CHAPTER 7

Bots, Trolls, Elves, and the Information War in Lithuania: Theoretical Considerations and Practical Problems

Asta Zelenkauskaite

Recently, the media sphere has been found to be regularly flooded with misinformation and disinformation.[1] Some scholars refer to such outbreaks as infodemics (Bruns et al., 2020). Clearly, we are at a point where we must challenge the assumption of online spaces as emblematic of democratic ideals where participation is assumed to solely foster healthy

[1] Misinformation refers to information that is factually inaccurate, i.e. "misleading information created or disseminated without manipulative or malicious intent" (UNESCO, 2018). Misinformation is often accompanied by disinformation—where individuals, groups, and organizations deliberately aim to create confusion and discord. Disinformation is defined by "deliberate (often orchestrated) attempts to confuse or manipulate" (UNESCO, 2018).

A. Zelenkauskaite (✉)
Drexel University, Philadelphia, PA, USA

Vilnius Tech, Vilnius, Lithuania
e-mail: az358@drexel.edu

J. Chakars, I. Ekmanis (eds.), *Information Wars in the Baltic States*, The Palgrave Macmillan Series in International Political Communication, https://doi.org/10.1007/978-3-030-99987-2_7

debate in an authentic marketplace of ideas. Prominent in this subverted media arena are bots, inauthentic participants in the exchange of information, and trolls, the "Ghostwriters," flamethrowers of the Internet, whether automated or not. This problem is felt acutely in Lithuania where there are fears that a Russian information war could turn into real war or other political disruption. It is critical to uncover how these inflamed efforts are manifested through various forms of computational propaganda, as argued by scholars like Woolley and Howard (2018).

Therefore, it is imperative to understand the process of this media manipulation. This chapter argues that we will have to adapt our theories and methodologies in order to most effectively do so by linking the theoretical groundings of mass media influence to the concept of information warfare (Cronin & Crawford, 1999), where social media chaos can warp civic participation (Zelenkauskaite, 2022).

Information War in Lithuania

Russia, since the collapse of the Soviet Union, stands as the adversary in an information war with the Baltic states (e.g., see Orenstein, 2019), where Russian disinformation and propaganda can be equated to the use of soft power (Simons, 2015). In the Baltic states, inauthentic participation in online spaces has been a particularly pressing manifestation of such power. Such inauthentic participation online—recently emerged through what is known as Russian trolling—has been attributed to Russia's attempts to provoke political instability, including forms of cyber coercion ranging from the hacking of online infrastructure to bot-driven social media posts (see, e.g., Valeriano et al., 2018) to online comments on news portals (see e.g., Zelenkauskaite & Niezgoda, 2017). In Lithuanian news portal comments, a segment of users was found to be masking their IP addresses when posting to appear anonymous. Similarly, the posting frequency of these users indicate either an unusually active dedication to online commenting or automated processing (Zelenkauskaite & Balduccini, 2017).

These findings can be contextualized globally as well. Russian trolling was evident during the 2016 US presidential election; study of the campaign also found that one-fifth of all tweets associated with the Democratic and Republican candidates were produced by nonhuman actors or bots (Bessi & Ferrara, 2016). This showcases how automated human and nonhuman efforts are shaping online platforms and reaching their readers (Jamieson, 2018).

Information influence has also been a critical question since the reestablishment of independence in Lithuania, but concerns have been heightened in parallel to those on the other side of the Atlantic. While Donald Trump and Hillary Clinton were vying for the presidency of their country, Lithuanian news portals published a report on how to recognize Russian trolls and the topics that they target (Garbačiauskaitė-Budrienė, 2016). These efforts were followed by a range of media literacy campaigns and investigative journalism initiatives by Delfi (Delfi.lt), with names like "Uncover," "Lie Detector," and "Media War and Propaganda". Delfi also partnered with researchers and "elves," citizen disinformation sleuths, to create Demaskuok (Debunk) to combat Russian disinformation. In the academic community, a center for media literacy was established between Vytautas Magnus University and UNESCO in 2019 (VDU, 2021).

As evidenced by articles covering Russian trolling—published, for example, by Delfi.lt editor-in-chief Monika Garbačiauskaitė-Budrienė already in 2016—Russian trolling has been found to systematically appear in the comments under the stories written by Lithuanian journalists who typically criticize Russian government. Such trolls particularly mock or criticize former Lithuanian President Dalia Grybauskaitė (who is an avid critic of Russian military operations and Russian President Vladimir Putin) or delegitimize Lithuanian military power and the NATO alliance. In response, Lithuanian "elves" have called out such claims in the same comments by trying to unmask the trolls and counter their claims, as found in the empirical analysis of news portal comments (Zelenkauskaite & Niezgoda, 2017). Yet, the Russian trolls were also found to engage in defense mechanisms through rebuttals, shifting blame to the news portals or shifting arguments to unrelated sensitive social issues, and thus attempting to diminish Lithuania's capacity to be perceived as a truly independent democratic country.

I refer to these narratives as part of information warfare, but taken at face value, one might argue that they are an expression of a typical democratic debate where people merely express their opinions or engage in activism by genuinely fighting for their ideals in the online public sphere. However, the manifestation of information warfare becomes more evident when these spaces are viewed from another perspective. In one systemic analysis of IP addresses associated with the messages, around half of the messages were found to originate from Lithuania, with 18% written from Russia. However, the third largest category (7.9%) of posts were written

with the intentional masking of their IP addresses (Zelenkauskaite & Balduccini, 2017). This indicates, that in addition to rhetorical tactics, a substantial presence of users deliberately choose to mask their IP addresses in these environments, clouding the nature of discourse. As a result, orchestrated messages dispatched to discredit are no longer part of the democratic debate, especially if they are not authentic.

Mechanics of Influence: Rise of the Bots

In 2019, the Lithuanian public broadcaster released a story citing a warning from the Ministry of Defense that fake news was targeting NATO's presence in Lithuania (LRT, 2019). Similarly, on July 30, 2020, *The Guardian* revealed that Russia-aligned hackers have been running an anti-NATO fake news campaign known as "Ghostwriter" since 2017, targeting Lithuania and Poland specifically (Sabbagh, 2020). Specific sociopolitical circumstances lend themselves to the exploitation of automation by authoritarian regimes, and studies of such regimes are filled with examples whereby online fora are disrupted (e.g., Bedford & Vinatier, 2018; Dukalskis, 2017; Pearce et al., 2018; Toepfl, 2018). The result of these tactics by authoritarian regimes is a challenge to content legitimacy and authenticity (Toepfl & Litvinenko, 2018).

While content creation must be coordinated by human efforts, nonhuman aids can become critical in amplifying or fabricating online content. Bots are a prime example of this challenge. But bots are not inherently evil; they serve important functions of automation. In the early days of the Internet, on platforms such as Internet Relay Chat (IRC), they were originally built to perform various mundane roles such as running statistics and trivia games. However, users soon started to leverage the automated power of bots for other purposes. For example, Canavan (2005) summarizes IRC bots' automated tasks as follows:

- Retrieve basic system information such as the operating system version, the computer name, and user information
- Retrieve login names and email addresses
- Retrieve dial-up networking settings, including usernames and passwords
- Update functionality

Similarly, on a different online platform, Wikipedia, bots were launched to systematically clean entries to adhere to standard formatting (see e.g., Geiger, 2018). However, these bots were later reprogrammed on IRC to infiltrate networks and spy on user's information or request ransoms. These sets of bots also evolved to infiltrate individual users' computer networks (Canavan, 2005). This shows how bots can be manipulated based on the needs of the users who create them.

The current landscape of online bots, typically found on social media, have been transformed into tools to persuade rather than penetrate individual computer networks. For example, Twitter bots have evolved to perform the following list of tasks online, as described by Daniel and Millimaggi (2020):

- Search users by keyword, account, ID
- Follow users and classify them based on predefined parameters, such as user types, trends, and keywords
- Like content, based on predefined parameters, such as user types, trends, and keywords
- Tweet and mention users and keywords, based on AI-generated content, fixed-templated
- Take content or cloned-content from other users
- Retweet users and trending content, and mass tweet based on specific parameters
- Chat to (reply) or chat with other users, use pauses to mimic application programming interface (API) or human expectations, store information for later use.

Machine-learning techniques can allow one to achieve specific content-based needs, for example, to support programmed Russian trolling through bots designed to express themselves with the appearance of authenticity. To achieve sustained authenticity, the programmer of a bot merely needs to constantly scan user-generated content that is posted and extract relevant content, be it social media or news portal comments, to construct authentic-looking content for further dispatch. Similarly, prepackaged messages are easily spread by automated means.

Influence as a Theoretical Consideration and the Internet

Automating content distribution is still highly dependent on the content that is already circulating and the mechanics of reaching targeted audiences that would be influenced. As a result, inauthentic participation calls for the need to revisit theories related to participation and influence online. Bots illustrate how media influence can be constructed through online means. In its most basic sense, influence is the capacity to have an effect on someone. Social influence, specifically, can be defined "as change in a person's cognition, attitude, or behavior, which has its origin in another person or group" (Raven, 1965, p. 1). There are at least three components relevant to influence: someone who influences, who they influence, and how influence takes place.

To understand the mechanics of inauthentic online behaviors, it is critical to contextualize them in the frameworks of influence. Seminal research on social influence in pre-online environments and mass media illuminated an interpersonal or a group dynamic (Katz & Lazarsfeld, 1955). In mass media, social influence is mediated through a two-step flow phenomenon (Lazarsfeld, et al., 1944). In this conception, particular people form a bridge connecting media and the public, and carry influence. An influencer thus becomes defined as someone "who exhibits some combination of desirable attributes—whether personal attributes like credibility, expertise, or enthusiasm, or network attributes such as connectivity or centrality—that allows them to influence a disproportionately large number of others" (Bakshy et al., 2011, p. 65). As in this definition, personal attributes only function relative to the structural and temporal aspects of media platforms—what van Dijk (1998) calls the context. Thus, social media posts or the bot personas behind them can constitute the second step in the two-step flow, presenting themselves at the bridge to a less invested citizenry looking to opinion leaders who appear most engaged.

Interpersonal Influence and Group Influence

Interpersonal and group influence has been typically analyzed through (1) individual traits of the influenced (Cohen, 1959; Dion & Stein, 1978; Raven, 1958; Wright, 1943); (2) individual traits of the source of influence, that is, the influencer (Katz & Lazarsfeld, 1955); (3) temporal dimensions (Allport, 1920; Hovland & Weiss, 1951; Fisher & Lubin,

1958); (4) social cues in communication (Katz & Lazarsfeld, 1955; Schein, 1960); and (5) organizational influence (Burt, 1976; Pelz, 1952).

Individual traits of the influenced include psychological variables such as self-esteem or levels of stress. For example, Cohen (1959) found that those with low self-esteem were more responsive to influence, compared to those with higher self-esteem; individuals with higher levels of self-esteem tended to repress impulses that contradicted their own views. Stress was found to exacerbate influence in interpersonal (Wright, 1943) and work contexts (Raven, 1958).

In their research, Katz and Lazarsfeld (1955) extensively measured relative traits of influencers. Variables relevant to influential power were sensitive to different social fields: fashion, politics, and so forth. However, generally, higher socioeconomic status (e.g., higher education levels, more affluence, and more social contacts) and gregariousness were found to be key traits in those who influenced. Gender was also found to play a role in interpersonal influence, where males were more likely to be influencers. Individual physical traits such as attractiveness were also found to have influential power. Dion and Stein (1978) found that attractive participants of the opposite sex were more prone to induce influence.

Hovland and Weiss (1951) found a temporal dimension ("sleeper effect") to function within circumstances of influence. Instead of immediate influence, opinion changes occurred after a lapse of time. They also found that over time the source of information (the influencer) decreased in relevance, becoming dissociated from the information itself. Furthermore, repetition over time was found to increase influence (Fisher & Lubin, 1958). Allport (1920) identifies the beneficial effects of group influence at the beginning and the end of a task. Social cues (verbal and nonverbal communicative attributes, e.g., social manners, dress-code, degree of difference) were also found to play a central role to the process of influence, whether positive or negative (Schein, 1960). Providing a participant with only negative or neutral social cues causes alienation and can make someone susceptible to external influence.

Katz and Lazarsfeld (1955) measured components of social structure to analyze influence. In their research social structural components included degree of mutual attraction, differences in the degree of interdependence, propinquity, and group climates (or cultures). They treat communication and influence to be context-sensitive.

Various authors have focused on the process of influence in group communication and at the organizational level. Influence was found to be

effectively exercised by a mere belonging to a given group. Deutsch and Gerard (1955) indicate that social influence can have a greater effect on those who are part of a specific group (with shared values or beliefs), compared to individuals who do not belong to a particular group supporting their beliefs. At the organizational level, influence includes aspects of uncertainty and its consequences to individuals, such as conditions that preclude interpersonal influence or fostering solidarity.

Structural Factors

While the theories described in the earlier paragraphs deal with the conceptual mechanics of influence, currently, they are mediated through sociotechnical online systems. The information infrastructure of online systems is yet another element that is pertinent. Online information systems have been designed through decentralized content distribution (compared to mass media where messages used to be dispatched centrally). Typically, content circulation can be broken down into two structural paradigms—network and threaded. They are detailed here to contextualize how influence and inauthentic participation can leverage such sociotechnical configurations.

Network Structure

Conceptualizations of online influence continue to entail elements proposed in pre-online contexts, but mediated by parameters pertinent to social construction of technology, that is, the ways in which technological properties can be used in information warfare. Influence in networks has been presumed to follow the interpersonal conceptualizations established by Katz and Lazarsfeld (1955), yet the presence of bots alters the character, flow, and, if revealed, authority of communicated messages.

Network structure is online communication that allows for diffusion of content, resulting in potential influence (Anagnostopoulos et al., 2008). Influence is the ability of individuals to induce their friends, enemies, strangers, and so forth, to act in a similar or particular way (Anagnostopoulos et al., 2008). Influencers act as opinion leaders, focal persons, who can function as either starters or connectors (Sun & Ng, 2013). They have strong connectivity and centrality within their networks (Bakshy et al., 2011). Influence entails not only strong and weak ties of passive and active participants (Baños et al., 2013), it has potential to identify hidden

influentials who can generate collective action (González-Bailón et al., 2013). To generate a small or large cascade, beyond one's immediate echo-chamber, individuals or groups must be sensitive, not only to temporal parameters (Baños et al., 2013; Sun & Ng, 2013) but also to the structural affordances embedded into the system, including names (Kalmar, 2010; Messias et al., 2013; Mislove et al., 2011), hashtags (Romero et al., 2011), and external linking (Bakshy et al., 2011).

Threaded Structure

While network structure is more predominant in social media, news portals that do not outsource user comments to social media usually organize communication in a threaded design.

Threaded communication is a form of asynchronous message-based communication. The user interface of threaded chat is structured in a reversed chronological sequential manner (or it can be at times sorted by "most popular" or "oldest" comments as well). Posts can be made either as stand-alone or as replies to an ongoing message wherein "turns are organized into turn and response structures called *threads* that can grow into any size" (Smith et al., 2000). Technological affordances of threaded interaction structure present communicative specificity, compared to face-to-face interaction management (Herring, 1999). When users send messages in a multi-participant chat, they are sent in a chronological order and thus lead to issues of disrupted turn adjacency, even if users perfectly make sense of it. Turn adjacency may be disrupted by mere technological issues: Even if a typed message is sent contemporarily by two given users, they are still displayed sequentially, where one follows the other. In a multi-participant chat, "a turn" is naturally disrupted by other people's messages that go between the time when one person responds to the other. Due to these technological constraints, users must find interpretative lenses to conceptualize the discourse (Herring, 1999).

One can argue that such structural features could be used strategically, given that users in these contexts do manage conversations with no major content comprehension difficulties. Moreover, in the threaded chat, new material can appear anywhere in the conversation tree. This requires frequent scanning to search for new turns (Smith et al., 2000) or be vigilant to introduce specific content threads to specific contexts. While influence can be foremost regarded as related to trust and reputation, structural features should also be given consideration.

Sociotechnical variables of online threaded environments (pertinent to online news portals) involve dimensions of content, time, and structure. Threaded interaction spaces variables related to influence can exploit the dialectic of self-replies (Smith et al., 2000) to provide a representation (or allusion) of self-dialogue mimicking interpersonal conversation (Herring, 1999). Similarly, one can effectively use the built in name/heading structures to draw user attention (Herring, 2003; Williams & Humphrey, 2007). One can send unlimited number of replies to silence responses of other users to subvert the conversation by using the speed of replies, typically used to make up for disruptive adjacency (Edelstein & Edwards, 2002; Khan & Jarvenpaa, 2010). Users can exploit the amount and types of comments by posting short versus long comments (Mandernach et al., 2006), few versus many comments (Kuk, 2006; Miller & Benz, 2008), and whether one's comments are condensed into one category of discussion or many (Kuk, 2006). Overlying issues include understanding participatory (Meyer, 2003) and regulatory practices Mandernach et al. (2006).

Influence Meets Structure

In order to understand soft power and disinformation campaigns, we must consider both the structure of online communication and patterns of influence. While content is the most relevant way to assess information warfare strategies, in the current media landscape, it is equally important to take into consideration the contexts, which are relevant to online environments. Thus, in order to understand information warfare online, it is critical to take into consideration content, structure, and mechanics. For example, structural components, relative to a specific environment—be it network-based (most social media) or threaded (e.g., news portal comments)—are examples of the context. Substantial theoretical work has recently advocated for social influence to be recognized, fundamentally, as a process mediated through structural components (see e.g., Friedkin, 2006).

Conclusion

In the context of Russian trolling, the main discussion is often about bots who are designed to push specific messages. These bots belong to the category of social bots "that automatically produces content and interacts with humans on social media" (Ferrara et al., 2016, p. 96). Thus, they can

appear like humans. Explicating typologies of bots (for example, for political agendas) becomes an important task to uncover and disambiguate some of the misconceptions of the online information and communication ecosystems (Gorwa & Guilbeault, 2018).

Dismissing bot activities online or treating them as genuine can be detrimental to the understanding of the online public sphere. Thus, it is necessary to at least consider them when discussing online spaces, even if bots can be difficult to detect, given that they project themselves as genuine (Hjorth & Adler-Nissen, 2019). Indeed, there are cases where even scholarly analyses are deceived into assuming that automated online behaviors are those of regular users. For example, an examination of the 2014 downing of Malaysian Airlines Flight 17 over Ukraine referred to disinformation as being done by "curators" and "involved" citizens who tweeted about it, without categorizing these actors as potentially linked to state influence (Golovchenko et al., 2018). Another closer analysis of the Russian Twittersphere revealed that its sample included a large number of bots who were acting on behalf of the "citizens" and appeared in the mask of an imposter (Stukal et al., 2019). Users in this case were actually masked automated bots who looked like highly active and involved posters, because of which it seemed as if it was regular citizens who spread disinformation. While not all platforms allow to check users if they are bots, such a tool, Botometer ("Botometer," n.d.) was developed for Twitter and it is a freely accessible tool available for anyone to use.

Thus, scholars, journalists, and the general public, when discussing online spaces and messages found on social media, are faced with a new challenge to discern which messages are genuine and which are produced by bots, "Ghostwriters," or authoritarian governments. All of it inevitably complicates the interpretation of our true understanding of the message—or as van Dijk (1998) states, what's in the context. Yet, one may argue, it depends on how audiences perceive those messages. And, if they do perceive them as real, then they are real for them.

Another complication is the perception of information warfare. My work centers on the political tension between the Lithuanian online public sphere and Russian ideological infiltration. Lithuanian news portals have long reported issues of Russian interference within online news comments. It has been exposed that the Russian government, through a government-based "agency," pays its workers to post messages tailored to evoke Soviet nostalgia in the Baltic states and silence the political critics who criticize Putin (Chen, 2015). Such conditions require us to adapt our

theories and methods of research to fully understand and confront this situation. Combatting disinformation and misinformation today is challenging, not only because there are insufficient resources to debunk falsehoods but also because there remains a need to identify flows of such information with an understanding of structural features and patterns of influence.

Much thought has been given to offensive and defensive strategies of sociopolitical communication in a military-strategic sense (e.g., Huhtinen et al., 2021), but looking at structural features in conjunction with the content can allow us to uncover communicative strategies, as well as the macro structures of information warfare. Given that in online environments there is a plethora of meta-data components (IP addresses), a mediatized (strategic) influence framework should consider all the elements available to construct the context—not only content, but also timing and sociotechnical structural elements, in addition to content manipulation that has been largely discussed in propaganda studies. In addition, scrutiny of the actors involved and content circulation can allow researchers to gain a more complete picture of online influence, in addition to a focus on the content that is distributed and shared. Crucially, we cannot assume that all participation online is genuine.

The Baltic states have been particularly targeted by Russian soft power. Thus, questions of misinformation and disinformation and targeted influence should be taken seriously. Yet, combatting disinformation is a hard task and typically involves pre-bunking (preemptive mechanisms) and debunking (through rebuttals). Recent study empirically demonstrates that debunking is meaningful for groups of populations that are embedded in the ideologies of entrenched oppositive beliefs (see e.g., Schmid & Betsch, 2019). Debunking efforts in Lithuania include grassroot initiatives such as Lithuanian elves. Lithuanian elves describe themselves as concerned citizens who monitor news portal comments and call out users whose comments suspiciously resemble Russian trolling posts (Debunk.eu, 2020). The media coverage of anti-Russian trolling initiatives has included interviews of some of these volunteer elves. One such elf, using the pseudonym “Hawk,” claimed that elves aim to act only defensively in online news portals (Sengupta, 2019). They neither engage in cyberattacks nor disseminate counter-propaganda. Although this story about Lithuanian elves was released in 2019, elf operations started earlier with the eruption of the ongoing Russia-Ukraine conflict that brought to light

the battlefield of facts in online spheres and its power to create chaos and foster disinformation.

In Lithuania, elves have emerged as a key opposition group combating Russian trolling. Yet, elves are placed at a disadvantage because they function as a reactive force—that is, a debunking force—rather than a proactive one. Their rhetorical tactics involving disinformation online are typically infused with affect. At times, they resemble typical trolling techniques that use dismissive and offensive tones, and their efforts have been subsequently critiqued by urging them to resist such responses (Vasiliauskaitė, 2021). Further, elves and regular citizens are not necessarily trained in the patterns or tactics of influence they wish to counter, while their adversaries have become increasingly sophisticated in this regard.

If one understands the structure of information and influence, one can utilize it to one's advantage. These days, in contrast to an era of analog disinformation, one can create chaos through comments on a news story or social media post by working with a decentralized information dispatch, compared to a centralized one, and focus on destabilizing common understanding by questioning existing information and tapping preexisting beliefs rather than working on the creation of new information. In addition, various actors can be employed to achieve these goals that go beyond human actors.

To combat disinformation in the current media landscape, it is not sufficient to rely on debunking opposition. One needs to understand how the flows of disinformation work and how to create messaging that is not merely reactive, but also proactive. Furthermore, choosing the right tone to express such proactive messaging remains the key to countering the "zero sum game" argument where trolls and bots are equated, by the skeptics, with elves. Such are the lessons from Lithuania. They remain relevant for researchers and strategists concerned with state-backed information campaigns emanating from Russia and elsewhere—and directed anywhere.

Acknowledgments Author expresses gratitude to research assistant Brandon Niezgoda who helped systemize some of the literature review presented in this chapter.

References

Allport, F. H. (1920). The influence of the group upon association and thought. *Journal of Experimental Psychology, 3*, 159–182.

Anagnostopoulos, A., Kumar, R., & Mahdian, M. (2008). Influence and correlation in social networks. In *Proceedings of the 14th ACM SIGKDD International Conference on knowledge discovery and data mining* (pp. 7–15).

Bakshy, E., Hofman, J. M., Mason, W. A., & Watts, D. J. (2011). Everyone's an influencer: Quantifying influence on Twitter. In *Proceedings of the fourth ACM International Conference on web search and data mining* (pp. 65–74).

Baños, R. A., Borge-Holthoefer, J., & Moreno, Y. (2013). The role of hidden influentials in the diffusion of online information cascades. *EPJ Data Science, 21*, 6.

Bedford, S., & Vinatier, L. (2018). Resisting the irresistible: 'Failed opposition' in Azerbaijan and Belarus revisited. *Government and Opposition, 54*(4), 686–714.

Bessi, A., & Ferrara, E. (2016). Social bots distort the 2016 U.S. presidential election online discussion. *First Monday, 21*(11). https://firstmonday.org/article/view/7090/5653

Botometer. (n.d.). Botometer. Accessed from https://botometer.iuni.iu.edu/#!/

Bruns, A., Harrington, S., & Hurcombe, E. (2020). 'Corona? 5G? Or both?': The dynamics of COVID-19/5G conspiracy theories on Facebook. *Media International Australia*. https://journals.sagepub.com/doi/full/10.1177/1329878X20946113

Burt, R. S. (1976). Positions in networks. *Social Forces, 55*(1), 93–122.

Canavan, J. (2005, October). The evolution of malicious IRC bots. *Virus Bulletin Conference*. https://www.semanticscholar.org/paper/The-evolution-of-malicious-IRC-bots-Canavan/4fb473e4741a5d9d157d075c6747a924eb22fa72

Chen, A. (2015, June 7). The agency. *The New York Times Magazine*. https://www.nytimes.com/2015/06/07/magazine/the-agency.html

Cohen, A. R. (1959). Some implications of self-esteem for social influence. In C. Hovland & I. L. Janis (Eds.), *Personality and persuasibility* (pp. 102–120). Yale University Press.

Cronin, B., & Crawford, H. (1999). Information warfare: Its application in military and civilian contexts. *Information Society, 15*(4), 257–263.

Daniel, F., & Millimaggi, A. (2020). On Twitter bots behaving badly: A manual and automated analysis of Python code patterns on GitHub. *Journal of Web Engineering, 18*(8), 1–36.

DebunkEU. (2020). *About elves*. https://debunk.eu/about-elves/

Deutsch, M., & Gerard, H. B. (1955). A study of normative and informational social influences upon individual judgment. *The Journal of Abnormal and Social Psychology, 51*(3), 629–636.

Dion, K. K., & Stein, S. (1978). Physical attractiveness and interpersonal influence. *Journal of Experimental Social Psychology, 14*(1), 97–108.

Dukalskis, A. (2017). *The authoritarian public sphere: Legitimation and autocratic power in North Korea, Burma, and China*. Routledge.

Edelstein, S., & Edwards, J. (2002). If you build it, they will come: Building learning communities through threaded discussions. *eLearn Magazine, 4*, 3.

Ferrara, E., Varol, O., Davis, C., Menczer, F., & Flammini, A. (2016). The rise of social bots. *Communications of the ACM, 59*(7), 96–104.

Fisher, S., & Lubin, A. (1958). Distance as a determinant of influence in a two-person serial interaction situation. *The Journal of Abnormal and Social Psychology, 56*(2), 230–238.

Friedkin, N. E. (2006). *A structural theory of social influence* (Vol. 13). Cambridge University Press.

Garbačiauskaitė-Budrienė, M. (2016, June 30). *Garbačiauskaitė-Budrienė. Atpažink Kremliaus trolį*. Delfi.lt. https://www.delfi.lt/news/ringas/lit/m-garbaciauskaite-budriene-atpazink-kremliaus-troli.d?id=71642580

Geiger, R. S. (2018). *The lives of bots*. arXiv preprint arXiv:1810.09590.

Golovchenko, Y., Hartmann, M., & Adler-Nissen, R. (2018). State, media and civil society in the information warfare over Ukraine: Citizen curators of digital disinformation. *International Affairs, 94*(5), 975–994.

González-Bailón, S., Borge-Holthoefer, J., & Moreno, Y. (2013). Broadcasters and hidden influentials in online protest diffusion. *American Behavioral Scientist, 577*, 943–965.

Gorwa, R., & Guilbeault, D. (2018). Unpacking the social media bot: A typology to guide research and policy. *Policy & Internet, 12*(2), 225–248.

Herring, S. C. (1999). Interactional coherence in CMC. *Journal of Computer-Mediated Communication, 4*(4).

Herring, S. C. (2003). Gender and power in on-line communication. In J. Holmes & M. Meyerhoff (Eds.), *The handbook of language and gender* (pp. 202–228). Blackwell.

Hjorth, F., & Adler-Nissen, R. (2019). Ideological asymmetry in the reach of pro-Russian digital disinformation to United States audiences. *Journal of Communication, 69*(2), 168–192.

Hovland, C. I., & Weiss, W. (1951). The influence of source credibility on communication effectiveness. *Public Opinion Quarterly, 15*(4), 635–650.

Huhtinen, A. M., Kotilainen, N., Särmä, S., & Streng, M. (2021). Information influence in hybrid environment: Reflexive control as an analytical tool for understanding warfare in social media. In *Research anthology on fake news, political warfare, and combatting the spread of misinformation* (pp. 243–259). IGI Global.

Jamieson, K. H. (2018). *Cyberwar: How Russian hackers and trolls helped elect a president, what we don't, can't, and do know*. Oxford University Press.

Kalmar, P. (2010). Bootstrapping websites for classification of organization names on Twitter. In *CLEF Notebook Papers/LABs/Workshops,* 2(6). http://clef2010.clef-initiative.eu/resources/proceedings/clef2010labs_submission_78.pdf

Katz, E., & Lazarsfeld, P. (1955). *Personal influence: The part played by people in the flow of mass communications.* The Free Press.

Khan, Z., & Jarvenpaa, S. L. (2010). Exploring temporal coordination of events with Facebook.com. *Journal of Information Technology, 25*(2), 137–151.

Kuk, G. (2006). Strategic interaction and knowledge sharing in the KDE developer mailing list. *Management Science, 52*(7), 1031–1042.

LRT. (2019, September 27). *More fake news target NATO's presence in Lithuania.* https://www.lrt.lt/en/news-in-english/19/1101632/more-fake-news-target-nato-s-presence-in-lithuania

Mandernach, B. J., Gonzales, R. M., & Garrett, A. L. (2006). An examination of online instructor presence via threaded discussion participation. *Journal of Online Learning and Teaching, 2*(4), 248–260.

Messias, J., Schmidt, L., Oliveira, R., & Benevenuto, F. (2013). You followed my bot! Transforming robots into influential users in Twitter. *First Monday, 18*(7). https://doi.org/10.5210/fm.v18i7.4217

Meyer, K. A. (2003). Face-to-face versus threaded discussions: The role of time and higher-order thinking. *Journal of Asynchronous Learning Networks, 7*(3), 55–65.

Miller, R. L., & Benz, J. J. (2008). Techniques for encouraging peer collaboration: Online threaded discussion or fishbowl interaction. *Journal of Instructional Psychology, 35*(1), 87–94.

Mislove, A., Lehmann, S., Ahn, Y. Y., Onnela, J. P., & Rosenquist, J. N. (2011). Understanding the demographics of Twitter users. *Proceedings of the International AAAI Conference on Web and Social Media, 5*(1). https://ojs.aaai.org/index.php/ICWSM/article/view/14168

Orenstein, M. A. (2019). *The lands in between: Russia vs. the West and the new politics of hybrid war.* Oxford University Press.

Pearce, K. E., Vitak, J., & Barta, K. (2018). Privacy at the margins| socially mediated visibility: Friendship and dissent in authoritarian Azerbaijan. *International Journal of Communication, 12*(22). https://ijoc.org/index.php/ijoc/article/view/7039

Pelz, D. C. (1952). Influence: A key to effective leadership in the first-line supervisor. *Personnel, 29*, 209–217.

Raven, B. H. (1958). Legitimate power, coercive power, and observability in social influence. *Sociometry, 21*, 83–97.

Raven, B. H. (1965). Social influence and power. In I. D. Steiner & M. Fishbein (Eds.), *Current studies in social psychology* (pp. 371–382). Holt, Rinehart & Winston.

Romero, D. M., Meeder, B., & Kleinberg, J. (2011, March). Differences in the mechanics of information diffusion across topics: Idioms, political hashtags, and complex contagion on Twitter. In *Proceedings of the 20th international conference on world wide web* (pp. 695–704). ACM.

Sabbagh, D. (2020, July 30). Russia-aligned hackers since 2017 have been running anti-NATO fake news campaign. *The Guardian*. https://www.theguardian.com/technology/2020/jul/30/russia-aligned-hackers-running-anti-nato-fake-news-campaign-report-poland-lithuania

Schein, E. H. (1960). Interpersonal communication, group solidarity, and social influence. *Sociometry, 23*(2), 148–161.

Schmid, P., & Betsch, C. (2019). Effective strategies for rebutting science denialism in public discussions. *Nature Human Behaviour, 3*(9), 931–939.

Sengupta, K. (2019, July). Meet the elves, Lithuania's digital citizen army confronting Russian trolls. *The Independent*. https://www.independent.co.uk/news/world/europe/lithuania-elves-russia-election-tampering-online-cybercrime-hackerskremlin-a9008931.html

Simons, G. (2015). Perception of Russia's soft power and influence in the Baltic States. *Public Relations Review, 41*(1), 1–13.

Smith, M., Cadiz, J. J., & Burkhalter, B. (2000, December). Conversation trees and threaded chats. In *Proceedings of the 2000 ACM conference on computer supported cooperative work* (pp. 97–105). ACM.

Stukal, D., Sanovich, S., Tucker, J. A., & Bonneau, R. (2019). For whom the bot tolls: A neural networks approach to measuring political orientation of Twitter bots in Russia. *SAGE Open, 9*(2). https://journals.sagepub.com/doi/full/10.1177/2158244019827715

Sun, B., & Ng, V. T. (2013). *Identifying influential users by their postings in social networks*. Springer Berlin Heidelberg.

Toepfl, F. (2018). Innovating consultative authoritarianism: Internet votes as a novel digital tool to stabilize non-democratic rule in Russia. *New Media & Society, 20*(3), 956–972.

Toepfl, F., & Litvinenko, A. (2018). Transferring control from the backend to the frontend: A comparison of the discourse architectures of comment sections on news websites across the post-Soviet world. *New Media & Society, 20*(8), 2844–2861.

UNESCO. (2018). *Journalism, 'Fake News' & disinformation: Handbook for journalism education and training*. https://unesdoc.unesco.org/ark:/48223/pf0000265552/PDF/265552eng.pdf.multi

Valeriano, B., Jensen, B. M., & Maness, R. C. (2018). *Cyber strategy: The evolving character of power and coercion*. Oxford University Press.

Van Dijk, T. A. (1998). *Ideology: A multidisciplinary approach*. Sage.

Vasiliauskaitė, N. (2021, January 11). *Nida Vasiliauskaitė. Laisvės propaganda.* Delfi.lt. https://www.delfi.lt/news/ringas/lit/nida-vasiliauskaite-laisves-propaganda.d?id=86197159

VDU (2021). *UNESCO-UNITWIN Medijų ir informacinio raštingumo tyrimų centras.* https://pmdf.vdu.lt/mokslas/mokslo-centrai/unesco-unitwin-mediju-ir-informacinio-rastingumo-tyrimu-centras/

Williams, R. S., & Humphrey, R. (2007). Understanding and fostering interaction in threaded discussion. *Journal of Asynchronous Learning Networks, 11*(2), 129–143.

Woolley, S. C., & Howard, P. N. (Eds.). (2018). *Computational propaganda: Political parties, politicians, and political manipulation on social media*. Oxford University Press.

Wright, M. (1943). The influence of frustration upon the social relations of young children. *Journal of Personality, 12*(2), 111–122.

Zelenkauskaite, A. (2022). *Creating chaos online: Disinformation and subverted post-publics*. University of Michigan Press.

Zelenkauskaite, A., & Balduccini, M. (2017). "Information warfare" and online news commenting: Analyzing forces of social influence through location-based commenting user typology. *Social Media + Society, 3*(3). https://journals.sagepub.com/doi/full/10.1177/2056305117718468

Zelenkauskaite, A., & Niezgoda, B. (2017). "Stop Kremlin trolls:" Ideological trolling as calling out, rebuttal, and reactions on online news portal commenting. *First Monday, 22*(5). https://doi.org/10.5210/fm.v22i5.7795