

ABSTRACT

Title of Thesis: EARLY AND OFTEN: CAN REAL-TIME INTERVENTION BY TRUSTED AUTHORITIES HELP STOP A TSUNAMI OF DISINFORMATION?

Karin Zeitvogel, Master of Arts, Journalism.
2018

Thesis directed by: Prof. Deborah Nelson, Journalism

A tsunami of disinformation is washing over the world, with social media helping it to spread quickly and widely. The purveyors of disinformation use it to press their agenda by adding untruths where previously there were none, fabricating stories, reporting them out of context, or doctoring images to promote their message. In the past, disinformation has been a prelude to and run concurrently with other attacks, including cyber and conventional warfare, and when officials reacted to disinformation, they successfully slowed its flow but did not entirely stop it, and may not have “won” cyber or conventional battles. Researchers say even multiple corrections don’t fully stop disinformation, and sowing skepticism by forewarning of a probable disinformation campaign is the most successful way of staunching the flow. Tools have been developed to help detect disinformation rapidly but officials often don’t have a plan to track, correct or refute it.

EARLY AND OFTEN: CAN REAL-TIME INTERVENTION BY TRUSTED
AUTHORITIES HELP STOP A TSUNAMI OF DISINFORMATION?

by

Karin Zeitvogel

Thesis submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Master of Arts
2018

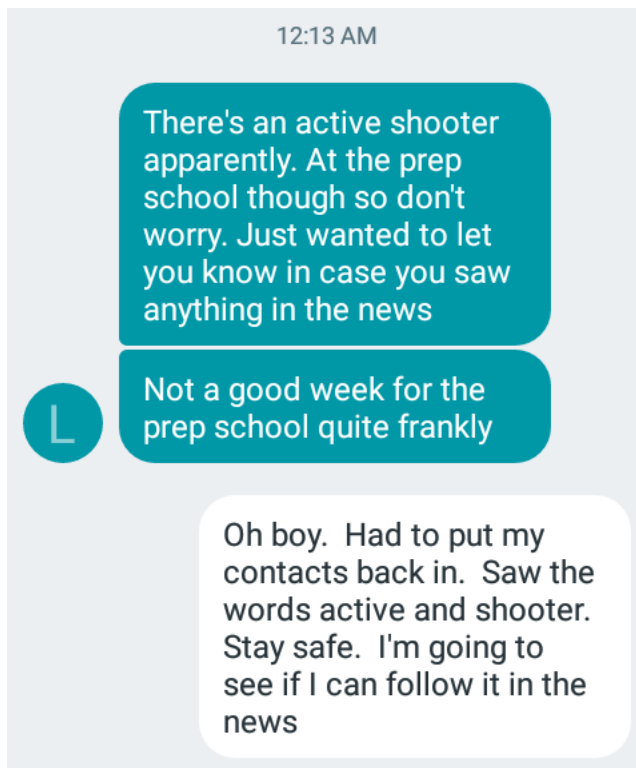
Advisory Committee:
Professor Deb Nelson, Chair
Senior Lecturer Christine Harvey
Professor Kevin Klose

© Copyright by
Karin Zeitvogel
2018

Preface

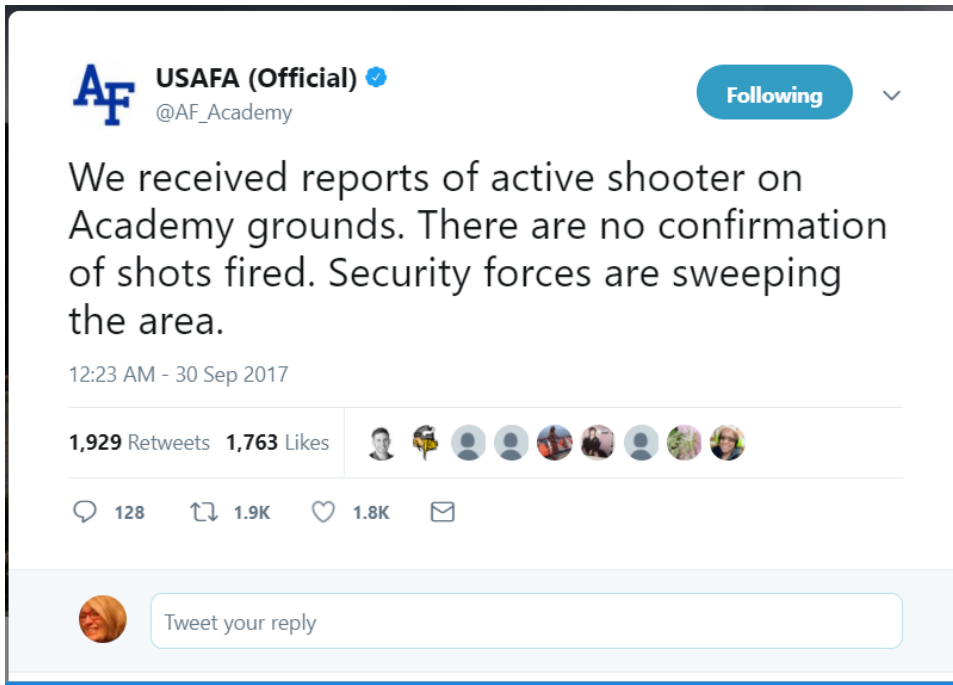
'Should be fine but I love you'

Just after midnight on September 29, 2017, I received this message from my son, 1,620 miles and two time zones away at the U.S. Air Force Academy (USAFA) in Colorado:



For the next three-and-a-bit hours, I sat in my living room, working my laptop with one hand and my phone with the other, trying to make sense for myself and my son of what was happening at USAFA.

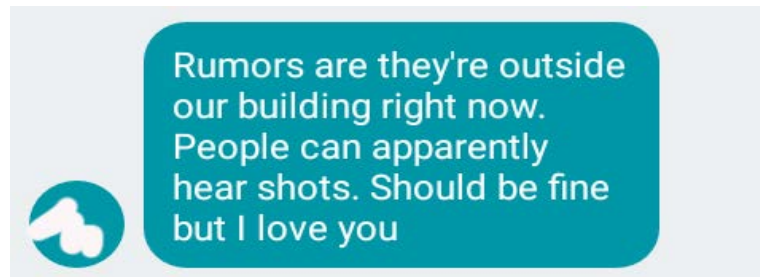
Ten minutes after that text from my son, the Air Force Academy tweeted about the incident.



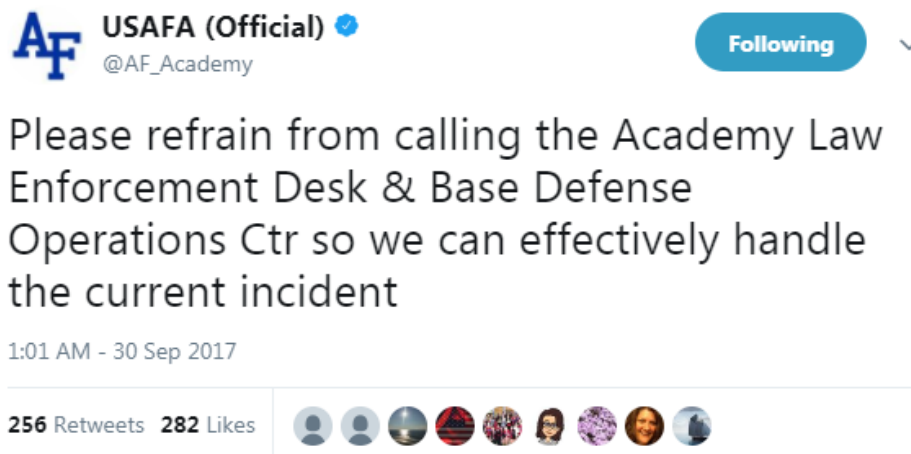
Within minutes, social media erupted. Parents had already been sending questions to @AF_Academy, asking if there really was a shooter – or shooters – after being alerted to the lockdown by their cadets, as students at the Academy are called. The cadets were following what was going on through online chat rooms and other social media, including Reddit and Twitter feeds associated with their squadrons.

It was in cyber space that the incident escalated. The “shooters” grew from one to three to eight within the space of a few tweets. They moved from the bowling alley to the Prep School and into the secure cadet area at lightning speed. Cadets tweeted that they'd heard shots fired, that people were jumping from dorm windows. A message on Reddit said the shooters were firing at windows, and, when questioned about the source for that piece of information, one cadet cited his "roommate looking out the fucking window." People who were off-base tweeted that one of the shooters had committed suicide, that the shooters were closing in on the dorm buildings, that

they were dressed as security forces and were trying to trick cadets into coming out of their rooms.



At the height of the incident, there was a frenzy of tweets and chatter on social media, making it impossible for the cadets to know what was going on, and very difficult for people like me, nearly 2,000 miles away, to make sense of what was happening and relay reliable information to those in the midst of the crisis. Air Force Academy officials remained silent, except for one additional tweet about an hour into the crisis.



Their next tweet came around two hours later and said no shooters had been found. It had all been a false alarm, reportedly triggered by a cadet candidate (a student at the prep school) mistaking the sounds of a Nerf gun battle taking place in

the hallways for live gunfire. None of the alarmist tweets about a shooter committing suicide, bad guys dressed as security forces, pot-shots aimed at the windows of cadets' rooms were true.

Experiencing this in real-time gave me an opportunity to watch as a crisis unfolded, observe the communications strategy of officials at USAFA and determine if it was effective at dispelling the disinformation that was doing the rounds and causing the crisis. Media reports said USAFA had assembled its crisis action team, a group of leaders who respond to emergencies.¹ Immediately after the incident had been resolved, Academy officials announced they were investigating what caused it. USAFA did not did not respond to my phone calls and follow-up emails about the investigation, including whether it was still ongoing.

The lack of a response from USAFA officials left my son and me and thousands of others² to try to make sense of what was going on, and allowed disinformation to spread unchecked. One report about the incident said the furor on social media over the false alarm “turned a minor incident into the ‘War of the Worlds.’”³ When I finally went to bed at 4 a.m. on September 30th, I wondered if the impact of the alarmist messages that I saw on Twitter during the three-hour incident at USAFA would have been different if Academy officials had been monitoring social media and had intervened with rebuttals to say the information we were

¹ Roeder, T. No evidence of shots fired at Air Force Academy; social media fuels 'hysteria.' Colorado Springs Gazette. Oct. 1, 2017. <http://gazette.com/no-evidence-of-shots-fired-at-air-force-academy-social-media-fuels-hysteria/article/1612291>

² There are 4,000 cadets and several thousand more service members living on-base at the U.S. Air Force Academy in Colorado Springs.

³ Roeder, T. <http://gazette.com/no-evidence-of-shots-fired-at-air-force-academy-social-media-fuels-hysteria/article/1612291>

receiving was supposition and rumor. When I had to find a topic for my thesis, I chose to focus on how disinformation spreads, and whether prompt action by authorities can affect our reception of the falsehoods doing the rounds. In the course of writing this thesis, I found that few organizations have a strategy to counteract disinformation on the same media – more often than not, social -- that is being used to spread it.

Dedication

To my dad in heaven, my mum in Florida, my son in Colorado, Serap for the temporary home and warm welcome at UMD, Jack for the offer of the cap and gown, and other friends and relatives around the world who've supported me over the years in all the whacko things I do – like writing a thesis in three months.

Acknowledgements

My thanks to Deb Nelson for agreeing to be my adviser, even though it involved so much more than she or I realized, and to Chris Harvey and Kevin Klose for rounding up the committee and helping me to get this thing done.

Table of Contents

Preface.....	ii
'Should be fine but I love you'	ii
Dedication	vii
Acknowledgements.....	viii
Table of Contents	ix
List of Illustrations	x
Chapter 1	1
Introduction.....	1
1.1 Methodology.....	1
1.2 Terminology.....	4
1.3 Literature review.....	11
Chapter 2.....	15
Disinformation and how it works	15
2.1. Characteristics of disinformation.....	15
2.2. How disinformation fools (some of) us	17
2.3 Russia.....	22
2.4 Who believes disinformation	24
Chapter 3.....	28
A brief history of disinformation	28
3.1 Disinformation in Ancient Rome.....	29
3.2 Cold War disinformation	32
Chapter 4.....	39
Case studies.....	39
4.1 Estonia.....	39
4.1.1. Disinformation and propaganda.....	43
4.1.2. The delivery method	47
4.2 How Georgia won a disinformation battle with Russia but lost the war	52
4.3 German newspaper reacts to disinformation.....	56
Chapter 5.....	64
How social media, new tech drive the disinformation tsunami	64
5.1 Twitter's role.....	64
5.2 Algorithms replace gatekeepers	72
Chapter 6.....	77
Tools to detect disinformation	77
6.1. TwitterTrails	77
6.2 Hoaxy.....	79
6.3. Hearsift and Rumor Gauge	80
6.4. Spotting photo fakery.....	82
6.5. Fake videos	86
Chapter 7.....	88
Conclusion	88
Bibliography	92

List of Illustrations

1. Screenshot of “active shooter” text message received from son, Sept. 30, 2017.....ii
2. First tweet from official U.S. Air Force Academy feed, Sept. 30, 2012.....iii
3. Screenshot of “should be fine but I love you” text message from son, 9/30/17.....iv
4. Second tweet from official U.S. Air Force Academy feed, 9/30/17.....iv
5. Vanessa Otero’s Media Bias Chart.....76
6. Results of a search on Hoaxy® for “USAFA” Table of speech acts used in rumor-detecting application, Hearsift.....80
7. Table of speech acts from MIT Ph.D dissertation by Soroush Vosoughi.....81
8. Doctored photo of Emma Gonzalez posted on Twitter by @LeighStewy.....83
9. Photo of boy in Alaska used in anti-DACA meme on Twitter.....85

Chapter 1

Introduction

When people find themselves in the midst of a crisis, their instinct is to try to make sense of what's happening. This is inherently difficult at the best of times, but more so when there is a large amount of disinformation about the crisis in circulation, and a lack of information to counter it from trusted authorities. During crises when danger is imminent, if official information is inconsistent, individuals often turn to unofficial channels to inform themselves and try to make sense of the situation.⁴ Today, those unofficial media are most often social media, where unverified information and rumors abound.⁵

1.1 Methodology

This thesis looks at the spread of disinformation during a crisis, and what impact the response to that disinformation by trusted authorities has on those involved in the crisis, including the authorities themselves and the institution they represent. It is a theoretical discussion paper based on published research and comments in the public domain. Case studies are used to show successful reactions to disinformation, and how they do not always mean officials are successful in “winning” in other arenas, such as cyber or conventional warfare. Studies by cognitive psychologist and University of Bristol professor Stephan Lewandowsky,⁶ University of Washington

⁴ Jones, N et al. Rumors during a campus lockdown. Proceedings of the National Academy of Sciences Oct 2017. 114 (44) 11663-11668. <http://www.pnas.org/content/114/44/11663>

⁵ Ibid.

⁶ Stephan Lewandowsky's website. <http://www.cogsciwa.com/>

assistant professor Kate Starbird,⁷ Christopher Paul and Miriam Matthews of the RAND Corporation,⁸ and University of California, Irvine doctoral student Nickolas Jones are cited extensively in this report, alongside scores of other reports looking at the characteristics of disinformation, whether it can be stopped, how technology and social media have enabled purveyors of disinformation, who the main purveyors are, and more. Several studies offer empirical advice on how to stop disinformation, or try to stop it.

The incidents I focused on are:

- the disinformation campaign, riots and subsequent cyber-attacks in Estonia in April and May 2007, triggered by the removal of a statue of a Red Army soldier from the center of the capital, Tallinn (Ch. 6.1);
- the brief conflict in the autonomous region of South Ossetia in Georgia, which Russia invaded and seized, and the longer disinformation conflict that preceded it, ran concurrently with it and continues to this day (Ch. 6.2);
- New Year’s Eve celebrations in Dortmund, Germany in 2016, which were hyperbolized and falsified by far-right media in the U.K. and Austria (Ch. 6.3).

I’ve also looked at the mass shooting in June 2016 at the Olympia shopping center in Munich, during which an anonymous Twitter user used “terror-indicating

⁷ <http://faculty.washington.edu/kstarbi/>

⁸ Paul, Christopher and Miriam Matthews, *The Russian "Firehose of Falsehood" Propaganda Model: Why It Might Work and Options to Counter It*. Santa Monica, CA: RAND Corporation, 2016. <https://www.rand.org/pubs/perspectives/PE198.html>.

news” to try to associate a violent act with Muslim extremists; the aftermath of the shooting at Marjorie Stoneman Douglas High School in Parkland, Florida, when groups opposed to the activism of the students and, separately, the sheriff’s department, used visual disinformation to attack them; disinformation during the U.S. elections in 2016; and historical cases of disinformation from ancient Rome and the Cold War.

I’ve focused primarily on Twitter as the means for conveying information and disinformation during a crisis. Many people turn to Twitter for information during a crisis,^{9 10 11 12} especially in the absence of updates from official sources.¹³ During the 2007 disinformation campaign, riots and cyber attacks in Estonia, and the 2008 conflict in Georgia, which was fought on the ground and in the disinformation realm, Twitter was in its infancy. The microblogging application has since come of age and proven to be a conduit for information and disinformation in emergency and disaster situations.¹⁴ Responses to disinformation during the Estonia and Georgia crises were made on the same media as the falsehoods were being spread.

⁹ Blair Thompson, Joseph P. Mazer, Holly J. Payne, Angela M. Jerome, E. Gail Kirby & William Pfohl (2016) Social Media and Active Shooter Events: A School Crisis Communication Challenge, *Qualitative Research Reports in Communication*, 18:1, 8-17. <https://www.tandfonline.com/doi/abs/10.1080/17459435.2016.1247111>

¹⁰ Vieweg, Sarah. Microblogged Contributions to the Emergency Arena: Discovery, Interpretation and Implications. *Computer Supported Collaborative Work*. Feb. 2010. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.495.520&rep=rep1&type=pdf>

¹¹ Westerman, D. et al. Social Media as Information Source: Recency of Updates and Credibility of Information. *Journal of Computer-Mediated Communication* 19 (2014) 171–183. <https://doi.org/10.1111/jcc4.12041>

¹² Zubiaga, A. et al. Analysing how people orient to and spread rumours in social media by looking at conversational threads. *PLoS ONE*. 11. <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0150989>

¹³ Jones, N et al. Rumors during a campus lockdown. *Proceedings of the National Academy of Sciences* Oct 2017. 114 (44) 11663-11668. <http://www.pnas.org/content/114/44/11663>

¹⁴ Vosoughi, S. Automatic Detection and Verification of Rumors on Twitter. Ph.D. dissertation.MIT. June 2015. p 19. http://lsm.media.mit.edu/papers/Soroush_Vosoughi_PHD_thesis.pdf

The active shooter incident at the U.S. Air Force Academy (USAFA) on Sept. 29-30 was the impetus behind this thesis. That crisis was characterized by the fact that it lasted several hours, giving ample time for disinformation to start to flow, gain momentum and eventually die; that the incident started out as *misinformation*, or unintentionally erroneous information, and became a crisis when *disinformation* was shared via social media; and that the authorities at USAFA did not react to the disinformation. Another unique element of the USAFA incident is that I was closely involved in it—my son is a cadet at the academy—and was watching in real-time and trying to filter disinformation from truth as it spread on Twitter.

By way of warning, I have not deleted expletives from quotes, tweets or book titles.

Lastly, although I was granted IRB approval, I had to withdraw my request to include human subjects (interviewees) and conduct a poll in my research for reasons outside of my control.

1.2 Terminology

Although many of the terms used in this thesis are part of the day-to-day lexicon, I will define several of them here.

- **Active measures**

A euphemistic term from Soviet times, broadly synonymous with disinformation today, “active measures” means the use of manipulation and media control, written and oral disinformation, clandestine radio broadcasting, and other

measures to “weaken the enemies of the USSR and create a favorable environment for advancing Moscow’s views and international objectives worldwide.”¹⁵ “Soviet active measures involved a wide range of techniques for influencing events in foreign countries including overt and covert propaganda... the use of front organizations, agents of influence, and forgeries; assistance to insurgencies and terrorist movements, and acts of sabotage and murder. The central objective was to create a psychological climate conducive to the advancement of Soviet interests, and to discredit and demoralize the West.”¹⁶

- **Agenda**

For the purposes of this thesis, agenda is defined as an underlying, often ideological plan or program. Disinformation is often use to further a government’s or group’s agenda, such as deepening political divides in a country or exacerbating ethnic tensions.

- **Botnet**

A botnet is a collection of internet-connected devices, which may include PCs, servers, mobile devices and internet of things devices that are infected and controlled by malware. Users are often unaware of a botnet infecting their system.¹⁷

- **Confirmation bias**

¹⁵ Boghart, T. Operation INFEKTION: Soviet Bloc Intelligence and Its AIDS Disinformation Campaign. *Studies in Intelligence* Vol. 53, No. 4 (December 2009)

¹⁶ Garrity, P.J. Strategy, Disinformation, and Deception in International Politics. <https://politicsandstrategy.com/strategy-disinformation-and-deception-in-international-politics-fb6ad8cf65f0>

¹⁷ <https://searchsecurity.techtarget.com/definition/botnet>

The tendency to look for information that confirms already held beliefs and ignore information that contradicts them, without concern for their truth or falsehood.¹⁸ The tendency to interpret new evidence as confirmation of one's existing beliefs or theories. When people would like a certain idea/concept to be true, they end up believing it to be true.¹⁹

- **Crisis**

A crisis is an unexpected occurrence that may have negative consequences, which can lead to uncertainty, instability, and, in many cases, a need for immediate action or response.²⁰ In the corporate world, a crisis is defined as: "Any situation that is threatening or could threaten to harm people or property, seriously interrupt business, damage reputation and/or negatively impact stock value."²¹ Taking elements from both definitions, a crisis is defined for the purposes of this report as a situation that causes or threatens to cause harm to people, property or a system – government, academia, financial, military base operations; interrupts normal activity, and holds the potential to damage a person's or institution's reputation.

- **Disinformation**

¹⁸ <http://www.animal-ethics.org/motivated-reasoning-confirmation-bias/>

¹⁹ Heshmat, S. What is confirmation bias? Psychology Today. April 23, 2015. Accessed online at <https://www.psychologytoday.com/us/blog/science-choice/201504/what-is-confirmation-bias>

²⁰ Stein, A. We Thought It Could Never Happen Here, *Journal of Promotion Management*, 12:3-4.Pp. 99-128, https://doi.org/10.1300/J057v12n03_07

²¹ Bernstein, J. Making a Crisis Worse, *Journal of Promotion Management*, 12:3-4, 39-45, 2006. https://doi.org/10.1300/J057v12n03_04

The Oxford Dictionaries website defines disinformation as “false information supplied by a government to the media with the intention of influencing the policies or opinions of those who receive it.” In “Disinformation: Former Spy Chief Reveals Secret Strategies for Undermining Freedom, Attacking Religion, and Promoting Terrorism,” published in 2013, former Romanian spy chief Ion Mihail Pacepa says disinformation is the spreading of an idea in a foreign nation, while making it appear as if the idea originated in that nation.”²² Kumar and Geethakumari say disinformation is “often the product of a carefully planned and technically sophisticated deceit process; may not come directly from the source that intends to deceive; and can include written or verbal communication, doctored photographs, fake videos etc.”²³

Many scholars believe disinformation to be the Anglicization of the Russian word *dezinformatsiya*. The term is said to have been coined by Soviet leader Joseph Stalin and to have come into use in the 1950s. Stalin reportedly wanted a word that sounded European (French) so that the original concept of disinformation would not be associated with the Soviet Union but with the West. Disinformation is, of course, used in the West as well as in the former Communist bloc. In the U.S. when it is used, its intent is most often to sway public opinion against a person or group, such as a political opponent of whomever has ordered the disinformation campaign.

- **Gatekeepers**

²² <http://russialist.org/russian-born-disinformation-breeds-terror-slanders-us-new-book/>

²³ Kumar, KPK & Geethakumari, Detecting misinformation in online social networks using cognitive psychology. *G. Hum. Cent. Comput. Inf. Sci.* (2014) 4: 14. <https://doi.org/10.1186/s13673-014-0014-x>

In the media world, gatekeepers are those who filter information before it is published, broadcast, posted on websites or social media or any other medium. Gatekeeping in communications/journalism used to be the remit of trusted journalists, professional editors and news agencies. Many citizen journalist sites have an editorial staff that is tasked with, among other things, gatekeeping. But the advent and popularization of social media have increased the number of stories that do not pass through gatekeepers.

- **Microblogging**

A form of lightweight chat that allows users to send short messages to people subscribed to their streams.²⁴

- **Misinformation**²⁵

Misinformation is information that is unintentionally false. Misinformation often arises out of misinterpretation of scientific results, such as a report that a flavonoid in hops, which is used to make beer, made young mice smarter.^{26 27} The intelligent beer-drinking mice story is also an example of how poor sourcing can contribute to the spread of misinformation: the writer of the story cites as his source

²⁴ Starbird et al. [Chatter on the red: what hazards threat reveals about the social life of microblogged information](#). Proceedings of the 2010 ACM conference on Computer supported cooperative work. 241-250.

²⁵ In some of the studies included in my research, the authors use misinformation where I have used disinformation.

²⁶ Yakowicz, W. Have a beer, it's good for your brain. Inc.com. Oct. 3, 2014. <https://www.inc.com/will-yakowicz/beer-improves-brain-function.html>

²⁷ Studies conducted in mice are not directly translatable to humans.

an article in The Daily Mail rather than the original source of the study, which he identifies as researchers at Oregon State University.²⁸

Misinformation often circulates in evolving crisis situations. Information that is considered “correct” at any given stage during a crisis can later turn out to be untrue. Examples of this sort of misinformation are changing death tolls, the cause of an explosion, or the whereabouts of a suspect. Providing timely news coverage of unfolding events is by its very nature piecemeal and requires occasional corrections of earlier statements, write Lewandowsky et al.²⁹

- **Motivated reasoning**

Motivated reasoning is a mechanism by which people access, construct and evaluate arguments and data in a biased way in order to arrive at a preferred conclusion that fits their beliefs, preferences and wishes.^{30 31 32}

- **Retweet**

A retweet is a message from another source that is forwarded by a Twitter user and becomes visible to followers.

²⁸ Zamzow, DR et al. Xanthohumol improved cognitive flexibility in young mice. Behavioural Brain Research. Volume 275, 15 December 2014, Pages 1-10. <https://doi.org/10.1016/j.bbr.2014.08.045>

²⁹ Lewandowsky, S. et al. Misinformation and Its Correction: Continued Influence and Successful Debiasing. Psychological Science in the Public Interest 13(3) 106–131. <http://journals.sagepub.com/doi/full/10.1177/1529100612451018>

³⁰ Weir, K. Why we believe alternative facts. American Psychological Association. May 2017, Vol 48, No. 5. <http://www.apa.org/monitor/2017/05/alternative-facts.aspx>

³¹ Stone, Daniel F. and Wood, Daniel H., Cognitive Dissonance, Motivated Reasoning, and Confirmation Bias: Applications in Industrial Organization (December 2, 2016). <http://dx.doi.org/10.2139/ssrn.2834072>

³² <http://www.animal-ethics.org/motivated-reasoning-confirmation-bias/>

- **Social media**

Social media are computer-mediated technologies that facilitate the creation and sharing of information, ideas, career interests and other forms of expression via virtual communities and networks.³³ Social media allow individuals and communities to share, create, discuss and modify user-generated content.³⁴

- **Troll**

An internet troll is a person who intentionally antagonizes others online by posting inflammatory, irrelevant, or offensive comments or other disruptive content.³⁵

- **Trusted authorities**

Trust is the firm belief in the reliability, truth or ability of someone or something.³⁶ An authority is a person or organization having political or administrative power and control, or a person with extensive or specialized knowledge about a subject; an expert.³⁷ Authority often appears in the plural form.³⁸ Trusted authorities are, therefore, a person or organization in a leadership position, with extensive knowledge about a subject, and whom we believe to be reliable and truthful.

³³ Wikipedia.

³⁴ Kietzmann, JH et al. Social media? Get serious! Understanding the functional building blocks of social media. *Business Horizons* (2011) 54, 241—251. <https://doi.org/10.1016/j.bushor.2011.01.005>

³⁵ <https://www.merriam-webster.com/dictionary/troll>

³⁶ <https://en.oxforddictionaries.com/definition/trust>

³⁷ <https://en.oxforddictionaries.com/definition/authority>

³⁸ Ibid.

- **Twitter**

A form or subset of social media that is sometimes referred to as microblogging, Twitter allows users to retrieve, produce and spread information using a maximum of 280 characters (previously limited to 140). It has been referred to as a real-time information network and a real-time news source created by everyday users.³⁹

- **User-generated content**

User-generated content (UGC) has been called the lifeblood of social media. It includes video, blogs, discussion forum posts, digital images, audio files, and other media created by consumers or end-users of an online system or service that is publicly available to others. The quality of UGC is threatened by the creation and diffusion of misleading information, including disinformation, manipulation, misinformation and rumors.⁴⁰

1.3 Literature review

Disinformation has been implicated in the outcomes of the Brexit vote in the U.K. in June 2016 and Donald Trump's surprise victory in the U.S. presidential election in November of the same year. Since then, it has been a hot topic for discussion, debate, analysis and research. Almost daily as I was writing this thesis, new media reports and studies were published about disinformation, which has

³⁹ Metaxas et al. Do Retweets indicate interest, trust, agreement? [arXiv:1411.3555v1 \[cs.SI\]](https://arxiv.org/abs/1411.3555v1)

⁴⁰ Pal, A. Reviewing the Landscape of Research on the Threats to the Quality of User-Generated Content. *Proceedings of the Association for Information Science and Technology*. 53:1. 2016. <https://onlinelibrary.wiley.com/doi/full/10.1002/pr2.2016.14505301077>

become a very hot topic. I have cited many of the studies, both contemporary and historical, that I have read while researching this thesis, to try to shed light on what disinformation is, whence it emanates, how we are trying to stop it, why some people fall for it, and why it is dangerous to them and to society.

Among the researchers that I've studied for this report are:

Stephan Lewandowsky, a cognitive scientist whose research includes deep dives into and analysis of how disinformation spreads and persists. In one of scores of articles he has had published on the topic, Lewandowsky writes that retracting a piece of disinformation often does not stop its influence.⁴¹ However, repeating retractions, corrections and refutations enhances their impact, he notes – just as it does for disinformation.⁴² But the impact of repeating disinformation is greater than that of repeating a retraction, says Lewandowsky.⁴³

Acceptance of disinformation as fact can be reduced if people are explicitly warned that information they will receive could be misleading or outright false.⁴⁴ Such warnings are more effective if they are made before the disinformation is received and registered by the public, who “by default expect information presented to be valid,”⁴⁵ writes Lewandowsky. Pre-empting disinformation with an a priori

⁴¹ Lewandowsky, S et al. Misinformation and its correction: continued influence and successful debiasing. *Psychological Science in the Public Interest*. 13(3) 106-131. 2012. Accessed online <http://journals.sagepub.com/doi/pdf/10.1177/1529100612451018>. (Note: Lewandowsky uses misinformation where I use disinformation.)

⁴² Lewandowsky. p 116

⁴³ Lewandowsky, p 117

⁴⁴ Lewandowsky, p. 116

⁴⁵ Ibid.

warning can change the expectation that all information a person is about to receive is valid, he writes.

Kate Starbird. A focus of Starbird's research is the use of social media during crisis events.⁴⁶ Among dozens of research papers that Starbird has published, one looked at the spread of disinformation via Twitter following the bombings at the 2013 Boston Marathon.⁴⁷ Starbird et al found that corrections to misinformation are muted compared with the propagation of the misinformation. The research also pointed to the huge volume of information that circulated after the bombings, of which the overwhelming majority was disinformation.

Nickolas Jones. While many studies focus on detecting falsehoods as they happen and on educating the public to recognize disinformation, and while social media outlets are cracking down on account holders who spread fake news, few studies look at what role officials who are involved in a disinformation attack should play to try to stop the flow of falsehoods. Among those that I found that do are a study by Jones et al. A doctoral student in the department of psychology and social behavior at the University of California, Irvine, Jones's research focuses on the ways in which people's interactions with technology, social media and other virtual environments impact psychological well-being and decision-making.

Jones and colleagues found that, in the absence of regular updates from trusted authorities during an active shooter lockdown on a university campus,

⁴⁶ <http://faculty.washington.edu/kstarbi/>

⁴⁷ Starbird, K. et al. Rumors, False Flags, and Digital Vigilantes: Misinformation on Twitter after the 2013 Boston Marathon Bombing. iConference 2014. PDF accessed online. https://faculty.washington.edu/kstarbi/Starbird_iConference2014-final.pdf

students and others involved in the crisis turned to unofficial sources – often Twitter – for information. Their study found that those who relied on social media for updates were exposed to more conflicting information than those who obtained information from traditional media, and often felt “helpless and without situational control, which can lead people to see patterns in the information obtained that are not present.”⁴⁸ The researchers concluded that, to reduce rumor exposure and mitigate distress, officials should monitor social media channels during a crisis and release substantive updates on social media at regular intervals during the event.

Christopher Paul is a senior social scientist at the RAND Corporation and professor at the Pardee RAND Graduate School. Paul collaborated with RAND Corporation behavioral and social scientist **Miriam Matthews** to produce a report titled “The Russian ‘Firehose of Falsehood’ Propaganda Model,”⁴⁹ which looks at the evolution in recent years of Russian disinformation/propaganda and what might be done to counter it.

⁴⁸ Jones, N et al. Rumors during a campus lockdown. Proceedings of the National Academy of Sciences Oct 2017. 114 (44) 11663-11668. <http://www.pnas.org/content/114/44/11663>

⁴⁹ Paul, Christopher and Miriam Matthews, The Russian "Firehose of Falsehood" Propaganda Model: Why It Might Work and Options to Counter It. Santa Monica, CA: RAND Corporation, 2016. pp 9-10 <https://www.rand.org/pubs/perspectives/PE198.html>

Chapter 2

Disinformation and how it works

Although disinformation has changed over time and continues to change with the times, and although technological advances make it easier to produce fake memes, doctored images and false stories, disinformation has retained some key characteristics over the ages.

2.1. Characteristics of disinformation

According to my research, the main features of disinformation are:

- **Omission** – Purveyors of disinformation leave out key elements of stories as published or broadcast by traditional media.
- **Non-contextuality** – Information is reported out of context.
- **Fabrication** – Entire stories or facts are made up and published as stand-alone “news” articles, or inserted into stories or disseminated via social media.
- **Loaded language** – Disinformation often uses language that appeals to emotions or stereotypes in a bid to influence an audience.
- **Lack of credible sourcing** – Disinformation often offers no source or a dubious source. In some cases, the media outlet/website carrying the disinformation is itself a dubious source, such as CNNchannel1.com, which carried fake and biased stories and videos during the 2017

Kenyan elections, intended to make viewer/readers believe that they were by CNN.

- **Repetition** – Repetition is considered to make a more persuasive argument for a fake story than hearing it from a single source. This persuasion of fact is underscored if each version of a story takes a different argument, but points to the same conclusion as the others.⁵⁰
- **“Terror-indicating news”** is a more recent addition to the disinformation toolbox. It’s often used by individuals or groups commenting in real-time during a crisis situation. Terror-indicating news is used almost exclusively to disparage/accuse migrant groups, particularly Muslims, of involvement in violent crises.⁵¹ For example, during the shooting at the Olympia shopping center in Munich in the summer of 2016, news spread on Twitter that the gunman had shouted “Allahu Akbar” or God is great in Arabic as he began his shooting rampage. The claim originated on an unverified Twitter account. It was quickly dispelled as false by the media and police—⁵² but not before an interview with someone claiming to have heard the gunman say it was carried by CNN.

⁵⁰ Paul, C and Matthews, M. The Russian “Firehose of Falsehood” Propaganda Model. Santa Monica, CA: RAND Corporation. 2016. Accessed online at

<https://www.rand.org/pubs/perspectives/PE198.html>

⁵¹ Several white Americans who have perpetrated mass attacks have ties to white supremacist or other racist movements, but saying that a white man who shouted out a white supremacist rally cry before driving a vehicle into a crowd or while shooting a group of church-goers is not considered terrorism.

⁵² Nguyen T.N., Li C., Niederée C. (2017) On Early-Stage Debunking Rumors on Twitter: Leveraging the Wisdom of Weak Learners. In: Ciampaglia G., Mashhadi A., Yasseri T. (eds) Social Informatics. SocInfo 2017. Lecture Notes in Computer Science. Vol 10540. Springer, Cham.

https://doi.org/10.1007/978-3-319-67256-4_13

2.2. How disinformation fools (some of) us

Creators and purveyors of disinformation use simple techniques to try to get as many people as possible to subscribe to the version of events that they're selling. For example, repeating true or false information leads to what's called the **illusion of truth effect**. Also called the illusory truth effect, this uses repetition of a statement to increase a person's belief in the truth of that statement.⁵³ The propagandist-in-chief in Nazi Germany, Joseph Goebbels, who has been credited with saying, "Repeat a lie often enough and it becomes the truth," understood and used the illusion of truth effect.

Information overload – spreading as many conflicting messages as possible – seeks to convince an audience that there are so many versions of a story that it is impossible to find the truth.⁵⁴ Christopher Paul and Miriam Matthews say in their report for the RAND Corporation that Russian propaganda today uses "high numbers of channels and messages and a shameless willingness to disseminate partial truths or outright fictions" to create a "firehose of falsehood."⁵⁵ Examples of information overload are below in the section on Russia and its involvement in disinformation.

⁵³ Hasher, L., Goldstein, D., & Toppino, T. (1977). Frequency and the conference of referential validity. *Journal of Verbal Learning and Verbal Behavior*, 16, 107- 112.

⁵⁴ EU East StratCom Task Force. Means, goals and consequences of the pro-Kremlin disinformation campaign. Istituto per gli Studi di Politica Internazionale (ISPI). Jan. 19, 2017. <https://www.ispionline.it/it/pubblicazione/means-goals-and-consequences-pro-kremlin-disinformation-campaign-16216>

⁵⁵ Paul, Christopher and Miriam Matthews. The Russian "Firehose of Falsehood" Propaganda Model: Why It Might Work and Options to Counter It. Santa Monica, CA: RAND Corporation. 2016 <https://www.rand.org/pubs/perspectives/PE198.html>

Most of the disinformation campaigns studied for this report involve groups or individuals affiliated with or supportive of the far-right, and/or who have ties to Russia. Russian disinformation campaigns, alleged and proven, during the 2016 U.S. elections, the Brexit vote in United Kingdom; the annexation of Crimea in eastern Ukraine, during a brief conflict with the former Soviet republic of Georgia in 2012, and the cyber attacks on another former Soviet republic, Estonia, in 2007, were all preceded by or run simultaneously with disinformation campaigns believed to have been orchestrated by Moscow.

The RAND report notes that contemporary Russian propaganda or disinformation is “rapid, continuous, and repetitive, and it lacks commitment to consistency.”⁵⁶ By circulating numerous versions of the same story, purveyors of disinformation seek to convince those receiving the information that “there are so many versions of events that it is impossible to find the truth,” says a report by the EU East StratCom Task Force published on the website of Italy’s Istituto per gli Studi di Politica Internazionale (ISPI).⁵⁷ Like the RAND report, the EU East StratCom Task Force report focuses on Russian disinformation strategies and tactics. It cites multiple versions posted by Russian media of the tragic downing of a Malaysian Airlines flight over Eastern Ukraine as an example of how Russia tries to “pollute the information space and make the objective facts seem unrecognizable.”⁵⁸ In July 2014,

⁵⁶ Ibid.

⁵⁷ EU East StratCom Task Force. Means, goals and consequences of the pro-Kremlin disinformation campaign. Istituto per gli Studi di Politica Internazionale (ISPI). Jan. 19, 2017. <https://www.ispionline.it/it/pubblicazione/means-goals-and-consequences-pro-kremlin-disinformation-campaign-16216>

⁵⁸ Unsigned. Nine ways to confuse us about MH17. EU vs Disinfo website. Sept. 30, 2016. <https://euvsdisinfo.eu/nine-ways-to-confuse-us-about-mh17/>

with the finger of blame pointing at Russia after the passenger plane was shot down, Russian media reports claimed that the images of the Malaysian Airlines Boeing actually showed a Ukrainian Air Force Antonov that had been shot down by pro-Kremlin separatists; that Ukraine was trying to shoot down Russian President Vladimir Putin's plane; that the passengers on the Malaysian Airlines flight might have been dead when the aircraft took off from Amsterdam; or that the Boeing was the same Malaysian Airlines plane that vanished on a flight from Kuala Lumpur to China in March 2014.⁵⁹

A more recent example of inundating the media-sphere with different fake versions of the same story is illustrated by the Russian response to accusations that Moscow was behind the attempted murder in the U.K. of former Russian spy Sergei Skripal and his daughter, using the nerve agent Novichok. According to the EU vs Disinfo online newsletter, more than 20 different narratives have been published or broadcast in Russia about the attempted murder.⁶⁰ “What was striking,” says an unsigned story on the EU-run website, “was how the stories – some appearing on the same state-controlled news outlet but completely contradicting one another – peddled multiple, unsubstantiated and often absurd conspiracy theories. This is perhaps the most classic of pro-Kremlin disinformation techniques: to confuse the wider audience

⁵⁹ Unsigned. Nine ways to confuse us about MH17. EU vs Disinfo website. Sept. 30, 2016. <https://euvdisinfo.eu/nine-ways-to-confuse-us-about-mh17/>

⁶⁰ Figure of the week: 20. Posted on EU vs Disinfo website, Mar. 27, 2018. <https://euvdisinfo.eu/figure-of-the-week-20/>

and at the same time to feed a loyal audience of disinformers ready to defend Russia with abundant lines to take.”⁶¹

Among the many different explanations for who tried to kill the Skripals were:⁶²

- The United Kingdom did it to fuel anti-Russian sentiment (source: Russia 1 TV channel)
- Ukraine did it to frame Russia (Russia 1)
- The United States did it to destabilize the world (Russia 1)
- Skripal’s future mother-in-law did it (Moskovsky Komsomolets: mk.ru)
- Terrorists did it (Russian ministry of foreign affairs spokesperson)
- A drone did it (Zvezda and Russian defense ministry)
- Theresa May helped orchestrate the attack because she is a friend of CIA director Gina Haspel (Zvezda)⁶³
- American-British financier Bill Browder, whose lawyer was Sergey Magnitsky,⁶⁴ did it (Russia 1)

⁶¹ Unsigned. Disinformation Review. EU vs Disinfo. March 15, 2018. <https://euvsdisinfo.eu/skripal-and-the-disinformation-swamp/>

⁶² Unsigned. Figure of the week: 20. EU vs Disinfo. March 27, 2018. <https://euvsdisinfo.eu/figure-of-the-week-20/>

⁶³ Gina Haspel was not CIA director in March 2018, but maybe the Russians knew something.

⁶⁴ Magnitsky died in a Russian prison after uncovering a massive tax fraud linked to the Kremlin. The U.S. passed a law that bears his name in 2012. The law, intended to punish Russian officials responsible for Magnitsky’s death, initially imposed U.S. visa bans on 18 Russian government officials, froze assets they held in U.S. banks and banned them from using the U.S. banking system. Since its passage in December 2012, it has been expanded to include dozens more Russians.

jailed in Russia after uncovering a massive fraud scheme involving the Kremlin. He died The U.S. passed a law in 2012

As for the nerve agent Novichok, which chemical weapons experts say was used in the attack on the Skripals,⁶⁵ the Russian media also sought to confuse through contradictory stories.⁶⁶ These claimed, among other things that:

- Russia has destroyed all its stockpiles of Novichok (Sputnik)
- Russia never developed Novichok (Interfax)
- There is no evidence that the nerve agent used against the Skripals was Novichok, and the British were struggling to identify the substance (RT).

The reason Russian media post stories that appear to contradict each other is because, “If one falsehood or misrepresentation is exposed or is not well received, the propagandists will discard it and move on to a new (though not necessarily more plausible) explanation,” write Paul and Matthews.⁶⁷ Furthermore, circulating multiple sources present a story on multiple outlets has been found to be more persuasive than having a single source present several different arguments, or multiple sources reporting the same version, according to a study by Stephen Harkins and Richard Petty.⁶⁸ “Given the large number of cognitive tasks that people confront daily and given peoples' limited cognitive resources, choices must be made about which stimuli

⁶⁵ Gross, J. Chemical weapons watchdog backs up U.K. view on spy poisoning. The Wall Street Journal. April 12, 2018. <https://www.wsj.com/articles/novichok-was-used-in-skripal-attack-chemical-weapons-watchdog-confirms-1523532806>

⁶⁶ Unsigned. Figure of the week: 20. EU vs Disinfo. March 27, 2018. <https://euvsdisinfo.eu/figure-of-the-week-20/>

⁶⁷ Paul, Christopher and Miriam Matthews. The Russian "Firehose of Falsehood" Propaganda Model: Why It Might Work and Options to Counter It. Santa Monica, CA: RAND Corporation. 2016 <https://www.rand.org/pubs/perspectives/PE198.html>

⁶⁸ Harkins, S. G., & Petty, R. E. (1987). Information utility and the multiple source effect. *Journal of Personality and Social Psychology*, 52(2), 260-268. <http://dx.doi.org/10.1037/0022-3514.52.2.260>

deserve careful scrutiny and which are less worthy of thought,” wrote Harkins and Perry in the Journal of Personality and Social Psychology in 1987. “If arguments from multiple sources are perceived as representing independent perspectives on an issue, then this information is more worthy of diligent consideration than information from only one perspective.”

2.3 Russia

“No one seems better resourced, more enthusiastic or as accomplished at [disinformation] as the Russians.”

*Carl Miller on Wired.*⁶⁹

Russian individuals and organizations have been accused of being behind the cyber attacks on Estonia in 2007,⁷⁰ of creating and feeding disinformation into the United States in the run-up to the 2016 elections,⁷¹ of using propaganda, fake news, media manipulation and forgeries ahead of and during the occupation of Crimea and the war against Ukraine in Donbas,⁷² of web page defacements, denial of service, and distributed denial of service attacks on Georgian government, media and financial institutions, as well as other public and private targets.⁷³

⁶⁹ Miller, C. The big business of disinformation. Wired.com. June 17, 2017. <http://www.wired.co.uk/article/carl-miller-disinformation-gets-commercialised>

⁷⁰ Wikileaks Cable 366. Dated May 11, 2007.

https://wikileaks.org/plusd/cables/07TALLINN366_a.html

⁷¹ Internet Research Agency Indictment. <https://www.justice.gov/file/1035477/download>

⁷² Fedchenko, Y. Kremlin propaganda: Soviet active measures by other means. Sojateadlane (Estonian Journal of Military Studies), Volume 2, 2016, pp. 141-170. PDF accessed online at www.ksk.edu.ee/publikatsioonid

⁷³ Nazario, J. Politically motivated denial of service attacks. NATO Cooperative Cyber Defence Centre of Excellence. Undated.

Russian military leaders have said information is part of their arsenal. "The day has come when we all have to admit that a word, a camera, a photo, the Internet, and information in general have become yet another type of weapon, yet another component of the armed forces," said Russian Defense Minister Sergei Shoigu in 2015 at an awards ceremony for a competition for journalists. "This weapon can be used in a good and in a bad way. This is a weapon that was involved in various events in our country in different years, both in our defeats and in our victories." he said.⁷⁴ The MEDIA-ACE competition for which the awards were being presented was instituted by the Russian defense ministry as a way of "encouraging" members of the media in Russia who have "made a significant contribution to strengthening the Russian armed forces' positive image," according to Russia's Interfax news agency.

However, when Russia is accused of being behind disinformation or cyber campaigns, it has dismissed the accusations as absurd and empty talk.⁷⁵ Even the word *dezinformatsiya* is part of the Russian subterfuge surrounding disinformation – Stalin wanted a word that sounded Western European so that the concept would not be traced back to Moscow.

The U.S. military considers that disinformation campaigns fall under the umbrella of information operations, which also include cyberwarfare.⁷⁶ Like other

http://www.ccdcoe.org/publications/virtualbattlefield/12_NAZARIO%20Politically%20Motivated%20DDoS.pdf

⁷⁴ Reported by Interfax news agency, March 28, 2015.

<http://www.interfax.com/newsinf.asp?id=581851>

⁷⁵ T. J. Kirkpatrick, Bloomberg News in The Moscow Times.

<https://themoscowtimes.com/news/russia-reacts-charges-election-meddling-mueller-investigation-60553>

⁷⁶ Wilson, Clay. 2007. Information operations, electronic warfare, and cyberwar capabilities and related policy issues. [Washington, D.C.]: Congressional Research Service, Library of Congress. <http://www.fas.org/sgp/crs/natsec/RL31787.pdf>

countries, including the U.S., Russia has used cyberwarfare, sometimes on its own and sometimes alongside conventional warfare, to achieve geopolitical goals. Cyberwarfare is an especially attractive way to influence the outcome of an event because “the global and anonymous nature of the internet makes it very difficult to prove a particular operation was state-sponsored.”⁷⁷ Russia is exploiting that anonymity when it disseminates disinformation, too. “Russia has a long history of propaganda and disinformation operations—techniques it continues to adapt to the online environment,”⁷⁸ E. J. Iasiello wrote in *The U.S. Army War College Quarterly*.

2.4 Who believes disinformation

“While any group can come to believe false information, misinformation is currently predominantly a pathology of the right...”

Report by Harvard Kennedy School Shorenstein Center on Media, Politics and Public Policy

Extremist groups on both sides of the political spectrum use disinformation and social media to spread their beliefs and secure greater control for themselves or others who share their ideology. These groups “leverage both the techniques of participatory culture and the affordances of social media to spread their various

⁷⁷ Dougherty, J. The mouse clicks of August: Hybrid warfare, nation-state actors and the future of cybersecurity. *Small Wars Journal*. Accessed online at <http://smallwarsjournal.com/jrnl/art/mouse-clicks-august-hybrid-warfare-nation-state-actors-and-future-cybersecurity>

⁷⁸ Iasiello, EJ. Russia’s improved information operations: From Georgia to Crimea. *The U.S. Army War College Quarterly*, Parameters 47. No. 2. p51. Summer 2017. <http://publications.armywarcollege.edu/pubs/3368.pdf>

beliefs. Taking advantage of the opportunity the internet presents for collaboration, communication, and peer production, these groups target vulnerabilities in the news media ecosystem to increase the visibility of and audience for their messages,” write Alice Marwick and Rebecca Lewis in a report⁷⁹ for the Data & Society research institute, which looks at the ways the same innovative tools that are positively transforming society can also be harmful.

Many of the groups that spread disinformation are on the far-right of the political spectrum, according to different groups of researchers.^{80 81 82} Right-wing disinformation most often originates in a country, like Hungary, or politically motivated organization⁸³ espousing right-wing ideologies.⁸⁴ For example, since Viktor Orban and the far right surged to power in Hungary in 2006, “at least 90 Hungarian-language fake news sites have emerged (mainly run by Hungarian far-right actors), most of them spreading lies and rumours about refugees,” writes Krisztian Simon, a doctoral candidate at the Freie Universität in Berlin, in the Green European Journal.⁸⁵

⁷⁹ Marwick, A & Lewis, R. Media manipulation and disinformation online. Data & Society Research Institute. Center for Media Justice. May 15, 202017. p. 3.

https://datasociety.net/pubs/oh/DataAndSociety_MediaManipulationAndDisinformationOnline.pdf

⁸⁰ Ibid, p.6.

⁸¹ Narayanan, V. et al. Polarization, Partisanship and Junk News Consumption over Social Media in the US. Computational Propaganda Project, University of Oxford. Feb. 6, 2018.

<http://comprop.oi.ox.ac.uk/wp-content/uploads/sites/93/2018/02/Polarization-Partisanship-JunkNews.pdf>

⁸² Lazer et al. Combating fake news: an agenda for research and action. Harvard Kennedy School, Shorenstein Center on Media, Politics and Public Policy. Conference report from Feb. 17-18, 2017. Accessed online at <https://shorensteincenter.org/combating-fake-news-agenda-for-research/>

⁸³ Lazer et al. Combating fake news: an agenda for research and action. Harvard Kennedy School, Shorenstein Center on Media, Politics and Public Policy. Conference report from Feb. 17-18, 2017. Accessed online at <https://shorensteincenter.org/combating-fake-news-agenda-for-research/>

⁸⁴ Ibid.

⁸⁵ Unsigned. Green Observatory: Fake News. Green European Journal. Undated. <https://www.greeneuropeanjournal.eu/green-observatory-fake-news/#Hungary>

In the U.S., an analysis conducted of more than 1.25 million stories published online between April 1, 2015, and Election Day 2016 found that a right-wing media network with Breitbart at its nucleus “developed as a distinct and insulated media system, using social media as a backbone to transmit a hyper-partisan perspective to the world.”⁸⁶ The media groups orbiting around Breitbart included Fox News, the Daily Caller, the Gateway Pundit, the Washington Examiner and Infowars, among others. All of these media outlets have published false information,⁸⁷ have a distinctly pro-Trump bias and have succeeded not only in setting the agenda for conservative media but “also strongly influenced the broader media agenda, in particular coverage of Hillary Clinton.”⁸⁸

A BuzzFeed analysis showed websites on the extremes of the political spectrum are causing the political divide in the U.S. to deepen. An analysis by BuzzFeed of three hyper-partisan right-wing Facebook pages—Eagle Rising, Right Wing News and Freedom Daily—found that 38 percent of all posts were either a mixture of true and false or mostly false information, compared to 19 percent of posts from three hyper-partisan left-wing pages—The Other 98%, Occupy Democrats and Addicting Info—that were either a mixture of true and false or mostly false.⁸⁹ The

⁸⁶ Benkler, Y et al. Study: Breitbart-led right-wing media ecosystem altered broader media agenda. *Columbia Journalism Review*. March 3, 2017. Accessed online at <https://www.cjr.org/analysis/breitbart-media-trump-harvard-study.php>

⁸⁷ PunditFact web pages on May 14, 2018 for Breitbart, Fox News, Daily Caller, InfoWars, Gateway Pundit, Washington Examiner <http://www.politifact.com/personalities/breitbart/> <http://www.politifact.com/search/?q=Fox+News> <http://www.politifact.com/search/?q=daily+caller> <http://www.politifact.com/search/?q=Infowars> <http://www.politifact.com/search/?q=gateway+pundit> <http://www.politifact.com/search/?q=washington+examiner>

⁸⁸ Benkler, *ibid.*

⁸⁹ Silverman, C et al. Hyperpartisan Facebook pages are publishing false and misleading information at an alarming rate. *Buzzfeed.News*. Oct. 20, 2016. https://www.buzzfeed.com/craigsilverman/partisan-fb-pages-analysis?utm_term=.cuQyewY4Y5#.tsAgjV878w

rapid growth of Facebook pages like these—some have more followers than mainstream political pages –and the fabricated content they publish suggest that “the best way to grow an audience for political content on the world’s biggest social network is to eschew factual reporting and instead play to partisan biases using false or misleading information that simply tells people what they want to hear,” BuzzFeed concluded.

Breitbart and other U.S. right-wing sites were among the most popular domains in 2017 out of 600 Twitter accounts that amplified messages seen as benefiting the Kremlin, according to the German Marshall Fund in the U.S.⁹⁰ Breitbart’s popularity may have been helped by the fact that Republicans are more likely than Democrats to believe disinformation, although a majority of followers of both parties buy into fake news.⁹¹

⁹⁰ Dorell, O. Breitbart, other ‘alt-right’ websites are the darlings of the Russian propaganda effort. USA Today. Aug. 24, 2017. <https://www.usatoday.com/story/news/world/2017/08/24/breitbart-other-alt-right-websites-darlings-russian-propaganda-effort/598258001/>

⁹¹ “On average, Clinton voters judged 58% of familiar fake news headlines as accurate, versus 86% for Trump voters.” Silverman, C & Singer-Vine, J. Most Americans Who See Fake News Believe It, New Survey Says. BuzzFeed News. Dec. 6, 2016. https://www.buzzfeed.com/craigsilverman/fake-news-survey?utm_term=.ag4PD0pOpQ#.wxwMR3VrVQ

Chapter 3

A brief history of disinformation

"I believe that the more you know about the past, the better you are prepared for the future."

– Theodore Roosevelt

Disinformation existed back in Roman times, during the Cold War, and the centuries in between. It was used for the same purposes then as now—to sway opinion and promote a viewpoint that the producer and purveyor of the disinformation wants to bring to the fore. For Maxim Eristavi, a former newsroom manager in Moscow and now Atlantic Council fellow and co-founder of an independent news outlet in Kiev, disinformation ‘went global’ in 2014, when Russian propaganda and falsehoods were part of the weaponry deployed by Moscow in the war in eastern Ukraine.⁹² Prior to that, however, eastern Europeans had been fighting Russian-driven disinformation for years. Understanding that disinformation is not new, how it has been used over the centuries to achieve goals, often political, and how the tools used to deploy it have evolved are key to being able to spot and potentially stop disinformation today.

⁹²Eristavi, M. Why the US Keeps Losing the Fight against Disinformation. Atlantic Council website. <http://www.atlanticcouncil.org/blogs/ukrainealert/why-the-us-keeps-losing-the-fight-against-disinformation>

A key change to disinformation is that it has increased in volume and can be shared at lightning speed today⁹³ versus in Ancient Roman times or, more recently, during the Cold War.

In the historical cases that follow, when a trusted authority intervened and refuted disinformation, as the U.S. did when the Soviets spread a false story about AIDS having been created at Fort Detrick in Maryland, the life and spread of disinformation was curtailed. But disinformation was not stopped entirely. These historical case studies show that, once a false story starts circulating, it's hard to get rid of every trace of it.

3.1 Disinformation in Ancient Rome

One of the first recorded disinformation campaigns may have happened some 2,050 years ago. There are some odd similarities with the political world of today.

It started after Julius Caesar found himself on the wrong side of republicans by declaring himself *dictator perpetuo* – often translated as dictator for life⁹⁴ -- which led to his murder on the Senate floor in 44 B.C. Caesar's death sparked a succession battle between his adopted son, Octavian, and Marc Antony, a general and close ally of the slain *dictator perpetuo*. This succession battle was marked, and possibly decided, by disinformation and the lack of a reaction to it.

⁹³ Lazer, D et al. Combating fake news: an agenda for research and action. Harvard Kennedy School, Shorenstein Center on Media, Politics and Public Policy. Conference report from Feb. 17-18, 2017. <https://shorensteincenter.org/combating-fake-news-agenda-for-research/>

⁹⁴ U.S. President Donald Trump reportedly told Republican donors at a fundraiser at his Mar-a-Lago resort in Florida in March 2018 that Chinese President Xi Jinping being named president for life was “great,” adding: “Maybe we’ll give that a shot some day.” <https://www.reuters.com/article/us-trump-china/trump-praises-chinese-president-extending-tenure-for-life-idUSKCN1GG015>

At first, Octavian and Antony were allies after Caesar's death. With Marcus Aemilius Lepidus, they formed a triumvirate, which sought to avenge Caesar's murder and stabilize the government. Lepidus was forced out of the ruling trio after repeatedly failing in battle, whereupon Octavian took control of the western half of the Roman Empire and Antony the east. It was around this time that Antony began a romantic liaison with the Egyptian queen, Cleopatra, who had also had a romantic liaison with Caesar and had borne him a son. Never at ease with the idea that Caesar had a son – and potential successor – with the Egyptian queen, and angry that Antony was lavishing attention on Cleopatra instead of his second wife back in Rome, who happened to be Octavian's sister, Octavian decided to launch a disinformation war, alongside a physical war, against Antony.^{95 96}

Octavian used disinformation to reduce Antony's chances of acceding to power. His tactics involved playing on Romans' fears of foreigners and strong women. Exploiting a fear of foreigners and strong women was also a hallmark of the disinformation campaign waged during the 2016 U.S. election. He used "short, sharp slogans written upon coins in the style of archaic tweets"⁹⁷ to spread a negative message about Antony. Historians say he also made public what he claimed was Antony's will after two deserters from Antony's camp told Octavian that "its

⁹⁵ Wasson, DL. Second Triumvirate. Accessed online at Ancient History Encyclopedia. https://www.ancient.eu/Second_Triumvirate

⁹⁶ The Russians do this today and call it hybrid or new-type war

⁹⁷ Kaminska, I. A lesson in fake news from the info-wars of ancient Rome. Financial Times. Jan 17, 2017. <https://www.ft.com/content/aaf2bb08-dca2-11e6-86ac-f253db7791c6>

contents, if published, would damn Antony's reputation." The deserters told Octavian that the will was in the safekeeping of the Vestal Virgins of Rome.⁹⁸

Some experts still question whether the will was a forgery,⁹⁹ noting that Antony was too astute and well-informed to have written a will that ignored some of the most basic Roman laws. In particular, they wonder why Antony would have named Cleopatra and her children as his heritors when Roman law forbade *peregrini*, or foreigners, from benefiting from the will of a citizen.

From Octavian's point of view, releasing the will achieved a goal. "[I]t was unimpeachable evidence as far as it went," that Antony had spent too much time in the eastern part of the Roman Empire with Cleopatra and had become un-Roman, "and this caused the Romans in their indignation to believe that the other reports in circulation were also true"¹⁰⁰ and "rekindled hatred"¹⁰¹ of Antony.

The emergence of the will was well-timed for Octavian – just as then FBI Director James Comey's announcement days before the 2016 presidential election that his agency was investigating another batch of Hillary Clinton emails was well-timed for Donald Trump. That some parts of the will may have been forged was immaterial – a reminder of the 2016 U.S. election where many voters in Trump's base overlooked the falseness of some circulating stories about Hillary Clinton because they strongly disliked her.

⁹⁸ Johnson, JR. The authenticity and validity of Antony's will. *L'Antiquité Classique*. T. 47, Fasc. 2 (1978), pp. 494-503. https://www.persee.fr/doc/antiq_0770-2817_1978_num_47_2_1908

⁹⁹ Forgery/fabrication is a feature of disinformation today.

¹⁰⁰ Sirianni, F. Was Mark Antony's will partially forged? *L'Antiquité Classique*. T. 53 (1984), pp. 236-241. https://www.persee.fr/doc/antiq_0770-2817_1984_num_53_1_2126

¹⁰¹ *Ibid.*

The story of Antony's will being offered to Octavian also recalls the incident where music promoter Rob Goldstone told Donald Trump Jr. in an email exchange that the "crown prosecutor of Russia" had "some official documents and information that would incriminate Hillary and her dealings with Russia and would be very useful to your father."¹⁰²

The Octavian and Antony rift used many of the same techniques as are used today to disseminate disinformation: falsification of at least part of Antony's will, repetition of negative claims, playing on prevalent fears and prejudices, opportunism and timing. Octavian shared the disinformation about Antony as widely as he could, given the communications means at his disposal. He read the will aloud in the Senate and had a decree issued that it be posted in the Forum and dispatched by messengers around the empire. Today's equivalent would be to have a story published, broadcast and shared on social media. We will probably never know whether Antony's inaction against the disinformation campaign helped to decide his fate.

3.2 Cold War disinformation

Fast forward to the 1980s and the Cold War. In 1987, a Department of State report accused the Soviets of using forged documents^{103 104} and a disinformation campaign to undermine the image of the U.S. The aim of the disinformation was to "convince the world that the AIDS virus ... had been "manufactured" as a result of

¹⁰² "The email exchange Trump Jr released, in chronological order." CNN <http://www.cnn.com/interactive/2017/07/politics/donald-trump-jr-full-emails/>

¹⁰³ United States Department of State, Soviet Influence Activities: A report on Active Measures and Propaganda, 1986-87. August 1987. Ch. 4, pp 29-32.

<https://www.globalsecurity.org/intell/library/reports/1987/soviet-influence-activities-1987.pdf>

¹⁰⁴ <http://www.nytimes.com/1987/06/03/world/fake-us-documents-appear-around-the-world.html>

genetic engineering experiments conducted at Fort Detrick, Maryland.”¹⁰⁵ Just as the goal in ancient Rome had been to discredit Antony in the eyes of Romans, the forged documents and fake news about HIV were an attempt to discredit the United States and generate anti-American feelings around the world; repeat Soviet propaganda that the U.S. was violating conventions that ban biological weapons and to counter U.S. accusations that the Soviets had done the same; to undermine defense arrangements between the U.S. and its allies; to pressure U.S. allies to seek the withdrawal of U.S. forces from their territory by linking the spread of AIDS to U.S. forces stationed abroad; and to discourage contact with all Americans including tourists, diplomats and businesspeople, by claiming that “American imperialism is responsible for a frightening disease that has made its way into the U.S.S.R. and Eastern Europe.”¹⁰⁶

The story started doing the rounds in 1983, the year AIDS was first attributed to a virus.¹⁰⁷ This was just four years after the 1979 Soviet invasion of Afghanistan, and two years after the inauguration in 1981 of Ronald Reagan as U.S. president. Reagan very rapidly boosted the budget for the U.S. Department of Defense, suspended arms reduction talks and reinstated work on missiles and bombers.^{108 109}

¹⁰⁵ Ibid. p33

¹⁰⁶ United States Department of State, Soviet Influence Activities: A report on Active Measures and Propaganda, 1986-87. August 1987. Ch. 4, pp 29-32.
<https://www.globalsecurity.org/intell/library/reports/1987/soviet-influence-activities-1987.pdf>

¹⁰⁷ Geissler, E., & Sprinkle, R. Disinformation squared: Was the HIV-from-Fort-Detrick myth a Stasi success? *Politics and the Life Sciences*, 32(2), 2-99. Retrieved from
<http://www.jstor.org/stable/43287281>

¹⁰⁸ Boghart, T. Operation INFEKTION: Soviet Bloc Intelligence and Its AIDS Disinformation Campaign. *Studies in Intelligence* Vol. 53, No. 4. December 2009. <https://www.cia.gov/library/center-for-the-study-of-intelligence/csi-publications/csi-studies/studies/vol53no4/soviet-bloc-intelligence-and-its-aids.html>

These moves caused Soviet leader Yuriy Andropov, who was reportedly a little paranoid to begin with, to be convinced the U.S. was actively planning for nuclear war. He ordered KGB operatives around the world to look for signs of a first strike. On September 30, 1982, KGB agents in the U.S. were “instructed to counterattack Washington’s aggressive stance with active measures.”¹¹⁰

The Soviets knew that active measures worked best when they blended in with politics, culture and specific events.¹¹¹ In the case of the AIDS story, Moscow knew that chemical/biological warfare was of great concern to Westerners and could be exploited for disinformation purposes. So a letter, purportedly from a well-known American scientist who wished to remain anonymous,¹¹² was sent to a pro-Moscow Indian newspaper, *The Patriot*. The letter said AIDS had been isolated and spread by U.S. scientists who were engaged in biological warfare preparations.¹¹³

The allegations in that letter were repeated numerous times in other publications. Two years later, for instance, in Soviet newspaper *Literaturnaya Gazeta*, writer Valentin Zapevalov said the United States had found AIDS while “seeking dangerous viruses.”¹¹⁴ A couple of years after that, in 1987, the Soviets began “repeating and embellishing allegations published throughout 1986 and adding

¹¹⁰ Ibid

¹¹¹ Ibid

¹¹² Unnamed sources are a red flag for journalists, less so for the general public.

¹¹³ Geissler E (2016) The AIDS Myth at 30. *Int J Virol AIDS* 3:017

¹¹⁴ Ibid.

distinct arms control themes as the 15th anniversary of the 1972 Biological Weapons Convention approached.”¹¹⁵

The U.S. refuted the allegations each time they appeared, but both the appearance and refutation of disinformation happened at a much slower pace than today – via press release, diplomatic cables, articles in journals, etc. It was only in the late 1980s when HIV began spreading through Russia and the Eastern Bloc that Moscow began to tone down the anti-U.S. AIDS disinformation campaign to allow Soviet scientists to collaborate with their U.S. counterparts. A disavowal by the Soviet Academy of Sciences¹¹⁶ that AIDS had been created artificially was published in the Soviet official newspaper, *Izvestia*. But on the same day, the *Sovetskaya Rossiya* newspaper published an article repeating the AIDS disinformation and defending the right of the Russian press to “report different views.”¹¹⁷

The AIDS disinformation campaign had largely ceased inside the Soviet Union by the summer of 1988, but the story continued to appear in other countries’ newspapers. In 1991, after the demise of the Soviet Union and amid a resurgence of hardliner ideology in Moscow, a newspaper in Zimbabwe reported that HIV was manufactured in a U.S. government lab. That report added a new accusation—that the

¹¹⁵ United States Department of State, *Soviet Influence Activities: A report on Active Measures and Propaganda, 1986-87*. August 1987. Ch. 4, pp 29-32.

<https://www.globalsecurity.org/intell/library/reports/1987/soviet-influence-activities-1987.pdf>

¹¹⁶ Boghardt, T. Soviet bloc intelligence and its AIDS disinformation campaign. *Studies in Intelligence*. Vol 53. No. 4. Dec. 2009. Accessed online at [https://www.cia.gov/library/center-for-the-study-of-](https://www.cia.gov/library/center-for-the-study-of-intelligence/csi-publications/csi-studies/studies/vol53no4/soviet-bloc-intelligence-and-its-aids.html)

[intelligence/csi-publications/csi-studies/studies/vol53no4/soviet-bloc-intelligence-and-its-aids.html](https://www.cia.gov/library/center-for-the-study-of-intelligence/csi-publications/csi-studies/studies/vol53no4/soviet-bloc-intelligence-and-its-aids.html)

¹¹⁷ Report to Congress by the U.S. Information Agency (USIA). *Soviet active measures in the era of glasnost*. March 1988. p 11.

<http://insidethecoldwar.org/sites/default/files/documents/Soviet%20Active%20Measures%20in%20the%20Era%20of%20Glasnot%20March%201988.pdf>

U.S. had exported “AIDS-oiled condoms” to the Soviet Union and Third World to spread HIV.¹¹⁸

In the early 2000s, the Soviet disinformation campaign about AIDS was again given new life, this time on U.S. home turf when Dr. Boyd E. Graves, a lawyer who became an HIV/AIDS activist after he was diagnosed with the virus in the 1990s,¹¹⁹ said AIDS was developed by the U.S. government to “thin the Black Population.”¹²⁰ By this time, what started as disinformation had grown into a conspiracy theory, which over the years has endured and caused damage to those who have bought into it. For example, in 2005, nearly half of African Americans surveyed for a study¹²¹ said they believed HIV was a man-made virus; nearly a quarter said they thought AIDS was produced in a U.S. government laboratory, and more than four in 10 said people who take antiretrovirals, which at the time were relatively new and untested, were guinea pigs for the government.

The authors of the study about beliefs about HIV/AIDS also found that men who believed AIDS conspiracy theories had more negative attitudes toward condoms and were less likely to consistently use them than those who did not buy into

¹¹⁸ From U.S. Information Agency report Soviet Active Measures in The 'Post-Cold War' Era 1988-1991, prepared for U.S. House of Representatives Committee on Appropriations. June 1992. Accessed online at http://intellit.muskingum.edu/russia_folder/pcw_era/sect_09a.htm. Full report at http://intellit.muskingum.edu/russia_folder/pcw_era/index.htm#Contents

¹¹⁹ Dr. Boyd E. Graves died in 2009 at the age of 57. This obituary appears to have been written by Graves himself. <http://boydgraves.blogspot.com/2009/06/dr-boyd-ed-graves-7752-61809.html>

¹²⁰ Graves, BE. The smoking gun of AIDS: A 1971 flowchart. Dec. 6, 2000.

https://www.bibliotecapleyades.net/sociopolitica/esp_sociopol_depapu40.htm

¹²¹ Bogart LM & Thorburn, S. Are HIV/AIDS Conspiracy Beliefs a Barrier to HIV Prevention Among African Americans? *Journal of Acquired Immune Deficiency Syndromes (JAIDS)*. 38(2):213-218, Feb. 2005.

https://journals.lww.com/jaids/Fulltext/2005/02010/Are_HIV_AIDS_Conspiracy_Beliefs_a_Barrier_to_HIV.14.aspx

conspiracies about the virus. Another study, published in 2008,¹²² found that black South Africans who believed HIV was manufactured by white people to reduce or control the black population were less likely to be tested for the virus, a key step to stemming the spread of HIV/AIDS. Yet another study,¹²³ conducted among African-American HIV-infected men who have sex with men, found that those who endorsed AIDS conspiracy theories were more likely to be non-adherent to antiretroviral therapy. That study was published in 2010. The disinformation put forward by the Soviets about AIDS being manufactured by the Americans had endured, in one form or another, for more than 35 years, even surviving the fall of the USSR and several refutations by U.S. and international scientists, diplomats and other trusted officials.

The AIDS disinformation campaign waged by the Soviets and carried on for years after the end of the USSR illustrates one way that disinformation is dangerous. Belief in the disinformation may have endured in a small minority of people, but the potential consequences of the disinformation on that minority were large and potentially serious. If a sector of the population subscribes to a false narrative, disinformation can form the basis for political and societal decisions they make – such as not getting tested for HIV, not using condoms, voting to leave the European Union which could run counter to a society’s best interest.¹²⁴ “If individuals are

¹²² Bogart LM, et al. Endorsement of a genocidal HIV conspiracy as a barrier to HIV testing in South Africa. *Journal of Acquired Immune Deficiency Syndromes (JAIDS)* 2008 Sep 1;49(1):115-6. https://journals.lww.com/jaids/fulltext/2008/09010/Endorsement_of_a_Genocidal_HIV_Conspiracy_a_s_a.20.aspx

¹²³ Bogart LM, Wagner G, Galvan FH, Banks D. Conspiracy Beliefs about HIV Are Related to Antiretroviral Treatment Nonadherence among African American Men with HIV. *Journal of acquired immune deficiency syndromes* (1999). 2010;53(5):648-655. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2845717/>

¹²⁴ Lewandowsky et al. p 107

misinformed, they may ... make decisions for themselves and their families that are not in their best interest and can have serious consequences,” write Lewandowsky et al.¹²⁵

¹²⁵ Ibid

Chapter 4

Case studies

Disinformation has played a role in numerous events and crises this century and last. Some of these crises are well known to Americans, while others happened without being noticed on this side of the Atlantic. For example, in 1991, as the Soviet Union was in the throes of collapse, Russian troops stormed the Vilnius television tower in Lithuania, killing 14 unarmed people. A newscast in Leningrad blamed the deaths on "traffic accidents" and "heart attacks,"¹²⁶ The newscast "showed a man with crazed eyes drinking from a flaming bowl" as the anchor said: "It seems the Lithuanians are prepared to drain the bitter cup of nationalism to its dregs."¹²⁷ Estonia was also accused of excessive nationalism and Nazism by the Russians when the small Baltic state became the first nation to come under cyber attack by another nation. Some have called the weeks'-long cyber and disinformation attacks Web War I.¹²⁸

4.1 Estonia

More than a decade ago, Estonia fought off months of disinformation, physical violence and cyber attacks prompted by a decision to move the statue of a Red Army soldier from the center of Tallinn – a decision it made good on in late

¹²⁶ From U.S. Information Agency report *Soviet Active Measures in The 'Post-Cold War' Era 1988-1991*, prepared for U.S. House of Representatives Committee on Appropriations. June 1992. http://intellit.muskingum.edu/russia_folder/pcw_era/sect_09a.htm.

¹²⁷ Ibid.

¹²⁸ Blank, S. *Web War I: Is Europe's First Information War a New Kind of War?*, *Comparative Strategy*, 27:3, 227-247. July 2008 <https://www.tandfonline.com/doi/pdf/10.1080/01495930802185312>

April 2007. Leading up to, during and after the removal of the Bronze Soldier statue, the rioting that ensued and the cyber attacks that disabled much of the country, including the government, banking and services sectors, the flow of disinformation from Moscow and its surrogates was almost unrelenting. When the cyber attacks began, mainstream media were among the first targets. Twitter was still in its infancy in 2007, so the Estonian authorities were attacked from and fought their disinformation battles mainly in online chatrooms and forums. They also fought to protect access to the internet in the highly wired Baltic country to allow domestic users of the web to continue to access it and go about their daily lives.

Vladimir Putin, who had come to power in Russia six years before the first major clashes over the statue with Estonia in 2006, set the tone to be taken in the disinformation campaign in Estonia. Putin said he wanted to “restore a version of history that was very similar to the earlier Soviet version of history.”¹²⁹ Russian propaganda and disinformation during the Bronze Soldier crisis in Estonia followed the logic that those who do not accept that the Red Army liberated Europe from fascism support fascism and do not deserve respect.¹³⁰ The Russians posted videos during the crisis, showing Estonian police arresting rioters, who were portrayed as law-abiding citizens, while Nazi music played in the background; depicting the Estonian ambassador to Moscow with a Hitler moustache; and writing the name of Estonian Prime Minister Andrus Ansip with two S’s – AnSSip – as in the Nazi

¹²⁹ Unsigned. Russia’s involvement in the Tallinn disturbances. Report by International Centre for Defence and Security (ICDS) think-tank. May 11, 2007.

<https://www.icds.ee/publications/article/russias-involvement-in-the-tallinn-disturbances/>

¹³⁰ Pääbo, H. War of Memories: Explaining “Memorials War” in Estonia. *Baltic Security & Defence Review* Volume 10, 2008.

German *Schutzstaffel*, to drive home their message that Estonia was behaving like World War II fascists.¹³¹ Russian media also martyred 20-year-old Russian Dmitriy Ganin, the only person to die during the rioting. Some reports in Russian media said Ganin died trying to protect the Bronze Soldier statue,¹³² while others said he was tortured by the Estonian police.¹³³ Estonian media immediately counteracted the reports coming out of Russia but the refutations in the Estonian press fell on deaf ears in the country's Russian population, who don't speak Estonian and get their information from Russian media, most of which are controlled by the state.

To this day, the Ganin case remains a hot topic in Russia and is revived regularly.¹³⁴ ¹³⁵ The Estonians continue to refute the accusations¹³⁶ each time new ones arise. Estonian prosecutors last year closed their investigation into the Russian youth's death.¹³⁷

The Estonians may appear to be the losers in what was “the first cyber attack in history that affected a country nation-wide,” according to Helen Popp, counselor for cyber issues at the Estonian Embassy in Washington, D.C. But by taking quick,

¹³¹ Pääbo, H

¹³² ICDB report.

¹³³ Pääbo, H

¹³⁴ Russian Ministry of Foreign Affairs spokeswoman's comment on investigation into the death of Ganin. Sept. 2016. http://www.mid.ru/en_GB/foreign_policy/news/-/asset_publisher/cKNonkJE02Bw/content/id/2435063

¹³⁵ Laruelle, M. (ed). Russian Nationalism, Foreign Policy and Identity Debates in Putin's Russia. Ibidem Press, Stuttgart. 2014. Pp. 133-134

¹³⁶ Unsigned. Bronze Night's only death still unsolved. Estonian Public Media website. April 26, 2017. <https://news.err.ee/592217/bronze-night-s-only-death-still-unsolved>

¹³⁷ Unsigned. Estonian prosecutors close investigation into Dmitri Ganin's death. Estonian Public Media website. June 15, 2017 <https://news.err.ee/602231/estonian-prosecutors-close-investigation-into-dmitri-ganin-s-death>

decisive steps to counter the cyber attacks on its infrastructure, Estonia avoided a complete, costly shutdown of the entire country, and impressed on its NATO partners the need to beef up cyber security capabilities.

Furthermore, by resisting the Russian disinformation offensive, Estonia was able to prevent Russia from achieving its agenda – to deepen the divide between Estonians and Russians who live in the country. The aim of Russia’s “formidable propaganda machine” in Estonia over the years has been “to shape, in overt and covert ways, the perceptions of ethnic non-Estonians [i.e. Russian-speakers]—a task at which it largely succeeds.”¹³⁸ By portraying Estonians as violators of human rights who desecrated the memory of Russian soldiers who fought the Nazis,¹³⁹ Moscow used disinformation to try to deepen the divide in Estonia between Russian-speakers and native Estonians, just as it tried in the run-up to the 2016 U.S. elections to deepen racial and political divides between Americans.¹⁴⁰ Estonia has stepped up efforts to integrate its Russian-speakers, who make up around 25 percent of the population. Estonia’s public broadcaster, ERR, announced in May 2018 that it will expand the reach of its Russian-language channel to the entire country,¹⁴¹ and Prime Minister Jüri

¹³⁸ Kivirahk, J. Integrating Estonia’s Russian-Speaking Population: Findings of National Defense Opinion Surveys. ICDS report. P 3. December 2014.

https://www.icds.ee/fileadmin/media/icds.ee/failid/Juhan_Kivirahk_-_Integrating_Estonias_Russian-Speaking_Population.pdf

¹³⁹ Lucas, E & Pomeranzev, P. Winning the information war. Center for European Policy Analysis report. p.23. August 2016. https://cepa.ecms.pl/files/?id_plik=2706

¹⁴⁰ Ioffe, J. The history of Russian involvement in America’s race wars. The Atlantic. Oct. 21, 2017. <https://www.theatlantic.com/international/archive/2017/10/russia-facebook-race/542796/>

¹⁴¹ Unsigned. Estonian pubcaster expanding reach of Russian-language channel. Digital TV Europe. May 1, 2018. <https://www.digitaltveurope.com/2018/05/01/estonian-pubcaster-expanding-reach-of-russian-language-channel/>

Ratas has pledged to ease citizenship requirements and boost economic development in parts of the country where Russians live.

4.1.1. Disinformation and propaganda

Estonian lawmakers voted months before things came to a head with Russia over the Bronze Soldier statue to have it removed from Tallinn city center. From the time the Estonian government began mulling moving the statue until months after it actually did, Estonia became the target of disinformation and propaganda campaigns.

One of the early anti-Estonian propaganda campaigns linked to the statue came in 2006 when the Embassy of the Russian Federation helped produce a propaganda film named “Estonia – the Crossroads of History.” A statement issued by the Russian embassy in Tallinn to coincide with the film’s release in December of 2006 said it depicted “Russia's principled position on the inadmissibility of revising the outcome of World War II.” Again, Estonians saw things differently. According to the Baltic Times newspaper, the Russian version of history portrayed in the movie tried to “make Estonians somehow culpable for fascism and the war crimes of National Socialist Germany.”¹⁴²

When the film was released, Estonian lawmakers were already mulling moving the statue. A bill presented in the Estonian parliament prior to the elections in March 2007, which saw Andrus Ansip’s Estonian Reform Party emerge as the largest in the legislature, proposed that the statue be moved to a military cemetery in the

¹⁴² Unsigned. "Estonia - Crossroads of History" expresses Russia's position. The Baltic Times. Dec. 15, 2016. <https://www.baltictimes.com/news/articles/17039/>

suburbs of Tallinn, and those buried underneath the statue be exhumed, identified and reburied at the statue's new site. The Russian media called these plans disrespectful and Estonians blasphemous.

When the statue was moved in April 2007, a website called EnglishRussia.com wrote that the Bronze Soldier, "...from Russian point of view commemorates the glory of Soviet forces liberating Estonia and Europe from the Nazi invasion. From the Estonian official point of view it is the monument of the invaders because Estonians were fighting together with German army against Soviet and thus it can't stand anymore in Estonia."¹⁴³ The site said the statue was demolished, not dismantled and moved, and that all that was left of it were "just two pair of bronze boots." Media in Russia reported distortions, half-truths and outright lies alongside images from Tallinn.¹⁴⁴ One Russian journalist reportedly had Russian activists stage a demonstration at the site of the Bronze Soldier, and Russian media "portrayed Russian youth gangs who went on a rampage of vandalism" as peaceful demonstrators¹⁴⁵ and studiously avoided airing footage of the vandalism in the Estonian capital. "By combining disinformation with footage from the city (even if the footage was staged), the programs gave a veneer of veracity to their content," write Lucas and Pomeranzev.¹⁴⁶

Correcting the disinformation about the statue being demolished was easy – Russian lawmakers attended the same ceremony as I did, within days of the statue

¹⁴³ Unsigned. The Bronze Soldier. <http://englishrussia.com/2007/04/27/the-bronze-soldier/>

¹⁴⁴ Lucas, E & Pomeranzev, P. Winning the information war. Center for European Policy Analysis report. p. 28. Aug. 26. https://cepa.ecms.pl/files/?id_plik=2706

¹⁴⁵ Ibid.

¹⁴⁶ Ibid.

being moved, at a military cemetery on the outskirts of Tallinn where the statue was reinstalled. As for fighting alongside the Nazis, Estonians note that some in the Baltic state welcomed the invading Germans with open arms and flowers after nearly two years under the Soviets, from 1939-1941, during which tens of thousands of Estonians were deported to Siberia or executed. But fight alongside the Germans, they did not, the Estonians say. In fact, they note, by the time the Soviets returned to re-occupy Estonia in 1944, the Germans had already withdrawn and the Red Army soldiers lowered the Estonian flag, not the Nazi swastika, when they entered Tallinn. Eight times more Estonians were killed under Soviet occupation than German, according to Estonian nationalist Tiit Madison, who spent several years in a Russian prison camp for anti-Soviet propaganda.¹⁴⁷

In 2006, a group with the Russian name Nochnoy Dozor, or Night Patrol, was set up in Estonia to “defend” the Bronze Soldier. The Russians said Nochnoy Dozor were anti-fascist and, therefore, patriotic. Prior to the statue’s removal, Nochnoy Dozor made several public statements, often calling Estonian politicians Nazis. Nochnoy Dozor was suspected by the Estonian media and officials of helping to organize the rioting that erupted in Tallinn after the statue was moved.

Russian TV channel RTR on April 27, 2007 – immediately after the statue’s removal – carried a phone interview with a Nochnoy Dozor member who claimed he was in a “place of detention” following the rioting. The RTR interviewee said one person had been beaten to death and his body taken away “in view of the other

¹⁴⁷ Blomfield, A. War of words over Bronze Soldier. The Telegraph. Feb. 5, 2007.
<https://www.telegraph.co.uk/news/worldnews/1541641/War-of-words-over-bronze-soldier.html>

detainees.” To refute these rumors, Estonia allowed a group of Russian lawmakers to visit the warehouses by Tallinn’s waterfront where the youth were being held. I was with the Duma representatives when they visited the warehouse area, but for legal and safety reasons, I was not allowed into the warehouse itself. When I spoke to the Russian members of parliament after they’d visited the warehouse-turned-detention-center, one of them, Leonid Slutsky, did not describe anything resembling mistreatment or inhumane conditions in the warehouses. The allegations of abuse and extrajudicial killings of Russians by the Estonians ended up fading away.

But the disinformation campaigns were not done. Six months after the statue had been moved and the bodies buried under it exhumed, identified, and reburied, Nochnoy Dozor posted disinformation on Russian language news sites in Estonia, saying the Estonian currency would be devalued by 50 percent at 4 p.m. and bank ATMs would cease to operate at 11 p.m. The Bank of Estonia quickly issued a denial and assured the public that the Kroon, Estonia’s pre-Euro currency, was not going to be devalued and ATMs would not go dark. But that didn’t stop many Russian-speakers in Estonia from panicking and making a run on exchange bureaus, emptying them of foreign currencies. A large part of the problem in this disinformation campaign was that residents of Estonia who do not speak Estonian do not follow Estonian media, including Russian-language media.¹⁴⁸ Television was, at the time, the main source of information for people over 20 years of age, but nearly 75 percent

¹⁴⁸ Unsigned. Russia’s involvement in the Tallinn Disturbances. ICDS report. May 11, 2007. <https://www.icds.ee/publications/article/russias-involvement-in-the-tallinn-disturbances/>

of Russians in Estonia watched programs broadcast by Russian state TV channels, “which are conspicuous by virtue of their anti-Estonian attitudes.”¹⁴⁹

4.1.2. The delivery method

Cyber attacks, disinformation campaign and physical violence or conventional warfare are elements of Russia’s 21st century *modus operandi* for conflict – hybrid warfare. Hybrid warfare has been defined as a strategy that employs conventional military force supported by irregular and cyber warfare tactics.¹⁵⁰ In testimony to the U.S. House of Representatives’ Committee on Armed Services, the RAND Corporation’s Christopher Chivvis defined hybrid warfare as a distinctly Russian thing where a broad range of subversive instruments, many of them nonmilitary, are deployed to further Russian national interests, often at the expense of the U.S. and its allies.¹⁵¹

Social media was still in its infancy in 2007, so the disinformation attacks that encouraged people to take down the online presence of what was at the time the most wired nation in Europe played out in closed, online Russian chat rooms and public online forums. Detailed instructions posted online in Russian explained how to launch an attack, who and what to target, and when to launch a coordinated attack. For the most part, these instructions were written in simple terms so that ordinary, non-

¹⁴⁹ ICDS report.

¹⁵⁰ Stowell, J. What is hybrid warfare? Global Security Review. Feb. 2, 2018. <https://globalsecurityreview.com/hybrid-and-non-linear-warfare-systematically-erases-the-divide-between-war-peace/>

¹⁵¹ Chivvis, C. Understanding Russian hybrid warfare and what can be done about it. Testimony presented before the House Armed Services Committee on March 22, 2017. Accessed online at https://www.rand.org/content/dam/rand/pubs/testimonies/CT400/CT468/RAND_CT468.pdf

computer-savvy individuals and their *babushka* could help to incapacitate Estonia, where 60 percent of the population relied on the Internet for “crucial” services every day, and nearly all banking transactions occurred online.¹⁵²

The cyber attacks were waged at three levels, or on three fronts, to maintain the military analogy: by individuals, dubbed “script kiddies” who copied programs, line for line, off hacker sites; by networks of computers around the world, called botnets, that had been hijacked by hackers and began flooding certain internet addresses with floods of data; and by hackers who deleted content on individual websites and replaced it with their own messages¹⁵³

Disinformation played a vital role in the whipping the human attackers – particularly the script kiddies – into a frenzy. As the script kiddies used ping attacks – a request, repeated hundreds of times per second, for a response from a web server – to overwhelm Estonian servers, they were “stoked into a fervor on Russian-language chat rooms.”¹⁵⁴

“First they were goaded by overheated rhetoric about the April 27 removal of the statue. A week later, hundreds of posts called for a coordinated attack at the stroke of midnight on May 9, the day Russia celebrates its World War II victory. ‘You do not agree with the policy of eSStonia???’¹⁵⁵ demanded a user named Victoris on a

¹⁵² Richards, J. Denial-of-Service: The Estonian Cyberwar and Its Implications for U.S. National Security. *International Affairs Review*. <http://www.iar-gwu.org/node/65>

¹⁵³ Davis, J. Hackers Take Down the Most Wired Country in Europe. *Wired Online*. Aug. 21, 2007. <https://www.wired.com/2007/08/ff-estonia/>

¹⁵⁴ Davis, J. *Ibid.*

¹⁵⁵ A reference to the oft-used Russian line that Estonians collaborated with the Nazis

Russian online forum. ‘You may think you have no influence on the situation??? You CAN have it on the Internet!’”¹⁵⁶

At the same time, botnets bombarded key Estonian websites, including government officials’ and agencies’ sites, political parties, banks, media, and communications companies, with distributed denial of service (DDoS) attacks – designed to cause a site to crash. Hackers took over selected sites and changed their content. Like the script kiddies, the hackers communicated among themselves and posted their intent on public forums.¹⁵⁷ An investigation by ICDS found that calls to attack Estonia’s connectivity were also being put out on Russian chat forums, 2ch.ru and forum.xaker.ru, and discussed and coordinated in Internet Relay Chat environments, which facilitate group discussions. The Estonian authorities were monitoring the chatrooms throughout and were able to mount a defense of Estonia’s internet connectedness, without which the country would grind to a halt.

At the end of April 2007, three days after the attacks had begun, the Estonian authorities blocked most internet traffic arriving in Estonia from IP addresses with the extension .ru (Russia) and from other foreign IP addresses. In the first few days of May, two telecommunications companies (Elion and Elisa) that provide internet services and the Postimees media group, publisher of the country’s flagship newspaper Postimees, were also targeted by the attackers. The newspaper’s IT manager decided as the attacks ramped up to block international requests to the paper

¹⁵⁶ Davis, J.

¹⁵⁷ Davis, J. Hackers Take Down the Most Wired Country in Europe. Wired Online. Aug. 21, 2007. <https://www.wired.com/2007/08/ff-estonia/>

in order to preserve Estonians' access. The tactic worked: Postimees' website "became accessible again within Estonia, but at a cost. Estonia's leading news outlet could not tell the world what was going on in its own country."¹⁵⁸

Prime Minister Andris Ansip told the new U.S. ambassador to Tallinn, Stanley Davis Philips, that the well-organized attacks sought to undermine government institutions in Estonia. Ansip added that when Estonian websites were down, Russian websites filled the void and spread disinformation.¹⁵⁹ Estonian Defense Minister Jaak Aaviksoo said he believed Russia's strategic objective in mounting the cyber attacks and spreading disinformation about Estonia was to create tensions within Europe and destabilize the EU.¹⁶⁰

The attacks ended on May 19 after Estonia "pulled the plug" and largely disappeared from the internet.¹⁶¹ Some considered this a victory for whomever had perpetrated the attacks – Estonians and others in the EU and NATO believed Russia was responsible,¹⁶² but Moscow denied having anything to do with the cyber onslaught. But Russia failed to advance its agenda of creating tensions within Estonia, and between Estonia and Europe. It also failed to destabilize the EU and NATO. In fact, the opposite happened. A research paper for NATO said the attacks "were a true wake-up call for NATO, offering a practical demonstration that cyber attacks could

¹⁵⁸ Davis, J.

¹⁵⁹ Wikileaks Cable 297, par. 5. https://wikileaks.org/plusd/cables/07TALLINN310_a.html

¹⁶⁰ Ibid, par 3.

¹⁶¹ Grassegger, H & Krogerus, M. Fake news and botnets: How Russia weaponized the web. The Observer. Dec. 2, 2017. <https://www.theguardian.com/technology/2017/dec/02/fake-news-botnets-how-russia-weaponised-the-web-cyber-attack-estonia>

¹⁶² Traynor, I. Russia accused of unleashing cyberwar to disable Estonia. Russia accused of unleashing cyberwar to disable Estonia. The Guardian. May 16, 2017. <https://www.theguardian.com/world/2007/may/17/topstories3.russia>

now cripple an entire nation which is heavily dependent on IT networks.”¹⁶³ A year after the attacks, NATO established a cyber defense research center in Tallinn and beefed up its cyber war capabilities. Furthermore, because they were not very sophisticated, the cyber attacks had little effect on Estonia’s IT network, and because Estonia is a relatively small country with highly trained IT experts and a Computer Emergency Response Team (CERT), its cyber experts were able to mobilize cyber defense teams and take speedy action to protect Estonia’s networks.¹⁶⁴

Israeli cyber security expert Gadi Evron, who traveled to Estonia to lend a hand during the 2007 attacks, said in a post-mortem of the crisis, “It took Hillar (Aareleid) and the others about three weeks to start getting things under control, but they did. Persistence and a focus on hardening the nation’s defenses slowed the momentum of the attackers, slowing them down before finally turning them back.”¹⁶⁵ The attacks on Estonia made clear that cyber and information warfare are “about people. Machines were just what people were using,” said Evron.¹⁶⁶ “The goal is to disarm the enemy by causing them to disarm themselves,” he said, citing Russian strategist S. P. Rastorguev,¹⁶⁷ who liked to explain information warfare by telling the fable of a fox that wanted to eat a turtle but couldn’t until he convinced her to take off her shell. Information warfare, Rastorguev would say, is the purposeful training of an enemy on how to remove its own shell. Russia has understood there is a human side

¹⁶³ Joubert, V. Five years after Estonia’s cyber attacks: lessons learnt for NATO? NATO Research Paper. No. 76. May 2012.

¹⁶⁴ Joubert, V.

¹⁶⁵ Evron, G. Estonia 10 Years Later: Lessons learned from the World’s First Internet War. Securityledger.com. April 28, 2017. <https://securityledger.com/2017/04/estonia-10-years-later-lessons-learned-from-the-worlds-first-internet-war/>

¹⁶⁶ Ibid.

¹⁶⁷ Ibid.

of information warfare, while the West, and the U.S. in particular, “concentrate on the technical aspects alone,”¹⁶⁸ said Evron.

“Estonia taught me that ... the human element should not be separated from the computerized aspects,” Evron said.¹⁶⁹

4.2 How Georgia won a disinformation battle with Russia but lost the war

A year after Estonia’s cyber war, conflict broke out between Russia and Georgia over the autonomous region of South Ossetia in Georgia. The conflict lasted less than a week and marked the first time that cyber and disinformation attacks were synchronized with major combat actions.^{170 171} More than 50 websites in Georgia were “attacked by rogue elements within Russia” immediately prior to the start of fighting.¹⁷² “With Georgian government websites shut down, the Russian story of its army coming to the defense of South Ossetia in the face of Georgian assault gained currency,” Robert Cutler, an analyst at the Central Asia Caucasus Institute (CACI) wrote around the time of the conflict.¹⁷³

¹⁶⁸ Evron, G. A Russian strategist’s take on information warfare. Darkreading.com. March 26, 2010. <https://www.darkreading.com/risk/a-russian-strategists-take-on-information-warfare/d/d-id/1133301>

¹⁶⁹ Ibid

¹⁷⁰ Hollis, D. Cyberwar case study: Georgia 2008. Small Wars Journal. <http://smallwarsjournal.com/blog/journal/docs-temp/639-hollis.pdf>

¹⁷¹ Cohen, A & Hamilton, R. The Russian Military and the Georgian War: Lessons and Implications. Strategic Studies Institute. Sept. 1, 2011. p. 44. <http://ssi.armywarcollege.edu/pdffiles/PUB1069.pdf>

¹⁷² Shachtman, N. Top Georgian Official: Moscow Cyber Attacked Us – We Just Can’t Prove It. Wired Magazine. March 11, 2009. Cited in Hollis, D. Cyberwar case study: Georgia 2001. Small Wars Journal. Available in full at <http://www.wired.com/dangerroom/2009/03/georgia-blames/>

¹⁷³ Cutler, R. Russia’s disinformation campaign over South Ossetia. CACI Analyst. August 20, 2008. <https://www.cacianalyst.org/publications/analytical-articles/item/11678-analytical-articles-caci-analyst-2008-8-20-art-11678.html>

Russia and Georgia recognized the importance of and sought to control the flow of information to the global community during the conflict. Russia's main message to the international community was that Georgia and its president, Mikheil Saakashvili, were the aggressors in the conflict and Russia was merely trying to defend its citizens living in South Ossetia.¹⁷⁴ The Russian media and "the highest Russian leaders"¹⁷⁵ accused Saakashvili of war crimes, repeatedly citing reports that 2,000 civilians had been killed and more than half the population of 70,000 of the main city of South Ossetia, Tskhinvali, had fled when Georgia launched an attack on the city.¹⁷⁶

Initially, Russia appeared to have the upper hand in the information war. The narrative that Russia was merely responding to Georgian aggression to protect its citizens took hold and continued to dominate discussions even after evidence came to light that Russia was in South Ossetia before the conflict broke out.¹⁷⁷ The overwhelming majority of respondents to a CNN poll taken at the time – 92 percent – said they believed Russia's intervention in South Ossetia was justified,¹⁷⁸ although sources say the poll results were skewed after it was amplified by Russian bloggers who circulated it widely among Russians.¹⁷⁹

¹⁷⁴ <https://freedomhouse.org/report/freedom-world/2009/south-ossetia>

¹⁷⁵ Ibid.

¹⁷⁶ Cutler, R. <https://www.cacianalyst.org/publications/analytical-articles/item/11678-analytical-articles-caci-analyst-2008-8-20-art-11678.html>

¹⁷⁷ Cohen, A & Hamilton, R. *The Russian Military and the Georgian War: Lessons and Implications*. Published by Strategic Studies Institute. Sept. 1, 2011. p. 47

¹⁷⁸ Iasiello.

¹⁷⁹ Levine, Y. *The CNN effect: Georgia schools Russia in information warfare*. Aug. 13, 2008. <http://exiledonline.com/the-cnn-effect-georgia-schools-russia-in-information-warfare/>

The Georgians quickly launched an information counter-attack, relying on the most widely used media sources of the day, television and websites, to share their version of events. Saakashvili's official website, which had been taken down in the cyber attacks on Georgia, was re-established as a page on the website of Polish President Lech Kaczynski, and other Georgian government websites reappeared as blogs "behind the protection of google.com."¹⁸⁰ In dozens of appearances on Western television and in discussions at Western think-tanks during the five-day conflict, the Georgians asserted they were the victims and Russia the aggressor.¹⁸¹

"Saakashvili's presence on CNN and other Western stations, and his and the Georgian leadership's command of foreign languages, have enabled them to get their message across effectively," wrote Cutler.¹⁸² In contrast, Russian leaders held "no spontaneous interviews, even with representatives of Russian media," he wrote.

Georgia also put at least one E.U. public relations firm to work spreading the Georgian version of events in the South Ossetia conflict.¹⁸³ "Almost hourly over the five-day war, press releases landed on foreign news desks," reported Peter Wilby in The Guardian in August 2008.¹⁸⁴ The press releases included trigger terms designed to pique the interest of international copy tasters and news editors: "Russia continues

¹⁸⁰ Cohen, A & Hamilton, R. The Russian Military and the Georgian War: Lessons and Implications. Published by Strategic Studies Institute. Sept. 1, 2011. p. 46

¹⁸¹ Hahn, G. The 2008 Georgian-Ossetian/Russian war: Saakashvili's strategic communications/disinformation campaign. Revised July 30, 2017. <https://gordonhahn.com/2017/07/30/the-2008-georgian-ossetianrussian-war-saakashvilis-strategic-communications-disinformation-campaign/>

¹⁸² Cutler, R. <https://www.cacianalyst.org/publications/analytical-articles/item/11678-analytical-articles-caci-analyst-2008-8-20-art-11678.html>

¹⁸³ Wilby, P. Georgia has won the PR war. The Guardian. Aug. 17, 2008 <https://www.theguardian.com/media/2008/aug/18/pressandpublishing.georgia>

¹⁸⁴ Wilby, P. The Guardian. <https://www.theguardian.com/media/2008/aug/18/pressandpublishing.georgia>

to attack civilian population' ... Tbilisi was 'intensively' bombed. A downed Russian plane turned out to be 'nuclear'... A 'humanitarian wheat shipment' was blocked ... 'invading Russian forces' began 'the occupation of Georgia,'" read the releases' headlines, according to Wilby.

The Georgian information strategy achieved the desired result. "Most [Western] newspapers hadn't a clue what was going on and lacked sufficient resources to find out," Wilby wrote. "So skilfully presented PR was at a premium. Most journalists treated it with at least some scepticism, but it inevitably had an effect. If there was a military war, there was also an information one, and Georgia got the better of it."¹⁸⁵

A review of the Russian military's performance in the Georgia incursion concurred that Georgia had won the information battle. Russia had shown "deficiencies in both the information-technical and information-psychological domains" and Georgia had "won the hearts and minds of the global community even though Russia won the physical battlespace."¹⁸⁶ Russia is still present in South Ossetia and is one of four countries that have recognized its independence.

¹⁸⁵ Ibid.

¹⁸⁶ Iasiello, EJ. Ibid. Jan. 6, 2011.

4.3 German newspaper reacts to disinformation

“We at Breitbart News ... take the utmost care to present facts.”

- London Editor in Chief Raheem Kassam

In January 2017, alt-right online publication Breitbart News’ London bureau published a story that claimed “a mob of more than 1,000 men chanted ‘Allahu Akhbar’” and burnt down Germany’s “oldest” church during New Year’s celebrations in Dortmund on Dec. 31, 2016.¹⁸⁷ The Dortmund police¹⁸⁸ and the reporter who covered the New Year’s Eve celebrations¹⁸⁹ for the Ruhr Nachrichten newspaper immediately took to Twitter and the online edition of the paper, and gave interviews and released statements refuting Breitbart’s assertions. Among other things, they noted that the church that caught fire was not targeted by Muslim extremists, did not burn to the ground, and is not Germany’s oldest. There is no mention in the Ruhr Nachrichten story, published as a live feed during the rowdy celebrations, of a 1000-strong mob chanting “Allahu Akbar” in the town center of Dortmund, and in the video posted online by Breitbart and Ruhr Nachrichten, it is hard to discern any chants of Allahu Akbar.

¹⁸⁷ Hale, V. Revealed: 1,000-Man Mob Attack Police, Set Germany’s Oldest Church Alight on New Year’s Eve. Breitbart.com. Jan. 3, 2017. Accessed online at <http://www.breitbart.com/london/2017/01/03/dortmund-mob-attack-police-church-alight/>

¹⁸⁸ Report on Ruhr Nachrichten’s website [https://www.Ruhr Nachrichten.de/Nachrichten/Polizei-reagiert-auf-Fake-News-zu-Dortmunder-Silvester-60951.html](https://www.RuhrNachrichten.de/Nachrichten/Polizei-reagiert-auf-Fake-News-zu-Dortmunder-Silvester-60951.html)

¹⁸⁹ [https://www.Ruhr Nachrichten.de/Staedte/Dortmund/Silvesterrakete-trifft-Obdachlosen-64703.html](https://www.RuhrNachrichten.de/Staedte/Dortmund/Silvesterrakete-trifft-Obdachlosen-64703.html)

Breitbart's version of the story was published on January 3, 2017. It appears to have been based not as much on the live feed published on Ruhr Nachrichten's website, but from the website of Wochenblick.at,¹⁹⁰ which was set up by former members of Austria's far-right Freiheits Partei (FPÖ) and is populated with mis- and disinformation. The first sentence of the opening paragraph of the report on the Austrian website reads:

“A mob of more than 1,000 men reportedly shot fireworks at police, Syrians shouted the Islamic battle cry ‘Allahu Akbar’ and the fire department had to extinguish a fire at a church.”

The lead of the Breitbart story reads:

“At New Year's Eve celebrations in Dortmund a mob of more than 1,000 men chanted ‘Allahu Akhbar’, launched fireworks at police, and set fire to a historic church.”

The Ruhr Nachrichten story was updated constantly throughout the evening of December 31, 2016. At around midnight, the newspaper wrote on its website: “In Leeds Place, the fireworks lasted a good 45 minutes, even though it's forbidden to shoot off fireworks near the church. The Reinhold's church is next to the square. Fire fighters had to access the church roof from Klepping Street after netting covering some scaffolding was set on fire by fireworks. The small fire was quickly put out.

¹⁹⁰ Wochenblick's Dortmund New Year's Eve 2016/17 story accessed online.
<https://www.wochenblick.at/silvester-in-dortmund-allahu-akbar-und-kirchenbrand/>

The firework that started the fire was probably launched from Klepping Street, where many people were also celebrating.

A group of Syrians sang as they celebrated the ceasefire in Syria.”¹⁹¹

Wochenblick says in its article that it used the photo gallery put together by Ruhr Nachrichten reporter Peter Bandermann and a report on WDR television to reach their conclusion that most of the mob of 1,000 men were migrants and that they targeted police with fireworks.¹⁹²

In the Ruhr Nachrichten live-feed from the new year’s celebrations, which Breitbart and Wochenblick cite in their articles, the German news website reported that the celebrations were rowdy and sometimes dangerous, but nowhere in its text or video reports did Ruhr Nachrichten say that Syrians were shouting “Allahu Akbar” or that “migrants” were responsible for shooting fireworks at the church, police or into the crowd. In fact, Ruhr Nachrichten cites police spokesman Volker Stall who blamed “these dangerous actions on young men, with and without a migration background.”¹⁹³

“At 11:45 pm, a crowd of at least 1,000, made up mainly of young men, gathered in Leeds Place in downtown Dortmund. Fireworks were often handled in a

¹⁹¹ <https://www.ruhrnachrichten.de/Staedte/Dortmund/Silvesterrakete-trifft-Obdachlosen-64703.html>

¹⁹² “Bei dem Mob von mindestens 1000 überwiegend jungen Männern handelt es sich größtenteils um Migranten, wie auch die Fotostrecke der Ruhr Nachrichten beweist. Diese hätten die Polizeikräfte gezielt mit Feuerwerkskörpern beworfen, meldet der Fernsehsender WDR.”

¹⁹³ “Verantwortlich für diese gefährlichen Aktionen seien ‘junge Männer mit und ohne Migrationshintergrund’.” From New Year’s Wrap, Ruhr Nachrichten.de. Aktualisierung Montag, 11.23 Uhr: Schwerverletzter am Hauptbahnhof (Update, Monday 11:23 am. Badly injured man at main train station). <https://www.ruhrnachrichten.de/Staedte/Dortmund/Silvesterrakete-trifft-Obdachlosen-64703.html>

risky manner – rockets held with bare hands as they were shot off, instead of placing them on the ground. Several unidentified persons threw bangers at police and shot fireworks into the crowd, in which there were families with children. The police ordered the crowd not to shoot bangers and fireworks into the crowd but their announcements over loudspeakers were ignored. Several men were followed, detained and searched.”¹⁹⁴

Bandermann reacted quickly after the Breitbart and Wochenblick articles were published. Going point by point through the Breitbart and Wochenblick reports, Bandermann corrected what the two websites had reported and confirmed what he had seen. While Bandermann agreed that the groups moving through downtown Dortmund on New Year’s Eve were made up mainly of young foreign men, he refuted reports in Breitbart and Wochenblick that the “mob” in Leeds Place had deliberately shot fireworks at the nearby Reinhold’s church and set it alight.

“The fact is the fire department spoke of a small fire on some scaffolding netting, which was quickly put out,” he said in his refutation, echoing what he had reported on New Year’s Eve.^{195 196}

Responding to reports that young men were shouting “Allahu Akbar” in the town center, Bandermann replied, “‘Allahu Akbar’ means ‘God is great.’ In the

¹⁹⁴ Original German live feed from Ruhr Nachrichten accessed at [https://www.Ruhr Nachrichten.de/Staedte/Dortmund/Silvesterrakete-trifft-Obdachlosen-64703.html](https://www.RuhrNachrichten.de/Staedte/Dortmund/Silvesterrakete-trifft-Obdachlosen-64703.html)

¹⁹⁵ Bandermann, P. Fragen und Antworten: Wie die Medien die Silversternacht für Hass-Aufbau nutzen (Questions and answers: How the media use the New Year’s celebrations to build hate). Ruhr Nachrichten.de. Jan 4, 2017. <https://www.ruhrnachrichten.de/Staedte/Dortmund/Wie-Medien-die-Silvesternacht-fuer-Hass-Aufbau-nutzen-55127.html>

¹⁹⁶ <https://www.ruhrnachrichten.de/Staedte/Dortmund/Silvesterrakete-trifft-Obdachlosen-64703.html>

Muslim religion, this expression is as normal as ‘amen’ is in the church. While it’s indeed true that terrorists misuse ‘Allahu Akbar’ before they launch attacks, fake news producers are linking a group of people in Leeds Place with the attackers in Munich (Olympia shopping center)¹⁹⁷ and Würzburg,¹⁹⁸ and the killer of a priest in France.¹⁹⁹ The fact is, there was no indication that there were terrorists celebrating in Dortmund.”

The Dortmund police issued a statement on Jan. 5, 2017, saying that the New Year celebrations in the city were “average to peaceful” and that police had to intervene far fewer times than the previous year.²⁰⁰

The editor of Breitbart defended the journalist who wrote the piece and the article itself. Raheem Kassam slammed U.S. mainstream media and the Ruhr Nachrichten for misreporting the incident in Dortmund and Breitbart’s coverage of it. The rebuttal article claims, “Much of the media’s outcry over Breitbart News reports emanates from the scale of the fire at St. Reinold’s Church;” that “outlets such as Mediaite, TeleSur, Sputnik, HuffPo, Media Matters, Deutsche Welle, The Washington Post and others decided to use words like ‘Muslim,’ ‘migrant,’ ‘Islam,’ ‘Arab,’ and ‘immigration’ in their headlines or reporting on our story ... to stoke fears and division.” Breitbart accused German police and politicians of colluding with

¹⁹⁷ Bandermann had himself fallen prey to repeated disinformation – the Munich Olympia shopping center shooter was found to be a far-right sympathizer.

¹⁹⁸ A 17-year-old Afghan asylum seeker in Germany attacked people with a knife and hatchet on a commuter train in Würzburg in July 2016. The so-called Islamic State claimed he was one of their fighters and an IS flag was found in the youth’s room.

¹⁹⁹ In July 2016, two men burst into a church near Rouen, in northern France, slit the 84-year-old priest’s throat and took four parishioners hostage. The attackers said they were linked to IS.

²⁰⁰ Statement by Nordrhein-Westfalen Dortmund Polizei, Jan. 5, 2017.

<https://www.presseportal.de/blaulicht/pm/4971/3527791>

the media to “mask a true version of events that are inconvenient at best, or institutionally damning at worst.” The rebuttal ended with the statement that “we at Breitbart News ... take the utmost care to present facts.”²⁰¹ Kassam said he would leave it to readers to decide whom to believe on the Dortmund story. Whether they will look at other sources is open to question. A study by Benkler et al has found that readers of far-right websites such as Breitbart tend to consume news mainly from sources at the same end of the political spectrum.²⁰²

It could be that right-wing media outlets such as Breitbart and Wochenblick were on alert for a repetition of what happened the previous year in Germany, when some 1,200 women reported being sexually assaulted by primarily Arab and North African men at New Year’s Eve celebrations in German cities. The bulk of the 2015 assaults happened in Cologne. Germany has taken in hundreds of thousands of migrants and refugees since 2015. Analysts have said Germany’s open-doors policy toward migrants led to the rise of the far-right Alternative for Germany (AfD) party, which finished third in elections in September 2017.²⁰³

The events in Dortmund would likely have garnered only local coverage but for the attention focused on Ruhr Nachrichten by the Breitbart article. Ruhr Nachrichten does not show how many times stories on its website are shared, but

²⁰¹ Breitbart rebuttal by London Editor in Chief Raheem Kassam.

http://www.breitbart.com/london/2017/01/08/fake-news-fake-news-media-sow-division-with-dishonest-attack-on-breitbart-allahu-akbar-church-fire-story/#disqus_thread

²⁰² Benkler, Y et al. Study: Breitbart-led right-wing media ecosystem altered broader media agenda. Columbia Journalism Review. March 3, 2017. Accessed online at

<https://www.cjr.org/analysis/breitbart-media-trump-harvard-study.php>

²⁰³ Erixon, F. Angela Merkel has created Germany’s far-right. The Spectator. February 24, 2018.

<https://www.spectator.co.uk/2018/02/angela-merkel-has-created-germanys-far-right/>

according to a story by Agence France-Presse (AFP) published in The Guardian newspaper, the Breitbart story about New Year's Eve in Dortmund was clicked on and shared tens of thousands of times.²⁰⁴ Breitbart's site shows it was shared over 17,000 times on Facebook alone. But shares of rebuttals far outstrip those figures. The AFP story cited above was shared more than 33,000 times after it was published in The Guardian. Another AFP story published by Ireland's The Journal has been viewed more than 16,000 times.²⁰⁵

In March 2017, Breitbart put on indefinite hold its plans to launch German and French editions,^{206 207} which had been announced after Donald Trump's victory in the U.S. in 2016. The aim of the bureaus would have been to capitalize on the anti-immigrant sentiment sweeping Europe with a view to creating Trump-style upsets in 2017 elections in the two European countries.^{208 209} Breitbart's Rome correspondent told Germany's Junge Freiheit website that the expansion into Europe had been delayed indefinitely due to "the many complications involved" in opening new

²⁰⁴ Agence France-Presse. German police quash Breitbart story of mob setting fire to Dortmund church. Published in The Guardian. Jan. 7, 2017. <https://www.theguardian.com/world/2017/jan/07/german-police-quash-breitbart-story-of-mob-setting-fire-to-dortmund-church>

²⁰⁵ Unsigned. Breitbart ran a fake story about a terrorist outrage in Dortmund, and Germans aren't one bit happy. Thejournal.ie.2016. <http://jrnl.ie/3173688>

²⁰⁶ Unsigned. Breitbart drags feet on German expansion. Handelsblatt Global. March 7, 2017.

<https://global.handelsblatt.com/companies/breitbart-drags-feet-on-german-expansion-720136>

²⁰⁷ Unsigned. Keine deutschsprachige Breitbart-Ausgabe (No German language Breitbart). Junge Freiheit website. March 3, 2017. <https://jungefreiheit.de/kultur/medien/2017/keine-deutschsprachige-breitbart-ausgabe/>

²⁰⁸ Mohdin, A. What's "alt-right" in German? Breitbart News is expanding in Europe. Quartz. Nov. 20, 2016. <https://qz.com/841959/breitbart-news-is-coming-to-france-and-germany-ahead-of-elections/>

²⁰⁹ Flitter, E. Exclusive: Riding Trump wave, Breitbart News plans U.S., European expansion. Reuters. Nov. 9, 2016. <https://www.reuters.com/article/us-usa-election-trump-strategy-idUSKBN1342TP>

bureaus.²¹⁰ Among the complications was the fact that “Breitbart has a poor reputation in Germany,” according to the Handelsblatt newspaper.²¹¹ “A recent Breitbart article about an alleged attack in Dortmund reported false and highly-exaggerated facts, leading to a tongue-lashing from various German news sources,” Handelsblatt reported, referring to the New Year’s Eve story run by Breitbart.

Note: Bandermann inadvertently showed the staying power of disinformation when he linked the shootings at Munich’s Olympia shopping center in July 2016 to Islamic extremists. During the Munich shootings, a woman who identified herself as Lairetta claimed on Twitter that the shooter had shouted “Allahu Akbar” before opening fire. “With her statements, she planted the seed of Islamic terror in people’s minds, a thought that many probably already entertained as the shootings were being reported,” wrote Nina Probst in the Miesbacher Merkur newspaper. “Lairetta told (CNN) by phone, ‘I heard ‘Allahu Akbar, Allahu Akbar, Allahu Akbar.’ She recognized it because she is herself Muslim, she said. Her 8-year-old son was in the restroom with the shooter and saw as he loaded his weapon. Only CNN carried a report with this witness. Later, her statement was shown to be false.” An investigation in Germany found that the Munich shopping center shooter had been motivated by far-right beliefs.

²¹⁰ Unsigned. Keine deutschsprachige Breitbart-Ausgabe (No German language Breitbart). Junge Freiheit website. March 3, 2017. <https://jungefreiheit.de/kultur/medien/2017/keine-deutschsprachige-breitbart-ausgabe/>

²¹¹ Unsigned. Breitbart drags feet on German expansion. Handelsblatt Global. March 7, 2017. <https://global.handelsblatt.com/companies/breitbart-drags-feet-on-german-expansion-720136>

Chapter 5

How social media, new tech drive the disinformation tsunami

Twitter was a year old when Estonia came under disinformation and cyber attack from Russia, and two years old when Russia launched similar attacks, along with a conventional warfare assault, on Georgia. Both Estonia and Georgia fought back against the attacks on the same media as they were occurring, largely the internet (but not social media) and conventional media. Today, many disinformation attacks occur on social media, and Twitter in particular.

5.1 Twitter's role

Launched in March 2006, Twitter had 330 million active users worldwide by 2018.²¹² Twitter is one of the largest mediums easily accessible to ordinary citizens, i.e. not trained journalists, on which they can receive, create and spread real and fake news. Like many other social media, Twitter allows news and other information to be posted publicly and immediately, but unlike some social media, particularly Facebook, Twitter users can follow another user without that user's agreement and without the other user following them.²¹³ ²¹⁴ This makes more information available to them. The use of hashtags on Twitter (# followed by a word or phrase) allows users

²¹² <https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/>

²¹³ Broersma, T. & Graham, Twitter as a news source. *Journalism Practice*, 7:4, 446-464. May 30, 2013. <https://doi.org/10.1080/17512786.2013.802481>

²¹⁴ Kwak, H. et al. What is Twitter, a social network or a news media? Proc. of the 18th international conference on World Wide Web. Pp 591-600. April 2009. <https://dl.acm.org/citation.cfm?id=1772751>

to categorize tweets by keyword, which helps to make tweets easier to find in a Twitter search.

Twitter has been called “the par excellence social media to catch up on recent news and events.”²¹⁵ The medium has grown into a real-time news source created by everyday users.²¹⁶ Three-quarters of Twitter users got their news on the microblogging site in 2017, according to a survey conducted in August 2017 by the Pew Research Center.²¹⁷

Twitter has on many occasions beaten traditional media with a breaking news story. For example, Twitter users announced 15 minutes before the mainstream media that a U.S. Airways flight had ditched on the Hudson River in New York in 2010.²¹⁸

²¹⁵ Zubiaga, A. et al. Curating and contextualizing twitter stories to assist with social newsgathering. Proceedings of IUI 2013, pp 213–224. March 19–22, 2013.

<https://dl.acm.org/citation.cfm?id=2449424>

²¹⁶ Metaxas, PT. Retweets indicate Agreement, Endorsement, Trust: A Meta-Analysis of Published Twitter Research. Computer Science Department, Wellesley College. Accessed online at <http://cs.wellesley.edu/~pmetaxas/WorkingPapers/Retweet-meaning.pdf>

²¹⁷ Shearer, E & Gottfried, J. News use across social media platforms 2017. Pew Research Center. Sept. 7, 2017. <http://www.journalism.org/2017/09/07/news-use-across-social-media-platforms-2017/#>

²¹⁸ Beaumont, C. New York plane crash: Twitter breaks the news, again. The Telegraph (UK). Jan. 16, 2009.

<https://www.telegraph.co.uk/technology/twitter/4269765/New-York-plane-crash-Twitter-breaks-the-news-again.html>

The public often turns to Twitter for information during a crisis.^{219 220 221 222}

Twitter served as a “crucial channel of communication” between the Kenyan government, emergency responders and the public during the four-day Westgate Mall siege in Nairobi in 2013,²²³ carried blow-by-blow updates from official and unofficial sources during and after the Boston Marathon bombings in 2013,²²⁴ and was used as a “hub of timely information provision to help people stay informed and safe” when Hurricane Sandy hit New York and New Jersey in 2012.²²⁵

Twitter users accept that, in a crisis especially, tweets will often not provide sources, which could be used to confirm the accuracy of information being shared.²²⁶ Users have also come to accept that the restricted length of tweets (280 characters, formerly 140) means tweets rarely, if ever, include supporting arguments.²²⁷ As a

²¹⁹ Blair Thompson, Joseph P. Mazer, Holly J. Payne, Angela M. Jerome, E. Gail Kirby & William Pfohl (2016) Social Media and Active Shooter Events: A School Crisis Communication Challenge. *Qualitative Research Reports in Communication*, 18:1, 8-17. <https://www.tandfonline.com/doi/abs/10.1080/17459435.2016.1247111>

²²⁰ Vieweg, Sarah. Microblogged Contributions to the Emergency Arena: Discovery, Interpretation and Implications. *Computer Supported Collaborative Work*. Feb. 2010. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.495.520&rep=rep1&type=pdf>

²²¹ Westerman, D. et al. Social Media as Information Source: Recency of Updates and Credibility of Information. *Journal of Computer-Mediated Communication* 19 (2014) 171–183. <https://doi.org/10.1111/jcc4.12041>

²²² Zubiaga, A. et al. Analysing how people orient to and spread rumours in social media by looking at conversational threads. *PLoS ONE*. 11. <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0150989>

²²³ Simon, T. et al. Twitter in the Cross Fire—The Use of Social Media in the Westgate Mall Terror Attack in Kenya. *PLoS ONE* 9.8 (2014). <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0104136>

²²⁴ Sargent, H. April 15, 2013: How Twitter informed us. April 15, 2014. <https://www.boston.com/news/local-news/2014/04/15/april-15-2013-how-twitter-informed-us>

²²⁵ Panagiotopoulos, P. et al. Social media in emergency management: Twitter as a tool for communicating risks to the public. <https://doi.org/10.1016/j.techfore.2016.06.010>

²²⁶ Chamberlain, P. Twitter as a vector for disinformation. *Journal of Information Warfare*. Vol. 9, Issue 1. 2010. <https://www.jinfowar.com/journal-issue/volume-9-issue-1> (Subscriber access only. Draft version available at <http://arkem.org/disinfo-twitter.pdf>)

²²⁷ Chamberlain, P. Twitter as a vector for disinformation. *Journal of Information Warfare*. Vol. 9, Issue 1. 2010. <https://www.jinfowar.com/journal-issue/volume-9-issue-1> (Subscriber access only. Draft version available at <http://arkem.org/disinfo-twitter.pdf>)

result, many users do not become suspicious of unreferenced assertions on Twitter, which makes it an excellent vector for disinformation.²²⁸

During rioting in London in August 2011, reports on Twitter said tanks were massing in the City of London, the London Eye Ferris wheel on the banks of the Thames was on fire, a tiger was roaming the city streets, and rioters had broken into a MacDonald's and started cooking their own food,²²⁹ The Guardian newspaper reported. Many of the fallacious tweets included pictures, and it was this "photographic evidence – even if doctored, or indeed even if of a completely separate event – [that] made believers of many tweeters," The Guardian said.

All of those tweets were shown to be false. Some false stories were shut down by other Twitter users, according to The Guardian story, which was based on data compiled for an analysis by The Guardian and the London School of Economics of the riots, and included a chapter on Twitter and other social media.²³⁰

People will turn to social media for updates during a crisis if the authorities do not post updates and other information about the emergency in a timely and regular manner, write Nickolas Jones et al in the Proceedings of the National Academy of Sciences (PNAS).²³¹ Individuals who have to rely on social media for updates are exposed to more conflicting information than those who obtain information from

²²⁸ Ibid.

²²⁹ Richards, J & Lewis, P. How Twitter was used to spread – and knock down – rumours during the riots. The Guardian. Dec. 7, 2011. <https://www.theguardian.com/uk/2011/dec/07/how-twitter-spread-rumours-riots>

²³⁰ Reading the riots. Investigating England's summer of disorder. The Guardian/London School of Economics. [http://eprints.lse.ac.uk/46297/1/Reading%20the%20riots\(published\).pdf](http://eprints.lse.ac.uk/46297/1/Reading%20the%20riots(published).pdf) pp. 30-33.

²³¹ Jones, NM et al. Distress and rumor exposure on social media during a campus lockdown. Proceedings of the National Academy of Sciences of the United States of America. 114(44). 11663–11668. <http://doi.org/10.1073/pnas.1708518114>

traditional media, and often feel “helpless and without situational control,” Jones et al say. This can “lead people to see patterns in the information obtained that are not present.” Jones et al suggest that officials should monitor social media channels during a crisis, and regularly release updates on those same social media during the event. This could reduce individuals’ exposure to mis- and disinformation and mitigate distress among people trying to follow and make sense of the crisis.²³²

Twitter makes spreading disinformation and rumors easy²³³ -- users simply hit “retweet” to pass on a message. During breaking news situations, when new information is released piecemeal – including by trusted authorities and journalists – and often starts off as unconfirmed information,²³⁴ more misinformation and disinformation are likely to be shared by users. Zubiaga et al found that unverified information that was shared on social media and later was shown to be false took seven times longer to leave the news stream (they call this “being resolved”) than unverified information that turned out to be true.²³⁵

A study by Kate Starbird et al. of how disinformation spread on Twitter after the 2013 Boston Marathon bombings found that retweets of false information outnumbered retweets of corrections by 44 to one.²³⁶ Starbird et al cited an

²³² Ibid.

²³³ Lewandowsky, S. et al. Misinformation and its correction: continued influence and successful debiasing. *Psychological Science in the Public Interest*. 13(3), 106-131. 2012.

<http://journals.sagepub.com/doi/10.1177/1529100612451018>

²³⁴ Zubiaga, A. et al. Analysing how people orient to and spread rumours in social media by looking at conversational threads. *PLoS ONE*. 11.

<http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0150989>

²³⁵ Zubiaga, A. et al. <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0150989>

²³⁶ Starbird, K. et al. Rumors, False Flags, and Digital Vigilantes: Misinformation on Twitter after the 2013 Boston Marathon Bombing. *iConference 2014*. PDF accessed online.

https://faculty.washington.edu/kstarbi/Starbird_iConference2014-final.pdf

@NBCNews tweet that said an eight-year-old spectator had been killed in the bombings. About 45 minutes later, someone else tweeted that the young victim was a girl and a competitor in the race:

@TylerJWalter (April 15, 2013 7:15pm): An eight year old girl who was doing an amazing thing running a marathon, was killed. I can't stand our world anymore

Four minutes later, another user added a fake picture and purposefully spread the false rumor:

@_Nathansnicely (April 15, 2013 7:21pm): The 8 year old girl that sadly died in the Boston bombings while running the marathon. Rest in peace beautiful x <http://t.co/mMOi6clz21>²³⁷

Starbird and her colleagues identified nearly 93,000 tweets that were related to the original erroneous tweet from @TylerJWalter. Of those tweets, nearly 91,000 were disinformation and just 2,046 were corrections.²³⁸ An analysis conducted by MIT researchers of 126,000 stories tweeted by 3 million users over more than 10 years found that false news narratives spread “farther, faster, deeper, and more broadly”²³⁹ than the real thing.

The study by Starbird et al. also illustrates the unmanageable volume of information that can circulate during a crisis. Filtering with the words “Boston,”

²³⁷ Ibid.

²³⁸ Ibid.

²³⁹ Vosoughi, S et al. The spread of true and false news online. *Science*. Mar 9, 2018. Vol. 359, Issue 6380, pp. 1146-1151. <http://science.sciencemag.org/content/359/6380/1146>

“bomb,” “explosion,” “marathon,” and “blast,” Starbird and her colleagues collected over a one-week period starting the day of the race (April 15, 2013) and ending on April 21 some 10.6 million tweets related to the marathon. More than half of those tweets were retweets.²⁴⁰

Blair Thompson, an associate professor of communications at Western Kentucky University whose research includes crisis communication at schools, looked at the role of social media in active shooter incidents in schools. Interviews with crisis team members from 21 school districts in three states found most schools don’t have a social media strategy to deploy during crises, and crisis teams said they felt unprepared to cope with the challenges presented by social media, particularly misinformation and disinformation generated by students and their parents.²⁴¹

Although many Twitter users question disinformation when they see it,²⁴² spreading misleading information during a crisis impedes sense-making and pushes some recipients of disinformation to take dangerous or harmful action. For instance, the armed man who showed up at a D.C. pizzeria a month after Donald Trump was elected president had been misled by disinformation that grew out of a hack of the Democratic National Committee’s emails, since blamed on Russia. The shooter

²⁴⁰ Ibid

²⁴¹ Thompson, B et al. Social Media and Active Shooter Events: A School Crisis Communication Challenge, *Qualitative Research Reports in Communication*, 18:1, 8-17. <https://doi.org/10.1080/17459435.2016.1247111>

²⁴² Mendoza, M. et al. Twitter Under Crisis: Can We Trust What We RT? In *Proc. of 1st Workshop on Social Media Analytics (SOMA 2010)*. http://snap.stanford.edu/soma2010/papers/soma2010_11.pdf

believed Hillary Clinton and her campaign manager, John Podesta, were running a child sex ring in the basement of the Comet Ping Pong pizzeria.

“As wait staff went table to table, whispering to customers to get out, [the gunman, Edgar Maddison] Welch maneuvered into the restaurant's kitchen. He shot open a lock and found cooking supplies. He whipped open another door and found an employee bringing in fresh pizza dough. Welch did not find any captive children – Comet Ping Pong does not even have a basement – but he did prove, if there were any lingering doubts after the election, that fake news has real consequences,” said an article published in Rolling Stone.²⁴³

The false story that drove Welch to travel to D.C. to “save” the children being held at the pizzeria had been circulating for weeks on social media, but got a boost and spread like wildfire when it was posted on Lt. Gen. Michael Flynn’s Twitter feed by his son. To my knowledge, there were few rebuttals to the tweet alleging that the pizzeria housed a child sex ring.

The improbability of Pizzagate makes many people ask: How did we get to the point where we swallow any piece of information that’s flung at us? Even the Pizzagate gunman’s friends, whom he tried to convince to go along on his trip from North Carolina to D.C. to liberate the children, reportedly told him it sounded as if they would be liberating “some oppressed pizza from the hands of an evil pizza joint.”²⁴⁴

²⁴³ Robb, A. Anatomy of a Fake News Scandal, Rolling Stone Magazine, Nov. 16, 2017. <https://www.rollingstone.com/politics/news/pizzagate-anatomy-of-a-fake-news-scandal-w511904>

²⁴⁴ Ibid, Rolling Stone.

Twitter has also been used to amplify disinformation. The Atlantic Council's Digital Forensic Research Lab (DFRLab) reported in March 2018 that a Twitter poll posted by a U.K. account, asking if British Prime Minister Theresa May's government had provided enough evidence to convince other Twitter users that Russia was behind the poisoning of former Russian spy Sergei Skripal and his daughter, had been amplified by numerous Russian-language accounts, many of which were pro-Kremlin trolls.²⁴⁵ DFRLab "tracks global disinformation campaigns, fake news stories, covert military developments, and subversive attempts against democracy while teaching the public skills to identify and expose attempts to pollute the information space."²⁴⁶ More than three-quarters of respondents to the Twitter poll voted no – in other words, they thought the British government had not produced enough evidence to point the finger at Russia for the nerve agent attack. DFRLab found that many of the 2,808 retweets of the poll were from Russian-language Twitter accounts or accounts that "systematically post pro-Kremlin content."²⁴⁷ Researchers at DFRLab concluded that pro-Kremlin Twitter users had influenced the outcome of the poll "to create the appearance of greater hostility towards the UK government than UK users themselves showed."²⁴⁸

5.2 Algorithms replace gatekeepers

²⁴⁵ Unsigned. #PutinAtWar: Trolls on Twitter. Battles on Twitter over the Skripal poisoning case. Digital Forensic Research Lab on Medium. March 24, 2018. <https://medium.com/dfrlab/putinatwar-trolls-on-twitter-5d0bb3dc30ae>

²⁴⁶ About page. DFRLab website. <https://www.digitalsherlocks.org/dfrlab>

²⁴⁷ Unsigned. #PutinAtWar: Trolls on Twitter. Battles on Twitter over the Skripal poisoning case. Digital Forensic Research Lab on Medium. March 24, 2018. <https://medium.com/dfrlab/putinatwar-trolls-on-twitter-5d0bb3dc30ae>

²⁴⁸ Ibid.

The role of the gatekeeper – individuals or organizations who filter information to determine if it’s true and decide how widely it should be shared – has been transferred in the social media world to anyone who produces a piece of information and has an internet connection.²⁴⁹ Anyone who posts, retweets or otherwise shares information on their blog, Facebook page or Twitter feed assumes the role once held by the likes of Walter Cronkite – once voted the most trusted man in America (although he was the only journalist among the choices and the rest were politicians) – and professional editors in traditional media. While some users of social media check what they publish before they post it, others “repost or retweet information without reading it carefully, much less doing any due diligence for accuracy,”²⁵⁰ writes Aly Colon, Knight professor for journalism and ethics at Washington and Lee University.

Left-wing political and internet activist Eli Pariser argued in a 2011 TED talk²⁵¹ that traditional gatekeepers have been replaced not by other humans but by computer algorithms. These algorithms are designed to show individuals stories that are in line with their interests, political leanings and personal beliefs. They push consumers into “filter bubbles,” defined by Wikipedia as “the intellectual isolation that can occur when websites make use of algorithms to selectively assume the information a user would want to see and then gives the user that information.”²⁵²

²⁴⁹ Allcott, H & Gentzkow, M. Social Media and Fake News in the 2016 Election. National Bureau of Economic Research working paper No. 23089. Jan 2017, revised April 2017.

<http://www.nber.org/papers/w23089>

²⁵⁰ Colon, A. You are the keeper of the news. The Conversation. Feb. 7, 2017.

<https://theconversation.com/you-are-the-new-gatekeeper-of-the-news-71862>

²⁵¹ https://www.ted.com/talks/eli_pariser_beware_online_filter_bubbles/transcript#t-356702

²⁵² https://en.wikipedia.org/wiki/Filter_bubble#cite_note-twsTechnoP101-1

“So if algorithms are going to curate the world for us, if they're going to decide what we get to see and what we don't get to see, then we need to make sure that they're not just keyed to relevance. We need to make sure that they also show us things that are uncomfortable or challenging or important ... other points of view,”²⁵³ Pariser said in his TED talk. Since he gave the talk, people have even less incentive to burst out of their filter bubble to see information that disagrees with their world view. In 2018, there were 300 moderately to highly partisan websites on the left and 200 on the right,^{254 255} according to Media Bias/Fact Check, a website dedicated to educating the public on media bias and deceptive news practices.^{256 257} Having a greater number of partisan media outlets available makes it easier for information consumers to read, hear or see only news that agrees with their world view. There are, however, plenty of opportunities to read unbiased reports: according to Media Bias/Fact Check, over 300 media outlets in the U.S. fall under the “least biased” category.²⁵⁸

Benkler et al²⁵⁹ say that liberal consumers of media are more likely to read both traditional media – the New York Times, Washington Post, CNN, etc. – and more partisan left-wing media such as the Huffington Post, MSNBC or the Daily

²⁵³ Ibid.

²⁵⁴ From Media Bias / Fact Check website. www.mediabiasfactcheck.com/left and www.mediabiasfactcheck.com/right

²⁵⁵ As a small aside, Cambridge Analytica is listed among the sources with a right-wing media bias.

²⁵⁶ <https://mediabiasfactcheck.com/about>

²⁵⁷ Media Bias / Fact Check looks not just at political bias but also at how factual the information published by a medium is and if links are provided to credible, verifiable sources. Like movie-ratings site, Media Bias/Fact Check also invites readers to vote on what they see as the publication's or media outlet's bias.

²⁵⁸ <https://mediabiasfactcheck.com/center/>

²⁵⁹ Benkler, Y et al. Study: Breitbart-led right-wing media ecosystem altered broader media agenda. Columbia Journalism Review. March 3, 2017. <https://www.cjr.org/analysis/breitbart-media-trump-harvard-study.php>

Beast, while conservative consumers of media, whom the researchers call Trump followers, are “Breitbart-centered” and less connected to conservative mainstream media, such as the Wall Street Journal. A study by Benkler et al of more than 1.25 million stories published online between April 1, 2015 and Election Day 2016 found that “a right-wing media network anchored around Breitbart developed as a distinct and insulated media system, using social media as a backbone to transmit a hyper-partisan perspective to the world.” But the researchers disagree with Pariser that this is likely to have been caused by algorithms. “If technology were the most important driver towards a ‘post-truth’²⁶⁰ world, we would expect to see symmetric patterns on the left and the right. Instead, different internal political dynamics in the right and the left led to different patterns in the reception and use of the technology by each wing,” they write in the Columbia Journalism Review.²⁶¹

A small study conducted by The Guardian newspaper showed how reluctant people are to leave their filter bubbles. In the run-up to the 2016 elections in the U.S., The Guardian asked a group of Americans on the right to read left-wing news sites and people on the left to read right-wing ones. Not all of the 10 participants in the study saw it through to the end. One participant said after visiting a right-wing website, “You might as well have been waterboarding a brother.”²⁶²

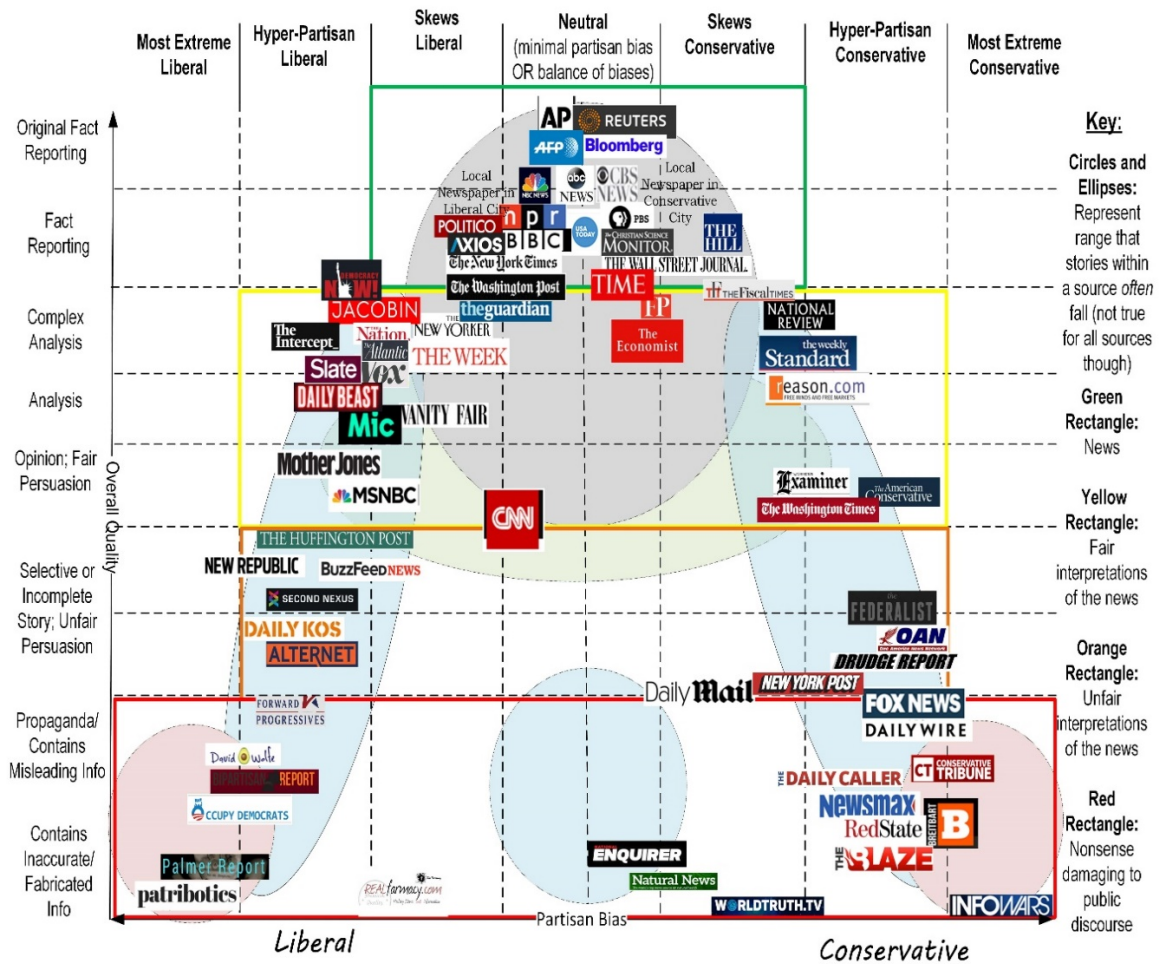
²⁶⁰ “Relating to or denoting circumstances in which objective facts are less influential in shaping public opinion than appeals to emotion and personal belief. Oxford Dictionaries.

²⁶¹ Benkler, Y et al. Study: Breitbart-led right-wing media ecosystem altered broader media agenda. Columbia Journalism Review. March 3, 2017. <https://www.cjr.org/analysis/breitbart-media-trump-harvard-study.php>

²⁶² Wong, JC et al. Bursting the Facebook bubble: We asked voters on the left and right to swap feeds. The Guardian. Nov. 16, 2016. <https://www.theguardian.com/us-news/2016/nov/16/facebook-bias-bubble-us-election-conservative-liberal-news-feed>

Denver-based lawyer Vanessa Otero has produced a helpful chart that assesses publications and media outlets on the basis of ideology – shown on the chart’s X-axis – and quality/credibility, depicted on the Y-axis. In addition to the Media Bias/Fact Check website, Otero’s Media Bias chart can be used to determine the political leanings of different media and how reliable their reporting is.

Vanessa Otero’s Media Bias Chart



Copyright © 2018 mediabiaschart.com
 Licensed for educational use

Denver-based lawyer Vanessa Otero plotted different media on a graph according to the quality of their reporting and credibility of their sources (x-axis), and their political leanings (y-axis) to produce the Media Bias Chart.

Reproduced with permission of Vanessa Otero

Chapter 6

Tools to detect disinformation

The proliferation of disinformation has led researchers to develop tools to help detect it.²⁶³ ²⁶⁴ These tools are intended to address the imbalance between the high volume of fake news and the number of humans working to debunk it. Some of the tools attempt to verify already published articles while others check breaking news or tweets, such as those sent out during a crisis via Twitter. Most are still in the beta stage of development. Many of the tools were developed for journalists.

The following tools developed in the United States focus on detecting disinformation, rather than correcting and stopping it.

6.1. TwitterTrails

TwitterTrails is a tool that allows members of the media to track the trustworthiness of stories shared on Twitter.²⁶⁵ By analyzing the origin and spread of a story on Twitter, TwitterTrails provides information that “a critically thinking person can use to examine how a Twitter audience reacts to the spreading of the story.”²⁶⁶

²⁶³ Vosoughi, Soroush. Automatic detection and verification of rumors on Twitter. Ph.D. dissertation. MIT. June 2015

²⁶⁴ Metaxas, P.T., Finn, S., & Mustafaraj, E. (2015) DEMO: Using TwitterTrails.com to Investigate Rumor Propagation. WebScience15, Oxford, UK.

<https://repository.wellesley.edu/cgi/viewcontent.cgi?article=1166&context=scholarship>

²⁶⁵ <http://twittertrails.com/>

²⁶⁶ Metaxas, P.T., Finn, S., & Mustafaraj, E. (2015) DEMO: Using TwitterTrails.com to Investigate Rumor Propagation. WebScience15, Oxford, UK.

<https://repository.wellesley.edu/cgi/viewcontent.cgi?article=1166&context=scholarship>

Using the URL of a tweet and keywords, TwitterTrails generates a propagation graph of a tweet's life, showing who broke the story on Twitter and who picked it up. It identifies when the story broke, how many times it was retweeted, which retweets came from verified accounts, how many followers each tweeter of the story has. The Wellesley College creators of TwitterTrails say "these are key elements in gauging the visibility of the tweet, as well as the degree of credibility other users will assign to the tweet and the amount of trust in the user as a source of information."²⁶⁷ Through visualizations, the tool also shows who was mainly responsible for sharing the information.

TwitterTrails uses an algorithm to evaluate two metrics in stories: how much they have been passed around and how much the public doubts the validity of the story. Higher skepticism and lower spread mean the story is more likely to be false. The other way around means it's more likely to be true. Like the other tools described below, TwitterTrails can help journalists to judge if a story is true or false but can't make them believe – at least not immediately – that they should stop reporting the story because it is untrue. In one example used to show how TwitterTrails works – the story of a rumored plane crash off the Canary Islands, which turned out to be false – the developers note that even after "the 9-1-1 type service @112canarias" announced that the crash was a false alarm, "news organizations kept propagating the false news for more than an hour."²⁶⁸

²⁶⁷ Metaxas, P.T., Finn, S., & Mustafaraj, E. (2015) DEMO: Using TwitterTrails.com to Investigate Rumor Propagation. WebScience15, Oxford, UK.

<https://repository.wellesley.edu/cgi/viewcontent.cgi?article=1166&context=scholarship>

²⁶⁸ Ibid

TwitterTrails requires only a few minutes to an hour to generate a report, depending on how many tweets there are to retrieve. It pulls Tweets for analysis from Twitter's search API, and can only analyze recent tweets, but the developers have requested access from Twitter for older events of significance.

6.2 Hoaxy²⁶⁹

Hoaxy® is open-source software that visualizes the spread of claims and related fact checking online. A claim may be a fake news article, hoax, rumor, conspiracy theory, satire, or even an accurate report. Hoaxy® allows users to see how a claim spreads on Twitter, plots the cumulative number of Twitter shares over time, and shows how claims spread from person to person.

Hoaxy® is fully automated (no human input involved beyond developing the software). It does not claim to decide whether a story is true or false but visualizes the spread of claims and fact checking.

The creators of Hoaxy® intend for it to be used to observe how unverified stories and the fact-checking of those stories spread on social media. It's up to users to gather evidence about a claim and refute it, if they deem that necessary.

Developed by researchers at Indiana University, Hoaxy only gathers data from public messages sent over Twitter (which would include most tweets sent in a crisis), and appears to be intended to track already published stories. When I input USAFA in Hoaxy's search bar, most of the stories returned were from websites with right-wing

²⁶⁹ <https://hoaxy.iuni.iu.edu/faq.html>

biases. Two others were on conspiracy-pseudoscience sites; one site had an extreme liberal bias, publishing fake news or overt propaganda with poor or no sourcing; and one published stories ranging from “moderately credible to tin-foil hat conspiracies,” provided by individuals without credentials. The list, copied below, of stories generated only includes those that were tweeted or retweeted, according to Hoaxy.

The screenshot shows the Hoaxy website interface. At the top, there is a navigation bar with the Hoaxy logo and links for Help, Stats, Data, Source, and OSoMe. Below the navigation bar, there is a search bar with the text "Select claims and fact checking articles to visualize." and a "Visualize" button. To the right of the search bar are "Tweet" and "Share" buttons. Below the search bar, there is a list of 10 articles, each with a checkbox, a title, a source, a date, and the number of tweets and Facebook shares.

Article Title	Source	Date	Tweets	Facebook Shares
US Air Force Academy Finds the Guy Who Wrote Racist Graffiti On Dorm Walls	redstate.com	Nov 8, 2017	93	71
Lesbian Colonel with "Wife" & Kids to Take Charge of USAF Academy	dcclthesline.com	Apr 2, 2017	42	405
HATE HOAX: Black Air Force Academy Cadet Wrote Racial Slurs On His Own Dorm Room	infowars.com	Nov 9, 2017	25	63
War on All Fronts - Be Ready Folks Politics	beforeitsnews.com	Sep 28, 2016	25	45
US Air Force general delivers powerful address on racist treatment of black cadets: 'Get out!'	theblaze.com	Sep 29, 2017	20	245
Lt. General Issues Statement After Black Air Force Cadette Is Called A N*gger (VIDEO)	bipartisanreport.com	Sep 28, 2017	17	3644
Another hate crime reported, another fake story exposed	wnd.com	Jan 8, 2017	13	59
The United States Air Force Academy Doesn't Train Warriors Anymore Global Unrest	beforeitsnews.com	Jan 22, 2017	2	0
HATE HOAX: Black Air Force Academy Cadet Wrote Racial Slurs On His Own Dorm Room	prisonplanet.com	Nov 9, 2017	1	7

Results of a search for “USAFA” on Hoaxy, conducted March 31, 2018.

6.3. Hearsift and Rumor Gauge²⁷⁰

Researchers at the MIT Media Lab have developed a two-part system designed to detect rumors (Hearsift) and verify them (Rumor Gauge).

Hearsift works by inputting raw tweets about an event and filtering them to exclude any tweets that don’t make an assertion – defined by the MIT researchers as

²⁷⁰ Vosoughi, S. Ph.D. dissertation. MIT

an utterance that commits the speaker to the truth of the expressed proposition. The filtered tweets are then clustered into groups that make similar assertions about an event. Tweets using other forms of speech including recommendations, expressions, questions, requests and what the MIT researchers call “miscellaneous” are not included in the clusters.

Speech Act	Example Tweet
Assertion	authorities say that the 2 boston bomb suspects are brothers are legal permanent residents of chechen origin - @nbcnews
Recommendation	If you follow this man for updates and his opinions on #Ferguson I recommend you unfollow him immediately.
Expression	Mila Kunis and Ashton Kutcher are so adorable
Question	Anybody hear if @gehrig38 is well enough to attend tonight? #red-sox
Request	rt @craigyh999: 3 days until i run the london marathon in aid of the childrens hopsice @sschospices . please please sponsor me here
Miscellaneous	We'll continue to post information from #Ferguson throughout the day on our live-blog

Table 3.1: Example tweets for each speech act type.

Table of speech acts. MIT Ph.D. dissertation by Soroush Vosoughi. P. 39

Clusters of assertion-containing tweets are called rumors. These are input into the Rumor Gauge algorithm, which produces a “veracity curve” for each rumor to indicate the likelihood that it is true over time.

The aim set by the MIT team was to be able “to predict the veracity of rumors about real-world emergencies faster than any other public source.”²⁷¹ The developers admit that automatic detection and verification of rumors in Twitter are very difficult tasks, because they require “extracting weak signals from very noisy environments.” They say their system is only 75 percent accurate, and that it still requires human/manual input of data. But the aim of automated programs like Rumor Gauge is not to replace humans but to make the job of sorting disinformation from information easier by markedly “reducing the bandwidth” of information that someone trying to make sense of a crisis situation in real-time has to sift through to sort what’s true from what’s false.

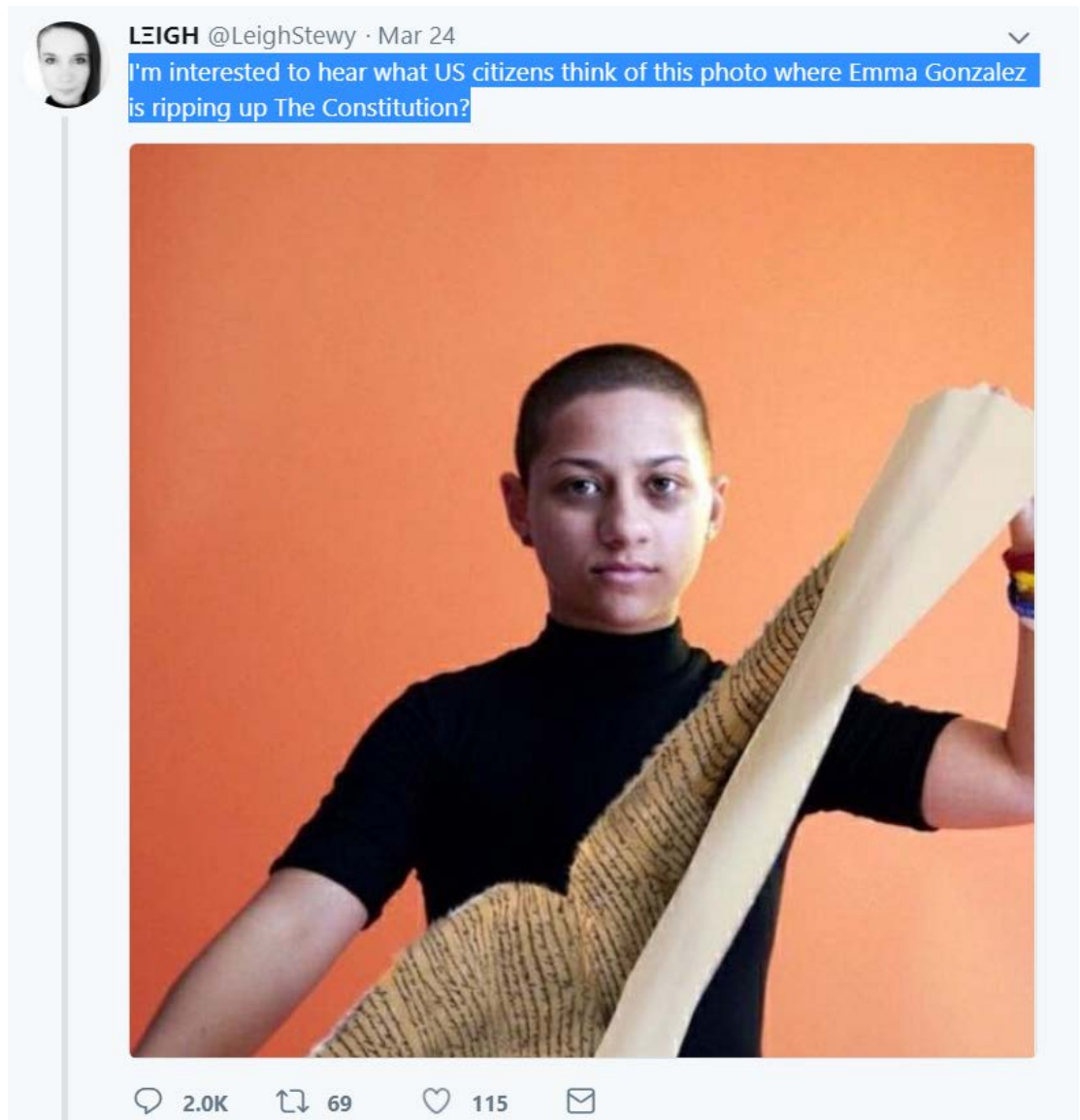
6.4. Spotting photo fakery

Emma Gonzalez became one of the main public voices of students who survived the shootings at Marjory Stoneman Douglas High School in Parkland, Florida. Not long after the shootings, Gonzalez and several other students took part in a photo shoot for Teen Vogue. One of the pictures that came out of the shoot showed Gonzalez tearing up a paper shooting target.

A far-right group Photoshopped the picture to show Gonzalez tearing up the U.S. Constitution. In addition to swapping the paper target for the Constitution, the Photoshoppers also made Gonzalez paler and put dark circles under her eyes. The doctored image was posted on Twitter by user @LeighStewy with the message: “I’m

²⁷¹ Vosoughi, S et al. Rumor Gauge : Predicting the veracity of rumors on Twitter. ACM Transactions on Knowledge Discovery from Data, Vol. 11, No. 4, Article 50, July 2017. Accessed online at

interested to hear what US citizens think of this photo where Emma Gonzalez is ripping up The Constitution?”



Doctored photo of Emma Gonzalez posted on Twitter by @LeighStewy.

Many of the 2,000-plus replies to the post pointed out that the picture was a fake. The poster of the original Tweet replied to one of them, “Hi Chris? I'm not sure

where I claimed it was real in the first place?”²⁷² Denial is often used by purveyors of disinformation when they are called out for their actions.

Like text and announcements read out in the Roman forum, photography lost its innocence many years ago.²⁷³ The first instance of photographic fakery is thought to date from around 1860, when Abraham Lincoln’s head was cut and pasted onto the body of Southern politician John Calhoun.²⁷⁴ Stalin, Adolf Hitler and present-day autocrat Kim Jong-un airbrush or Photoshop people they’ve fallen out with or had killed²⁷⁵ from pictures.²⁷⁶

Today, it has not only become easier to doctor images than it was in the 1860s but there are also several tools for detecting doctored photos, or for finding the origin of a photo. Photographers and photo editors can often spot a doctored image by checking whether the light source on all subjects is from the same direction or by using their knowledge of other photos. The Ukrainian website StopFake.org, which was set up in 2014, the year Russia annexed Crimea, to verify and refute disinformation and propaganda about events in Ukraine being circulated in the media,

²⁷² Leigh’s use of question marks, not mine.

²⁷³ Farid, H. Digital doctoring: how to tell the real from the fake. Royal Statistical Society. Nov. 2006. <https://rss.onlinelibrary.wiley.com/doi/full/10.1111/j.1740-9713.2006.00197.x>

²⁷⁴ Farid, *ibid.*

²⁷⁵ Leon Trotsky was removed from official photos in the USSR after he fell out with Joseph Stalin in the 1920s. He was deported from the USSR in 1928 and assassinated in 1940 in Mexico, where he was granted asylum. Joseph Goebbels was removed from a 1937 photo with Adolf Hitler and filmmaker Leni Riefenstahl, for reasons that are still unclear. Kim Jong-un’s uncle Jang Song-thaek was removed from a documentary of his nephew’s military inspection trips, shortly after Kim had had him executed.

²⁷⁶ Stanger, M. 6 people who were literally erased from history. Business Insider. Feb 4, 2014. <http://www.businessinsider.com/people-who-were-erased-from-history-2013-12>

has put together a list of tools to detect doctored photos.²⁷⁷ Fotoforensics.com uses algorithms to help researchers determine if a picture is real or computer graphics, if it's been modified and even how it's been modified.²⁷⁸ Reverse image searches (<https://reverse.photos/>, <https://www.tineye.com/>, inter alia) can be used to check for the origins of a picture. The picture below, for instance, was used on an anti-DACA meme circulating on Twitter in January, saying we should feed hungry American-born children before looking after children brought to the U.S. illegally by their parents.

Here's a dreamer. He dreams about food and he was born in America. We have 14 million kids living in poverty but disgusting Dems are fighting for 800k illegals.



Photo of boy in Alaska used in anti-DACA meme on Twitter.

A reverse image search showed that the picture was taken by Natalie Fobes on the Yukon River in Alaska. It was licensed by Getty Images for editorial use. Other

²⁷⁷ 13 online tools that help to verify the authenticity of a photo. StopFake.org. July 29, 2014. <https://www.stopfake.org/en/13-online-tools-that-help-to-verify-the-authenticity-of-a-photo/>

²⁷⁸ <http://fotoforensics.com/faq.php?show=General&c=guidelines>

pictures in the series taken by Fobes indicate that the child is probably not hungry. The original image can be viewed [here](#).

6.5. Fake videos

Technological advances allow fake videos to be convincingly real. For example, Professor Matthias Niessner's Visual Computing Group at the Technical University Munich²⁷⁹ has developed a program that allows the facial expressions of someone in a video to be altered in real-time using a webcam that shoots images of a source actor and transfers the actor's facial movements to the target person, with very realistic results. The potential for hacking livestreams and other broadcasts is frightening.

Even without software, videos can be faked. One way to do this is to post a video or part of a video out of context. This happened during the USAFA shooter alert when a cadet shot a video through the spyhole of his dorm room door, showing security forces conducting a sweep of the building. While some posted the video in its correct context, other Twitter users said the [brief cellphone video](#) showed the alleged shooters dressed as security forces. Those videos have been taken down.

Completely fabricated videos were used to attack Broward County Sheriff Scott Israel after the mass shooting at Marjory Stoneman Douglas High School in Parkland, Florida, on Feb. 14, 2018. The videos show a young woman saying the sheriff had impregnated her when she was 17 and then told her to have an abortion.²⁸⁰ The videos were first posted in 2012, when Israel successfully ran for sheriff against

²⁷⁹ <http://niessnerlab.org/projects/thies2016face.html>

²⁸⁰ The videos are still available on YouTube. I've chosen not to include a URL for it.

the Republican incumbent, Al Lamberti. They were left on YouTube over the years, in spite of Israel's request that they be taken down because they were fake and a rebuttal video by Israel's wife. After the mass shooting at Marjory Stoneman Douglas High School, the videos resurfaced and very rapidly had more than 130,000 views – until Politico reporter Marc Caputo found the woman in the videos, who told him they were made when she was 17 for \$25 each, and were completely fake.²⁸¹ Caputo's investigative work succeeded in pressuring Google, which owns YouTube, to take down the videos, although the Politico report says they initially refused.

Although the original videos are no longer accessible on YouTube, I was able to find one of the series posted on an individual's YouTube page, with opening credits that read "Fair Use Act 1976." Furthermore, several right-wing sites have not issued corrections to indicate that the videos are fabrications. This illustrates how disinformation can endure online, particularly in the universe of social media, even after it's been conclusively proven to be false.

It's also worth noting that these false videos were detected through "manual" investigations by a journalist. Caputo found the woman in the videos by "combing internet videos and social media."²⁸²

²⁸¹ Caputo, Marc. 'Oh my God...It's fake': Far right falls for hoax about Broward County sheriff. Politico. March 23, 2018. <https://www.politico.com/story/2018/03/23/florida-school-shooting-sheriff-hoax-482170>

²⁸² Caputo.

Chapter 7

Conclusion

I set out to test the hypothesis that refuting rumors and disinformation early in a crisis can help to stop disinformation. This was driven by a personal experience – an active shooter alert at my son’s school during which updates from the authorities were few and far between while disinformation ran riot.

For this research I looked at several cases, including the disinformation campaigns in Estonia, Georgia and Germany. My aim in analyzing disinformation events in these countries was to see how quickly officials responded to disinformation attacks, and whether their responses were effective in shutting down the attacks. My conclusions are that their responses were quick, and they did prevent disinformation from achieving its presumed goals, usually to promote the agenda of the suspected or known perpetrator. But even rapid intervention and refutation did not stop disinformation completely.

When disinformation targets a country, it seeks to “impact national decision-making.”²⁸³ In Estonia, its aim was to convince the Estonians to reverse a decision to relocate a Soviet-era statue of a Red Army soldier to a military cemetery. It wanted to convince Georgians to not pursue NATO membership. In Germany, it sought to sow anti-immigrant sentiments with an article about New Year’s Eve celebrations in Dortmund. Quick and decisive responses from all three targets helped to stem the

²⁸³ Kinstler, L. How to survive a Russian hack. The Atlantic. Feb. 2, 2017.
<https://www.theatlantic.com/international/archive/2017/02/russia-disinformation-baltics/515301/>

flow of disinformation and shift support away from the purveyors of disinformation to the local authorities' version of events. More research is needed, however, to be able to attribute the success of these responses to their timeliness.

While many institutions have plans in place for physically managing a crisis – ordering a lockdown, evacuating a building, etc – few appear to have a strategy for dealing with disinformation, usually on social media, during a crisis. Few researchers have looked at the disinformation/social media response of organizations, but those who have, such as Thompson et al,²⁸⁴ say school districts are “typically underprepared” for mounting a social media response during a crisis situation. “Communication elements such as how to monitor and address social media traffic are often not considered until after the shooting, resulting in a lack of preparedness. Yet, scant research exists which could directly inform the management of these events (e.g., managing emotions of those directly involved, identifying impacted students, continuing student learning) with both internal (administration, teachers, student body) and external (media, parents, law enforcement, cooperation between

²⁸⁴ Thompson et al cite: Barker, G. G., & Yoder, M. E. The Virginia tech shootings: Implications for crisis communication in the educational setting. *Journal of School Public Relations*, 32, 78–101.2012. http://scholar.google.com/scholar_lookup?hl=en&publication_year=2012&pages=78-101&author=G.+G.+Barker&author=M.+E.+Yoder&title=The+Virginia+tech+shootings%3A+Implications+for+crisis+communication+in+the+educational+setting& ; Carr, N. Times have changed. *American School Board Journal*, 196(4), 57–58. 2009. http://scholar.google.com/scholar_lookup?hl=en&publication_year=2009&pages=57-58&issue=4&author=N.+Carr&title=Times+have+changed& ; Mazer, J. et al. Communication in the face of a school crisis: Examining the volume and content of social media mentions during active shooter incidents. *Computers in Human Behavior*, 53, 238–248. 2015. <https://www.tandfonline.com/servelet/linkout?suffix=CIT0013&dbid=128&doi=10.1080%2F17459435.2016.1247111&key=000361921100024>

agencies, hospitals, counseling services, coordination of the command center) constituents.”²⁸⁵

This lack of social media preparedness on the part of organizations occurs in spite of the fact that more and more people turn to social media for information during a crisis. An analysis by Jones et al of five hours of Twitter data gathered during a school shooting found that when there were gaps in official communication during the crisis, more rumors were shared via Twitter and negative emotions among students rose.²⁸⁶ The advice they have for officials involved in a crisis is to monitor social media channels to mitigate the negative impact of rumors during collective traumas by posting official updates early and often on the same social media channel as the disinformation is circulating.

This is key to dispelling disinformation. Lewandowsky et al note that once disinformation is received by the public, it becomes very difficult to correct or retract.²⁸⁷ False information that lingers can have a severe impact on society, they say. For example, after unsubstantiated claims of a vaccination-autism link, many parents decided not to immunize their children, which led to an increase in preventable diseases “and hence preventable hospitalizations, deaths, and the unnecessary

²⁸⁵ Blair Thompson, Joseph P. Mazer, Holly J. Payne, Angela M. Jerome, E. Gail Kirby & William Pfohl (2016) Social Media and Active Shooter Events: A School Crisis Communication Challenge, *Qualitative Research Reports in Communication*, 18:1, 8-17. 2016. <https://doi.org/10.1080/17459435.2016.1247111>

²⁸⁶ Jones et al.

²⁸⁷ Lewandowsky, S et al. Misinformation and its correction: continued influence and successful debiasing. *Psychological Science in the Public Interest*. 13(3) 106-131. 2012. Accessed online <http://journals.sagepub.com/doi/pdf/10.1177/1529100612451018>

expenditure of large amounts of money for follow-up research and public-information campaigns aimed at rectifying the situation,” the researchers say.²⁸⁸

Most of the applications being developed today aim to detect disinformation (Chapter 6), but my observations are that few organizations have policies on what officials should do once the disinformation alarms have sounded. Few have a rapid response strategy that is poised and ready to be deployed when disinformation spreads on social media – which is where most disinformation circulates. These issues need to be studied in greater detail, and organizations should be helped to devise strategies to fight disinformation if it is to be nipped in the bud.

²⁸⁸ Ibid.

Bibliography

- Agence France-Presse. German police quash Breitbart story of mob setting fire to Dortmund church. Published in The Guardian. Jan. 7, 2017.
<https://www.theguardian.com/world/2017/jan/07/german-police-quash-breitbart-story-of-mob-setting-fire-to-dortmund-church>
- Allcott, H & Gentzkow, M. Social Media and Fake News in the 2016 Election. National Bureau of Economic Research working paper No. 23089. Jan 2017, revised April 2017. <http://www.nber.org/papers/w23089>
- The Baltic Times. Unsigned. "Estonia - Crossroads of History" expresses Russia's position. Dec. 15, 2016. <https://www.baltictimes.com/news/articles/17039/>
- Barker, G. G., & Yoder, M. E. The Virginia tech shootings: Implications for crisis communication in the educational setting. *Journal of School Public Relations*, 32, 78–101.2012.
http://scholar.google.com/scholar_lookup?hl=en&publication_year=2012&pages=78-101&author=G.+G.+Barker&author=M.+E.+Yoder&title=The+Virginia+tech+shootings%3A+Implications+for+crisis+communication+in+the+educational+setting&
- Benkler, Y et al. Study: Breitbart-led right-wing media ecosystem altered broader media agenda. *Columbia Journalism Review*. March 3, 2017. Accessed online at <https://www.cjr.org/analysis/breitbart-media-trump-harvard-study.php>
- Bernstein, J. Making a Crisis Worse, *Journal of Promotion Management*, 12:3-4, 39-45, 2006. https://doi.org/10.1300/J057v12n03_04
- Beaumont, C. New York plane crash: Twitter breaks the news, again. *The Telegraph* (UK). Jan. 16, 2009.
<https://www.telegraph.co.uk/technology/twitter/4269765/New-York-plane-crash-Twitter-breaks-the-news-again.html>
- Blank, S. Web War I: Is Europe's First Information War a New Kind of War?, *Comparative Strategy*, 27:3, 227-247. July 2008
<https://www.tandfonline.com/doi/pdf/10.1080/01495930802185312>
- Blomfield, A. War of words over Bronze Soldier. *The Telegraph*. Feb. 5, 2007.
<https://www.telegraph.co.uk/news/worldnews/1541641/War-of-words-over-bronze-soldier.html>

- Bogart LM & Thorburn, S. Are HIV/AIDS Conspiracy Beliefs a Barrier to HIV Prevention Among African Americans? Journal of Acquired Immune Deficiency Syndromes (JAIDS). 38(2):213-218, Feb. 2005.
https://journals.lww.com/jaids/Fulltext/2005/02010/Are_HIV_AIDS_Conspiracy_Beliefs_a_Barrier_to_HIV.14.aspx
- Bogart LM, et al. Endorsement of a genocidal HIV conspiracy as a barrier to HIV testing in South Africa. Journal of Acquired Immune Deficiency Syndromes (JAIDS) 2008 Sep 1;49(1):115-6.
https://journals.lww.com/jaids/fulltext/2008/09010/Endorsement_of_a_Genocidal_HIV_Conspiracy_as_a.20.aspx
- Bogart LM, Wagner G, Galvan FH, Banks D. Conspiracy Beliefs about HIV Are Related to Antiretroviral Treatment Nonadherence among African American Men with HIV. Journal of acquired immune deficiency syndromes (1999). 2010;53(5):648-655. <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2845717/>
- Boghart, T. Operation INFEKTION: Soviet Bloc Intelligence and Its AIDS Disinformation Campaign. Studies in Intelligence Vol. 53, No. 4. December 2009.
<https://www.cia.gov/library/center-for-the-study-of-intelligence/csi-publications/csi-studies/studies/vol53no4/soviet-bloc-intelligence-and-its-aids.html>
- Breitbart rebuttal by London Editor in Chief Raheem Kassam.
http://www.breitbart.com/london/2017/01/08/fake-news-fake-news-media-sow-division-with-dishonest-attack-on-breitbart-allahu-akbar-church-fire-story/#disqus_thread
- Broersma, T. & Graham, Twitter as a news source. Journalism Practice, 7:4, 446-464. May 30, 2013. <https://doi.org/10.1080/17512786.2013.802481>
- Caputo, Marc. ‘Oh my God...It’s fake’: Far right falls for hoax about Broward County sheriff. Politico. March 23, 2018.
<https://www.politico.com/story/2018/03/23/florida-school-shooting-sheriff-hoax-482170>
- Carr, N. Times have changed. American School Board Journal, 196(4), 57–58. 2009.
http://scholar.google.com/scholar_lookup?hl=en&publication_year=2009&pages=57-58&issue=4&author=N.+Carr&title=Times+have+changed&
- Chamberlain, P. Twitter as a vector for disinformation. Journal of Information Warfare. Vol. 9, Issue 1. 2010. <https://www.jinfowar.com/journal-issue/volume-9->

[issue-1](#) (Subscriber access only. Draft version available at <http://arkem.org/disinfo-twitter.pdf>)

- Chivvis, C. Understanding Russian hybrid warfare and what can be done about it. Testimony presented before the House Armed Services Committee on March 22, 2017. Accessed online at https://www.rand.org/content/dam/rand/pubs/testimonies/CT400/CT468/RAND_CT468.pdf
- Cohen, A & Hamilton, R. The Russian Military and the Georgian War: Lessons and Implications. Strategic Studies Institute. Sept. 1, 2011. p. 44. <http://ssi.armywarcollege.edu/pdffiles/PUB1069.pdf>
- Colon, A. You are the keeper of the news. The Conversation. Feb. 7, 2017. <https://theconversation.com/you-are-the-new-gatekeeper-of-the-news-71862>
- Cutler, R. Russia's disinformation campaign over South Ossetia. CACI Analyst. August 20, 2008. <https://www.cacianalyst.org/publications/analytical-articles/item/11678-analytical-articles-caci-analyst-2008-8-20-art-11678.html>
- Davis, J. Hackers Take Down the Most Wired Country in Europe. Wired Online. Aug. 21, 2007. <https://www.wired.com/2007/08/ff-estonia/>
- Digital Forensic Research Lab. Unsigned. #PutinAtWar: Trolls on Twitter. Battles on Twitter over the Skripal poisoning case. March 24, 2018. <https://medium.com/dfrlab/putinatwar-trolls-on-twitter-5d0bb3dc30ae>
- Digital TV Europe. Unsigned. Estonian pubcaster expanding reach of Russian-language channel. May 1, 2018. <https://www.digitaltveurope.com/2018/05/01/estonian-pubcaster-expanding-reach-of-russian-language-channel/>
- Dorell, O. Breitbart, other 'alt-right' websites are the darlings of the Russian propaganda effort. USA Today. Aug. 24, 2017. <https://www.usatoday.com/story/news/world/2017/08/24/breitbart-other-alt-right-websites-darlings-russian-propaganda-effort/598258001/>
- Dougherty, J. The mouse clicks of August: Hybrid warfare, nation-state actors and the future of cybersecurity. Small Wars Journal. Accessed online at <http://smallwarsjournal.com/jrn/art/mouse-clicks-august-hybrid-warfare-nation-state-actors-and-future-cybersecurity>

- Englishrussia.com. Unsigned. The Bronze Soldier. <http://englishrussia.com/2007/04/27/the-bronze-soldier/>
- Eristavi, M. Why the US Keeps Losing the Fight against Disinformation. Atlantic Council website. <http://www.atlanticcouncil.org/blogs/ukrainealert/why-the-us-keeps-losing-the-fight-against-disinformation>
- Erixon, F. Angela Merkel has created Germany's far-right. The Spectator. February 24, 2018. <https://www.spectator.co.uk/2018/02/angela-merkel-has-created-germanys-far-right/>
- Estonian Public Media website. Unsigned. Bronze Night's only death still unsolved. April 26, 2017. <https://news.err.ee/592217/bronze-night-s-only-death-still-unsolved>
- Estonian Public Media website. Unsigned. Estonian prosecutors close investigation into Dmitri Ganin's death.. June 15, 2017 <https://news.err.ee/602231/estonian-prosecutors-close-investigation-into-dmitri-ganin-s-death>
- EU East StratCom Task Force. Means, goals and consequences of the pro-Kremlin disinformation campaign. Istituto per gli Studi di Politica Internazionale (ISPI). Jan. 19, 2017. <https://www.ispionline.it/it/publicazione/means-goals-and-consequences-pro-kremlin-disinformation-campaign-16216>
- EU vs Disinfo website Unsigned. Nine ways to confuse us about MH17. Sept. 30, 2016. <https://euvsdisinfo.eu/nine-ways-to-confuse-us-about-mh17/>
- EU vs Disinfo website. Figure of the week: 20. Posted on Mar. 27, 2018. <https://euvsdisinfo.eu/figure-of-the-week-20/>
- EU vs Disinfo website. Unsigned. Disinformation Review. March 15, 2018. <https://euvsdisinfo.eu/skripal-and-the-disinformation-swamp/>
- Evron, G. Estonia 10 Years Later: Lessons learned from the World's First Internet War. Securityledger.com. April 28, 2017. <https://securityledger.com/2017/04/estonia-10-years-later-lessons-learned-from-the-worlds-first-internet-war/>
- Evron, G. A Russian strategist's take on information warfare. Darkreading.com. March 26, 2010. <https://www.darkreading.com/risk/a-russian-strategists-take-on-information-warfare/d/d-id/1133301>

- Fedchenko, Y. Kremlin propaganda: Soviet active measures by other means. Sojateadlane (Estonian Journal of Military Studies), Volume 2, 2016, pp. 141-170. PDF accessed online at www.ksk.edu.ee/publikatsioonid
- Farid, H. Digital doctoring: how to tell the real from the fake. Royal Statistical Society. Nov. 2006. <https://rss.onlinelibrary.wiley.com/doi/full/10.1111/j.1740-9713.2006.00197.x>
- Flitter, E. Exclusive: Riding Trump wave, Breitbart News plans U.S., European expansion. Reuters. Nov. 9, 2016. <https://www.reuters.com/article/us-usa-election-trump-strategy-idUSKBN1342TP>
- Garrity, PJ. Strategy, Disinformation, and Deception in International Politics. <https://politicsandstrategy.com/strategy-disinformation-and-deception-in-international-politics-fb6ad8cf65f0>
- Geissler, E., & Sprinkle, R. Disinformation squared: Was the HIV-from-Fort-Detrick myth a Stasi success? Politics and the Life Sciences, 32(2), 2-99. Retrieved from <http://www.jstor.org/stable/43287281>
- Grassegger, H & Krogerus, M. Fake news and botnets: How Russia weaponized the web. The Observer. Dec. 2, 2017. <https://www.theguardian.com/technology/2017/dec/02/fake-news-botnets-how-russia-weaponised-the-web-cyber-attack-estonia>
- Graves, BE. The smoking gun of AIDS: A 1971 flowchart. Dec. 6, 2000. https://www.bibliotecapleyades.net/sociopolitica/esp_sociopol_depopu40.htm
- Green European Journal. Unsigned. Green Observatory: Fake News. Undated. <https://www.greeneuropeanjournal.eu/green-observatory-fake-news/#Hungary>
- Gross, J. Chemical weapons watchdog backs up U.K. view on spy poisoning. The Wall Street Journal. April 12, 2018. <https://www.wsj.com/articles/novichok-was-used-in-skripal-attack-chemical-weapons-watchdog-confirms-1523532806>
- The Guardian/London School of Economics. Reading the riots. Investigating England's summer of disorder. [http://eprints.lse.ac.uk/46297/1/Reading%20the%20riots\(published\).pdf](http://eprints.lse.ac.uk/46297/1/Reading%20the%20riots(published).pdf) pp. 30-33.
- Hahn, G. The 2008 Georgian-Ossetian/Russian war: Saakashvili's strategic communications/disinformation campaign. Revised July 30, 2017. <https://gordonhahn.com/2017/07/30/the-2008-georgian-ossetianrussian-war-saakashvilis-strategic-communications-disinformation-campaign/>

- Hale, V. Revealed: 1,000-Man Mob Attack Police, Set Germany's Oldest Church Alight on New Year's Eve. Breitbart.com. Jan. 3, 2017. Accessed online at <http://www.breitbart.com/london/2017/01/03/dortmund-mob-attack-police-church-alight/>
- Handelsblatt Global. Unsigned. Breitbart drags feet on German expansion. March 7, 2017. <https://global.handelsblatt.com/companies/breitbart-drags-feet-on-german-expansion-720136>
- Harkins, S. G., & Petty, R. E. (1987). Information utility and the multiple source effect. *Journal of Personality and Social Psychology*, 52(2), 260-268. <http://dx.doi.org/10.1037/0022-3514.52.2.260>
- Hasher, L., Goldstein, D., & Toppino, T. (1977). Frequency and the conference of referential validity. *Journal of Verbal Learning and Verbal Behavior*, 16, 107- 112.
- Heshmat, S. What is confirmation bias? *Psychology Today*. April 23, 2015. Accessed online at <https://www.psychologytoday.com/us/blog/science-choice/201504/what-is-confirmation-bias>
- Hollis, D. Cyberwar case study: Georgia 2008. *Small Wars Journal*. <http://smallwarsjournal.com/blog/journal/docs-temp/639-hollis.pdf>
- Iasiello, EJ. Russia's improved information operations: From Georgia to Crimea. *The U.S. Army War College Quarterly, Parameters* 47. No. 2. p51. Summer 2017. <http://publications.armywarcollege.edu/pubs/3368.pdf>
- Interfax news agency, March 28, 2015. <http://www.interfax.com/newsinf.asp?id=5818511>
- International Centre for Defence and Security (ICDS). Unsigned. Russia's involvement in the Tallinn disturbances. May 11, 2007. <https://www.icds.ee/publications/article/russias-involvement-in-the-tallinn-disturbances/>
- Internet Research Agency Indictment. <https://www.justice.gov/file/1035477/download>
- Ioffe, J. The history of Russian involvement in America's race wars. *The Atlantic*. Oct. 21, 2017. <https://www.theatlantic.com/international/archive/2017/10/russia-facebook-race/542796/>

- Johnson, JR. The authenticity and validity of Antony's will. *L'Antiquité Classique*. T. 47, Fasc. 2 (1978), pp. 494-503. https://www.persee.fr/doc/antiq_0770-2817_1978_num_47_2_1908
- Jones, N et al. Rumors during a campus lockdown. *Proceedings of the National Academy of Sciences* Oct 2017. 114 (44) 11663-11668. <http://www.pnas.org/content/114/44/11663>
- Joubert, V. Five years after Estonia's cyber attacks: lessons learnt for NATO? NATO Research Paper. No. 76. May 2012.
- Thejournal.ie. Unsigned. Breitbart ran a fake story about a terrorist outrage in Dortmund, and Germans aren't one bit happy. 2016. <http://jrnl.ie/3173688>
- Junge Freiheit website. Unsigned. Keine deutschsprachige Breitbart-Ausgabe (No German language Breitbart). March 3, 2017. <https://jungefreiheit.de/kultur/medien/2017/keine-deutschsprachige-breitbart-ausgabe/>
- Kaminska, I. A lesson in fake news from the info-wars of ancient Rome. *Financial Times*. Jan 17, 2017. <https://www.ft.com/content/aaf2bb08-dca2-11e6-86ac-f253db7791c6>
- Kietzmann, JH et al. Social media? Get serious! Understanding the functional building blocks of social media. *Business Horizons* (2011) 54, 241—251. <https://doi.org/10.1016/j.bushor.2011.01.005>
- Kinstler, L. How to survive a Russian hack. *The Atlantic*. Feb. 2, 2017. <https://www.theatlantic.com/international/archive/2017/02/russia-disinformation-baltics/515301/>
- Kirkpatrick, TJ. Bloomberg News in The Moscow Times. <https://themoscowtimes.com/news/russia-reacts-charges-election-meddling-mueller-investigation-60553>
- Kivirahk, J. Integrating Estonia's Russian-Speaking Population: Findings of National Defense Opinion Surveys. ICDS report. P 3. December 2014. https://www.icds.ee/fileadmin/media/icds.ee/failid/Juhan_Kivirahk_-_Integrating_Estonias_Russian-Speaking_Population.pdf
- Kumar, KPK & Geethakumari, Detecting misinformation in online social networks using cognitive psychology. *G. Hum. Cent. Comput. Inf. Sci.* (2014) 4: 14. <https://doi.org/10.1186/s13673-014-0014-x>

- Kwak, H. et al. What is Twitter, a social network or a news media? Proc. of the 18th international conference on World Wide Web. Pp 591-600. April 2009. <https://dl.acm.org/citation.cfm?id=1772751>
- Laruelle, M. (ed). Russian Nationalism, Foreign Policy and Identity Debates in Putin's Russia. Ibidem Press, Stuttgart. 2014. pp. 133-134
- Lazer et al. Combating fake news: an agenda for research and action. Harvard Kennedy School, Shorenstein Center on Media, Politics and Public Policy. Conference report from Feb. 17-18, 2017. Accessed online at <https://shorensteincenter.org/combating-fake-news-agenda-for-research/>
- Levine, Y. The CNN effect: Georgia schools Russia in information warfare. Aug. 13, 2008. <http://exiledonline.com/the-cnn-effect-georgia-schools-russia-in-information-warfare/>
- Lewandowsky, S. et al. Misinformation and Its Correction: Continued Influence and Successful Debiasing. Psychological Science in the Public Interest 13(3) 106–131. <http://journals.sagepub.com/doi/full/10.1177/1529100612451018>
- Lucas, E & Pomeranzev, P. Winning the information war. Center for European Policy Analysis report. p.23. August 2016. https://cepa.ecms.pl/files/?id_plik=2706
- Marwick, A & Lewis, R. Media manipulation and disinformation online. Data & Society Research Institute. Center for Media Justice. May 15, 202017. p. 3. https://datasociety.net/pubs/oh/DataAndSociety_MediaManipulationAndDisinformationOnline.pdf
- Mazer, J. et al. Communication in the face of a school crisis: Examining the volume and content of social media mentions during active shooter incidents. Computers in Human Behavior, 53, 238–248. 2015. <https://www.tandfonline.com/servlet/linkout?suffix=CIT0013&dbid=128&doi=10.1080%2F17459435.2016.1247111&key=000361921100024>
- Media Bias / Fact Check website. Right- and left-wing sites. Centrist sites. www.mediabiasfactcheck.com/left www.mediabiasfactcheck.com/right <http://mediabiasfactcheck.com/center/>
- Mendoza, M. et al. Twitter Under Crisis: Can We Trust What We RT? In Proc. of 1st Workshop on Social Media Analytics (SOMA 2010). http://snap.stanford.edu/soma2010/papers/soma2010_11.pdf

- Metaxas et al. Do Retweets indicate interest, trust, agreement?
[arXiv:1411.3555v1 \[cs.SI\]](https://arxiv.org/abs/1411.3555v1)
- Metaxas, P.T. Retweets indicate Agreement, Endorsement, Trust: A Meta-Analysis of Published Twitter Research. Computer Science Department, Wellesley College. Accessed online at <http://cs.wellesley.edu/~pmetaxas/WorkingPapers/Retweet-meaning.pdf>
- Metaxas, P.T., Finn, S., & Mustafaraj, E. (2015) DEMO: Using TwitterTrails.com to Investigate Rumor Propagation. WebScience15, Oxford, UK.
<https://repository.wellesley.edu/cgi/viewcontent.cgi?article=1166&context=scholarship>
- Miller, C. The big business of disinformation. Wired.com. June 17, 2017.
<http://www.wired.co.uk/article/carl-miller-disinformation-gets-commercialised>
- Mohdin, A. What's "alt-right" in German? Breitbart News is expanding in Europe. Quartz. Nov. 20, 2016. <https://qz.com/841959/breitbart-news-is-coming-to-france-and-germany-ahead-of-elections/>
- Narayanan, V. et al. Polarization, Partisanship and Junk News Consumption over Social Media in the US. Computational Propaganda Project, University of Oxford. Feb. 6, 2018. <http://comprop.oii.ox.ac.uk/wp-content/uploads/sites/93/2018/02/Polarization-Partisanship-JunkNews.pdf>
- Nazario, J. Politically motivated denial of service attacks. NATO Cooperative Cyber Defence Centre of Excellence. Undated.
http://www.ccdcoe.org/publications/virtualbattlefield/12_NAZARIO%20Politically%20Motivated%20DDoS.pdf
- Nguyen T.N., Li C., Niederée C. (2017) On Early-Stage Debunking Rumors on Twitter: Leveraging the Wisdom of Weak Learners. In: Ciampaglia G., Mashhadi A., Yasseri T. (eds) Social Informatics. SocInfo 2017. Lecture Notes in Computer Science. Vol 10540. Springer, Cham. https://doi.org/10.1007/978-3-319-67256-4_13
- Nordrhein-Westfalen Dortmund Polizei, Jan. 5, 2017.
<https://www.presseportal.de/blaulicht/pm/4971/3527791>
- Pääbo, H. War of Memories: Explaining "Memorials War" in Estonia. Baltic Security & Defence Review Volume 10, 2008.
- Pal, A. Reviewing the Landscape of Research on the Threats to the Quality of User-Generated Content. Proceedings of the Association for Information Science and

Technology. 53:1. 2016.

<https://onlinelibrary.wiley.com/doi/full/10.1002/pra2.2016.14505301077>

- Panagiotopoulos, P. et al. Social media in emergency management: Twitter as a tool for communicating risks to the public. <https://doi.org/10.1016/j.techfore.2016.06.010>
- Paul, Christopher and Miriam Matthews, The Russian "Firehose of Falsehood" Propaganda Model: Why It Might Work and Options to Counter It. Santa Monica, CA: RAND Corporation, 2016.
<https://www.rand.org/pubs/perspectives/PE198.html>.
- PunditFact web pages on May 14, 2018 for Breitbart, Fox News, Daily Caller, InfoWars, Gateway Pundit, Washington Examiner
<http://www.politifact.com/personalities/breitbart/>
<http://www.politifact.com/search/?q=Fox+News>
<http://www.politifact.com/search/?q=daily+caller>
<http://www.politifact.com/search/?q=Infowars>
<http://www.politifact.com/search/?q=gateway+pundit>
<http://www.politifact.com/search/?q=washington+examiner>
- Richards, J. Denial-of-Service: The Estonian Cyberwar and Its Implications for U.S. National Security. International Affairs Review. <http://www.iar-gwu.org/node/65>
- Richards, J & Lewis, P. How Twitter was used to spread – and knock down – rumours during the riots. The Guardian. Dec. 7, 2011.
<https://www.theguardian.com/uk/2011/dec/07/how-twitter-spread-rumours-riots>
- Robb, A. Anatomy of a Fake News Scandal, Rolling Stone Magazine, Nov. 16, 2017.
<https://www.rollingstone.com/politics/news/pizzagate-anatomy-of-a-fake-news-scandal-w511904>
- Roeder, T. No evidence of shots fired at Air Force Academy; social media fuels 'hysteria.' Colorado Springs Gazette. Oct. 1, 2017. <http://gazette.com/no-evidence-of-shots-fired-at-air-force-academy-social-media-fuels-hysteria/article/1612291>
- Ruhr Nachrichten. [https://www.Ruhr Nachrichten.de/Nachrichten/Polizei-reagiert-auf-Fake-News-zu-Dortmunder-Silvester-60951.html](https://www.RuhrNachrichten.de/Nachrichten/Polizei-reagiert-auf-Fake-News-zu-Dortmunder-Silvester-60951.html)
- Ruhr Nachrichten. [https://www.Ruhr Nachrichten.de/Staedte/Dortmund/Silvesterrakete-trifft-Obdachlosen-64703.html](https://www.RuhrNachrichten.de/Staedte/Dortmund/Silvesterrakete-trifft-Obdachlosen-64703.html)
- Ruhr Nachrichten. Bandermann, P. Fragen und Antworten: Wie die Medien die Silversternacht für Hass-Aufbau nutzen (Questions and answers: How the media use

the New Year's celebrations to build hate). Ruhr Nachrichten.de. Jan 4, 2017. <https://www.ruhrnachrichten.de/Staedte/Dortmund/Wie-Medien-die-Silvesternacht-fuer-Hass-Aufbau-nutzen-55127.html>

- Russian Ministry of Foreign Affairs spokeswoman's comment on investigation into the death of Ganin. Sept. 2016. http://www.mid.ru/en_GB/foreign_policy/news/-/asset_publisher/cKNonkJE02Bw/content/id/2435063
- Sargent, H. April 15, 2013: How Twitter informed us. April 15, 2014. <https://www.boston.com/news/local-news/2014/04/15/april-15-2013-how-twitter-informed-us>
- Shachtman, N. Top Georgian Official: Moscow Cyber Attacked Us – We Just Can't Prove It. Wired Magazine. March 11, 2009. Cited in Hollis, D. Cyberwar case study: Georgia 2001. Small Wars Journal. Available in full at <http://www.wired.com/dangerroom/2009/03/georgia-blames/>
- Shearer, E & Gottfried, J. News use across social media platforms 2017. Pew Research Center. Sept. 7, 2017. <http://www.journalism.org/2017/09/07/news-use-across-social-media-platforms-2017/#>
- Silverman, C et al. Hyperpartisan Facebook pages are publishing false and misleading information at an alarming rate. BuzzFeed.News. Oct. 20, 2016. https://www.buzzfeed.com/craigsilverman/partisan-fb-pages-analysis?utm_term=.cuQyewY4Y5#.tsAgjV878w
- Simon, T. et al. Twitter in the Cross Fire—The Use of Social Media in the Westgate Mall Terror Attack in Kenya. PLoS ONE 9.8 (2014). <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0104136>
- Sirianni, F. Was Mark Antony's will partially forged? L'Antiquité Classique. T. 53 (1984), pp. 236-241. https://www.persee.fr/doc/antiq_0770-2817_1984_num_53_1_2126
- Stanger, M. 6 people who were literally erased from history. Business Insider. Feb 4, 2014. <http://www.businessinsider.com/people-who-were-erased-from-history-2013-12>
- Starbird et al. **Chatter on the red: what hazards threat reveals about the social life of microblogged information.** Proceedings of the 2010 ACM conference on Computer supported cooperative work. 241-250.

- Starbird, K. et al. Rumors, False Flags, and Digital Vigilantes: Misinformation on Twitter after the 2013 Boston Marathon Bombing. iConference 2014. PDF accessed online. https://faculty.washington.edu/kstarbi/Starbird_iConference2014-final.pdf
- [Statista.com. https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/](https://www.statista.com/statistics/272014/global-social-networks-ranked-by-number-of-users/)
- Stein, A. We Thought It Could Never Happen Here, Journal of Promotion Management, 12:3-4.Pp. 99-128, https://doi.org/10.1300/J057v12n03_07
- Stone, Daniel F. and Wood, Daniel H., Cognitive Dissonance, Motivated Reasoning, and Confirmation Bias: Applications in Industrial Organization (December 2, 2016). <http://dx.doi.org/10.2139/ssrn.2834072>
- StopFake.org. 13 online tools that help to verify the authenticity of a photo. July 29, 2014. <https://www.stopfake.org/en/13-online-tools-that-help-to-verify-the-authenticity-of-a-photo/>
- Stowell, J. What is hybrid warfare? Global Security Review. Feb. 2, 2018. <https://globalsecurityreview.com/hybrid-and-non-linear-warfare-systematically-erases-the-divide-between-war-peace/>
- TED.com. Eli Pariser TED talk. https://www.ted.com/talks/eli_pariser_beware_online_filter_bubbles/transcript#t-356702
- Traynor, I. Russia accused of unleashing cyberwar to disable Estonia. Russia accused of unleashing cyberwar to disable Estoniahe Guardian. May 16, 2017. <https://www.theguardian.com/world/2007/may/17/topstories3.russia>
- Thompson, B. et al. (2016) Social Media and Active Shooter Events: A School Crisis Communication Challenge, Qualitative Research Reports in Communication, 18:1, 8-17. <https://www.tandfonline.com/doi/abs/10.1080/17459435.2016.1247111>
- United States Department of State, Soviet Influence Activities: A report on Active Measures and Propaganda, 1986-87. August 1987. Ch. 4, pp 29-32. <https://www.globalsecurity.org/intell/library/reports/1987/soviet-influence-activities-1987.pdf>

- U.S. Information Agency (USIA). Soviet active measures in the era of glasnost. March 1988.
<http://insidethecoldwar.org/sites/default/files/documents/Soviet%20Active%20Measures%20in%20the%20Era%20of%20Glasnot%20March%201988.pdf>
- U.S. Information Agency report Soviet Active Measures in the 'Post-Cold War' Era 1988-1991, prepared for U.S. House of Representatives Committee on Appropriations. June 1992. Accessed online at
http://intellit.muskingum.edu/russia_folder/pcw_era/sect_09a.htm. Full report at
http://intellit.muskingum.edu/russia_folder/pcw_era/index.htm#Contents
- Vieweg, Sarah. Microblogged Contributions to the Emergency Arena: Discovery, Interpretation and Implications. Computer Supported Collaborative Work. Feb. 2010.
<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.495.520&rep=rep1&type=pdf>
- Vosoughi, S. Automatic Detection and Verification of Rumors on Twitter. Ph.D. dissertation.MIT. June 2015. p 19.
http://lsm.media.mit.edu/papers/Soroush_Vosoughi_PHD_thesis.pdf
- Vosoughi, S et al. The spread of true and false news online. Science. Mar 9,2018. Vol. 359, Issue 6380, pp. 1146-1151.
<http://science.sciencemag.org/content/359/6380/1146>
- Vosoughi, S et al. Rumor Gauge: Predicting the veracity of rumors on Twitter. ACM Transactions on Knowledge Discovery from Data, Vol. 11, No. 4, Article 50, July 2017.
- Wasson, DL. Second Triumvirate. Accessed online at Ancient History Encyclopedia.
https://www.ancient.eu/Second_Triumvirate
- Weir, K. Why we believe alternative facts. American Psychological Association. May 2017, Vol 48, No. 5. <http://www.apa.org/monitor/2017/05/alternative-facts.aspx>
- Westerman, D. et al. Social Media as Information Source: Recency of Updates and Credibility of Information. Journal of Computer-Mediated Communication 19 (2014) 171–183. <https://doi.org/10.1111/jcc4.12041>
- Wikileaks Cable 297, par. 5.
https://wikileaks.org/plusd/cables/07TALLINN310_a.html
- Wikileaks Cable 366. Dated May 11, 2007.
https://wikileaks.org/plusd/cables/07TALLINN366_a.html

- Wilby, P. Georgia has won the PR war. The Guardian. Aug. 17, 2008
<https://www.theguardian.com/media/2008/aug/18/pressandpublishing.georgia>
- Wilson, Clay. 2007. Information operations, electronic warfare, and cyberwar capabilities and related policy issues. [Washington, D.C.]: Congressional Research Service, Library of Congress. <http://www.fas.org/sgp/crs/natsec/RL31787.pdf>
- Wochenblick's Dortmund New Year's Eve 2016/17 story accessed online.
<https://www.wochenblick.at/silvester-in-dortmund-allahu-akbar-und-kirchenbrand/>
- Wong, JC et al. Bursting the Facebook bubble: We asked voters on the left and right to swap feeds. The Guardian. Nov. 16, 2016. <https://www.theguardian.com/us-news/2016/nov/16/facebook-bias-bubble-us-election-conservative-liberal-news-feed>
- Yakowicz, W. Have a beer, it's good for your brain. Inc.com. Oct. 3, 2014.
<https://www.inc.com/will-yakowicz/beer-improves-brain-function.html>
- Zamzow, DR et al. Xanthohumol improved cognitive flexibility in young mice. Behavioural Brain Research. Volume 275, 15 December 2014, Pages 1-10.
<https://doi.org/10.1016/j.bbr.2014.08.045>
- Zubiaga, A. et al. Analysing how people orient to and spread rumours in social media by looking at conversational threads. PLoS ONE. 11.
<http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0150989>
- Zubiaga, A. et al. Curating and contextualizing twitter stories to assist with social newsgathering. Proceedings of IUI 2013, pp 213–224. March 19–22, 2013.
<https://dl.acm.org/citation.cfm?id=2449424>