

Ecosystem or Echo-System? Exploring Content Sharing across Alternative Media Domains

Kate Starbird[†], Ahmer Arif[†], Tom Wilson[†], Katherine Van Koevering[†],
Katya Yefimova[†] & Daniel Scarnecchia[§]

[†] University of Washington; [§] Harvard T.H. Chan School of Public Health
{kstarbi, ahmer, ktvank, tomwi, yekaty}@uw.edu, dscarnecchia@hsph.harvard.edu

Abstract

This research examines the competing narratives about the role and function of Syria Civil Defence, a volunteer humanitarian organization popularly known as the White Helmets, working in war-torn Syria. Using a mixed-method approach based on seed data collected from Twitter, and then extending out to the websites cited in that data, we examine content sharing practices across distinct media domains that functioned to construct, shape, and propagate these narratives. We articulate a predominantly alternative media “echo-system” of websites that repeatedly share content about the White Helmets. Among other findings, our work reveals a small set of websites and authors generating content that is spread across diverse sites, drawing audiences from distinct communities into a shared narrative. This analysis also reveals the integration of government-funded media and geopolitical think tanks as source content for anti-White Helmets narratives. More broadly, the analysis demonstrates the role of alternative newswire-like services in providing content for alternative media websites. Though additional work is needed to understand these patterns over time and across topics, this paper provides insight into the dynamics of this multi-layered media ecosystem.

Introduction

In September 2016, a documentary featuring the Syrian Civil Defense, a volunteer response group in Syria who are also known as the White Helmets (WH), was released to wide acclaim. Later, the film would win an Oscar for Best Documentary. It was set within the context of the then five-year-old civil war in Syria, troubling claims about the brutality of Syrian President Bashar al-Assad who had been accused of bombing civilians and medical workers (Fouad et al. 2017), the rise of the Islamic State (IS) and other extremist groups in areas outside the government’s control, and a massive exodus of refugees escaping the

violence (BBC News 2016). The documentary and subsequent sympathetic articles by mainstream media outlets worldwide brought global awareness to the plight of Syrian people, especially those resisting the Assad government.

This narrative, which promoted the role of the WH as a humanitarian response organization in Syria, functioned to grow solidarity between many Western audiences and the WH—as well as Syrian people from rebel-held areas who were seen as victims of the Assad regime. However, this narrative was not aligned with other views of the complex geopolitical landscape of the conflict. In particular, representatives and supporters of the Syrian government and its allies in that conflict (including Russia and Iran) resisted this sympathetic perspective. In response, a counter-narrative took shape and eventually spread to other online communities. Critics argued that the group, which is funded by Western governments, was a propaganda construct, supported by mainstream media, and employed as a tool of NATO interests in Syria (RT 2017). Some claimed that the WH aided and in some cases were themselves active in terrorist organizations (Beeley 2017). Supporters of the WH, in turn, accused their critics of orchestrating a propaganda campaign to spread a “conspiracy theory” about the group (e.g. Grey Ellis 2017). These contested narratives, which are still active, are spread through and shaped by various media—including mainstream news articles, alternative media articles, blog posts, social media posts and interactions, etc. The resulting information space is an evolving and multi-layered one.

This paper explores the dynamics of a subsection of the media ecosystem that was active around Twitter conversations about the WH during a three-month period in the summer of 2017. Using a mixed-method approach based on seed data collected from Twitter, and then extending out to the websites cited in that data, we examine content sharing practices across distinct media domains—articulating an alternative media, and, to a lesser extent, a mainstream

media “echo-system” of websites that repeatedly share content about this topic. Among other findings, our analysis reveals a small set of source domains (and authors) generating content that is spread across diverse domains, drawing audiences from distinct communities into a common narrative. This analysis also reveals the integration of government-funded media (RT, SputnikNews) and geopolitical think tanks (GlobalResearch) as source content for anti-WH narratives. More broadly, the analysis demonstrates the role of alternative “newswire” services in providing content for alternative media websites. Though more study is needed to understand these patterns over time and across topics, this paper provides insight into the dynamics of this multi-layered media ecosystem.

Background

The White Helmets and the Syrian Civil War

The ongoing conflict in Syria has taken more than 400,000 lives and displaced millions more. Armed conflict began during the 2011 Arab Spring, when anti-government protests calling for President Bashar al-Assad to step down escalated into full scale civil war between the Syrian government and those opposing Assad’s rule (ICRC 2012). It has since evolved into an internationalized, multi-sided conflict. Militants from the IS took advantage of the conflict to capture much of Eastern and Northern Syria (Whewell 2013). Russia and Iran initially supported the Assad government with materiel and financial assistance, and later committed troops, targeting Syrian opposition forces and IS militants (Associated Press 2013, Kramer & Barnard 2015). In 2014, the United States and Gulf League states began bombing IS in Syria and providing assistance to Kurdish forces fighting IS and Syrian opposition groups (Gordon 2014). UK joined this coalition in 2015.

The WH are a group of trained volunteer rescuers that operate throughout Syria’s opposition-controlled areas to assist civilians affected by the violence. According to the organization’s website (<http://syriacivildefense.org>), they abide by the fundamental principles of the Red Cross and work in accordance to Article 61 of the Additional Protocol I, which defines the activities that constitute civil defence: protecting civilians from hostilities or disasters, aiding recovery in the aftermaths of such events, and providing conditions for the survival of the civilian population. WH activities include search and rescue, evacuating buildings, firefighting, medical aid, providing emergency shelter and supplies (Pictet 1979.)

Online Propaganda & Disinformation in 2017

This research took place during a period (June 2017-January 2018) of global attention to the threat of misinfor-

mation, disinformation, and political propaganda, and the role of technology in facilitating their spread. In prior years, researchers optimistically noted the rise of citizen journalism (Gilmor 2006), and also acknowledged a related “crisis in journalism” (Fuller 2010) as economic, production and distribution models for “news” became disrupted. As traditional journalists and media outlets attempted to adapt to these new conditions, citizen journalists and media outlets outside the mainstream worked to establish their legitimacy—resulting in an information space where new voices were heard in new ways. Simultaneously, it became increasingly difficult for information consumers to assess the validity of the information they saw. The rise of partisan “fake news” sites and the subsequent appropriation of that term to challenge “mainstream” outlets (Qiu 2017) corresponded with record-low levels of trust in media and information (Swift 2016; Barthell & Mitchel 2017).

Additionally, for decades critics have explicated systemic biases in “mainstream” media¹—e.g. towards neo-liberal, pro-Western, and colonialist/imperialist ideologies (e.g. Herman & Chomsky 1988). As a prominent example, relevant here, were claims within the New York Times (Narasimhan 2005) that Saddam Hussein had weapons of mass destruction, a premise used to garner public support for the 2003 U.S. invasion of Iraq. These examples and the arguments constructed around them likely contribute to diminished trust in mainstream media.

Information Warfare and Online Disinformation

While these claims suggest strategies tied to traditional means of news production, recent evidence suggests others—including state and non-state actors—are working to leverage online technologies to forward their geopolitical goals (e.g. Weedon, Nuland & Stamos 2017). These actors are using a mix of automation and human curation to intentionally spread misleading information using online technologies such as social media (Woolley & Howard 2017).

In particular, Russia has been accused of conducting an “information war” that extends long-standing tactics of disinformation to new Internet-enabled channels (Pomerantsev & Weiss 2014). Though researchers and intelligence communities are still working to understand these strategies, evidence suggests that Russia and others are utilizing social media in conjunction with other channels to spread their messages (Weedon, Nuland, & Stamos 2017; Paul & Matthews 2016). The Russian government also utilizes its own media apparatus, including RT and SputnikNews, to

¹ Footnote: The term “mainstream media” has historically been used as a pejorative, especially by those who identify with “alternative” perspectives or media, to criticize the agenda-setting power of mass media. However, the “mainstream” and “alternative” terms have also become a common way to distinguish between media types. We employ these terms here for efficiency, but also with acknowledgment of their political roots and the tension between them.

forward their geopolitical aims. RT receives the vast majority of its funding from its government (Moscow Times 2014), and its editor-in-chief has argued that RT uses that funding to support “information warfare ... against the Western world” (DFRLab 2018).

Pomerantsev & Weiss (2014) define Russian disinformation as intended not to simply convince, but to confuse—to sow doubt and distrust across a society. The idea is that doubt can act to reduce agency. In other words, if we are not sure about what the truth is, we cannot choose the best action to take, and therefore will take no action. For this reason, disinformation campaigns do not need to rely on a single narrative or counter-narrative, but can work by presenting diverse and even contradictory narratives.

Information Operations and Non-Governmental Organizations (NGOs)

The Russian government has previously taken issue with—and actively worked to undermine the mission of—NGOs which they see as a threat to its geopolitical interests (Ambrosio 2007). In the early 2000s, Western NGOs supported “pro-democracy” civil society groups in Russia and its neighboring states and played a role in facilitating a shift away from Russia-aligned governments and policies. Following the color revolutions that took place in the former Soviet states, the Russian government argued that these activities represented unfair interference. Vladimir Putin specifically called out and criticized “pseudo-NGOs” funded by foreign governments and corporations for their role in destabilizing other countries (Putin 2012). In this study, we can see an extension of that criticism to a humanitarian response organization working—both through its efforts to assist affected people and to garner attention for their cause—against the geopolitical interests of Russia and its ally, the Syrian government.

Conducting Research on Information Operations

This paper is a small component of a larger research effort examining contested narratives involving the WH. Over the course of several months, our team spent hundreds of hours analyzing this data at multiple levels. This information space can be intensely *disorienting*. Our researchers repeatedly use this word to describe how the qualitative analyses affect us. The arguments and evidence presented in support of narratives on both sides are often compelling. Despite, or because of, deep engagement with this content, our researchers are often left in a state of confusion about what and whom to believe. In this study, we do not speak directly to this question. Instead, we focus on describing the media ecosystem surrounding these conversations—especially the dynamics of content-sharing practices—with the goal of gaining insight into how these narratives and counter-narratives are produced and disseminated.

Methods

Data Collection and Processing

Our White Helmets dataset (WH dataset) consists of tweets posted between May 27 and September 9, 2017. We created this collection using the Twitter Streaming API, initially tracked various keyword terms related to the Syrian conflict including geographic terms of affected areas. Later, we scoped this data to tweets that contained “white helmet” or “whitehelmet”, resulting in 135,827 tweets.

To understand the role played by external websites in Twitter conversations about the WH, we examined the links embedded within these tweets. 52,903 tweets contained a URL link. To process this data, we expanded shortened links, removed HTML parameters, and filtered out duplicates. We also removed links (approximately 35% of the total) that resolved to social media domains (i.e. Twitter, Facebook, Youtube) and newsreaders (feed-proxy.google.com).

The resulting set of 3410 distinct URLs was used to extract articles in a structured and automated fashion via a tool built using Newspaper, a python library designed for full-text and article metadata retrieval (Ou-Yang 2017). Due to how some web servers and content delivery networks were configured, we were unable to automatically scrape content from 111 domains. Subsequent analysis suggested that some of these domains play a prominent role in this information space (e.g. GlobalResearch), so we manually captured content from the top-10 most tweeted of these missing domains by traversing the URL to the article and copy-pasting its content. However, 322 URLs from the other 101 domains were omitted from the analysis.

Next, we passed the articles through an algorithm for detecting article similarity. This was done by computing the term frequency-inverse document frequency (tf-idf) statistic (Salton, Fox & Wu 1983) for each article and obtaining the cosine distance between the tf-idf vector for each pair of articles. The resulting matrix of similarity scores was used to identify duplicate articles across domains. We selected a threshold of $\geq 85\%$ to identify two articles as containing shared content. This level of similarity generally captured identical articles without being overly sensitive to small changes in image captions and article bylines.

After collapsing links to similar articles within the same domain, there were 1680 distinct news articles. From there, we identified 558 articles that had significant ($\geq 85\%$) overlap with another article within another domain in our set. Interestingly, nearly two-thirds (63%) of tweets with URLs cited one of these 558 articles that appeared on more than one domain.

Next, we constructed “paths” for each article—tracing all URLs in our dataset where that article appeared. We identified 135 paths, which provide insight into how con-

tent was shared across domains in the media ecosystem. We used these 135 paths to construct a network graph (Figures 1-3) where two nodes (domains) are connected if they appear in the same path—e.g. if one domain hosted an article that had $\geq 85\%$ similarity with an article in the other domain. The edge weight represents the number of similar articles shared by the two domains. These edges do not encode directionality, but merely reflect similar content. Using manual analysis to determine the original source of each article, we labeled each node as primarily a *source* (publishes original content), an *amplifier* (republished content from others), or a *hub* (published original and borrowed content). Nodes are sized by the number of tweets in the WH data that link to that domain, therefore representing the salience of this domain—and its articles—in the Twitter conversation. We use the ForceAtlas2 algorithm to determine the visual layout of the graphs and the Louvain method to detect communities for Figure 1.

Interpretative, Mixed-Method Analysis

We conducted interpretive, mixed method analysis of this data, expanding upon methods developed for the investigation of online rumors in the context of crisis events (Maddock et al. 2015). This approach iteratively blends quantitative and qualitative analyses—in this case generating a network graph to see larger patterns of content sharing across domains, and then using that representation as an entry point for a closer examination of both the practices of content sharing and the influential domains within this ecosystem. For the qualitative analysis, we focused primarily on the content within each domain, including its home page, about page, and the content-sharing practices visible within the specific articles cited in the WH dataset.

Note on Data and Privacy

In this paper, we identify a small number of prominent authors within the alternative media ecosystem. We considered anonymizing these names, but chose to publish real names because these authors are self-identified journalists and their patterns of activity—both within our data set and before/after—are important for understanding the nature of content sharing in this ecosystem. Several articles from both “sides” of this conversation are cited in the references.

Findings

Figure 1 shows the complete content-sharing domain network graph for the entire WH dataset (nodes sized by tweet volume and colored by Louvain-detected community). From a high level, this graph has several key features: two large clusters—Cluster A in pink on the left and Cluster B in blue on the right, with some connective tissue between them; and a small distinct community (Cluster C, in yellow)

low) loosely connected with Cluster B. There are also a large number of small, distinct clusters that are unconnected to the other clusters (in grey). One of these (Cluster D, in red) is interesting because it contains a highly tweeted, but disconnected, domain (see Table 1).

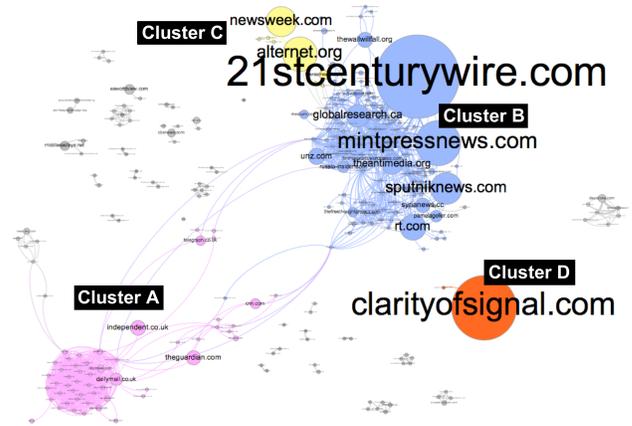


Figure 1: WH Content Sharing Domain Network Graph

Cluster A: An Associated Press News Cluster

Cluster A (Figure 2) is a relatively large component of 40 nodes, consisting of several western “mainstream” media sites (i.e. wsj.com, dailymail.co.uk, apnews.com), news outlets from the Arab world (aljazeera.com, arabnews.com, english.alarabiya.com), and other local and alternative media outlets from around the world. There are also a few news aggregators in this cluster, including castwb.com. Most of the edges in this cluster have weight of one, representing $\geq 85\%$ overlap of a single WH-related article within both web domains.

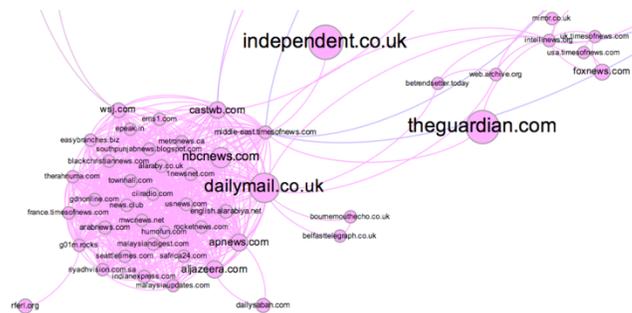


Figure 2: Close-up of Cluster A

Almost all of the connections in this cluster were related to a single article published in August 2017 (Mroue 2017), describing the murder of seven WH volunteers at their office in Idlib, Syria. This article was sympathetic to the WH, presenting them as “first-responders who have been known to risk their lives to save people from the civil war.”

This article constituted the largest path in our data—its content appeared in 44 different domains. Its original

source, almost always cited in the downstream articles, was the Associated Press (AP). The AP operates as an international, non-profit news cooperative and allows its partner news outlets to reuse its content. These partners pay to use the AP’s content in their own newspaper or website. This content-sharing model allows media outlets to provide coverage of diverse topics across the globe. Another similar agency, Agence France-Presse (AFP), appears within a small isolated cluster elsewhere in the graph. This model, which is a long-standing one that pre-exists the Internet (Fenby 1986), results in content sharing across news outlets in our data, including many that are considered “mainstream”. However, the relative scarcity of paths other than this AP path suggests that intensive content sharing about the WH was not observed in this set of websites.

A small number of nodes serve to connect—over one to three degrees—Cluster A (the AP cluster) to Cluster B (on the right of the graph). Within this connective tissue are a few “mainstream” media domains including the Telegraph, Independent, CNN, and the Guardian. These domains’ content was re-published by news aggregators (i.e. intellinews.org, f3nws.com) that connected those websites to other mainstream and alternative media domains. Most edges in this section have an edge weight of one—a single article, of which the mainstream media domain was the original source. The articles featured in this “connective tissue” area were generally supportive of the WH—promoting narratives that featured the WH as courageous volunteers who were risking their lives to rescue and provide medical assistance to Syrians who were injured by Syrian government and Russian military operations.

Cluster B: The Alternative Media Ecosystem

Most of the volume represented in the complete content-sharing network graph (Figure 1)—in terms of tweets, articles, and distinct domains—resides in Cluster B (Figure 3). The articles cited within these domains were highly critical of the WH. This cluster contains 110 nodes or web domains. 10,821 tweets in the WH data included a URL that linked to Cluster B, compared to only 2526 in Cluster A. Unlike Cluster A, which exhibits a consistent, nearly symmetrical structure, Cluster B is more heterogeneous in terms of both node size and edge weight. Edges vary in strength from one to seven articles. Thicker edges represent more consistent content sharing patterns over time.

Structural analysis on the domain graph in combination with content analysis of the articles within the content-sharing paths and the domains that hosted them reveal a few salient categories of domains: a small set of prominent alternative media “hub” domains that produce source content for the rest of the graph and occasionally re-publish each other’s articles (21stCenturyWire, MintPressNews, GlobalResearch); two Russian government-funded outlets

(RT, SputnikNews) that provide source content and occasionally amplify articles from the prominent hub domains; and a diverse set of alternative news aggregators that consistently amplify content from the peripheral sources.

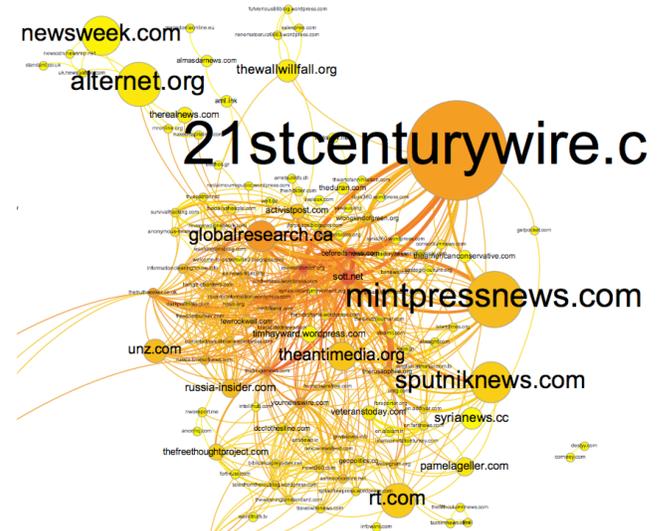


Figure 3: Close-up, expanded view of Clusters B and C; Colored by degree, Yellow=lower; Red=higher

As evidenced by their size in the graph, Cluster B includes many of the most highly tweeted domains in the data. Table 1 lists Top 10 domains within the WH collection in terms of tweet volume. Seven are located in Cluster B, and each of these was cited for multiple articles that were critical of the WH. Table 1 also provides the number of WH tweets that link to each domain and the “degree” of each domain in the graph—e.g. the number of other domains in the graph that are cited in the WH tweets for an article that has high similarity to one of the articles cited from this domain. We also note whether the domain is a source, a hub, or primarily an amplifier of content in this ecosystem.

Domain	Tweets	Degree	Network Role
21stcenturywire.com	3119	30	Central Hub
clarityofsignal.com	2391	1	Isolated
mintpressnews.com	1630	22	Central Hub
alternet.org	1219	6	Peripheral Hub
sputniknews.com	1110	16	Central Source
newsweek.com	1046	2	Peripheral Source
rt.com	879	17	Central Source
globalresearch.ca	707	33	Central Hub
theantimedia.org	682	19	Amplifier
unz.com	512	22	Central Source

Table 1. Top 10 Most Tweeted Domains in WH Dataset

Central Hubs in the Content Re-sharing Ecosystem

One key finding is that three of the most-tweeted domains (21stCenturyWire, MintPressNews, and GlobalResearch)

generated the majority of source content for the re-sharing practices reflected in Cluster B. Interestingly, these domains were not exclusively source domains, but also borrowed content from each other, and published original content from some of the same authors. One author, Vanessa Beeley—a British journalist and leading critic of the WH—had original articles on each of the three domains, content which later appeared on one or more of the others. These three domains are central hubs of content re-sharing in the anti-WH conversation.

21stCenturyWire is by far the most cited domain in the WH collection—3119 tweets link to 26 distinct articles within this web domain. 13 of these articles appear elsewhere in the ecosystem, republished in all or large part on other domains. Most of these articles were written by Beeley. 21stCenturyWire was founded by Patrick Henningsen, whose Guardian byline includes affiliations to RT and alternative news site Infowars. The website positions itself as grassroots, independent media that provides “news for the waking generation.” In the WH data, 21stCenturyWire is often the source of content that spreads across other prominent domains in the cluster (i.e. MintPressNews and theAntiMedia) and across several less trafficked domains (i.e. BeforeItsNews, YourNewsWire, and JewWorldOrder). 21stCenturyWire also re-publishes content that originally appears elsewhere in the graph, including articles from GlobalResearch, MintPressNews, and RT.

MintPressNews (MPN) is the third most-tweeted domain in the WH data—1630 tweets, 12 distinct articles. This domain is connected to 22 different domains in the graph, including many of the same domains as 21stCenturyWire. MPN describes itself as an “independent watchdog journalism organization” that features original reporting “through the lens of social justice and human rights.” Its home office is located in Minnesota (USA), but the website covers both national topics and foreign affairs. In our data, their content is strongly pro-Syrian government and critical of the WH. Apart from their original content, MPN has multiple news and syndication partners whose content they frequently re-publish. Their original articles appear on other prominent domains (21stCenturyWire and GlobalResearch) as well as common amplifiers (i.e. theAntiMedia.org, YourNewsWire, BeforeItsNews, Sott.net, and JewWorldOrder). They also re-share content from other domains, including 21stCenturyWire, ActivistPost, and TheAmericanConservative.

GlobalResearch is an influential hub within the content sharing network that appears 8th on our most tweeted list, being tweeted 707 times for 17 different articles. Seven of these had high similarity with other articles in the WH data. GlobalResearch is operated by the Centre for Research on Globalization, “a non-profit independent research and media organization” that describes itself as a think tank on economic and geopolitical issues. The center

is operated by Michel Chossudovsky, Professor Emeritus at the University of Ottawa. Chossudovsky has previously published claims of conspiracies related to world events—including that the September 11, 2001 attacks were not perpetrated by Islamic terrorists (Chossudovsky 2015). In our data, GlobalResearch is both a source and an amplifier. Underscoring its role in supporting this information ecosystem, of the top-10 most-tweeted domains, GlobalResearch has the highest degree, sharing articles whose content overlaps with 33 different domains. Though it has a strong, multi-article connection to 21stCenturyWire and shares several overlapping amplifiers with 21stCenturyWire and MintPressNews, its content also reaches a subset of domains outside of that subnetwork (i.e. LewRockwell.com, WashingtonsBlog, and FreedomBunker).

Government-Funded News Outlets

Two government-funded news outlets (SputnikNews and RT) are also within the Top 10 most-tweeted domains and Cluster B. **RT** is a Russian government-funded media outlet that provides content to international audiences. Founded in 2005 with the stated purpose of improving Russia’s image abroad, it has been accused of spreading disinformation and its U.S.-based affiliate has been forced to register as a foreign agent (Stubbs & Gibson 2017). Its current tagline is “Question More,” and its content often encourages readers to question western and mainstream narratives of world events. RT.com was tweeted 879 times for 16 different articles which all take a critical perspective of the WH. Within our content-sharing paths, RT is primarily a source domain. Its content is re-shared entirely or in large excerpts across 17 other domains. Interestingly though, all of its edges have the weight of one (article). The graph shows a large number of domains borrowing a single RT WH article (not always the same one), rather than consistently re-publishing their content. In addition to content-sharing that we can see through the similarity graph, several articles within other domains embed videos from RT in their content.

SputnikNews, founded in 2013 as the replacement for the “Voice of Russia”, is another Russian government-funded media outlet that features radio, television, and online content. Like RT, they have also been accused, primarily by western governments and media, of spreading disinformation and political propaganda that is favorable to the current Russian government (Dearden 2017). They are slightly more highly tweeted than RT in our WH data—1110 tweets for 15 articles, all critical of the WH. They are also primarily a source domain in this set. Their content is re-shared across 16 domains. Their most common amplifiers are Sott.net, theRussophile.org, and en.Addiyar.com (a Lebanese news outlet with a pro-Syrian government lean-

ing). Each of those websites re-shared multiple articles from SputnikNews in our data.

Central Source Domains

In addition to RT and SputnikNews, there are two other central source domains—UNZ.com and ActivistPost.com. UNZ is an alternative media outlet founded by Ron Unz, a former (conservative) political candidate in California. The outlet’s national security editor, Philip Giraldi, was the author of an article arguing that the WH are “a fraud”. This article, hosted on the UNZ website, was tweeted 512 times within our data. It was also re-published on 21 other domains that appear in the WH data. UNZ therefore performed as a central source domain, though solely through sharing of this single article.

Another central source domain is **ActivistPost**, an alternative independent media outlet whose tagline is “propaganda for peace, love, and liberty.” Their WH-related content is consistently critical, echoing many of the common narratives, claiming that they are a propaganda construct of mainstream media and western government interests. In terms of tweet count and compared to the more visible hubs, ActivistPost is relatively small—the domain was only tweeted 213 times. However, their role in the content sharing is significant. They are the source domain for eight different “paths” in the graph—e.g. eight of their original articles were re-published in all or part by other domains in the graph. In total, their content appeared in 18 domains, including central hub domains GlobalResearch and MintPressNews. All of their articles were authored by Brandon Turbeville, and include a Creative Commons license that enables the free distribution of the work.

Alternative News Aggregators

Another core component of Cluster B is a large number of Alternative News Aggregators that repeatedly share content that originally appears elsewhere in the graph. The most prominent of these aggregator domains, in terms of tweet volume, is theAntiMedia.org. This website positions itself as the “homepage for the independent media movement”, claiming to be a “non-partisan, anti-establishment news publisher and crowd-curated media aggregator.” **TheAntiMedia** functions in part as a news aggregator, pulling in articles from other alternative and independent media outlets and mixing those with its own original articles. In the WH data, theAntiMedia was tweeted 682 times for one original and three borrowed articles (from MintPressNews and 21stCenturyWire). Most of these tweets link to a single article, re-shared from MintPressNews.

A large number of other domains in the graph function exclusively as amplifiers. In Figure 3, domains are colored by their degree (number of edges), from yellow (few edges) to red (many edges). Many of the most connected web domains (in red) are primarily content borrowers that repeatedly republish content from other websites in the

graph. **Sott.net**, **theRussophile.org**, **JewWorldOrder.org**, and **BeforeItsNews.com**, are the domains with the highest degree in the graph, which also have thick, multi-article edges with the three central hub sites. All are exclusively amplifiers in this conversation—serially reposting content that first appeared elsewhere. Two other domains, **YourNewsWire** and **FringeNews**, are slightly less connected, but serve similar roles in a subnetwork in the lower-left-center of Cluster B. Many of these exclusively amplifier domains receive far fewer tweets for this content than other domains in our graph.

Another core component of the graph are the small (in terms of tweet volume) domains that are connected via thin edges to a relatively small number of domains. These domains typically appear in the graph for re-publishing one or two WH-related articles. Their relative positioning, near certain hubs and not others, may reflect a particular type of ideological targeting. For example, in the lower-left of the graph, near RT, TheFreeThoughtProject (a small source domain) and YourNewsWire, are domains like **ASheep-NoMore**, **GovtSlaves**, and **HumansAreFree** which promote content questioning many mainstream narratives and suggesting large-scale geopolitical conspiracies. And in the upper left, near GlobalResearch and UNZ are a collection of libertarian-leaning domains (**LewRockwell**, **Freedom-Bunker**, **HangTheBankers**). Most of these political and/or ideology-centered domains appear in the graph for a single article re-shared from one of the source or hub domains. The domains do not necessarily amplify everything in the ecosystem, but may pick and choose content to re-share, as their focus is not necessarily on the WH, but on a specific worldview that these anti-WH narratives reflect.

Cluster C: A Peripheral Hub: Reframing Mainstream Content for the Alternative Ecosystem

Cluster C is a small (in terms of number of domains), distinct community that is loosely connected to Cluster B. Two of the Top 10 domains are in this peripheral cluster: Newsweek and Alternet. Newsweek, a “mainstream” media outlet, was cited for six articles in the WH data. However, we only found evidence of one of these articles being re-shared on other domains. Rather, Newsweek appears in Cluster C, and is peripherally connected to Cluster B, through the Alternet domain—due to one highly-tweeted article that described how a WH volunteer was caught on video (and subsequently fired for) disposing of the mutilated bodies of Syrian soldiers. Alternet, an alternative news site that both aggregates content and posts its own articles, re-published this article with attribution. Alternet’s version, however, uses a different title, re-framing the original content to suggest that this incident was part of an ongoing pattern of misbehavior by WH volunteers, which aligns with other Alternet content critical of the WH. In

this case, Alternet functioned as a peripheral hub, borrowing source content from “mainstream” media and re-framing it to fit the predominant narrative of Cluster B.

Content Remixing Practices and Echo Effects

Though there are hundreds of distinct URLs in our tweet data, a significant percentage of the linked-to content is authored by a small number of prolific authors whose content is often re-shared and re-mixed elsewhere in the ecosystem. Beeley is the author of record for at least a dozen articles that appear in “paths” within the WH data—shared across several domains in Cluster B. In our tweet data, she was cited for original content in at three source domains: 21stCenturyWire, MintPressNews, and TheWallWillFall. She also repeatedly appears as a secondary source in articles by other authors through quotes, excerpts, and embedded videos of interviews. Similarly, Turbeville, who primarily publishes in ActivistPost, authored eight articles that were source articles for multiple content-sharing paths across the ecosystem.

The following example traces a single, short path that includes both Beeley and Turbeville and illustrates several of the diverse content remixing practices and echo effects that manifest in this ecosystem. On May 2 2017, ActivistPost published an article by Turbeville titled “Photos from Syria Show White Helmets and Nusra/Qaeda Are The Same Organization” (Turbeville 2017). This article used photos and videos that Beeley captured while in Syria and posted on her Facebook account. Later in the article there is a textual excerpt, citing Beeley, that describes the content in one of the videos. At the foot of the article there are seven links to other articles about the WH: four authored by Turbeville and published on ActivistPost, and three authored by Beeley published on 21stCenturyWire.

This same article (including Beeley’s photos, videos, and excerpt) is published on MintPressNews on the same day, citing Turbeville and ActivistPost, but removing the links to related content on 21stCenturyWire and ActivistPost. Thirteen days later, on May 15 2017, Beeley publishes the same ActivistPost article on TheWallWillFall, her personal blog. In this version the article is titled “WHITE HELMETS: Living next door to Al Qaeda in Aleppo” and Beeley is listed as the author. However, below an additional image that did not appear in the original version, Turbeville at ActivistPost is cited as the author, followed by the original article in its entirety—including the photos, videos and the quote from Beeley. There are now 13 links to other related White Helmet-related articles—ten of them on 21stCenturyWire, plus TruthDig, WrongKindofGreen, and Wikipedia. These circular citations and remix practices create another kind of echo effect within this system.

Discussion: Alternative Media Echo-System

In this research, we explored content sharing practices across media domains, using URL links in tweets to capture domains that were active in an online conversation, and an article similarity metric to determine domains that shared articles with high similarity. The conversation we focused on—views of the WH in relation to the ongoing the civil war in Syria—is a highly contested one with geopolitical significance. Using content similarity, we generated a network graph of shared content, and utilized that network graph to conduct a mixed-method, interpretative analysis of the structure and dynamics of content sharing across active domains.

Our analysis uncovered sharing practices among both mainstream and alternative media domains, as well as a few aggregator domains that bridged the two. Articles that originated (or echoed) within mainstream media (Cluster A) were largely supportive of the WH, reflecting some of the critique (from those in Cluster B) that the WH are favored by Western, mainstream media. Our graph shows a couple of clear examples of content sharing of and by mainstream media. In particular, Cluster A represents (primarily) a single “path” of an article originally posted by the AP and re-published by dozens of domains, including global and local mainstream media outlets and news aggregators. This activity is not insignificant, and clearly demonstrates that 1) mainstream media were participating in the WH conversation, primarily through the production and diffusion of pro-WH narratives; and 2) content-sharing is a component of mainstream news distribution.

However, for the WH conversation happening on Twitter during the summer of 2017, the vast majority of content production and amplification occurred on and through the alternative media domains represented in Cluster B. The content shared within these domains was strongly critical of the WH, promoting several related narratives that framed them as a propaganda construct and accused them of aiding, working with, or being terrorists. Using tweets as seed data, we were able to unwind trajectories of content-sharing across domains, articulating an alternative media ecosystem—or “echo-system”—of 130 distinct domains that provided source content for, or re-published existing content from, another domain in Cluster B. Analysis of this content-similarity structure, the domains that played significant roles within it, and the practices of content-sharing and remixing across domains, provides several interesting insights and leads to additional questions about how and why this echo-system has these properties.

Explicit Critique of Mainstream Media

One widespread theme within the domains that constituted the alternative media ecosystem in Cluster B is criticism of

mainstream media and skepticism or outright rejection of its narratives. Messaging across many domains suggests that the mainstream media is lying to members of the public who should come to this website to get the truth. While some of these domains are extremely conspiratorial in nature (a synergistic worldview to the anti-media arguments), others are more focused on questioning the motives of western governments, in some cases specifically around conflicts in the Middle East, and positioning mainstream media as tools of those governments in those conflicts.

Support of Russian Government

Perhaps not surprising, considering the position of Russia as an ally of the Syrian government (which views the WH as assisting rebel forces), many of the domains in this ecosystem are explicitly supportive of the Russian government. Beyond RT and Sputnik, there are a few others sites that focus on Russia-related topics from a point of view favorable to the current regime: Russia-Insider.com, Rus-sophile.org (Russia News Now), and Fort-Russ.com. Many of the other domains in Cluster B feature content supportive of Russian geopolitical positions (abroad) and specifically resistant to accusations that Russia had impact on recent elections in the U.S. and elsewhere.

Shared Content across Ideologically Diverse Sites

But perhaps more interesting—or more impactful—than the commonalities across domains are the differences between them. Superficially, many of the domains in the alternative media echo-system articulated here appear to promote different ideologies. Consider a selection of domains that appear in this echo-system—i.e. MintPressNews, JewWorldOrder, LewRockwell, FreedomBunker, UprootedPalstinians, TruePatriot, TheDailySheeple, TheFringeNews, Anonymous-News, MakeWarHistory, ActivistPost, and TheRussophile. This list includes websites with strong political themes reflecting distinct (and in some cases, seemingly conflicting) ideologies—including anti-imperialist left, libertarian, conservative and alt-right; as well as other more niche ideological leanings, including explicit anti-Semitism. These websites are publishing the same content, but inside very different wrappers. The content itself is not necessarily tailored for each community (though each domain may select the articles most likely to resonate with its audience), but it is packaged up for them, appearing within a domain that features other material that may appeal to a reader’s existing ideology. The effects of this kind of sharing may be to draw people from diverse, niche, political and ideological communities into a set of common narratives. We may think of these niche communities as being isolated and distinct, but here they are connected (in terms of common content) with other quite different communities.

Future Work to Understand the Drivers and Impacts of the Information Echo-System

Although in our Twitter seed data we found hundreds of domains and dozens of articles, we discovered that a small number of authors are responsible for a large proportion of highly-cited articles. We uncovered instances of circular attributions when authors cite themselves from other sources. We also demonstrated how similar content is repeated across in some cases vastly different domains. While such practices could reflect a more coordinated strategy, our evidence suggests that this complex ecosystem both has organic properties and is strongly influenced by a small set of politically, ideologically, and financially motivated actors and organizations.

Prior work has found that exposure to repetition of misinformation (which is not necessarily intentional) leads to a fluency effect—as people become familiar with claims they are more likely to judge them as true (Nyhan and Reifler 2012). By disseminating unverified or falsified stories to audiences through various channels, and from multiple sources, people may begin to assume they are true, regardless of the credibility of the individual sources (Paul and Matthews 2016). Although this research does not provide evidence of a coordinated strategy, the distribution of content across these seemingly distinct domains resembles a kind of intentional “astroturfing” campaign (Ratkiewicz et al. 2011) meant to exploit these cognitive biases. Future work is needed both to better understand the mechanisms underlying these patterns of content sharing and their effects on online audiences.

Limitations

One limitation of this work is the use of tweet data to “seed” the investigation of the surrounding ecosystem. This method had the advantage of allowing us to measure the impact of these domains on the online conversation, but the disadvantage of having our view of the content-sharing shaped by the contours of a single social media platform. The vast majority of tweets with URLs in our dataset link to domains that were critical of the WH and appear in Cluster B in the graph—the area that we have termed the alternative media echo-system. Additionally, and perhaps consequently, the majority of content-sharing “paths” across domains are in this area of the graph as well. It is likely that this tweet-seeded method resulted in a better view of content-sharing practices across alternative media than across mainstream media. Other less significant limitations include the loss of some URLs (and the articles/domains they pointed to) that we were unable to resolve, and the exclusion of article content that was not publicly available when we completed our automatic and manual scraping (in November 2017).

Acknowledgments

This research was supported by U.S. NSF grant #1715078 and ONR grants N00014-17-1-2980 and N00014-18-1-2012. We wish to thank students who provided significant assistance to this project, including Gordon, Ostin, and Vera.

References

- Ambrosio, T. 2007. Insulating Russia from a Colour Revolution: How The Kremlin Resists Regional Democratic Trends. *Democratisation*, 14(2), 232-252.
- Associated Press. 2013. Russia to Keep Helping Syria If It's Attacked. *Associated Press Financial Wire*, September 6, 2013
- Barthel, M. and Mitchell, A., 2017. Americans' attitudes about the news media deeply divided along partisan lines. *Pew Research Center*, 10.
- BBC News. 2016. Syria: The Story of the Conflict. *BBC*, March 11, 2016
- Beeley, V. 2017. EXCLUSIVE: The REAL Syria Civil Defence Exposes Fake 'White Helmets' as Terrorist-Linked Imposters. *21st Century Wire*, September 23, 2016
- Chossudovsky, M. 2015. Saudi Arabia's Alleged Involvement in the 9/11 Attacks and the 28 Pages: "Red Herring", Propaganda Ploy. *Global Research*.
- Dearden, L. 2017. NATO Accuses Sputnik News of Distributing Misinformation as Part of 'Kremlin Propaganda Machine'. *The Independent*, February 11, 2017
- Question That: RT's Military Mission. Assessing Russia Today's role as an "information weapon". Digital Forensic Research Lab (DFRLab) 2018. Medium. January 8, 2018
- Fenby, J. 1986. *The International News Services*. New York: Schocken Books.
- Fouad, F.M., Sparrow, A., Tarakji, A., Alameddine, M., El-Jardali, F., Coutts, A.P., El Arnaout, N., Karroum, L.B., Jawad, M., Roborgh, S. and Abbara, A. 2017. Health Workers and the Weaponisation of Health Care in Syria: A Preliminary Inquiry for The Lancet-American University of Beirut Commission on Syria (p. 2518). *The Lancet*.
- Fuller, J. 2010. *What is Happening to News: The Information Explosion and the Crisis in Journalism*. University of Chicago Press.
- Gillmor, D. 2006. *We the Media: Grassroots Journalism By the People, For the People*. O'Reilly Media, Inc.
- Gordon, M.R. 2014. U.S. Strikes Follow Plea by Syrian Opposition Leader on Behalf of Kurds. *The New York Times*. September 22, 2014
- Grey Ellis, E. 2017. Inside the Conspiracy which Turned Syria's First Responders into Terrorists. *Wired*. April 30, 2017
- Herman, Edward S, and Noam Chomsky. 1988. *Manufacturing Consent: The Political Economy of the Mass Media*. New York: Pantheon Books.
- International Committee of the Red Cross (ICRC). 2012. "Syria: ICRC and Syrian Arab Red Crescent Maintain Aid Effort amid Increased Fighting," Operational Update (Geneva: International Committee of the Red Cross, July 17, 2012).
- Kramer, A.E., and Barnard A. 2015. Russian Soldiers Join Syria Fight. *The New York Times*, October 6, 2015.
- Moscow Times. 2014. Looking West: Russia Beefs Up spending on global media giants. *Moscow Times*, September 23, 2014
- Maddock, J., Starbird, K., Al-Hassani, H.J., Sandoval, D.E., Orand, M. and Mason, R.M., 2015. Characterizing online rumoring behavior using multi-dimensional signatures. In *CSCW 2015*.
- Mroue, B. 2017. 7 White Helmets Medics Killed in Syria's Idlib. *Associated Press*, August 12, 2017
- Narasimhan, R. 2005. Looking Beyond Flawed Journalism: How National Interests, Patriotism, and Cultural Values Shaped the Coverage of the Iraq War. *Harvard International Journal of Press/Politics* 10.1:45-62.
- Nyhan, B. and Reifler, J.A., 2012. *Misinformation and Fact-Checking: Research findings from Social Science*. New America Foundation.
- Ou-Yang, L. 2017. Newspaper [Python library]. Available at: <https://github.com/codelucas/newspaper>
- Paul, C and Matthews, M. 2016. The Russian "Firehose of Falsehood" Propaganda Model: Why It Might Work and Options to Counter It. Santa Monica, CA: RAND Corporation, 2016
- Pictet, J. 1979. *The Fundamental Principles of the Red Cross*: commentary. Geneva: International Committee of the Red Cross
- Pomerantsev, P., & Weiss, M. 2014. *The Menace of Unreality: How the Kremlin Weaponizes Information, Culture and Money*. New York: Institute of Modern Russia.
- Putin, Vladimir. 2012. Russia and the Changing World, *Valdai*, February 22, 2012
- Qiu, L. 2017. Fact Check: Trump Blasts "Fake News" and Repeats Inaccurate Claims at CPAC. *The New York Times*, February 24, 2017
- Ratkiewicz, J., Conover, M., Meiss, M. R., Gonçalves, B., Flammini, A., & Menczer, F. 2011. Detecting and Tracking Political Abuse in Social Media. *ICWSM*, 11, 297-304.
- RT. 2017. UK's Shadowy £1bn Conflict Fund Being Kept Secret from MPs. *RT*, March 7, 2017
- Salton, G., Fox, E.A. and Wu, H., 1983. Extended Boolean information retrieval. *Comm. of the ACM*, 26(11), pp.1022-1036.
- Starbird, K. 2017. Examining the Alternative Media Ecosystem Through the Production of Alternative Narratives of Mass Shooting Events on Twitter. In *ICWSM* (pp. 230-239).
- Stubbs, J., and Gibson, G. 2017. Russia's RT America Registers as 'Foreign Agent' in U.S. *Reuters*. November 13, 2017
- Swift, A. 2016. Americans' Trust in Mass Media Sinks to New Low. *Gallup*. <http://news.gallup.com/poll/195542/americans-trust-mass-media-sinks-new-low.aspx>
- Turbeville, B. 2017. Photos From Syria Show White Helmets And Nusra/Qaeda Are The Same Organization. *ActivistPost*. May 2, 2017
- Weedon, J., Nuland, W., & Stamos, A. 2017. Information Operations and Facebook. Facebook, Version, 1, 27
- Whewell, T. 2013. BBC News. Inside Syrian City Run by Al-Qaeda. *BBC*. November 13, 2013
- Woolley, S. C., & Howard, P. N. 2017. Computational Propaganda Worldwide: Executive Summary. Working Paper 2017.11, *Oxford Internet Institute*. <http://comprop.oii.ox.ac.uk/>