Deadly Lies: How Twitter Fuels Misinformation Campaigns with Lethal Impacts

A Research Paper submitted to the Department of Engineering and Society

Presented to the Faculty of the School of Engineering and Applied Science

University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements for the Degree

Bachelor of Science, School of Engineering

Maxwell John Penders

Spring 2025

On my honor as a University Student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments

Advisor

Caitlin D. Wylie, Department of Engineering and Society

Introduction

On March 11th, 2020, the World Health Organization officially declared COVID-19 to be a pandemic. As millions around the world rushed to get the vaccination, another pandemic made itself clear: that of COVID-19 misinformation. In the United States, nearly three-quarters of the population was exposed to COVID-19 misinformation, with the more highly exposed groups being far less likely to get vaccinated for the virus (Neely et al., 2022). The cost of people not getting vaccinated is estimated to be more than 230 thousand lives in the United States alone (Jia et al., 2023). Misinformation is commonly understood to be false information, but it is important here to understand its more specific meaning, contrasted with disinformation. While misinformation is unintentionally false information, disinformation is intentionally false, often with malicious intent. As the pandemic progressed, hundreds of thousands more people would die due to a refusal to get the recommended vaccinations (Jia et al., 2023). Given that mis- and dis-information have the ability to impact our lives in such powerful ways, this paper aims to review existing literature to determine how mis- and dis-information are formed, spread through the internet by observing the Twitter social media platform, and are adapted into the belief systems of everyday people. From that, possible vectors with which to prevent this misinformation can be identified. To best do this, this paper uses Actor-Network Theory, which organizes the relevant actors to the problem and arranges them as a network which communicates, changes, and orders itself (Law, 1992).

Sources of Mis- and Dis-information

Where does a false Tweet originate? By the time you see it on Twitter, it is likely a real person reposting it, often without any malintent. This is especially true in the case of memes and other images which can be easily screenshotted or copied without linking back to the original

account (Dupuis & Williams, 2019). However, understanding the role hidden key originators play in the spread of mis- and dis-information offers methods to counter their work.

Foreign States

Powerful foreign states, notably US adversaries Russia and China, can be found behind many of the stories and ideas which end up as Tweets. As China reportedly pumps "billions" of dollars into pro-Russian disinformation in the US and anti-Taiwanese messaging elsewhere, Russia spends around 1.5 billion dollars yearly on its state-owned and state-funded media which act as polarizing forces around the world (*APRSA 2024 | Chapter 5*, 2024; Michałowska & Kubś, 2022; Wintour, 2023).

In the case of Russia, the source propaganda is direct, coherent, and personalized to the target group. They take a "two-armed" approach to the disinformation they choose to spread: longer term strategic aims (often general pro-Russia, anti-West, and anti-NATO sentiments) combined with specific tactical goals (for example, painting Ukraine as fascist and Nazified) (Two Years On, 2024). Both of these approaches are what military strategists call "breaking the kill chain" – the best deterrence for the American defense of Ukraine is directing the sympathies of the American people away from Ukraine (McGeehan, 2018). Years before it annexed Crimea, the Kremlin had already started ramping up funding towards its state media. Once it began eyeing Ukraine, the funding increased further. As it became clear that the success of Ukraine in sustaining itself during the Russo-Ukrainian war was due to American support, spending increased further, with more of it being invested in the American dis-information campaign according to spending analysis by Lithuanian fact-checking group Debunk.org (Michałowska & Kubś, 2022; Two Years On, 2024).

While the state funding of media allowed the production of a multitude of propaganda, those media groups never obtained mainstream use outside of Russia. Instead, the Kremlin's media groups utilized personalization to increase their impact size. In Germany, Russia focused on three main groups: the far-right, particularly those associated with the AfD (a German political party), the far-left, who in Germany have some crossover with the far-right in terms of their anti-NATO, anti-establishment, and anti-West stances, especially due to the historical prevalence of Communism in Russia, and the Russian-German diaspora, who are often Post-Soviet migrants or descendants of Germans who had migrated to Russia then migrated back (called the Russlanddeutsche) and may have cultural or linguistic ties to Russia, who are told scare stories about the failures and strife of Germany as well as allegations of Germanic Russophobia (Make Germany Great Again, 2017).

Increasingly, these Russian state media outlets worked with the AfD to obtain mutual goals. Combined with the Russian state media apparatus, the tweeting of Russian news articles by "bot-like" accounts, and the use of Russian media news articles by prominent online AfD influencers, the AfD dominated the 2017 election on Twitter. Despite its small membership, the AfD achieved the most Twitter traffic of any party that election, and converted it into a massive swing of parliamentary power in their favor (*Bundeswahlleiterin*, 2017; Make Germany Great Again, 2017).

In the United States, Russia has utilized similar methods. Prior to the invasion of Ukraine, Russian media efforts focused on fringe but motivated groups: the far-right, the far-left (although to a lesser degree than in Germany), conspiracy theory-oriented communities, and anti-vaxxers. With the advent of the COVID19 pandemic, Russia added on the denialist communities (Two Years On, 2024). As the Russo-Ukrainian war progressed, Russia targeted isolationism and

conspiracies of NATO aggression, even secretly funding a group of six well-known conservative personalities whose talking points were already aligned with the Kremlin – including massively popular influencers Tim Pool and Dave Rubin (Suderman & Swenson, 2024). Traditional American conservative organizations have been targeted, too. Maria Butina was arrested and sentenced to 18 months after she pleaded guilty for one count of conspiracy to act as a Russian agent in the U.S. without registering with the Justice Department. She utilized her "professed interest in gun rights and American politics to establish relationships with senior members of the National Rifle Association and Republican party" (Lucas, 2019). While these instances are ones known to the public, it is highly likely that Russia still has many tendrils that have not yet been exposed. The goal of connections like these is to associate with a party, particularly radical elements of it, and unofficially create a symbiotic relationship between the success of the party and the spread of Russian state media.

In contrast with the deliberate success of Russian state dis-information, China has failed to establish a strong foothold on the US (concerns regarding TikTok-specific content are outside the scope of this paper). Rather than creating and establishing state media groups among personalized radical political groups, China has used "Spamoflage", which is a network of fake accounts on Twitter and other services pretending to be American voters and U.S. soldiers. Rather than focus on a pre-existing political group, China has tried to create a new one (with slight overlap with the authoritarian far-left) which criticizes both Democrats and Republicans, pretending to be U.S. citizens "frustrated by American politics and the West" (Thomas, 2024). Notably, China also seems to much more heavily rely on AI-generated images than Russia, attempting to create convincing images and memes to target a specific narrative, although Russia has attempted utilizing Deepfake technology in their anti-Ukraine propaganda (Allyn, 2022).

This approach has net far less success than Russian-style information warfare, with the "Spamouflage" accounts receiving low impression amounts, but there are likely far more fake accounts then were found, meaning even low-impact accounts could sum to have a meaningful impact. It is also important to note that while Russia certainly had a pro-Trump bias in the 2024 U.S. election, U.S. intelligence believed China had no plans to influence it (Thomas, 2024). This means that a more specifically targeted Chinese dis-information campaign of a similar nature may be more impactful.

Individuals

In an area where follower counts equate to power and profit, some individuals weaponize falsehoods for maximum reach. In common parlance, "clout" means influence and/or power, especially in regard to social media audience, and as a proxy, number of followers. With 75% of young social media users having some desire to become an influencer, it is no surprise that many people would take the fast route to popularity if they had the ability (Fetter et al., 2023). It's not just the clout, either. Though pay numbers vary too widely to give an exact amount, people have made a significant sum through ad-filled websites (such as blogs) that garner significant impressions (Wasilczuk & Heinonen, 2023, p. 9).

Due to these incentives, some individuals have created fake news websites, social media accounts, and blogs which consistently post factually false, highly provocative, and trendy articles with eye-catching headlines. Some of these "mis-influencers" focus on the highly emotional topic of politics: Jestin Coler ran a fake news website called the *Denver Guardian*, where he published a report entitled "FBI Agent Suspected In Hillary Emails Leaks Found Dead In Apparent Murder-Suicide". This fully fabricated report net 1.6 million views (Sydell, 2016; Verbruggen, n.d.). Others focus on pop culture: several YouTube channels utilizing a mix of AI-

generated and human-manipulated media published videos with fake information regarding Black celebrities, including Steve Harvey and Denzel Washington (Tenbarge, 2024). The median number of combined views for these channels was 21 million.

Mis-influencers aren't generally concerned with the external impact of their messaging, the subjects they discuss, or broader internet safety. Instead, they are concerned with some combination of clout and money and do not seem to be dissuaded by the inherent risk of lawsuits. While the ability of any single mis-influencer to have an impact is limited by the size of their operation, the sheer ease with which any individual may generate and spread misinformation makes it a legitimate source of concern in defeating misinformation's spread.

Twitter

Once a false story has been initially Tweeted, it is not guaranteed to succeed. However, unlike a discussion within a classroom or among friends, Twitter (as with many other social media platforms) is highly effective at building networks of people with homogenous beliefs. If each Twitter user tends to follow and interact with other users with similar beliefs and therefore see Tweets which highly align with their belief systems (regardless of the veracity of the Tweet). As a result, Twitter's algorithm has a tendency to promote low-credibility content with high-engagement generated by the highly influential thought leaders within the targeted community (Corsi, 2024).

Twitter's primary motivation is fairly obvious: they built a highly engaging and possibly addictive algorithm (Zakon, 2020) to increase the time users spend on the platform and therefore the time spent viewing ads. A monetary incentive continues to be present in Twitter's opposition to proposed legal counters to dis-information. Elon Musk and Twitter filed a lawsuit against

California's AB 2655, which would make "social media companies responsible for... election-related disinformation" (*Elon Musk, X Sue*, 2024). This is not necessarily due to any philosophical disagreements with the proposal (although tech executives famously tend towards libertarianism) but with the ramifications of the law: if Twitter is responsible for the election disinformation it hosts, the platform must then exponentially increase spending on AI recognition of disinformation and humans to review possible misleading posts, or face the financial consequences of disobeying the law. It is vital that Twitter is not to be viewed as an enemy, traitor, or "bad guy." Like any other company, it is responding to market incentives in order to provide value to its shareholders. Rather, it is a stakeholder in the problem and in a market economy needs to be worked with as a part of the solution.

Despite the prevalence of misinformation on Twitter, it has attempted to make a positive impact (or at least improve its image) through the rollout of the Community Notes system in January 2021. This system uses what is called a "bridging algorithm", which attempts to discount the impact of, for example, a progressive voting for a note on a right-leaning post, in order to reduce the partisanship of which notes are approved (Stafford, 2025). However, it is working under the assumption that political leaning is the primary polarizer, and if the site were to become politically homogeneous, the system may run the risk of instead splitting along the lines of naïve user and expert. The other goal of Community Notes, exemplified by the recent rollout of Lightning Notes (Community Notes [@CommunityNotes], 2024), is to label false Tweets as rapidly as possible.

Due to the nature of the Twitter algorithm, the vast majority of impressions and interactions with a Tweet happen within a short period of time. The half-life of a Tweet, meaning the amount of time it takes until half of the expected lifetime impressions of the Tweet occur, is

about 80 minutes (Pfeffer et al., 2023). Each small improvement in time to note could mean exponential reductions in overall impact of the false Tweet.

Users

It is important to understand that the users who spread false information, once deliberately created, are typically regular users with no malintent. Bots tend to spread both true and false news at the same rate (Vosoughi et al., 2018) and there is no evidence of any organization or government having sufficient resources to meaningfully contribute to the spread so directly. Understanding *why* these users believe and subsequently spread misinformation is highly important in understanding *how* to limit the spread of misinformation in this section of the network.

As the world spends more and more time online, the way adolescents discover their identity becomes social media and internet oriented. For example, as Farah Pandith discusses in her book *How We Win*, "Driven by their identity crisis, Muslim youth have gone online for answers, seeking information about Islam and opportunities to share their religious beliefs online" (2019, p. 193). What is true for young Muslims in Azerbaijan is true for online "incels" (involuntary celibates) in the United States. A struggle with one's own personal identity in the age of the internet leaves them susceptible to groups with rigid rules, extreme views, and large levels of in-group vs out-group sentiment (Pandith, 2019; Van Bavel et al., 2024). As these uncertain youth solidify themselves within these right-wing extremist circles, they are being exposed to social media posts, Russian state news articles, and political narratives from their in-group peers.

Even beyond the identity-seeking of the youth, older adults have still found themselves enjoying the affirmation provided by social media. Those over the age of 30 tend to stay fairly rigid in their political dispositions for the rest of their lives (Sears & Funk, 1999), but that makes them all the more susceptible to misinformation when it is in their favor. Regardless of age, the social in-group yearning is still a major factor, one that has been documented countless times (Beauvais, 2022; Herrero-Diz et al., 2020; O'Connor & Weatherall, 2019; Pandith, 2019; Van Bavel et al., 2024). Despite a slightly higher resistance to fake news (Soetekouw & Angelopoulos, 2024), adults often act as the respected leaders within these communities, acting as the local source for the spread. Although older generations have been slower to incorporate technology into their lives it is unquestionable that the adoption of technology is shared by all age groups (Wang et al., 2018). As social communication moves online, exposure to high levels of algorithmic bias is nearly guaranteed.

Defending Against Mis- and Dis-information

It should be clear by this point that the problem of mis- and dis-information is one that needs to be addressed on a scale far beyond the scope of any single Western country. Because foreign adversaries are known to be behind many of these false stories, it is not unreasonable to consider the problem one of international security and defense. While the US military has written about the issue from one of national defense (McGeehan, 2018), it perhaps is more meaningfully addressed as a covert offensive against NATO. If it is addressed as such, there are some incredible examples of success among the allies: the founding member Norway and the recent additions of Finland and Sweden.

The least flashy but perhaps most effective weapon against mis- and dis-information in the NATO toolkit is that of media and information literacy (MIL), a term borrowed from the

Swedish government's "media och informationskunnighet" (Giannetto, 2023). Although the three countries integrate the study MIL into education in different ways, some elements of their success remain constant. First, the curricula "explicitly facilitate [critical thinking (CT)] skills as a separate course while embedding CT into subject coursework", as compared to the United States, where the curriculum "implicitly embeds CT into subject coursework only" (Horn & Veermans, 2019, p. 24). Second, media literacy is a shared effort across many levels, from multinational organizations such as the EU all the way down to the local libraries (Giannetto, 2023). This allows each level of education to focus on the timely problems faced at that level, from national messaging to specific messaging targeting a local minority group. Finally, this education is not purely targeted at youth, but across all stages of life, from elementary schools to universities to NGOs to senior community centers (Giannetto, 2023). While there are more country-specific details which may be promising avenues for instantiation within the United States (in Sweden, for example, media literacy is pitched as a civil defense used to strengthen the population against foreign attacks) (Giannetto, 2023), it is more important here to discuss the several other less proven defense vectors and their efficacy.

Despite the recent outsized role of Twitter in the spread of mis- and dis-information, extremist groups have shown a high willingness to shift platforms when faced with pressure, making it ineffective to target any platform singularly (Bateman & Jackson, 2024). To deal with the algorithmic amplification of mis- and dis-information as a whole, there has been recent advocacy for the creation of a stakeholder-managed government body tasked with rendering "black box" social media algorithms which contribute to the amplification of disinformation "appropriately intelligible to users" (Sun, 2022, p. 371). While the organization would be more effective if it remained aligned with the general goals of the MIL agenda discussed above, it

would also remain specific enough to deal with social media-specific threats such as algorithmic addiction.

A less powerful strategy attempting to combine the MIL and algorithmic regulatory goals is the labeling and fact-checking of social media posts. As discussed before, Twitter has had a long-standing program, Community Notes, designed to put community-created fact checks on popular misleading Tweets before reaching the critical mass of views. This program has seemed to have beneficial impacts on certain metrics of fake Tweets such as deletion rates, spread rates, and retweets, but recent analyses have found a failure to intervene early enough within the lifespan of the Tweet (Chuai, Pilarski, et al., 2024; Chuai, Tian, et al., 2024; Drolsbach & Pröllochs, 2023). Despite its significant efficacy, it is clearly not sufficient: despite being introduced in early 2021, roughly a year prior to the Russian invasion of mainland Ukraine (as compared to the 2014 Crimean invasion and annexation), Kremlin misinformation has still done its job in increasing U.S. polarization and support for Russia. Furthermore, as Twitter shifts to become more widely right-leaning (Faverio, 2023), the Community Notes algorithm becomes more and more unable to label right-leaning disinformation.

Conclusion

The intense level at which mis- and dis-information have ingrained themselves into the online Twitter experience should be clear at this point, and the not-so-distant loss of life during the COVID-19 pandemic points towards the importance of addressing the issue. Based on what we know about this problem, the best way to address the threat is through democratic advocacy. In order to implement media literacy programs from elementary school to the retirement home, introduce government oversight of black box algorithms, and reduce our political polarization, individuals must care and advocate for these improvements to occur.

There is also more work to be done. While Twitter is certainly a large actor in the problem of algorithmic spread of mis- and dis-information, other social media platforms need addressing, both for their unique algorithms and for the unique users and communities being targeted on the platform. And, if some platforms have been willing to break the black box, what can we learn from it?

References

Allyn, B. (2022, March 16). Deepfake video of Zelenskyy could be “tip of the iceberg” in info war, experts warn. *NPR*. https://www.npr.org/2022/03/16/1087062648/deepfake-video-zelenskyy-experts-war-manipulation-ukraine-russia

APRSA 2024 | Chapter 5: Driving Wedges: China’s Disinformation Campaigns in the Asia-Pacific. (2024, May). IISS. https://www.iiss.org/publications/strategic-dossiers/asia-pacific-regional-security-assessment-2024/chapter-5/

Bateman, J., & Jackson, D. (2024). *Countering Disinformation Effectively: An Evidence-Based Policy Guide* (p. 130). https://carnegieendowment.org/research/2024/01/countering-disinformation-effectively-an-evidence-based-policy-guide?lang=en

Beauvais, C. (2022). Fake news: Why do we believe it? *Joint Bone Spine*, *89*(4), 105371. https://doi.org/10.1016/j.jbspin.2022.105371

Chuai, Y., Pilarski, M., Renault, T., Restrepo-Amariles, D., Troussel-Clément, A., Lenzini, G., & Pröllochs, N. (2024). *Community-based fact-checking reduces the spread of misleading posts on social media* (arXiv:2409.08781). arXiv. https://doi.org/10.48550/arXiv.2409.08781

Chuai, Y., Tian, H., Pröllochs, N., & Lenzini, G. (2024). *Did the Roll-Out of Community Notes Reduce Engagement With Misinformation on X/Twitter?* https://doi.org/10.1145/3686967

Community Notes [@CommunityNotes]. (2024, October 29). *Introducing Lightning Notes ⚡ People want fast notes. So we re-architected the system to score notes faster. Since rolling this out, we’ve seen notes go live in as little as 14m33s after being written, and 18m20s after the post itself was written. Ludicrously fast? Now reality! Https://t.co/CuNsGEIKJY* [Tweet]. Twitter. https://x.com/CommunityNotes/status/1851337944822325253

Corsi, G. (2024). Evaluating Twitter's algorithmic amplification of low-credibility content: An observational study. *EPJ Data Science*, *13*(1), Article 1. https://doi.org/10.1140/epjds/s13688-024-00456-3

Drolsbach, C. P., & Pröllochs, N. (2023). Diffusion of Community Fact-Checked Misinformation on Twitter. *Proc. ACM Hum.-Comput. Interact.*, *7*(CSCW2), 267:1-267:22. https://doi.org/10.1145/3610058

Dupuis, M. J., & Williams, A. (2019). The Spread of Disinformation on the Web: An Examination of Memes on Social Networking. *2019 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)*, 1412–1418. https://doi.org/10.1109/SmartWorld-UIC-ATC-SCALCOM-IOP-SCI.2019.00256

Elon Musk, X Sue to Halt CA's New Flagship Anti-Disinformation Law. (2024, December 9). Common Cause California. https://www.commoncause.org/california/press/elon-musk-x-sue-to-halt-cas-new-flagship-anti-disinformation-law/

Ergebnisse Deutschland—Die Bundeswahlleiterin. (2017). https://www.bundeswahlleiterin.de/bundestagswahlen/2017/ergebnisse/bund-99.html

Faverio, M. (2023, May 17). Majority of U.S. Twitter users say they've taken a break from the platform in the past year. *Pew Research Center*. https://www.pewresearch.org/short-reads/2023/05/17/majority-of-us-twitter-users-say-theyve-taken-a-break-from-the-platform-in-the-past-year/

Fetter, S., Coyne, P., Monk, S., & Woodruff, S. J. (2023). An Exploration of Social Media Users' Desires to Become Social Media Influencers. *Media Watch*, *14*(2), 200–216. https://doi.org/10.1177/09760911231159690

Giannetto, M. (2023). *An analysis on the media literacy efforts of Finland, Sweden, and Norway* [Master Thesis]. Forsvarets høgskole.

Herrero-Diz, P., Conde-Jiménez, J., & Reyes de Cózar, S. (2020). Teens' Motivations to Spread Fake News on WhatsApp. *Social Media + Society*, *6*(3), 2056305120942879. https://doi.org/10.1177/2056305120942879

Horn, S., & Veermans, K. (2019). Critical thinking efficacy and transfer skills defend against 'fake news' at an international school in Finland. *Journal of Research in International Education*, *18*(1), 23–41. https://doi.org/10.1177/1475240919830003

How We Win. (2019). In F. Pandith, *How we win: How cutting-edge entrepreneurs, political visionaries, enlightened business leaders, and social media mavens can defeat the extremist threat* (First edition, pp. 1–4). Custom House.

Jia, K. M., Hanage, W. P., Lipsitch, M., Johnson, A. G., Amin, A. B., Ali, A. R., Scobie, H. M., & Swerdlow, D. L. (2023). Estimated preventable COVID-19-associated deaths due to non-vaccination in the United States. *European Journal of Epidemiology*, *38*(11), 1125. https://doi.org/10.1007/s10654-023-01006-3

Law, J. (1992). Notes on the theory of the actor-network: Ordering, strategy, and heterogeneity. *Systems Practice*, *5*(4), 379–393. https://doi.org/10.1007/BF01059830

Lucas, R. (2019, April 26). Russian Agent Maria Butina Sentenced To 18 Months Following Guilty Plea. *NPR*. https://www.npr.org/2019/04/26/716799929/russian-agent-maria-butina-to-be-sentenced-in-federal-court-on-friday

"Make Germany Great Again" – Kremlin, Alt-Right and International Influences in the 2017 German Elections. (2017). ISD Global. https://www.isdglobal.org/isd-publications/make-germany-great-again-kremlin-alt-right-and-international-influences-in-the-2017-german-elections/

McGeehan, T. P. (2018). Countering Russian Disinformation. *The US Army War College Quarterly: Parameters*, *48*(1). https://doi.org/10.55540/0031-1723.2850

Michałowska, A., & Kubś, J. (2022, August 8). *Coining lies. Kremlin spends 1.5 Billion per year to spread disinformation and propaganda*. Debunk.Org. https://www.debunk.org/coining-lies-state-budget-financing-of-russian-propaganda

Neely, S. R., Eldredge, C., Ersing, R., & Remington, C. (2022). Vaccine Hesitancy and Exposure to Misinformation: A Survey Analysis. *Journal of General Internal Medicine*, *37*(1), 179–187. https://doi.org/10.1007/s11606-021-07171-z

O'Connor, C., & Weatherall, J. O. (2019). *The Misinformation Age: How False Beliefs Spread*. Yale University Press.

Pandith, F. (2019). *How we win: How cutting-edge entrepreneurs, political visionaries, enlightened business leaders, and social media mavens can defeat the extremist threat* (First edition). Custom House.

Pfeffer, J., Matter, D., & Sargsyan, A. (2023). *The Half-Life of a Tweet* (arXiv:2302.09654). arXiv. https://doi.org/10.48550/arXiv.2302.09654

Sears, D. O., & Funk, C. L. (1999). Evidence of the Long-Term Persistence of Adults' Political Predispositions. *The Journal of Politics*, *61*(1), 1–28. https://doi.org/10.2307/2647773

Soetekouw, L., & Angelopoulos, S. (2024). Digital Resilience Through Training Protocols: Learning To Identify Fake News On Social Media. *Information Systems Frontiers*, *26*(2), 459–475. https://doi.org/10.1007/s10796-021-10240-7

Stafford, T. (2025, January 14). Do Community Notes work? *Impact of Social Sciences*. https://blogs.lse.ac.uk/impactofsocialsciences/2025/01/14/do-community-notes-work/

Suderman, A., & Swenson, A. (2024, September 5). *Right-wing influencers were duped to work for covert Russian operation, US says*. AP News. https://apnews.com/article/russian-interference-presidential-election-influencers-trump-999435273dd39edf7468c6aa34fad5dd

Sun, H. (2022). Regulating Algorithmic Disinformation. *Columbia Journal of Law & the Arts*, *46*(4), 367–418.

Sydell, L. (2016, November 23). We Tracked Down A Fake-News Creator In The Suburbs. Here's What We Learned. *NPR*. https://www.npr.org/sections/alltechconsidered/2016/11/23/503146770/npr-finds-the-head-of-a-covert-fake-news-operation-in-the-suburbs

Tenbarge, K. (2024, January 30). *Fake news YouTube creators target Black celebrities with AI-generated misinformation*. NBC News. https://www.nbcnews.com/tech/misinformation/ai-deepfake-fake-news-youtube-black-celebrities-rcna133368

Thomas, E. (2024, April 1). Pro-CCP Spamouflage campaign experiments with new tactics targeting the US [Think Tank]. *ISD Global Digital Dispatches*. https://www.isdglobal.org/digital_dispatches/pro-ccp-spamouflage-campaign-experiments-with-new-tactics-targeting-the-us/

Two Years On: An Analysis of Russian State and Pro-Kremlin Information Warfare in the Context of the Invasion of Ukraine. (2024). ISD Global. https://www.isdglobal.org/isd-publications/two-years-on-an-analysis-of-russian-state-and-pro-kremlin-information-warfare-in-the-context-of-the-invasion-of-ukraine/

Van Bavel, J. J., Rathje, S., Vlasceanu, M., & Pretus, C. (2024). Updating the identity-based model of belief: From false belief to the spread of misinformation. *Current Opinion in Psychology*, *56*, 101787. https://doi.org/10.1016/j.copsyc.2023.101787

Verbruggen, Y. (n.d.). *Fake news*. International Bar Association. Retrieved March 3, 2025, from https://www.ibanet.org/article/0adbdb24-c0c2-4cc8-bef8-e9b172dcf12a

Vosoughi, S., Roy, D., & Aral, S. (2018). The spread of true and false news online. *Science*, *359*(6380), 1146–1151. https://doi.org/10.1126/science.aap9559

Wang, K. H., Chen, G., & Chen, H.-G. (2018). Understanding Technology Adoption Behavior by Older Adults. *Social Behavior and Personality: An International Journal*, *46*(5), 801–814. https://doi.org/10.2224/sbp.6483

Wasilczuk, J. E., & Heinonen, J. (2023). *Chapter 9: I love to write and create. Can I earn money doing it? Entrepreneurial process of bloggers*. https://www.elgaronline.com/edcollchap/book/9781035320684/book-part-9781035320684-17.xml

Wintour, P. (2023, February 28). China spends billions on pro-Russia disinformation, US special envoy says. *The Guardian*. https://www.theguardian.com/world/2023/feb/28/china-spends-billions-on-pro-russia-disinformation-us-special-envoy-says

Zakon, A. (2020). Optimized for Addiction: Extending Product Liability Concepts to Defectively Designed Social Media Algorithms and Overcoming the Communications Decency Act Comment. *Wisconsin Law Review*, *2020*(5), 1107–1146.