modes of communication, and communities of those who make up the broader

QAnon community, in order to proactively curtail the spread of misinformation. Perhaps the key is not to find an individual or group of individuals who are Q, but rather recognize that Q is a concept, mapping the theories related to it in a semantic space as we do. Such a distillation may lead us more efficiently toward characterizing and tracing the elusive Q.

REFERENCES

[1] R. Hofstadter and S. Wilentz, *The Paranoid Style in American Politics*. Vintage, Knopf Doubleday Publishing Group, 2012.

[2] M. Lofgren, *The Party Is Over: How Republicans Went Crazy, Democrats Became Useless, and the Middle Class Got Shafted*. Penguin Publishing Group, 2012.

[3] C. Reichert, "5g coronavirus conspiracy theory leads to 77 mobile towers burned in uk, report says," May 2020.

[4] M. Fisher, J. W. Cox, and P. Hermann, "Pizzagate: From rumor, to hashtag, to gunfire in dc," *Washington Post*, vol. 6, 2016.

[5] R. L. Doty, "States of exception on the mexico–us border: Security,"decisions," and civilian border patrols," *International political sociology*, vol. 1, no. 2, pp. 113–137, 2007.

[6] W. Ahmed, F. López Seguí, J. Vidal-Alaball, and M. S. Katz, "Covid-19 and the "film your hospital" conspiracy theory: Social network analysis of twitter data," *J Med Internet Res*, vol. 22, p. e22374, Oct 2020.

[7] M. S. Islam, T. Sarkar, S. H. Khan, A.-H. M. Kamal, S. M. Hasan, A. Kabir, D. Yeasmin, M. A. Islam, K. I. A. Chowdhury, K. S. Anwar, *et al.*, "Covid-19–related infodemic and its impact on public health: A global social media analysis," *The American Journal of Tropical Medicine and Hygiene*, vol. 103, no. 4, p. 1621, 2020.

[8] E. Dubé, C. Laberge, M. Guay, P. Bramadat, R. Roy, and J. A. Bettinger, "Vaccine hesitancy: an overview," *Human vaccines & immunotherapeutics*, vol. 9, no. 8, pp. 1763–1773, 2013.

[9] W. Quattrociocchi, A. Scala, and C. R. Sunstein, "Echo chambers on facebook," *Available at SSRN 2795110*, 2016.

[10] M. Injadat, F. Salo, and A. B. Nassif, "Data mining techniques in social media: A survey," *Neurocomputing*, vol. 214, pp. 654–670, 2016.

[11] K. Bontcheva, L. Derczynski, A. Funk, M. Greenwood, D. Maynard, and N. Aswani, "TwitIE: An open-source information extraction pipeline for microblog text," in *Proceedings of the International Conference Recent Advances in Natural Language Processing RANLP 2013*, (Hissar, Bulgaria), pp. 83–90, INCOMA Ltd. Shoumen, BULGARIA, Sept. 2013.

[12] Z. Ashktorab, C. Brown, M. Nandi, and A. Culotta, "Tweedr: Mining twitter to inform disaster response," in *ISCRAM*, 2014.

[13] C. Xia, R. Schwartz, K. Xie, A. Krebs, A. Langdon, J. Ting, and M. Naaman, "Citybeat: real-time social media visualization of hyper-local city data," *Proceedings of the 23rd International Conference on World Wide Web*, 2014.

[14] K. Shu, A. Sliva, S. Wang, J. Tang, and H. Liu, "Fake news detection on social media: A data mining perspective," *SIGKDD Explor. Newsl.*, vol. 19, p. 22–36, Sept. 2017.

[15] K. Shu, D. Mahudeswaran, S. Wang, D. Lee, and H. Liu, "Fakenewsnet: A data repository with news content, social context, and spatiotemporal information for studying fake news on social media," *Big Data*, vol. 8, no. 3, pp. 171–188, 2020. PMID: 32491943.

[16] J. Shin, L. Jian, K. Driscoll, and F. Bar, "The diffusion of misinformation on social media: Temporal pattern, message, and source," *Computers in Human Behavior*, vol. 83, pp. 278–287, 2018.

[17] L. Wu and H. Liu, "Tracing fake-news footprints: Characterizing social media messages by how they propagate," in *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining*, WSDM '18, (New York, NY, USA), p. 637–645, Association for Computing Machinery, 2018.

[18] K. Shu, D. Mahudeswaran, and H. Liu, "Fakenewstracker: a tool for fake news collection, detection, and visualization," *Computational and Mathematical Organization Theory*, vol. 25, pp. 60–71, Mar 2019.

[19] "Cnn's facts first searchable database."

[20] S. C. R. Gangireddy, D. P, C. Long, and T. Chakraborty, "Unsupervised fake news detection: A graph-based approach," in *Proceedings of the 31st ACM Conference on Hypertext and Social Media*, HT '20, (New York, NY, USA), p. 75–83, Association for Computing Machinery, 2020.

[21] S. Yang, K. Shu, S. Wang, R. Gu, F. Wu, and H. Liu, "Unsupervised fake news detection on social media: A generative approach," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, pp. 5644–5651, Jul. 2019.

[22] A. Amarasingam and M.-A. Argentino, "The qanon conspiracy theory: A security threat in the making," *CTC Sentinel*, vol. 13, no. 7, pp. 37–44, 2020.

[23] E. Cambria, Y. Li, F. Z. Xing, S. Poria, and K. Kwok, "Senticnet 6: Ensemble application of symbolic and subsymbolic ai for sentiment analysis," in *Proceedings of the 29th ACM International Conference on Information &; Knowledge Management*, CIKM '20, (New York, NY, USA), p. 105–114, Association for Computing Machinery, 2020.

[24] E. Cambria and A. Hussain, *Sentic Computing: A Common-Sense-Based Framework for Concept-Level Sentiment Analysis*. Springer Publishing Company, Incorporated, 1st ed., 2015.

[25] W. Xiong, T. Hoang, and W. Y. Wang, "DeepPath: A reinforcement learning method for knowledge graph reasoning," in *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, (Copenhagen, Denmark), pp. 564–573, Association for Computational Linguistics, Sept. 2017.

[26] G. Angeli, M. J. Premkumar, and C. D. Manning, "Leveraging linguistic structure for open domain information extraction," in *Association for Computational Linguistics (ACL)*, 2015.

[27] D. Vrandečić and M. Krötzsch, "Wikidata: A free collaborative knowledgebase," *Commun. ACM*, vol. 57, p. 78–85, Sept. 2014.

[28] T. Pellissier Tanon, G. Weikum, and F. Suchanek, "Yago 4: A reasonable knowledge base," in *The Semantic Web* (A. Harth, S. Kirrane, A.-C. Ngonga Ngomo, H. Paulheim, A. Rula, A. L. Gentile, P. Haase, and M. Cochez, eds.), (Cham), pp. 583–596, Springer International Publishing, 2020.

[29] T. R. Tangherlini, S. Shahsavari, B. Shahbazi, E. Ebrahimzadeh, and V. Roychowdhury, "An automated pipeline for the discovery of conspiracy and conspiracy theory narrative frameworks: Bridgegate, pizzagate and storytelling on the web," *PLOS ONE*, vol. 15, pp. 1–39, 06 2020.

[30] S. Shahsavari, P. Holur, T. Wang, T. R. Tangherlini, and V. Roychowdhury, "Conspiracy in the time of corona: automatic detection of emerging covid-19 conspiracy theories in social media and the news," *Journal of Computational Social Science*, vol. 3, pp. 279–317, Nov 2020.

[31] K. Leetaru and P. A. Schrodt, "Gdelt: Global data on events, location, and tone," *ISA Annual Convention*, 2013.

[32] L. v. Ahn, "Offensive/profane word list."

[33] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast unfolding of communities in large networks," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2008, p. P10008, oct 2008.

[34] M. Widenius, D. Axmark, and P. DuBois, *Mysql Reference Manual*. USA: O'Reilly &; Associates, Inc., 1st ed., 2002.

[35] L. M. Surhone, M. T. Tennoe, and S. F. Henssonow, *Node.Js*. Beau Bassin, MUS: Betascript Publishing, 2010.

[36] S. bin Uzayr, N. Cloud, and T. Ambler, *React*, pp. 507–521. Berkeley, CA: Apress, 2019.

[37] M. Bostock, V. Ogievetsky, and J. Heer, "D¡sup¿3¡/sup¿ data-driven documents," *IEEE Transactions on Visualization and Computer Graphics*, vol. 17, p. 2301–2309, Dec. 2011.

[38] K. Crowston, "Amazon mechanical turk: A research tool for organizations and information systems scholars," in *Shaping the Future of ICT Research. Methods and Approaches* (A. Bhattacherjee and B. Fitzgerald, eds.), (Berlin, Heidelberg), pp. 210–221, Springer Berlin Heidelberg, 2012.

[39] J. Gerstein and A. Restuccia, "Intelligence scholar sues cambridge academic, u.s. news outlets over reports on flynn links," May 2019.

[40] P. Holme, "Rare and everywhere: Perspectives on scale-free networks," *Nature communications*, vol. 10, pp. 1016–1016, Mar 2019. 30833568[pmid].

[41] G. U. Yule, "Ii.—a mathematical theory of evolution, based on the conclusions of dr. j. c. willis, f. r. s," *Philosophical Transactions of the Royal Society of London. Series B, Containing Papers of a Biological Character*, vol. 213, no. 402-410, pp. 21–87, 1925.

[42] S. Kokoska and D. Zwillinger, *CRC standard probability and statistics tables and formulae*. Crc Press, 2000.