THE UNIVERSITY OF CHICAGO

Watching Them, Watching Us: How the Misinformation Beat Redefines Journalism's Relationship with Conspiracy Theories

BY

REGINA KAYLA G. CATIPON

JUNE 2021

A paper submitted in partial fulfillment of the requirements for the Master of Arts degree in Computational Social Science

Faculty Advisor: Dr. Marshini Chetty, Computer Science

Preceptor: Pedro Alberto Arroyo, Sociology

Abstract

In 2020, conspiracies moved from the fringes of society, to the very center of socio-political life in the United States. In response, news sources have created an area of reporting called the misinformation beat. Covering conspiratorial content, however, poses unique epistemic challenges for journalists and researchers. One concern is that covering emerging conspiracy theories may inadvertently expose trusting readers to an especially seductive form of misinformation. Another concern is that savvy conspiracy theorists may be manipulating the media into unintentionally “spreading the word”. Given these dilemmas, this project seeks to quantify and qualify the coverage of conspiracy theories by digital news sources in the United States. The interaction between news and conspiratorial content is modeled through a computational content analysis and a network analysis of New York Times articles.

1. Introduction

In response to the chaotic information landscape that bloomed during Covid-19, news outlets like CNN, the Washington Post, and the New York Times dedicated reporting resources to cover the rise of misinformation, disinformation, and conspiracy theories. I will refer to this new dedicated subject of reporting throughout this thesis as it was named by McClure Haughey et al, the *misinformation beat* (2020). This project hypothesizes that news coverage of fringe groups and conspiracy theories may expose wider audiences to primary source misinformation. By doing so, journalists may inadvertently place readers on the path of conspiracy rabbit holes in which readers are led from one website to the next, thanks to the very structure of the Internet (Stewart, 1999). A reader's vulnerability to such conspiracies depends on the reader’s psychographic, epistemic, existential, and social backgrounds (Douglas, 2017). And though these factors influence dissemination and internalization of conspiracism, it is important to understand one of the mechanisms by which conspiracies find their way into the public’s attention, journalism.

This research project seeks to answer the following questions: how have news outlets like the New York Times (NYT) reporting practices been redefined by conspiracies in 2020? This question breaks down into three parts, 1) what is the misinformation beat and how is it unique, and 2) does the misinformation beat contain a greater risk of exposure to conspiracies than other types of news, and 3) what is the potential distance from readers to conspiratorial content because of the misinformation beat? This project relies on a mixed-methods approach wherein a theory of knowledge at the intersection of journalism and conspiracism drives computational analysis.

The interaction between journalism and conspiracies has become highly relevant for a number of reasons. The period around Covid-19 has seen two epidemics -- a physical, infectious disease and an *infodemic* (Starbird, 2020). As a result of conspiracies and false reports, individuals have ignored mask mandates and social distancing orders (Stewart, 2020). Others have falsely believed the virus to be nothing more than a flu or worse, a machination of evil global elites like Bill Gates (Wakabayashi, Alba, & Tracy, 2020). Theories like QAnon, that believed the US government was satanic and centered around a now discredited "insider" named Q, have even made it into the presidential debates (Kunzelman, 2020). And while conspiracy theories are not a new phenomenon, in more recent months conspiracy theorists have moved from theory to action, such as with the Storming of the Capitol, in which insurrectionists were fueled by conspiracies that the 2020 elections was stolen (Lazer et al, 2021). Understanding the interactions between conspiracies and journalism will not only contextualize current conspiracism within the historical cycles of event or crisis sense-making, but will also shed light on the re-defining of journalistic practices in the digital age.

2. Literature Review

2.1 Background

To be sure, conspiracies have long existed in the U.S. zeitgeist and in fact, are woven into the very beginnings of the country. One example is founding father Thomas Jefferson, who believed that, "...corrupt ministers of the king were secretly conspiring..." to alienate the rights of colonists (Webb, 2019). Such conspiracism and distrust influenced his writing and in turn, helped to sow the seeds of the American revolution. For each epoch of history and accompanying crisis(es), there has been conspiracism to channel anxieties and to help make sense of social changes (Prooijen and Douglas, 2017; Hofstaetder, 1966). Given the historical contexts for the emergence and dissipation of conspiracy beliefs, it is no wonder that the current anxieties around the pandemic have given way to conspiracism as a sense-making practice. These days, research estimates that, "...over half of the American population consistently endorses some kind of conspiratorial narrative about a current political event", further adding that such belief is, "...a trait that is remarkably widespread and stable across the usual ideological and sociological divides." (Maeyer, 2018; Oliver and Wood, 2014). In a polarized country, it seems, perhaps one of the few common practices is a growing belief in conspiracy.

This polarized political climate, combined with social unrest, and a global pandemic have created the perfect conditions for a re-emergence of conspiracy belief. According to Pew Research, about 25% of US adults see some truth in the conspiracy theory that the Covid-19 virus was actually planned (Schaefer, 2020). Before the pandemic, white Chrisitan Evangelicals were already moving towards conspiracy beliefs that Donald Trump was a messianiac figure that would help stop the rise of evil (Fea, 2020). Their embracing of this narrative, it has been theorized, was brought on by a deep anxiety over fear that the US has been declining as a Christian nation (Fea, 2020). Not all conspiracy messaging is misinformation, however, and not all news is truth. One can look at the case of Watergate for example, where distrust and suspicion of the government was indeed founded. Thus it may be worthwhile to delineate the boundaries between journalism and conspiracism and define the operationalization of both in theory.

2.2 Theoretical Constructs

a. <u>Conspiracies</u>

A long held understanding of conspiracism defines conspiracy theories as the belief that the hidden machinations of powerful and well connected groups are the true cause of major events. (McCauley & Jacques, 1979). Conspiracies as reflective of a collective paranoia of power structures has held a fascination in both pop culture and in social sciences, with recent theoretical work focused on internet conspiracies as an eschatology, epistemological, and philosophical construct of the postmodernist condition (Aupers, 2012; De Maeyer, 2018).

a.2. *Conspiracy as postmodernist condition*

Some have described conspiracy theories as the “a relentless ‘will to believe’ in a disenchanted world”, a sure sign of postmodernist angst (Aupers, 2012). As an eschatology, many conspiracies have a main narrative of an apocalyptic good versus evil battle waged in the shadows. This consistency of narrative, despite the rotating cast of “evil” characters ranging from Jewish people to the Deep State, is the reason why today’s conspiracies have deep roots in historical prejudices and remain adaptive to current events (Donovan, 2020). Because conspiracists prefer their own reality over facts presented by others, they participate in a particular "deconstructionist" epistemology in which there is no way of ever truly knowing and there is no way of ever truly disproving (De Maeyer, 2020). Instead the total subjectivity of conspiracism "overwhelms" journalism' epistemologies and rhetoric of "objectivity" (De Maeyer, 2020).

There has been much debate, however, over the true nature of conspiracy, anxiety, paranoia, and a postmodern understanding of the relationship between the self and society. An easy critique is that since conspiracies are not new, how can they be symptomatic of a postmodern condition? Or if postmodernism rejects grand narratives and questions knowledge claims as conditioned responses, then how can conspiracies which are the epitome of grand narratives be considered a condition? Aupers believes that the nature of conspiracies has actually transformed with modernity (2012). Authors like Melley would venture that the positioning of global governments and multinational corporations as the bad guy, reflects the modern fear of losing agency and the self, especially when faced with these monoliths. Melley claims that the loss of a "centered, autonomous subject" of postmodernism actually comes with a desire to conserve the self. A complete subjectivity means little structure to define oneself then, the self is romanticized and aggrandized (Melley, 2000) in response to a paranoia of "modern society itself" (Aupers, 2012). For example, we often see conspiracists on social media platforms rejecting the "objectivity" of news and instead defining truth for themselves. Thus these conspiracists place themselves as more capable and less conditioned than the media and society. A common phrase in conspiracy theory circles compares non-believers to obedient livestock, "Wake up, sheeple" (Jolley et al, 2018).

a.3. *Conspiracy as motivation for misinformation*

But how does conspiracy fit within the current model of information propagation online? What constitutes a conspiracy within the context of disinformation and misinformation? Conspiracies and misinformation are often mentioned together and yet remain separate. Many have worked to characterize and define misinformation and disinformation, and rather than dabble in the nebulous field of what is or is not intentionally misleading content, let us instead define conspiracy here as a sense-making practice that leverages the power of narratives to resist fact, even when presented with overwhelming evidence that contradicts the belief. When conspiracists resist truth and reconstruct reality through the sharing of "evidence" such as documents, images, or interpretations of events, they participate in misinformation dissemination. Therefore one theoretical supposition of this paper asserts that conspiracy as a narrative can be a motivation for the act of spreading misinformation.

Some of the spread of conspiracies like QAnon has been done intentionally by true believers online. At this point conspiracists become misinformation actors and these often actors show great sophistication and understanding of how information propagates online. They use social engineering practices like "platform filtering" to decontextualize information as they move messages across the web (Donovan, 2020). They intentionally create and craft content that is likely to go viral and to reach such a critical mass of interest that journalistic outlets pick up on the story, a form of media manipulation called "trading up the chain" (Donovan, 2020). As McClure Haughey states, reporters are left with an ethical dilemma in covering this new misinformation beat (2020) which is a new journalistic area of reporting that focuses on disinformation, misinformation, harassment, fake news, and conspiracy theories (McClure Haughey, 2020). While McClure Haughey focuses on the misinformation beat in the context of journalistic tracking online activity, the misinformation beat of newspapers such as the New York Times cover “daily distortions” that include conspiracies from offline activities as well, as seen in figure 1.

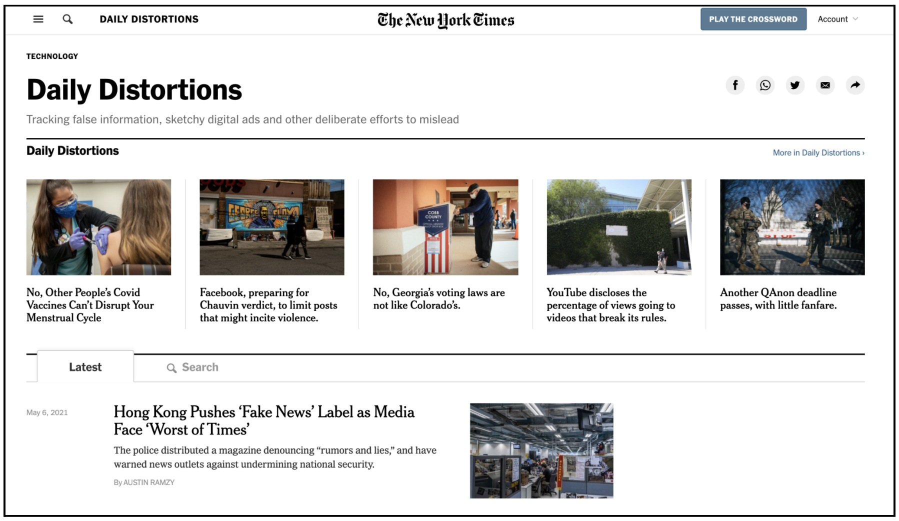

Fig 1. A screenshot of the NYT’s Daily Distortions page, an example of a media outlet’s misinformation beat, taken on 5/6/2021

Misinformation beat reporters are often faced with the difficult dilemma; do they debunk the misinformation and cover it as a news story or not? Despite the risk of exposure to the readers, journalists are compelled to report on misinformation because of its ever growing dangers. But if a story focuses on a conspiracy like Qanon, then readers may enter a virtual Deleuzian, “universe in which divergent series trace endlessly bifurcating paths, and give rise to violent discords and dissonances that are never resolved into a harmonic tonality” (Smith & Protovi, 2018). And this “chaosmos” has begun to define the Internet as a shared experience.

b. Journalism

Ironically, on face value, journalism and conspiracies share similar genetic makeup. Both worlds revolve around following clues and gathering evidence. As De

Maeyer wrote, “Conspiracism, after all, could be understood as the ultimate form of journalistic exposé.” (2018). But conspiracies lack a scientific rigor or the professionalism of journalism, and conspiracism very rarely reflects the reality of events. Media and communications studies would place journalism closer to rational science at this point, creating a separation from the unpurified, irrational science of conspiracy (Aupers, 2012; De Maeyer, 2018; La Tour, 1988). This project examines the potential for news readers to enter the conspiracy theory rabbit hole, but if we are interested in journalism as a potential amplifier of conspiracism, then the epistemology of journalism needs to be addressed.

b.1. Journalism epistemology as constructed versus assembled

Is journalism as Vorp would say, a constructed process or is it an assembled process, as De Maeyer asserts? For the purposes of operationalizing journalism in this study, this paper supposes that journalistic epistemology is a bit of both. Like Vorp, I take a constructionist view of narrative and discourse formation in news media ergo arguing that “individuals and groups actively create social reality from different information sources,” and that, “ Journalists are in the middle of this dynamic process” (2010). Journalists thus become the interpreters of events and issues, constructing meaning in news stories through their semantic choices (Vorp, 2010). But like De Maeyer states, journalism assembles facts, and therefore can not construct reality, an “*artificially created”* reality. Therefore one contribution of this project is the assertion that epistemic distinction lies in what journalism truly constructs. Journalism’s form of epistemology means it can not construct *information*, hence relying on “information sources”, but journalists do construct *meaning* from that information. It is the classic question of whether the artist can be separated from the art. In journalism knowledge production, objectivity is attempted through fact-finding, but by the very nature of fact interpretation, total objectivity is never quite reached. Arguably it is this epistemological insecurity that conspiracy exploits (Aupers, 2012).

c. *Conspiracy and Journalism, mutually defining*

One other theoretical contribution of this work is the recontextualization of the journalistic interaction with conspiracy theories in the Covid-19 era. In 2004, Bratich observed the beginnings of conspiracy theories and the web "mutually defining each other", with even earlier signs of a growing association between new technologies and conspiracies starting in 1996 (Bratich, 2004). Bratich examined the CIA-Crack-Contra conspiracy, in which drug trafficking by a Nicaraguan rebel group was linked to the Central Intelligence Agency (CIA), and noted how the web response to an article by the *San Jose Mercury News* on the CIA backing of the Contra group, outstripped "normal channels of news transmission" (Bratich, 2004). At the time, traditional and more mainstream news sources like the New York Times were slow to pick up on the story and also dismissed the published story as pure conspiracy, an example of theorizing and problematizing conspiracy theories (Bratich, 2004). Bratich asserts it was Mercury's choice to publish the article on an emerging medium, the Internet, that caused other journalists to cast doubt on the veracity of the reporting. Conspiracy theory was made "more 'foul' because it was circulated through an untrustworthy medium" (Bratich, 2004). Some of journalism's early, negative characterization of the Internet as sidestepping the gatekeepers runs through the discourse of today. Yet, almost two decades later, digital news is the norm. And in the past decade conspiracism as a culture has been "… increasingly normalized, institutionalized and commercialized." (Aupers, 2012). News adapted to the web, and now, news is adapting to conspiracies and misinformation.

In the past, conspiracies and rumors were potential sites of news-making (Bratich), but now conspiracy and misinformation is its own beat -- its own subject of dedicated reporting. Political polarization, partisanship, and media bias have contributed to distrust in journalism and helped to foster trust in the Internet as an alternative source of information (Watts, 2017; Suiter and Fletcher, 2020). In turn, Internet platforms like Reddit and 4chan help to fuel belief that mainstream media sell "Fake News". The Internet is where news is made and disseminated, picked apart by conspiracists, who are then covered by reporters, whose articles are then scrutinized by conspiracists, ad nauseum. Because as much as the US media has begun to report on conspiracy theories, conspiracists have long monitored the news for evidence of their theories. Such a simultaneous feedback loop is unique to the digital age; no other technology but the Internet could have facilitated and sped up the dissemination and consumption process to this degree. However, it is this constantly shifting information landscape, exacerbated by unequal power dynamics with social media companies on which they rely on for data access, that have left journalists feeling overwhelmed and outmatched by the pace of conspiracies, disinformation, and misinformation (McClure Haughey, 2020).

If more conspiracies were like Watergate, in which a sitting president was implicated in a political scandal, or the Contra affair, perhaps this feedback loop would be a positive step forward in the empowerment of everyday citizens, an upside of the Internet's democratization of information. Unfortunately, most conspiracies during

Covid-19 like the 5G theory, that the virus was caused by cell phone towers, or Plandemic, that the virus was intentionally planned for profit, have had tangible consequences on public infrastructures and health (Bruns, Harrington, & Hurcombe, 2020). The question at hand then is, given the theoretical understanding of both journalism and conspiracy theories, and their new (almost instantaneous) mutual coverage, how does modern journalism interact with conspiracy messaging?

2.3 Related Computational Work

This project relies on a mixed-methods approach to the analysis of the new misinformation beat. Previous research on misinformation and journalism will be divided into two categories, 1) social media network analysis and 2) computational content analysis of news and conspiracy narratives. As a potential contribution to the body of work, this project leverages social theory and media theory to drive computational analysis. The misinformation beat was initially identified and defined by Mclure Haughey et al through their semi-structured qualitative interviews with journalists (2020). As far as the author is aware, this thesis provides a computational analysis of the misinformation beat as well as the first formal attempt at a social theory and media theory of the misinformation beat.

Most prior computational methods have focused on social media networks as amplifiers and disseminators of misinformation and disinformation. Less have focused on traditional news sites as potential disseminators of misinformation. One such example of social network analysis came from Lazer, et al, who in 2019 found evidence that Fake News made up 6% of news on Twitter but was heavily concentrated and mostly shared amongst conservative voters. Other examples include, Ahmer Arif, Leo

Stewart, and Kate Starbird's research who applied an crisis-informatics theoretical framework to quantify the central positions of Russian information operation agents in a sample of Twitter conversations on #BlackLivesMatter (Arif, et al, 2018). Still more have looked at messaging platforms like WhatsApp and its cross-platform effects. In one case, authors were able to represent the spreading of images with misinformation, before, during, and after interaction with the platform as a network (Resende, et al, 2019). Reddit has also been researched as a network that codifies or amplifies misinformation beliefs and behaviors. In an example of combining network and content analysis, Tangherlini, et al modelled the complexities of Reddit conspiracy narratives as a network (2020). Their research used natural language processing to extract the text features of comments and posts. Then the authors combined qualitative and quantitative methods to represent the narrative as connections (real or imagined) between hero figures, evil political figures, and institutions. In each of the literature examples cited here, the authors focused on network representations of social data to follow the flow of misinformation or to reconstruct misinformation narratives. This project builds on this practice by creating citation networks from NYT articles to assess whether these articles share primary sources of misinformation.

A network analysis can give an overview of the structure and topology of information sharing, however, identification of conspiratorial content will require content analysis methods. Such methods rely on frequency and placement of words to trace semantic relationships. In an interdisciplinary literature review comparing news bias analysis approaches taken by social sciences versus computer science, the authors Hamborg, Donnnay, and Gipp, found that computational methods often lacked the theoretical nuances present in social science work (2018). Conversely, these authors also found that social science approaches lacked scalability and often relied on descriptive statistics such as word frequency (2018), such as with the example of the first large-scale computational analysis of political slant in the news was published by Genkowtz and Shapiro in 2010. Similarly to the research done on political slant, previous work on detecting conspiracy theories in the media have been extensive, exhaustive, but also manual. In one example, 104,803 published letters that US citizens sent to the *New York Times* and the *Chicago Tribune* between 1890 and 2010 were manually coded by researchers (Uscinski & Parent, 2014). These researchers found peaks and troughs in conspiracy belief through the manual annotation of letters to the editor. They surmised that there is an historical context for the emergence and dissipation of conspiracism (Uscinski & Parent, 2014; Prooijen & Douglas, 2017). To bridge this gap, this project seeks to marry social theory and computational methods. As this project relies on a grounded theory approach, I also leverage manual inspection and labelling, specifically when identifying primary versus secondary sources shared in articles. Though manual, this inspection and labelling becomes iterative through large-scale computing methods.

3. Data & Methods

This project seeks to model the information dissemination of conspiracy in the misinformation beat. In this section I will focus on the production of online news and the network effects of internet content. As such, the paper will limit its data collection and scope to US-based news sources for two reasons. While Covid-19 and the structure of many conspiracies relate to the larger globe, the conspiracies of interest are US-centric and far-right populist, which have particular cultural understanding in the States. As the earlier theorizing of conspiracies and journalism were focused on a US media marketplace, the paper will model the consumption of US news in order to stay internally consistent. There are of course, limitations and drawbacks to a US-focused theorization of conspiracies in a global pandemic, as the work will not be generalizable.

3.1 Data Overview

Data collection - As established in previous research (Termain, 2017), the New York Times (NYT) plays a central role in the US news market and continues to influence the newsmaking process of other outlets. The NYT has a self-reported subscriber count of over 7.5 million, compared to the Wall Street Journal's 2 million and Fox News prime time viewership at 3.387 million (NiemanLab, 2020). The data collection in this project therefore focuses on digital content from the NYT. Through a third-party package (Pietz, 2021), I scraped over 56,000 stories from the NYT website in year 2020 from Januart to December, which includes articles and web content such as multimedia slideshows. 2020 was chosen as a timeline because of Covid-19 and its "...inherent and persistent uncertainty" (Starbird, 2020). This uncertainty contributed to the rise in misinformation, disinformation, and conspiracy belief. After processing the dataset, the metadata features of identification number, publication date, keywords, text, html, news desk, and author were selected in order to extract temporal, linguistic, editorial, and web page features. I then merged a public dataset containing the number of comments per each NYT scraped in February 2021 by a data scientist student for a coding learning project and made available through Kaggel (Dornel, 2021).

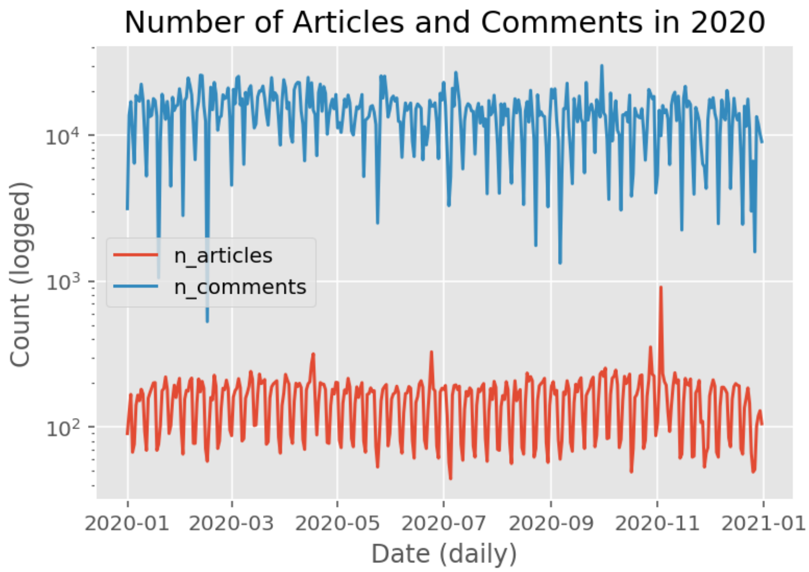

Fig. 2. Daily volume of articles from the NYT reflects weekly news cycle with an increase in articles during the 2020 presidential election.

Dataset features - As you can see in figure 2 above, the volume of articles and comments falls into a weekly news cycle, with a peak in article publication during the U.S. presidential election, November 2020.

Each article and story from the NYT was tagged with keywords selected by the editors. An exploration of the most popular keywords in 2020 is shown in figure 3, below.

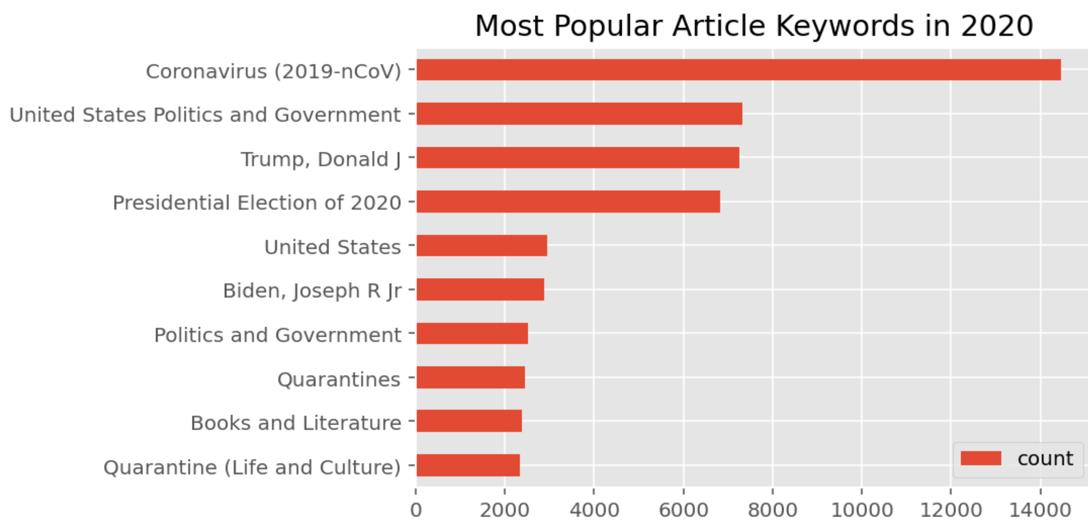

Fig 3. The top ten most tagged article keywords as labelled by NYT editorial staff.

Over 14,000 news stories were dedicated to the Coronavirus in 2020, and many of the top keywords reflect news around the pandemic such as “Quarantines” and “Quarantine (Life and Culture)”. The other most popular article subjects were related to U.S. politics, the election, and the presidential candidates former-president Donald Trump and president Joe Biden.

3.2 Q1: What is the misinformation beat and how is it unique?

Defining misinformation beat articles by NYT keywords **-** In order to identify the articles that pertain to the misinformation and conspiracy beat, I searched for stories containing the string “QAnon" as an example of a conspiracy theory that became popular and found that all of these articles in 2020 were tagged with the labels “Misinformation and Rumors” or “Fringe Groups”. The NYT also dedicated space on their site to “Daily Distortions: Tracking Viral Misinformation” and all of these articles were similarly tagged, containing at least one of those two keywords.

To compare stories from the misinformation beat versus non-misinformation beat stories, the dataset was filtered to observations matching both NYT created keywords “Misinformation and Rumors” and “Fringe Groups”, resulting in a final count of 721 stories. Comments for each of these articles were then scraped totalling almost 190,000 observations. The following graph (fig. 4) shows the number of articles and comments over time from the misinformation beat.

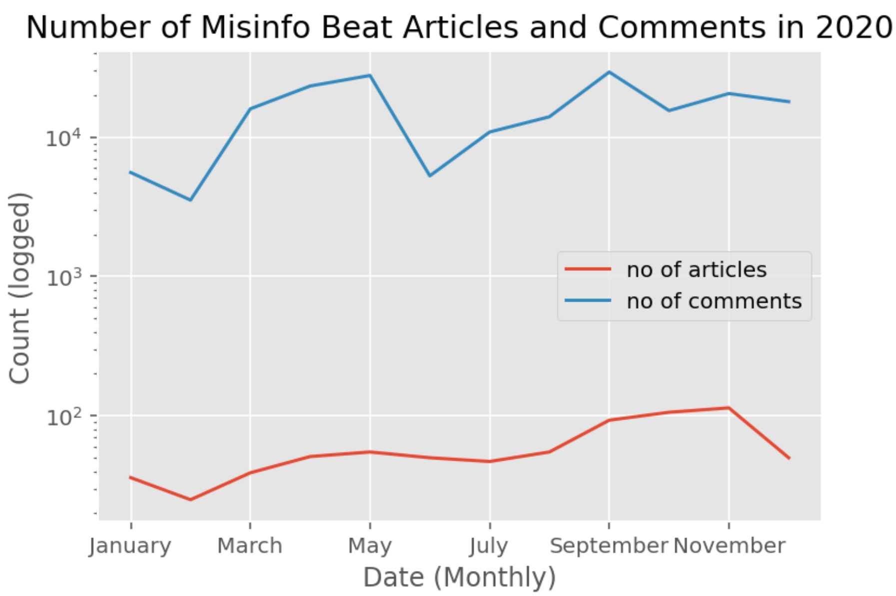

Fig 4. The number of comments and articles tagged as “Misinformation and Rumors” and “Fringe Groups” trended upwards in 2020.

The graph indicates a general positive trend in both the number of articles and the number of comments from the misinformation beat over time. While there was an earlier peak in comments during May 2020, the number of articles related to misinformation and conspiracies peaked during the election in November 2020.

If not identified as an article from the misinformation beat, or in this case, from the Daily Distortions section of the NYT, the article was labeled as “general news”. General news covers all other keywords except for “Misinformation and Rumors” and

“Fringe Groups”. Keywords in the general news group include anything from “Arts” to “Sports” with a total of 5,5428 articles . The dataset was thus divided into articles about misinformation (misinformation beat) and articles not directly related (general news). On average, the number of comments on Misinfo Beat articles were higher (mean = 262 comments) than the number of comments on general news stories (mean = 90 comments).

Comparing linguistic features to assess misinformation beat’s uniqueness through salient terms and topics models **-** Is the misinformation beat actually a unique area of reporting? One means of finding its unique features is through a comparison of linguistic features in the misinformation beat dataset versus the general news dataset. After processing and cleaning the text using the NLTK library, I employed a term-frequency, document-frequency comparison to find salient terms and to compare them across the two datasets. I also conducted a topic model comparison to get a high level overview of the discourse. Latent Dirichlet Allocation models leverage Bayesian probability to measure the likelihood terms will be seen together (Blei, 2012). I used coherence scores to find the optimal number of topics per model and then compared these models through a Jaccard distance metric to assess similarity. The Jaccard distance measures the probability of terms in one topic appearing in the probability distribution of another topic. By comparing the similarity or dissimilarity of topics from one dataset’s model to another, we can identify the relative uniqueness of the language employed by the misinformation beat.

3.3. Q2: Does the misinformation beat contain a greater risk of exposure to conspiracies than other types of news?

Quantifying opportunities to leave the NYT - In order to assess the risk of exposure to conspiratorial content, I first had to identify opportunities to leave the website. To do so I used BeautifulSoup to parse through the html of each observation, an article in this context, and extracted all links that were embedded in the body of the article. Domains were extracted from all shared links in a story using the python package TLDExtract. Links and domains were further filtered to identify all internal directs, i.e. website URLs that contain the phrase “nytimes.com” , “nytco.com”, and “tbrand”, which is the NYT branding studio. By searching for the string “nyt”, I was also able to detect any links that might have been shortened. The remaining links that directed readers off the NYT website were labelled as "external”. The number of external links out of the total number of links shared in an article became a new metric, a ratio that describes the relative opportunity a reader has to go off-site.

Comparing citations and patterns of information sharing in the misinformation beat to general news stories - To visualize the structure of information sharing, I generated citation networks for articles in the misinformation beat and for general news articles. The web URL of a NYT article was one node and an external domain was another. If a domain was shared in that article, an edge connected the domain to the article URL. Only external domains were connected to NYT nodes, so as not to inflate the network with the common practice of websites linking to their own content. I used the PageRank algorithm which accounts for inbound links with respect to the edge weight, in order to identify the relative influence of these sources and which domains were most highly cited (Heidemann et al, 2010).

Manually inspecting citations to categorize shared links as primary versus secondary sources of conspiratorial content - However, since conspiratorial content is most likely to be a rare phenomenon, the links had to be manually inspected in order to whether a shared external link contained primary source conspiratorial content or was an interpretation of that content by a secondary author. Because there were over 55, 000 articles in the general news dataset, I took a random sample of n = 50 from both the misinformation-related articles and non-related articles. The external links were then identified in each article, with 276 external links in the misinformation-related and 219 external links in the general news dataset. I then manually inspected the contents of these 475 total links.

After an initial reading of the links and their contents, I created a four point schema based on the type of information shared (primary versus secondary) and whether it was conspiracy related or not. With the focus on identifying primary versus secondary sources, this classification system relies on research standards of source selection. The table below (table 1.) offers a taxonomy of the external links shared.

Classification categories of links shared in news articles				
0 - Not related	1 - Primary source, conspiratorial content)	2 - Secondary source, news coverage of conspiracies	3 - Primary Source, tech reaction	4 - Primary source, conspiracy correction
The link is not directly related to conspiracy content.	The link contains primary source content such as documents, images, or first hand accounts of aspects of a conspiracy.	The links contains interpretation of primary sources, activities, or events related to conspiracy content	The link contains released statements and blog posts from technology companies in response to conspiracies.	The link contains content that directly contradicts conspiratorial content such as statements by health organizations and tweets by public officials.

Table 1. The criteria for the manual labeling of links focuses on primary versus secondary forms of information.

The taxonomy is first organized by information hierarchies of primary or secondary information, i.e. first hand accounts or original documents versus secondary analysis or descriptions of events. The second hierarchy of labels is then whether the linked web page contains conspiracy content. The first label, “not related” refers to links that did not refer to conspiratorial content as the subject of the page. For example, a general website like the homepage of Google would not be considered related and would therefore be given the label of zero. The second label, “Primary source, conspiratorial content” defines links to web pages that share first hand accounts of conspiracies or other pieces of evidence that support conspiratorial belief. For example, photos of night skies with a suspicious orb shared with the caption “E.T. is real” would be considered a primary source and related to alien conspiracies. The third label, “Secondary source, news coverage” could be any news article or journalistic content such as a reporter’s tweet that describes, analyzes, or comments on a conspiracy. A link to a Washington Post article that describes the rise of Qanon would be one such example. The fourth label, “Primary source, tech reaction” emerged from a second review of the links. As the emergence of misinformation, disinformation, and conspiratorial content intersects with socio-technical systems like Facebook and Twitter, many technology companies released statements on how their platform was dealing with the dissemination of information, hence the need for the fourth label. Finally, the last label is called “Primary source, conspiracy correction”. If, for example, a public health official offered corrections to a Covid-19 conspiracy, this would be an example of a primary source correction as government documents or officials speaking in the capacity of their roles become institutional evidence.

3.4. Q3. What is the potential distance from readers to conspiratorial content because of the misinformation beat?

Simulating a user's journey from NYT misinformation article to direct conspiracy content - So far the methods have compared the features, the structure of outbound links, and the content of misinformation-related stories to that of general news. One of the hypotheses of this project is that increased exposure to misinformation coverage may place readers on the path of a conspiracy rabbit hole. One way to operationalize this distance between journalism and conspiracy theory is to simulate a reader's journey from a NYT article to conspiracy content. Each article that links to media coverage of conspiracy theories or direct links to the conspiracy content (1 or 2 in the taxonomy) represents an opportunity for a reader to leave the site and engage with misinformation coverage further.

To test this theory I simulated semi-random walks through the Internet with the NYT articles as origin points. I took all of the articles that contained a conspiratorial link either as a primary or secondary source. I then followed and scraped that link for html content. From each of these links, I scraped and parsed the website, randomly selecting an external link from the page. Not all links on the page were thus weighted evenly as external links were highly overweight to better mimic a user's jump from one site to the next.

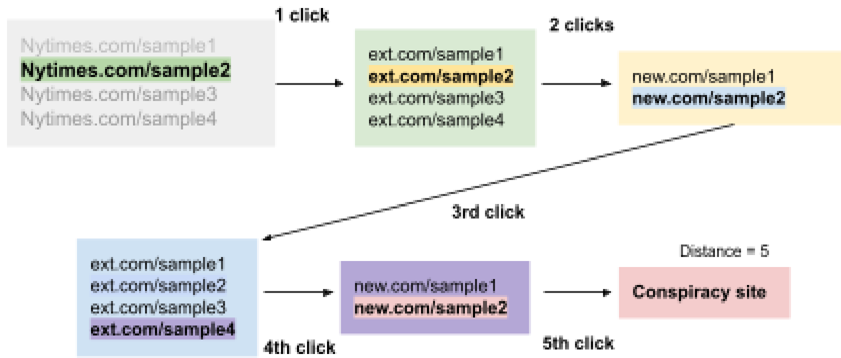

Fig 5. An illustration of the semi-random walk process wherein external links are overweighted.

While the aforementioned citations graphs above represent one-hop traversal, or one reader link-click on an article, a user's journey may take them multiple clicks away from the NYT website. In order to mimic this multiple click journey while also anticipating the growing potential for failure as the scraper moved away from the curated, this process of extracting links and scraping for more links was repeated up to five jumps, in an emulation of user "clicks", per story.

If the link destination was a social media platform, such as TikTok or Facebook, the simulation stops the journey for that origin story. Social media sites are stopping points because of their anti-scraping policy. Of course, there are a number of factors that may influence user behavior and how likely they are to click on a link. The purpose of reconstructing possible pathways in a semi-random manner where external links have an equal probability of reader interaction, is not to assume the behavior of users but to discern the distance between the NYT and fringe conspiracy sites as a measure of potential exposure.

The newly scrapped links were then also manually inspected for further identification of primary or secondary conspiratorial content. The total number of scrapped links successfully "walked to" and then manually labelled were 212 links. In total, 707 links were manually inspected in this iterative process.

4. Findings

4.1. Q1: What are the linguistic features of the misinformation beat and is it unique?

Firstly, the number of comments on Misinfo Beat articles were higher (mean = 262 comments) than the number of comments on general news stories (mean = 90 comments). Secondly, the salient terms in the misinformation dataset were found to be related to misinformation actors and conspiracies, especially around the pandemic (fig. 6.a. And 6.b.). For example, the salient terms in the Misinfo Beat set included “dr._yan” and “dr_zelenko”. The former of the conspiracies references Dr. Yan, a Chinese scientist that claims Covid-19 was lab made (Timberg, 2021). Dr. Zelenko originally claimed that the US Food and Drug Administration backed treating coronavirus with hydroxychloroquine, azithromycin and zinc (LaFraniere & Roose, 2020). The general news dataset was too large and disparate to produce comparable tf-idf results, as most of the term-xs were names without titles or context. Hence the need for topic models which could offer more interpretable results across the two datasets.

Top Salient Terms with Tf-IDF by Dataset

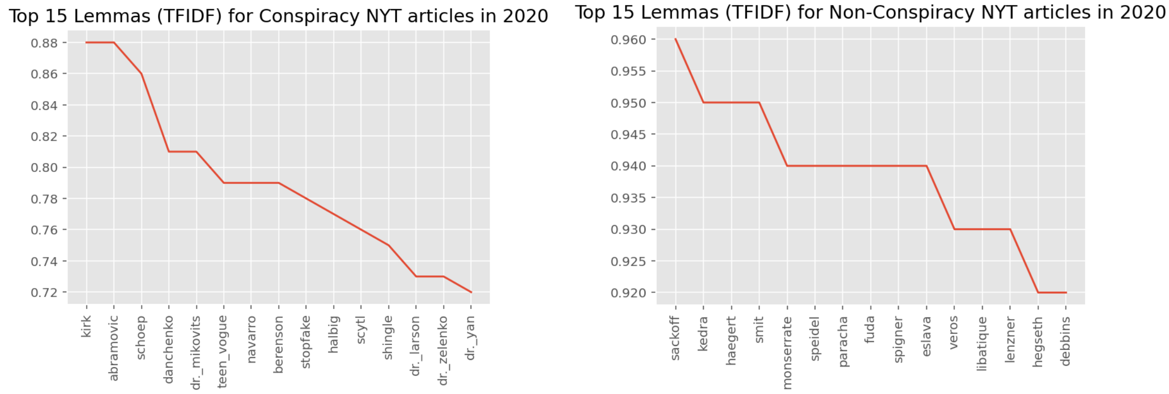

Fig 6.a. And 6.b. Side-by-side comparison of the terms with the highest tf-idf scores in the two datasets.

Coherence Score Models by Dataset

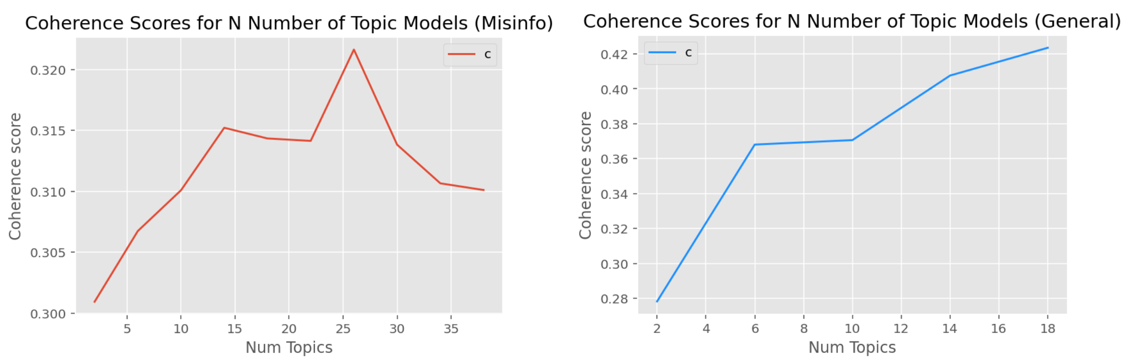

Fig 7.a. & 7.b. Coherence scores show 14 and 18 topics as the optimal number of topics for misinformation-related articles (7.a.) and non-misinformation articles (7.b.), respectively.

Topics models were generated for each dataset (Appendix 1.a and 1.b). For the misinformation beat dataset, though the coherence score peaks at 26 topics, however a total of six topics with a coherence value of 0.306 were more human-interpretable (Fig. 7a.). For the general news corpus the optimal number was 18 topics with a coherence value of 0.42 (Fig. 7b.). Upon a manual inspection, most of the topics in the general news corpus aligned with the various news sections such as U.S. politics, lifestyle, and sports, arts, and etc. The misinformation beat showed topics related to former

President Trump, social media, QAnon and the coronavirus.

I also used the Jaccard distance (Appendix 2.) to identify any overlapping topics in the two datasets in a pairwise fashion. The Jaccard distance uses set theory to measure whether a word in one topic is in the probability distribution of the other topic . Most of the topics diverged, however, the most correlated topics between general news and conspiracy news were about the US presidential election, perhaps indicative of how conspiracy theories during the pandemic were political in nature.

4.2. Q2: Does the misinformation beat contain opportunities for exposure to conspiracies?

Misinformation beat articles offer more opportunities to leave the NYT - Articles from the NYT have a median value of 18 links shared per article, and of those links shared, it has a median value of 2 non-NYT websites per article. This indicates that generally, the NYT directed user traffic to other parts of their website, keeping the attention on NYT articles. I took a random sample of n= 500 of both datasets, the general news and misinformation beat articles, and plotted the distribution of ratios, the number of external links is divided by the total number of links shared in an article, for the misinformation beat versus general news. The distribution indicates that articles about misinformation or conspiracies are more likely to share external links than the general news dataset.

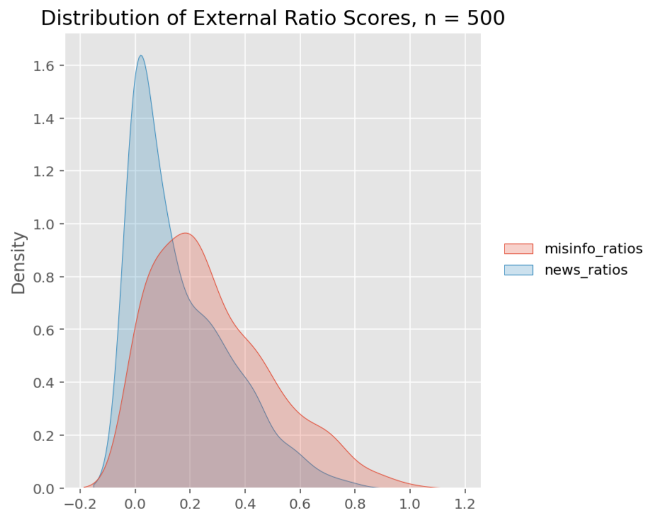

Fig 8. A distribution of the ratio of external to total links shared per article.

The difference in external website links between general news articles and conspiracy articles was found to be statistically significant. These findings indicate that the opportunities to leave the site are higher on Misinfo Beat articles.

Related Sample T-test, n = 500 (Misinfo vs General link ratios)	
T-Statistic	p-value
7.45	4.26E-13

Table 2. Statistical significance of the difference in ratios, the external to total links shared per article.

Most cited outbound links were social media and other news outlets - The top non-NYT domains shared in general news show social media as one of the top references, Twitter, Facebook, Reddit, and Whatsapp to name the top five scores. WashingtonPost, another legacy newspaper and Wall Street Journal were the other top referenced news outlets. CDC and the NIH were the top health organizations mentioned in the citation graph. In the Misinfo top pagerank scores, we see that there are more news sources referenced, such as WaPo, CNN, NBCNews, Politico, Guardian, and Buzzfeed.

Top 15 Domains by PageRank Influence

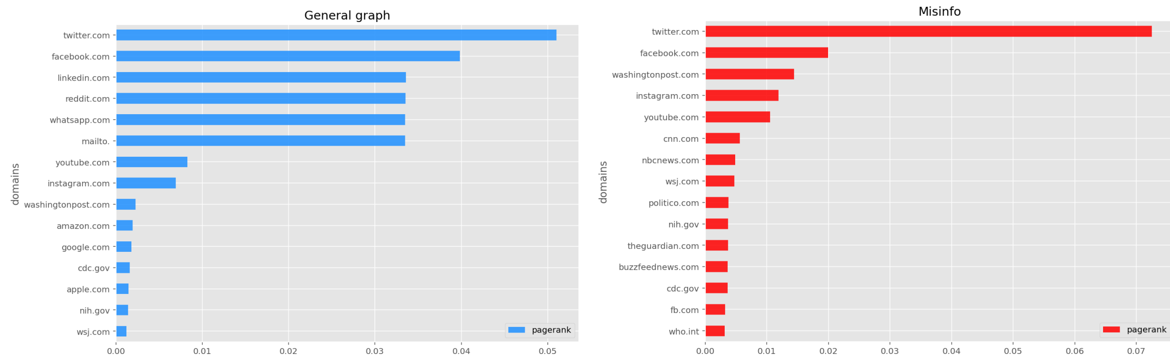

Fig 9.a. and 9.b. PageRank scores of cited domains in a NYT article demonstrate popular sources.

One interpretation is that in covering conspiracies and misinformation, the most highly cited sources are other reputable news outlets. However, these sources may embed conspiratorial content in their coverage and so it was necessary to conduct a manual inspection of the cited links in order to identify primary versus secondary content.

Manual inspection finds more primary and secondary sources of conspiratorial content in misinformation beat articles - From a random sample of 50 articles per dataset, over 400 external links were extracted and then manually inspected. The results of the inspection found that over 55.76% of the external links shared from the misinformation beat article were news sources covering conspiracy theories, and were therefore secondary sources. The percent of primary source conspiratorial content was .05% in the general dataset versus 14.5% in the misinformation beat dataset. Thus both primary and secondary sources of conspiracy related information were higher in the misinformation beat articles.

Manual Inspection Results

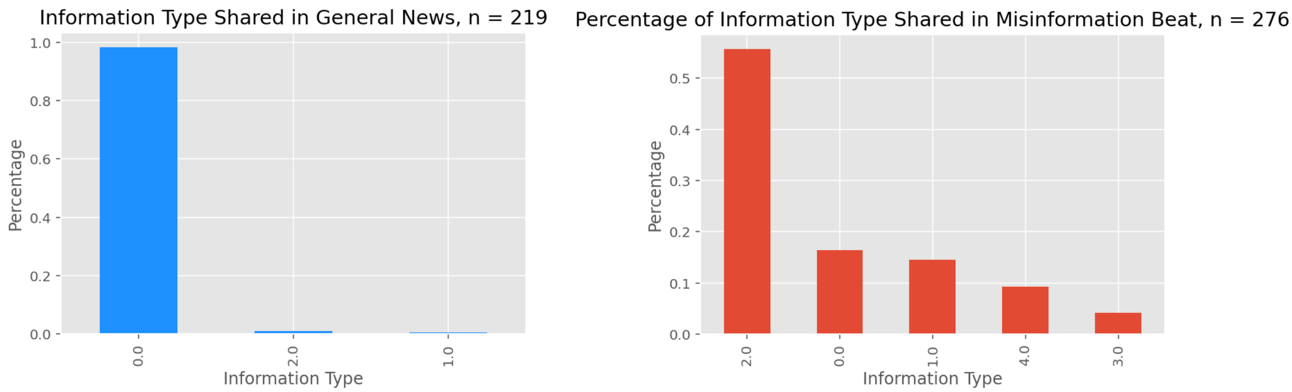

Fig 10.a. & 10.b. Information types include 1-Primary source and conspiratorial content, 2- secondary source and conspiracy-related, 0 - not conspiracy-related, 4 - primary source, tech response to conspiracies.

The following section will then take the articles containing primary and secondary conspiratorial content, i.e. primary sources that contain a first-person account of someone's belief in a conspiracy and secondary sources that report, describe, correct, or analyze the conspiracy phenomenon, as starting points in a simulated walk to identify the potential distance between the NYT and conspiracy theories. According to the results of the manual inspection then, a reader is more likely to come across primary sources of conspiratorial content because of the misinformation beat. A related sample t-test found the difference to be statistically significant between the two datasets.

Related Sample T-test, n = 150 (Misinfo vs General link ratios)	
T-Statistic	p-value
-17.87	6.76e-39

Table 2. Statistical significance of the difference in ratios, the external to total links shared per article.

4.2. Q3: What potential distance do misinformation beat readers have to conspiratorial content?

A limited simulation found distance between primary source conspiracy and NYT more likely to d = 1 - The semi-random simulation found that the distance between NYT and primary source conspiracies were likely to be equal to 1. This is, in the limitations of a simulated walk, the primary sources were more likely to be a direct citation within an article and were not found downstream.

Recall that 50 articles were randomly sampled from the two datasets. From these articles all external links were manually labelled. The links that were secondary or primary sources and also related to conspiracy content were then scraped for further links. Two citation networks were then created from the simulation results, where a NYT article was a starting node and all subsequent links were connected nodes.

Out of the dataset articles not identified as the misinformation beat, three out of the 219 links extracted from the n= 50 sample were labeled as news coverage of conspiracies (label 2) or a primary source (label 1). The three links originated from two articles. In one article, rumors around a Catholic bishop turned out to be true. The NYT article on the religious scandal referenced the news coverage by the sources Slate and the Washington Post. In the other, a NYT article linked to the tweet of a Black Lives Matter leader. This protester found evidence of a conspiracy to assassinate him when Google searches suggested the terms “shoot him” and “shoot in the head” as auto-complete text with his name. Though a semi-random walk was simulated for the three links found, the subsequent links did not contain conspiratorial content. These examples are cases that demonstrate some conspiracies are indeed true, and demonstrate the value of covering such stories. From the network, only one primary source was found (the BLM leader's tweet), and it was one traversal hop from the origin story.

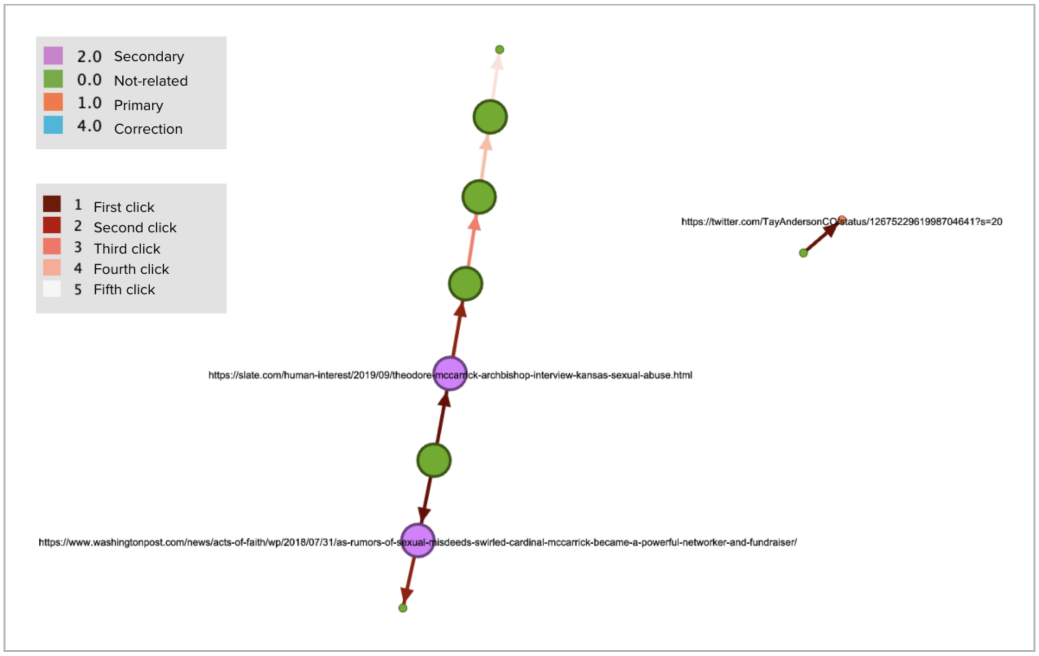

Fig. 11. The simulated semi-random walk from non-misinformation beat stories that contained conspiracy-related links.

A second citation network was generated by the misinformation beat sample. 189 out of 276 external links were either labelled as primary source or secondary source and related to conspiracy theories. From these links a semi-random walk was simulated, starting at the origin NYT stories. The links connected closest to the center node were direct references in the article. Each subsequent traversal hop represents a simulated click away from a NYT story so that the number of clicks equals the distance between nodes.

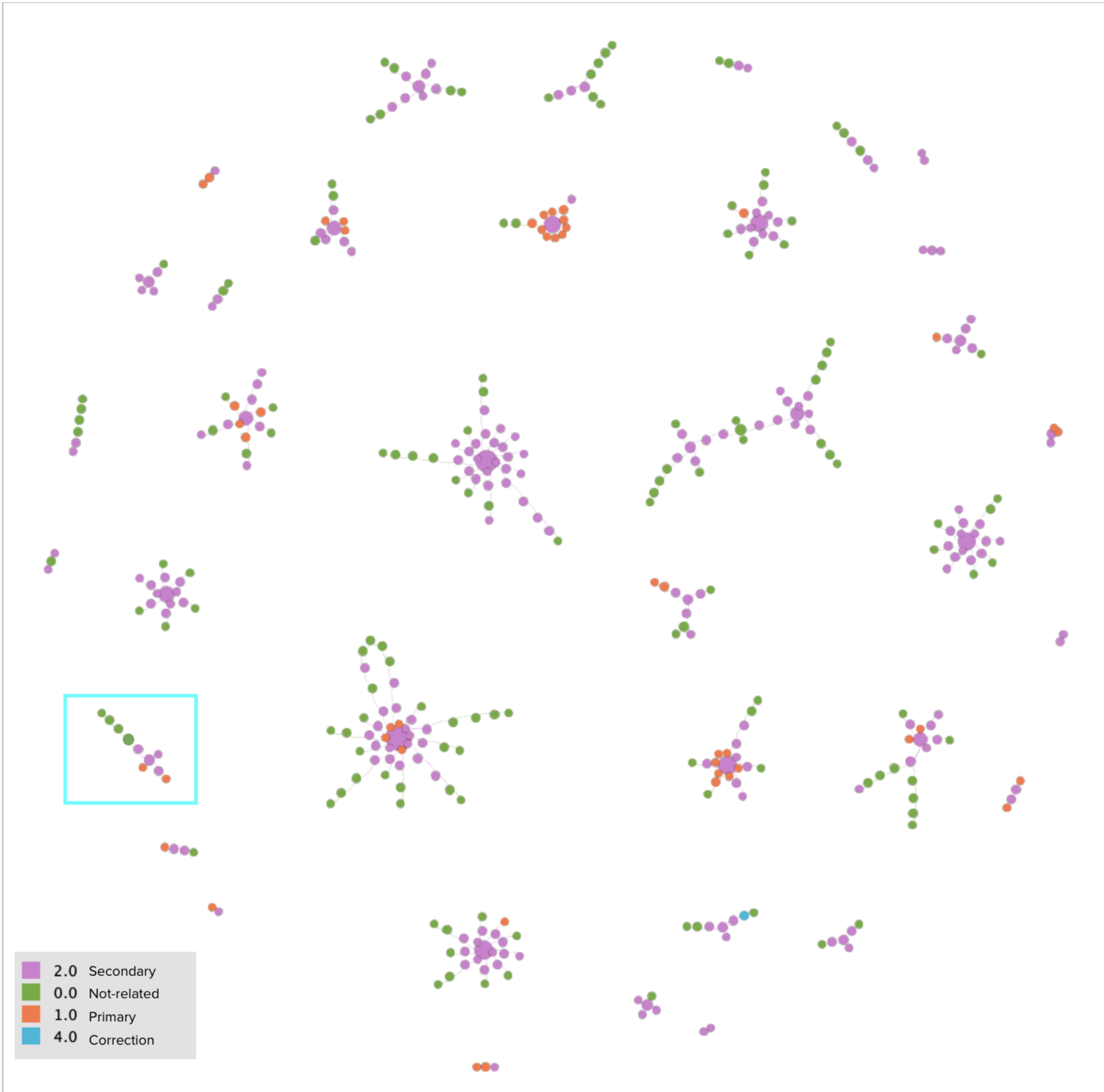

Fig. 11.a. The simulated semi-random walk from misinformation beat stories that contained conspiracy-related links.

In Fig. 11.a., 48 links were found to be primary sources and conspiracy-related. Of these direct examples of conspiracy content, six links were found at the second click. Two conspiracy primary sources were found via the third click. This means that 38 out of the 48 (about 79%) examples of primary source content were directly referenced in an article from the NYT misinformation beat.

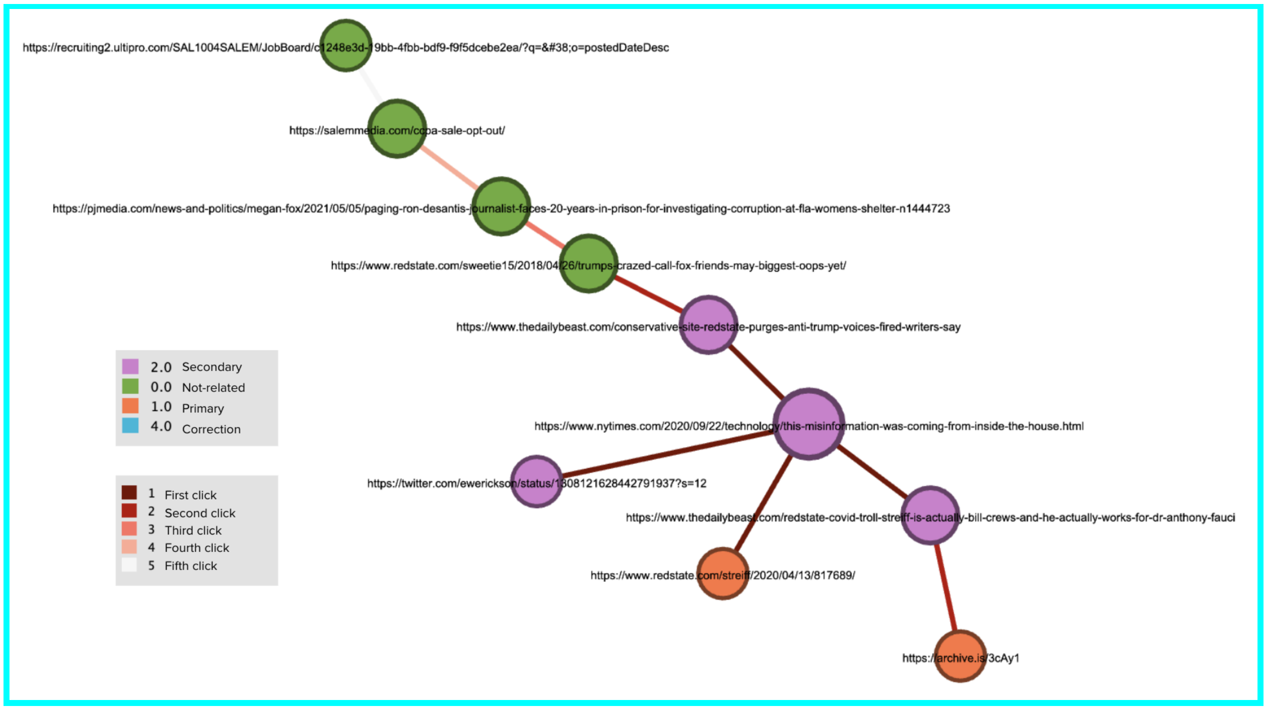

Fig. 11.b. A close look at one of the clusters in which the semi-random walk encountered a conspiracy primary source.

Let's take a closer look at one of the simulated semi-random walks. In the example cluster above (Fig. 11.b.), the NYT article (largest purple node) shared direct links to a RedState article (a right-leaning news blog), a tweet, and two Daily Beast articles (a left-leaning news blog). From one of the Daily Beast articles, the simulation found an since-archived RedState article that referred to Covid-19 as the Wuhan virus.

Because the power of the experiment is relatively small, and there are a number of limitations to the simulation approach, the results presented above should be taken as an initial exploration into the Internet space around the misinformation beat.

5. Discussion

Takeaways - The findings offer a few insights. The first, that stories covering misinformation and conspiracies grew in both volume and engagement over the course of 2020, and that engagement is on average higher than for general news, with an average score of 262 comments for misinformation beat versus an average comment score of 90 for other news. The second, that readers of conspiracy articles have more opportunities to leave the NYT site than readers of general news articles, with an average 0.2 ratio versus an average of 0.0, respectively. The third, that the content of general news versus the misinformation beat tends to diverge in topics with a decorrelation index ranging from 0.7 to 1.0. The highest similarity topic in both models were related to the election and Republicans, indicating that the beat is mostly unique in language but shares political subjects with other sections of the NYT. The fourth is that the misinformation news stories are more likely to link to other news outlets, than to social media accounts, as the domains with the highest PageRank in the misinformation beat were other media outlets. The fifth, Misinfo Beat articles are more likely to share primary source conspiratorial content (at 14.50% versus 0.005%) with statistical significance, suggesting readers are more likely to come across conspiracy content through interaction with Misinfo Beat stories than general news stories from the NYT. The fifth and final takeaway is that though an attempt at simulating semi-random walks from NYT articles encountered more conspiratorial content in explicitly referenced links, it is unclear if the results were dependent on journalistic practices of sharing examples of misinformation or on the performance of the scraper. Therefore the simulation may be seen as an exploratory first step in assessing the distance between conspiracy theories and the new misinformation beat.

Data limitations - There are, of course, a number of limitations with this study. One limitation is that, assuming a reader clicks on an external link from a story, all subsequent links were filtered to external links on the site, overweighting the possibility a user would leave the pge. The experiments are meant to be a proxy for and simulation of human behavior, however, such an approach to Internet interactions is incomplete as website design and content are likely to be major factors in influencing what a reader may click on next. Another limitation of the study is that linked content was categorized as conspiratorial or not through a manual inspection process. While other researchers have also utilized a grounded, mixed-methods approach, there remains potential for false negative or positive identification and general room for interpretation. Further research should improve on the power of the simulation experiment by scaling up the sample size, repeating trials, and finding the universe of all possible links.

In a chaotic information landscape, such as the one experienced during 2020, there are unique ethical considerations in covering conspiracies. How can news outlets share urgent information such as the rise of QAnon believers without exposing readers to potential harm? An analogous, though imperfect, parallel could be drawn with the choice to broadcast mugshots.

In recent years, newsrooms have moved away from showing the mugshots of persons who have been arrested and not convicted, because of the harmful long term effects on the subject (Blakinger, 2020). In this example, journalists have chosen to withhold some information and to restrain their practices from falling into “clickbait”. Conversely, conspiracy stories are no doubt sensational and could be seen as perfect clickbait by their very nature. The fact that the average number of comments are higher on conspiracy stories than on general news stories, speaks to their popularity. But a potential solution could be for newsrooms to scale back the depth of their coverage and to not reconstruct conspiracy theories for audiences. There should still be a misinformation beat, especially given its social prevalence, but there should be some friction to information that benefits trusting readers and prevents exposure to harm.

6. Conclusion

Due to the persistent presence and continuously ambiguous nature of the Covid-19 pandemic in the United States, collective anxieties have fed into conspiracy theories. But the question of how best to engage with these theories, from both a journalism and an academic research standpoint, remains open. That conspiratorial content like QAnon should so directly mingle with the public discourse that it becomes an oxymoron of a “mainstream” conspiracy theory, speaks to the extent that these conspiracies have commanded attention. At times the call to attention has been violent, both epistemically and physically. This study explored the mutual news making practices of conspiracies and the New York Times, and offers as a contribution, a formal social and media theory of the new misinformation beat and the first computational analysis of the emerging practice.

7. References (Works Cited)

Al-khateeb, Samer, Muhammad Nihal Hussain, and Nitin Agarwal. "Social Cyber Forensics Approach to Study Twitter's and Blogs' Influence on Propaganda Campaigns." In *Social, Cultural, and Behavioral Modeling*, edited by Dongwon Lee, Yu-Ru Lin, Nathaniel Osgood, and Robert Thomson, 108–13. Lecture Notes in Computer Science. Cham: Springer International Publishing, 2017. https://doi.org/10.1007/978-3-319-60240-0_13.

Ahmer Arif, Leo Graiden Stewart, and Kate Starbird. 2018. Acting the Part: Examining Information Operations Within #BlackLivesMatter Discourse. Proc. ACM Hum.-Comput. Interact. 2, CSCW, Article 20 (November 2018), 27 pages. DOI:https://doi.org/10.1145/3274289

American Conspiracy Theories. Oxford, New York: Oxford University Press, 2014.

AP NEWS. "AP Explains: What's behind Trump's Town Hall Answer on QAnon," October 16, 2020. https://apnews.com/article/trump-town-hall-qanon-explained-b4614edad1aae27a0cbcb9dfe0c0c44c.

Blei, David M. "Probabilistic topic models." Communications of the ACM 55, no. 4 (2012): 77-84.

Harper's Magazine. "[Article] The Paranoid Style in American Politics, By Richard Hofstadter," November 1, 1964. https://harpers.org/archive/1964/11/the-paranoid-style-in-american-politics/.

Aupers, Stef. "'Trust No One': Modernization, Paranoia and Conspiracy Culture." *European Journal of Communication* 27, no. 1 (March 1, 2012): 22–34. https://doi.org/10.1177/0267323111433566.

Bratich, Jack Zeljko. "Trust No One (On the Internet): The CIA-Crack-Contra Conspiracy Theory and Professional Journalism." *Television & New Media* 5, no. 2 (May 1, 2004): 109–39. https://doi.org/10.1177/1527476403255810.

Bruns, Axel, Stephen Harrington, and Edward Hurcombe. "'Corona? 5G? Or Both?': The Dynamics of COVID-19/5G Conspiracy Theories on Facebook." *Media International Australia* 177, no. 1 (November 1, 2020): 12–29. https://doi.org/10.1177/1329878X20946113.

Connolly, Jennifer, Joseph Uscinski, Casey Klofstad, and Jonathan West. "Communicating to the Public in the Era of Conspiracy Theory." *Public Integrity* 21 (May 15, 2019): 1–8. https://doi.org/10.1080/10999922.2019.1603045.

Donovan, Joan. MIT Technology Review. "Covid Hoaxes Are Using a Loophole to Stay Alive—Even after Content Is Deleted." Accessed October 23, 2020. https://www.technologyreview.com/2020/04/30/1000881/covid-hoaxes-zombie-content-wayback-machine-disinformation/.

Craft, Stephanie, Seth Ashley, and Adam Maksl. "News Media Literacy and Conspiracy Theory Endorsement." *Communication and the Public* 2, no. 4 (December 1, 2017): 388–401. https://doi.org/10.1177/2057047317725539.

De Maeyer, Juliette. "Taking Conspiracy Culture Seriously: Journalism Needs to Face Its Epistemological Trouble." *Journalism* 20, no. 1 (January 1, 2019): 21–23. https://doi.org/10.1177/1464884918807037.

Douglas, Karen M., Robbie M. Sutton, and Aleksandra Cichocka. "The Psychology of Conspiracy Theories." *Current Directions in Psychological Science* 26, no. 6 (December 1, 2017): 538–42. https://doi.org/10.1177/0963721417718261.

Douglas, Karen, Robbie Sutton, Mitchell Callan, Rael Dawtry, and Annelie Harvey. "Someone Is Pulling the Strings: Hypersensitive Agency Detection and Belief in Conspiracy Theories." *Thinking and Reasoning* 22 (January 2, 2016). https://doi.org/10.1080/13546783.2015.1051586.

Fea, John. *Believe Me: The Evangelical Road to Donald Trump*. Wm. B. Eerdmans Publishing, 2018.

Featherstone, Mark. "The Obscure Politics of Conspiracy Theory:" *The Sociological Review*, May 9, 2014. https://journals.sagepub.com/doi/10.1111/j.1467-954X.2000.tb03519.x?icid=int.sj-abstract.similar-articles.3.

Fee, Christopher R., and Jeffrey B. Webb. *Conspiracies and Conspiracy Theories in American History [2 Volumes]*. ABC-CLIO, 2019.

———. *Conspiracies and Conspiracy Theories in American History [2 Volumes]*. ABC-CLIO, 2019.

Gentzkow, Matthew and Shapiro, Jesse. "What Drives Media Slant? Evidence From U.S. Daily Newspapers." *Econometrica* 78, no. 1 (2010): 35–71. https://doi.org/10.3982/ECTA7195.

Grinberg, Nir, Kenneth Joseph, Lisa Friedland, Briony Swire-Thompson, and David Lazer. "Fake News on Twitter during the 2016 U.S. Presidential Election." *Science* 363, no. 6425 (January 25, 2019): 374–78. https://doi.org/10.1126/science.aau2706.

Guess, A., Nagler, J., & Tucker, J. (2019). Less than you think: Prevalence and predictors of fake news dissemination on Facebook. Science advances, 5(1), eaau4586. Retrieved from https://advances.sciencemag.org/content/5/1/eaau4586

Hamborg, Felix, Karsten Donnay, and Bela Gipp. "Automated Identification of Media Bias in News Articles: An Interdisciplinary Literature Review." International Journal on Digital Libraries 20, no. 4 (December 1, 2019): 391–415. https://doi.org/10.1007/s00799-018-0261-y.

Heidemann, Julia, Mathias Klier, and Florian Probst. "Identifying key users in online social networks: A pagerank based approach." (2010).

Hussein, Eslam, Prerna Juneja, and Tanushree Mitra. "Measuring Misinformation in Video Search Platforms: An Audit Study on YouTube." *Proceedings of the ACM on Human-Computer Interaction* 4, no. CSCW1 (May 28, 2020): 1–27. https://doi.org/10.1145/3392854.

Iyyer, Mohit, Peter Enns, Jordan Boyd-Graber, and Philip Resnik. "Political Ideology Detection Using Recursive Neural Networks." In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 1113–1122. Baltimore, Maryland: Association for Computational Linguistics, 2014. https://doi.org/10.3115/v1/P14-1105.

Jolley, D., Douglas, K.M. and Sutton, R.M. (2018), Blaming a Few Bad Apples to Save a Threatened Barrel: The System‐Justifying Function of Conspiracy Theories. Political Psychology, 39: 465-478. https://doi.org/10.1111/pops.12404

Kata, Anna. "A Postmodern Pandora's Box: Anti-Vaccination Misinformation on the Internet." *Vaccine* 28, no. 7 (February 17, 2010): 1709–16. https://doi.org/10.1016/j.vaccine.2009.12.022.

Kim, Minchul, and Xiaoxia Cao. "The Impact of Exposure to Media Messages Promoting Government Conspiracy Theories on Distrust in the Government: Evidence from a Two-Stage Randomized Experiment." *International Journal of Communication* 10, no. 0 (July 27, 2016): 20.

Konkes, Claire, and Libby Lester. "Incomplete Knowledge, Rumour and Truth

Seeking.” *Journalism Studies* 18 (October 1, 2015): 1–19. https://doi.org/10.1080/1461670X.2015.1089182.

Krafft, P. M., and Joan Donovan. “Disinformation by Design: The Use of Evidence Collages and Platform Filtering in a Media Manipulation Campaign.” *Political Communication* 37, no. 2 (March 3, 2020): 194–214. https://doi.org/10.1080/10584609.2019.1686094.

Krämer, Benjamin. “How Journalism Responds to Right-Wing Populist Criticism.” In *Trust in Media and Journalism: Empirical Perspectives on Ethics, Norms, Impacts and Populism in Europe*, edited by Kim Otto and Andreas Köhler, 137–54. Wiesbaden: Springer Fachmedien, 2018. https://doi.org/10.1007/978-3-658-20765-6_8.

Kramer, A. D., Guillory, J. E., & Hancock, J. T. (2014). Experimental evidence of massive-scale emotional contagion through social networks. Proceedings of the National Academy of Sciences, 111(24), 8788-8790. Retrieved from https://www.pnas.org/content/111/24/8788.short

LaFraniere, Sharon & Kevin Roose. (2020, May 1) “Doctor Who Promoted Malarial Drug Draws Scrutiny of Federal Prosectors”. The New York Times. Accessed May 7, 2021. https://www.nytimes.com/2020/05/01/us/coronavirus-doctor-zelenko-malaria-drug.html

Lazer, D., Baum, M., Ognyanova, K., Simonson, M., Green, J., Druckman, J., … Quintana, A. (2021, February 15). The COVID States Project #39: Public attitudes towards the storming of the Capitol building. https://doi.org/10.31219/osf.io/3qfs9

Levendusky, Matthew S. “Why Do Partisan Media Polarize Viewers?” American Journal of Political Science 57, no. 3 (2013): 611–23. https://doi.org/10.1111/ajps.12008.

Li, Chang, and Dan Goldwasser. "Encoding social information with graph convolutional networks forPolitical perspective detection in news media." In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pp. 2594-2604. 2019.

McClure Haughey, Melinda, Meena Devii Muralikumar, Cameron A. Wood, and Kate Starbird. “On the misinformation beat: Understanding the Work of Investigative Journalists Reporting on Problematic Information Online.” *Proceedings of the ACM on Human-Computer Interaction* 4, no. CSCW2 (October 14, 2020): 133:1–133:22. https://doi.org/10.1145/3415204.

Melley, Timothy. *Empire of Conspiracy: The Culture of Paranoia in Postwar America*. Cornell University Press, 2000.

NPR.org. “Michigan Authorities Charge 13 In Plot To Kidnap Governor Gretchen Whitmer.” Accessed October 23, 2020. https://www.npr.org/2020/10/09/922065759/michigan-authorities-charge-13-in-plot-to-kidnap-governor-gretchen-whitmer.

NW, 1615 L. St, Suite 800Washington, and DC 20036USA202-419-4300 | Main202-857-8562 | Fax202-419-4372 | Media Inquiries. “A Look at the Americans Who Believe There Is Some Truth to the Conspiracy Theory That COVID-19 Was Planned.” *Pew Research Center* (blog). Accessed November 20, 2020. https://www.pewresearch.org/fact-tank/2020/07/24/a-look-at-the-americans-who-believe-there-is-some-truth-to-the-conspiracy-theory-that-covid-19-was-p lanned/.

———. "A Look at the Americans Who Believe There Is Some Truth to the Conspiracy Theory That COVID-19 Was Planned." *Pew Research Center* (blog). Accessed November 20, 2020. https://www.pewresearch.org/fact-tank/2020/07/24/a-look-at-the-americans-who-believe-there-is-some-truth-to-the-conspiracy-theory-that-covid-19-was-planned/.

Prooijen, Jan-Willem van, and Karen M Douglas. "Conspiracy Theories as Part of History: The Role of Societal Crisis Situations." *Memory Studies* 10, no. 3 (July 1, 2017): 323–33. https://doi.org/10.1177/1750698017701615.

"Managing the COVID-19 Infodemic: Promoting Healthy Behaviours and Mitigating the Harm from Misinformation and Disinformation." Accessed October 14, 2020. https://www.who.int/news/item/23-09-2020-managing-the-covid-19-infodemic-promoting-healthy-behaviours-and-mitigating-the-harm-from-misinformation-and-disinformation.

NBC News. "Russian Operation Masqueraded as Right-Wing News Site to Target U.S. Voters." Accessed October 24, 2020. https://www.nbcnews.com/tech/tech-news/russian-operation-masqueraded-right-wing-news-site-n1241710.

Resende, Gustavo & Melo, Philipe & Sousa, Hugo & Messias, Johnnatan & Vasconcelos, Marisa & Almeida, Jussara & Benevenuto, Fabrício. (2019). (Mis)Information Dissemination in WhatsApp: Gathering, Analyzing and Countermeasures. 10.1145/3308558.3313688.

Shahsavari, Shadi, Pavan Holur, Timothy R. Tangherlini, and Vwani Roychowdhury. “Conspiracy in the Time of Corona: Automatic Detection of Covid-19 Conspiracy Theories in Social Media and the News.” *ArXiv:2004.13783 [Cs]*, April 28, 2020. http://arxiv.org/abs/2004.13783.

Smith, Daniel, and John Protevi. “Gilles Deleuze.” In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Spring 2020. Metaphysics Research Lab, Stanford University, 2020. https://plato.stanford.edu/archives/spr2020/entries/deleuze/.

Starbird, Kate, Kate Starbird, and Kate Starbird. “Why It’s so Hard to Discern between Conspiracies and Intentional Disinformation.” Fast Company, July 27, 2020. https://www.fastcompany.com/90532423/why-its-so-hard-to-discern-between-conspiracies-and-intentional-disinformation.

Stempel, Carl, Thomas Hargrove, and Guido Stempel. “Media Use, Social Structure, and Belief in 9/11 Conspiracy Theories.” *Journalism & Mass Communication Quarterly - JOURNALISM MASS COMMUN* 84 (June 1, 2007): 353–72. https://doi.org/10.1177/107769900708400210.

Suiter, Jane, and Richard Fletcher. “Polarization and Partisanship: Key Drivers of Distrust in Media Old and New?” *European Journal of Communication* 35, no. 5 (October 1, 2020): 484–501. https://doi.org/10.1177/0267323120903685.

Sutton, Robbie M, and Karen M Douglas. “Conspiracy Theories and the Conspiracy Mindset: Implications for Political Ideology.” *Current Opinion in Behavioral Sciences*, Emotion, Motivation, Personality and Social Sciences, 34 (August 1, 2020): 118–22. https://doi.org/10.1016/j.cobeha.2020.02.015.

Starbird, Kate. “How a Crisis Researcher Makes Sense of Covid-19 Misinformation”. (OneZero). (2020, March 8th). https://onezero.medium.com/reflecting-on-the-covid-19-infodemic-as-a-crisis-informatics-researcher-ce0656fa4d0a

Stewart, E. “Anti-maskers explain themselves” Vox. (2020, August 7). https://www.vox.com/the-goods/2020/8/7/21357400/anti-mask-protest-rallies-donald-trump-covid-19

Timberg, C. “Scientists said claims about China creating the coronavirus were misleading. They went viral anyway.” (2021, Feb. 12) https://webcache.googleusercontent.com/search?q=cache:laIW1DDBBcAJ:https://www.washingtonpost.com/technology/2021/02/12/china-covid-misinformation-li-meng-yan/+&cd=1&hl=en&ct=clnk&gl=us

Wakabayashi, D., Alba, D., & Tracy, M., (2020, April 17). “Bill Gates, at Odds With Trump on Virus, Becomes a Right-Wing Target”. The New York Times. Retrieved from https://www.nytimes.com/2020/04/17/technology/bill-gates-virus-conspiracy-theories

Waisbord, Silvio. “Truth Is What Happens to News.” *Journalism Studies* 19, no. 13 (October 3, 2018): 1866–78. https://doi.org/10.1080/1461670X.2018.1492881.

Waters, Anita M. “Conspiracy Theories as Ethnosociologies: Explanation and Intention in African American Political Culture:” *Journal of Black Studies*, July 26, 2016. https://doi.org/10.1177/002193479702800107.

Webb, Jeffrey B.. Conspiracies and Conspiracy Theories in American History [2 Volumes]. United States: ABC-CLIO, 2019.

Wood, Michael J. “Some Dare Call It Conspiracy: Labeling Something a Conspiracy Theory Does Not Reduce Belief in It.” *Political Psychology* 37, no. 5 (2016): 695–705. https://doi.org/10.1111/pops.12285.

Tambuscio, M., Ruffo, G., Flammini, A., & Menczer, F. (2015, May). Fact-checking effect on viral hoaxes: A model of misinformation spread in social networks. In Proceedings of the 24th international conference on World Wide Web (pp. 977-982). ACM. Retrieved from https://iris.unito.it/retrieve/handle/2318/1557768/128173/nostro.pdf

Tangherlini, Timothy R.. “An automated pipeline for the discovery of conspiracy and conspiracy theory narrative frameworks: Bridgegate, Pizzagate and storytelling on the web”. PLoS One. (June 16, 2020): doi: 10.1371/journal.pone.0233879

8. Appendix

A.1.a. Topic Model from Misinformation Beat articles

Topic Model - Misinfo, topn = 15					
QAnon - 0	Election - 1	Coronavirus - 2	Trump - 3	Facebook - 4	Coronavirus - 5
trump, president, election, post, state, call, qanon, twitter, video, new, group, claim, facebook, republican, official	trump, president, facebook, twitter, post, group, call, company, election, tweet, republican, state, report, official, biden	trump, facebook, election, president, call, state, information, coronavirus, group, biden, post, american, company, report, official	trump, facebook, post, company, president, election, group, american, show, video, virus, call, state, social, coronavirus	facebook, trump, post, group, company, twitter, election, president, call, republican, state, american, political, tweet, account	trump, facebook, election, twitter, post, company, president, group, coronavirus, new, american, video, virus, try, call

A.1.b. Topic Model from non-Misinformation Beat articles

Topic Model - General, topn = 15					
Sports - 0	Climate - 1	NYC - 2	Politics - 3	Voting - 4	Science - 5
team, game, player, season, league, play, sport, first, fan, last, two, coach, start, win, baseball	climate_change, dr., climate, fire, scientist, color_strong, color_visit, environmental, auto_overflow, hidden_strong, rotate_overflow, rotate_transform, hidden_transition, border_0, auto_focus	city, home, two, food, restaurant, house, park, street, new_york, open, back, place, building, area, water	trump, state, president, case, republican, law, official, democrat, court, decision, congress, call, washington, federal, election	state, voter, vote, election, republican, democrat, party, candidate, poll, support, race, johnson, campaign, political, voting	flight, mar, nasa, space, pilot, plane, mission, fly, launch, moon, rover, boeing, earth, spacecraft, modern_love
Race - 6	Foreign Policy - 7	Education - 8	Business - 9	Covid-19 - 10	Literature - 11
black, white, woman, history, protest, american, racism, community, church, black_life, america, right, name, change, movement	china, country, chinese, hong_kong, government, beijing, united_state, britain, new, world, museum, economy, international, japan, official	school, child, student, parent, family, dr., help, kid, university, learn, home, mother, teacher, think, college	company, business, new, employee, market, industry, facebook, million, worker, big, american, government, pandemic, sale, help	virus, state, coronavirus, test, pandemic, case, new, city, hospital, new_york, patient, covid-signup-module_margin-left, dr., reopen, health	book, write, novel, human, story, animal, author, woman, life, writes, read, writer, publish, world, dr.
Protests - 12	Rent - 13	Arts - 14	Film - 15	Election - 16	War - 17
police, officer, city, official, kill, protest, two, protester, police_officer, report, arrest, government, case, accord, call	pay, money, job, benefit, program, city, worker, home, cost, new_york, tax, rent, family, payment, help	artist, art, music, new, look, show, first, song, dance, new_york, feel, thing, life, start, call	show, life, first, film, think, love, write, thing, story, play, feel, movie, look, u, two	trump, biden, president, campaign, white_house, republican, president_trump, democrat, political, obama, candidate, former, event, sander, call	american, country, united_state, iran, war, political, military, america, leader, power, world, force, government, think_article, commit_publishing

A.2. Topic model dissimilarity index through the Jaccard Distance