

Manufacturing Perception:

An Exploration of the Global Landscape of Political Influence Operations on Social Media

By Giuliana Stephanie Treñas

Submitted to OCAD University in partial fulfillment of the requirements for the
degree of Master of Design in Strategic Foresight & Innovation

Toronto, Ontario, Canada, 2025

Copyright Notice

This work is licensed under CC BY-NC-SA 4.0. To view a copy of this license, visit <https://creativecommons.org/licenses/by-nc-sa/4.0/>

You are free to:

Share — copy and redistribute the material in any medium or format

Adapt — remix, transform, and build upon the material

The licensor cannot revoke these freedoms as long as you follow the license terms.

Under the following terms:

Attribution — You must give appropriate credit, provide a link to the license, and indicate if changes were made. You may do so in any reasonable manner, but not in any way that suggests the licensor endorses you or your use.

NonCommercial — You may not use the material for commercial purposes.

ShareAlike — If you remix, transform, or build upon the material, you must distribute your contributions under the same license as the original.

No additional restrictions — You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.

Notices:

You do not have to comply with the license for elements of the material in the public domain or where your use is permitted by an applicable exception or limitation.

No warranties are given. The license may not give you all of the permissions necessary for your intended use. For example, other rights such as publicity, privacy, or moral rights may limit how you use the material.



Abstract

Social media is an entity that has permeated the lives of almost every individual in the modern world. Over the last decade, the number of active social media users globally has more than doubled from 2 billion to 5 billion active users. This is only expected to grow as the internet becomes increasingly accessible worldwide. While much research has been done to uncover the negative effects of social media on mental health, addiction, and misinformation, little can be done to fully address the concerns brought about by various social media platforms, as they have become too complex and expansive to control.

What has become increasingly apparent in recent years is that what people see on social media can shape user perceptions. From online discourse in comments sections, to the widespread popularity of social media influencer marketing, to social phenomena like cancel culture, people are fed a constant stream of content that both consciously and unconsciously reshape how opinions are formed. This influx of information, when paired with social media's engagement-focused algorithms, creates the perfect breeding ground for the rapid spread of disinformation, and the agendas that they bolster.

This research project explores the systemic issues surrounding influence operations carried out on social media platforms. Through the use of a synthesis map, it will present the various stakeholders involved in the process, from top to bottom, and uncover different patterns in influence operation tactics around the world. Furthermore, the project will illustrate the dynamics that have led to inefficiencies in today's solutions. By introducing users to the different players within the system, their agendas, and the pinch points they cause, readers will be able to better understand the power imbalances and tensions that have led to systemic fragmentation. The aim of this systemic exploration is to provide awareness for users by shedding light on the social media ecosystem as it relates to influence and perception, and to aid individuals in making informed decisions as they traverse the internet.

Acknowledgements

I would like to thank the participants of this study for providing me with the insights that have shaped the trajectory of my research. Every conversation has been an invaluable learning experience that has continued to feed the passion behind this project.

I extend my gratitude to my research advisor, Michele Mastroeni, for your guidance and patience throughout our 8 months of collaboration. Thank you for allowing me to achieve a sense of balance by allowing me to explore this topic in depth while also setting me on the right course when I needed direction.

To my program cohort: congratulations on completing the program! Though SFI has taught me a lot, I've also gained an irreplaceable wealth of knowledge through all of you. I'm so lucky to have been granted the opportunity to collaborate and grow with such an amazing group of people.

Finally, I'd like to specially thank my family and friends—both from Toronto and Manila—for sticking with me throughout the chaos that my research year has been, for picking up my late night calls when I needed second opinions, and for providing me with an endless supply of love and support. No words could express how grateful I am to have all of you.

Table of Contents

Introduction.....	6
Scope and Context.....	6
Research Question.....	7
Methodology.....	7
Chapter I: Subject Matter	11
Social Media: A Brief History	11
Engagement and Paid Media Revenue.....	13
Actors, Influence, and Agency	15
Misinformation, Disinformation, and Malinformation.....	18
Influence Operations.....	18
Types of Regulation and Mitigation.....	22
Chapter Summary:.....	23
Chapter II: Influence Operations Around the World.....	24
Global Comparison of Social Media Influence Operations.....	24
Key Patterns and Observations	38
Chapter Summary	39
Chapter III: Illustrating Core Systemic Issues	40
Platform Perspective.....	40
Policy Perspective	44
Proponent Perspective	45
Partner Perspective	46
People Perspective	47
Chapter Summary	50
Chapter IV: Conclusion.....	51
Key Actor Groups and System Relationships.....	53
Systemic Pinch Points.....	53

Issues in Creating Change.....53
 Limitations of the Study54
 Further Research.....55
 Bibliography.....56
 Appendix A: Navigating the Synthesis Map.....72

List of Tables & Figures

Fig 1. Mechanisms behind recommender systems..... 13
 Fig 2. Integrated influence operation tactics..... 20
 Fig 3. Information control creating fear. 21
 Fig 4. Algorithmic amplification boosting content engagement revenue..... 41
 Fig 5. Shift to short-form leading to more misleading content online..... 42
 Fig 6. Resource-efficient community notes models heavily rely on users. 43
 Fig 7. Common approach used in narrative control influence operations..... 45
 Fig 8. Fact-checking initiatives limited support and reach also limits their impact..... 46
 Fig 9. Increased trust in content shared with a user’s social network. 48
 Fig 10. Disinformation leads to both user polarization but also to increased vigilance. ... 49
 Fig 11. Conflicting agendas and approaches leading to fragmented solutions..... 50
 Fig 12. Synthesis map on global political influence operations on social media..... 52
 Table 1. Influence Operations and Mitigations by Geographic Region 25

Introduction

Scope and Context

In the last few decades, social media has been perceived as a double-edged sword in the modern world. In its early forms, it was lauded as an opportunity to connect with likeminded people and share information with wider audiences. Now, platforms are heavily associated with addiction, psychological manipulation, and data privacy risks, and they receive constant criticism (Li & Wagner, 2022; Olaniran, 2020).

One facet of social media that has come under criticism is its ability to influence and skew user perceptions (Fitzpatrick, 2018). The income streams set up by different platforms, which revolve around increased user engagement and paid media revenue, allow for some voices to echo louder than others through advertisements and sponsored posts. While mostly unnoticed by users on an individual level, the aftereffects of these large information ecosystems become more apparent when user behavior patterns of larger audiences are observed. Within the last decade, social media is believed to have impacted the outcome of national elections, public sentiment about geopolitical issues, and perceptions of specific racial groups (Gordichenko, 2021). Different approaches have been proposed to mitigate the risks associated with different platforms, but as they continue to grow and develop, no real long-term solutions have been implemented that both protect users, while granting them the freedom that social media is supposed to provide. This is well illustrated in the case of Section 230 of the United States of America's Communications Decency Act of 1996, which has been met with mixed reception for either supporting free speech or allowing platforms to get away with spreading false information (Draper, 2022).

Because social media has grown to serve different purposes, audiences, and geographies, it is difficult to fully present the complex interrelations between the information ecosystem. However, this research aims to contextualize how it is utilized to sway influence, and how users respond to changes on the platforms they grow attached to. To do this, this research will look at social media influence operations that are carried out on social networking, media-sharing, and microblogging platforms, evaluating them on both a macro and micro level. The research will focus on these types of social media both due to their widespread use and the unencrypted nature of their information systems, which allows for a more consistent and transparent analysis of platform regulation.

This paper is comprised of a contextual review and horizon scan to identify patterns within global influence operations; a collection of existing and currently proposed measures to address various issues surrounding social media platforms; semi-structured interviews; and a map to uncover how these various actors and actants influence each other and fail to create effective change. Providing readers with an introduction to the inner workings of political influence operations carried out on social media, readers will be better informed of how information flows online, and will be more equipped to navigate online platforms.

Research Question

To guide the exploration of the social media space and the operations that are carried out within it, the following question will guide how the research will be carried out.

Primary Research Question: In what ways are political influence operations reshaping perception on social media platforms, and what factors contribute to their success?

Though there is existing research on different types of influence operations, existing cases of social media influence operations, and the policies, regulations, and mitigations that try to reshape the online information space, this research aims to examine the system from a macro lens. Though there is a wide range of ideologies and agendas that give operations their purpose, this paper will focus on political aspects due to the direct impact that the state can have on social media and vice versa. Through systemic exploration, this paper aims to explore the complex relationships between forces that strengthen or weaken the dynamics in social media influence operations around the world, and how platforms, policies, and individual people are responding in different geographic regions. By analyzing these different regions, observations may be drawn for what drives inefficacy in various approaches and why.

Methodology

This project utilizes a combination of systems thinking tools to contextualize influence operations as they are carried out on social media platforms. The Systemic Design Toolkit was used as a reference to perform an exploration of the social media space and identify relationships, patterns, and insights (Jones & Van Ael, 2023). To better represent these uncovered insights, the results of this systems exploration were integrated into a

synthesis map for a more comprehensive understanding of social media influence operations and their different moving parts.

Initially, this research project aimed to look at social media platforms and how they affect perceptions, credibility, and culture, but after performing an initial contextual review, the scope was narrowed to instead focus on influence operations, primarily in the area of local and geopolitical perceptions.

The resulting research will be a synthesis map explaining the systemic components, actors, and a global snapshot of how various regions are coping with their corresponding social media influence operations. Furthermore, this map will present dynamics between these moving parts, presenting the key issues that are inhibiting long-term and positive change in the social media information space. This written paper acts as a supplement to the synthesis map, expounding on research and providing readers with additional context.

The Systemic Design Toolkit

To answer the questions framing this project, the research used the first three steps of the Systemic Design Toolkit: Framing the System, Listening to the System, and Understanding the System. These three steps are integral to systems exploration, allowing the research to establish system boundaries, observe behaviors, and explore the forces that create these behaviors (Jones & Van Ael, 2023).

Framing the System

- *Literature Review*

A broad exploration of the existing system was performed to better define the scope and boundaries of the study. The literature review was done in two parts: an initial literature review to contextualize social media use and perceptions, and a secondary review that focused on real world examples of influence operations around the world. Both rounds of reviews gathered information from a variety of sources, including peer-reviewed journals, books on social theory, public datasets, news and media articles, documentaries, and opinion pieces.

Listening to the System

- *Horizon Scan*

Using information gathered during the literature review, the current landscape of social media and influence operations was contextualized using a STEEP-V lens,

evaluating social, technological, environmental, economic, political, and values-based information (Loveridge, 2002). This allowed for the identification of patterns and high-level trends that allowed for a better understanding of how the space and industry are changing, as well as how people, platforms, and policies are currently responding to this change.

Due to the global reach that social media and influence operations have, the information from the literature review and horizon scan were also sorted into tables by region to identify patterns for future analysis.

- *Expert Interviews*

One of the questions that remained since the beginning of the research period involved uncovering why the many approaches and existing mitigations to the issues surrounding social media and its information spaces have yet to fully address its negative effects. Though white papers and peer-reviewed journals have been written about this on a higher level, consulting with experts who are familiar with specific cases and examples allowed for a more targeted view that could guide the research process.

Participants are experts who work in the fields of social media studies, sociology, or information ecosystems. Interviews were semi-structured and focused on validating secondary research about the evolution of social media, their algorithms, and how perceptions and trust are established online. Existing influence operations were discussed to better understand how different types of information move across social media spaces to reach target audiences.

Additionally, experts discussed existing mitigations for the issues surrounding influence operations, as well as the challenges that they may sometimes bring about to different stakeholder groups. This first set of interviews illuminated both patterns and additional challenges that some pre-existing solutions create for various stakeholder groups but also uncovered individual user behaviors that could be expounded on.

- *User Interviews*

While the initial horizon scan looked at pre-existing research on user behavior and perceptions, this research was qualitative and limited to a specific geography. To better understand the system from top to bottom, six semi-structured interviews were also conducted with self-identifying active social media users. These

individuals were aged 20-30, based in 4 different countries, and were active on at least two social media platforms. Using the insights from both the horizon scan and the expert interviews, users were asked about how they both engage with and trust what they see on social media. A short activity was also held during the course of the interview to gauge users' perceptions of the different kinds of content they commonly find on different platforms. Through this second set of interviews, patterns in the level of trust that users have, as well as what intentions guide their social media usage were identified.

Understanding the System

- *Systems Mapping*

A combination of systems frameworks was used to synthesize data and identify patterns and relationships within the information collected during the previous phases of the research. These were translated into visual models that could be used to identify the systemic issues within the social media influence operations sphere.

To best illustrate the cause-and-effect reactions within influence operations, Causal Loop Diagrams were used to show the dynamics within the system (Barbrook-Johnson & Penn, 2022). Additionally, these loops were used to identify which relationships and influences are acting as pinch points and that are inhibiting long-term and positive change.

- *Insights*

Because the aim of the research study is to provide food for thought at the user level, and to hopefully spark conversation about change within social media systems, it is important that the way by which insights are presented is digestible. While a synthesis map is the best way to provide an overview of the system, explanations will be simplified to better match the intended audience.

Chapter I: Subject Matter

Since its early conception in the 1990's, social media has grown to become a mammoth paradoxical force that is correspondent to the modern world (Dhingra & Mudgal, 2019). While it's extremely successful at breaking down barriers and connecting people, it also has the power to deepen divides between its users. The heavy focus of social media platforms to increase engagement, when combined with the new technologies that allow them to be more effective in doing this, has led to their reputation as echo chambers that reinforce bias.

These echo chambers of social media are built upon the values of a post-truth society, wherein personal perceptions affect the levels of perceived trust and credibility (Fuentes & Alberto, 2022). On these platforms, truth is relative, and people are instead left to decide on what or whom they believe in. This highly complex ecosystem, where anything goes, has created the ideal information marketplace for larger organizations to churn out copious amounts of dis- or misinformation, push their agendas, and subconsciously manufacture "truths" that benefit them.

While the detrimental effects brought about by social media have long been known, policy takes more time to catch up to the constant changes platforms make to their ways of working (Rocheffort, 2020). Both restriction and regulation have been common approaches, but these have been met with mixed results. Additionally, policy is geographically bound, while social media platforms are highly globalized. Platforms have also been encouraged to take initiative and mitigate psychological harm and data privacy risks associated with social media through their content policies, but too much change could be detrimental to these engagement-driven systems (Peter, 2024).

Social Media: A Brief History

The rapid development of the cyberspaces containing social media is a natural progression due to the acceleration and integration of everything, which results in multiple technologies and industries converging into a singularity. This convergence theory is exemplified by the Internet, which has "absorbed" other forms of media, such as newspapers, magazines, television (Adams et al, 2018). As social media exist within the Internet, they are also subject to constant reinvention, expansion, and transformation.

While social media has existed since the 1980's with the online communities like UseNet (McIntyre, 2014), one of its biggest turning points was in 2006, two years after the creation of Facebook. Mark Zuckerberg, then falling behind existing social media platforms but not wanting to be bought out, was forced to pivot how his platform worked to stay afloat. Since then, many new platforms, like Instagram and TikTok have been introduced, each one answering to slightly different purposes—socializing, posting photos, sharing opinions, and more. Evidence continues to show that social media is becoming a force that permeates human identity and psychology, and can affect how people think, act, and relate to one another (Fisher, 2022).

Types of Social Media

As previously outlined in the project scope, this research would be considering three types of platforms in evaluating the social media system: online social networking platforms, media sharing platforms, and microblogging platforms. Though some social media sites have overlapping functions, they will be evaluated based on their primary or distinguishing function, guided by the following definitions outlined by Koukaras, Tjortjis, and Rousidis (2019).

- **Online social networking platforms**
These platforms primarily function to allow its users to connect with real world friends or acquaintances online. Interaction can take place through posts, status updates, media sharing, and messages. Facebook and LinkedIn are examples of online social networking platforms.
- **Media sharing platforms**
Platforms of this variety primarily function as a means of sharing video, audio, or photos to either a private following, or to a wider audience. TikTok, YouTube, and Instagram are all examples of media sharing platforms.
- **Microblogging platforms**
These platforms let users share short and concise textual and multimedia content, and these are usually maintained by an individual or by a community. Microblogs allow users to connect with a wider audience, as connections do not necessarily have to be with real world acquaintances. X and BlueSky are both microblogging platforms.

A key factor to mention in evaluating the functions of social media platforms is their secondary features. Over time, many platforms have added messaging or chat features, which are on their own, another form of social media. However, due to the varying

regulatory mechanisms for messaging and chat applications, in which content shared is private and encrypted.

Engagement and Paid Media Revenue

Because social media is widely accessible, metrics of success vary from older forms of media, like magazines and newspapers. Instead of tracking sales targets, most social media platforms capitalize on user engagement, or how much a person interacts with and spends time on a platform. To do this, social media sites like Facebook have increased the visibility of groups that are more active and are programmed to promote increased interaction by recommending content and targeted ads based on keywords and search terms (Fisher, 2022).

Social media's algorithms are primarily run by recommender systems, which filters information so that users see what is most relevant to them. There are three broad subclasses based on recommendation approaches, and they are further outlined in Figure 1 (Roy and Dutta, 2022).

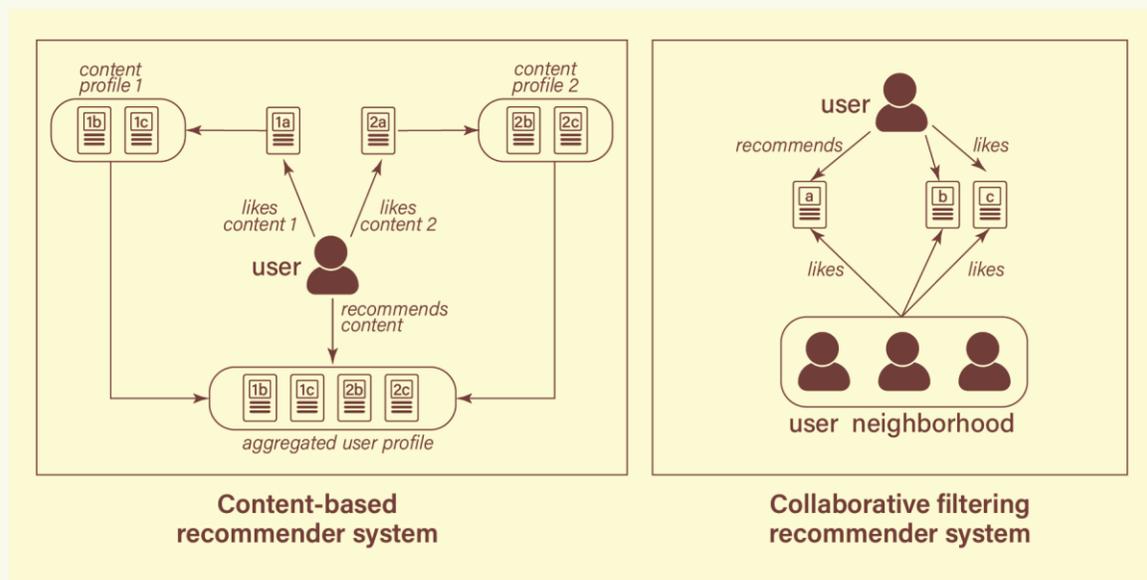


Figure 1. Mechanisms behind recommender systems.

- **Content-based recommender systems** sort item or post data into different profiles based on features or description. When users respond positively to a piece of content, similar posts are aggregated to build a user profile. These posts

bearing similar characteristics are presented to users. This classification of recommender system is based on the activity of a single user in relation to the content that they actively engage with.

- **Collaborative filtering** utilizes different user profiles that the system deems similar. The perceived likes and dislikes of user profiles are used to recommend overlapping content to users with similar profile. This allows users to potentially expand the range of content they view.
- **Hybrid approaches** aggregate two or more filtering techniques to counteract and balance some of the limitations of traditional content-based or collaborative filtering.

Another revenue stream for social media that overlaps more with traditional forms of media is the focused on paid advertising. Through paid media advertising, certain posts are boosted, sponsored, or promoted so that they appear on a user's feed or timeline (Voorveld et al., 2018). While this form of digital marketing on social media is primarily attributed to consumer products, it also has other applications, including politics, thought leadership, advocacies, education, and public service announcements.

When these two features of social media are utilized in tandem, various organizations can sponsor posts for increased visibility, and once they get users to engage, a post can organically reach a wider audience. This self-supporting system is what allows people and posts to go viral overnight, for trend to spread like wildfire, and for influence operations to grow and flourish.

Types of Engagement

Engagement can be described as measurement for how users are interacting with a piece of content, with increased attention and interaction translating as higher engagement (Chayka, 2024). Although social media platforms measure success through engagement, there is a range of methods by which engagement on social media can be measured. For this research project, types of engagement will be based on user activity, as outlined by Kaye (2021).

Interactive social media use

This is exhibited by an active exchange or sequential impact on another when engaging with content, such as a communicative exchange. These can be interactive threads of comments that involve actively replying.

Reactive social media use:

This type of engagement perceives the user as an audience to the information exchange, and they are primarily reacting to others. They can interact with content, through actions such as liking, sharing, or unilaterally commenting, but this does not largely impact the original content sharer.

Broadcasting social media use:

This type of interaction is the unilateral sharing of information with others, and this does not necessitate interactivity with others. Examples of this type of use are posting on one's timeline, or sharing Stories, which subject others to seeing what the user is sharing, but does not require a reply.

Passive social media use:

This final type of engagement also perceives the user as an audience to shared content, but unlike reactive social media use, does not involve any behavioral engagement. Instead, the user may access information by clicking in to access content, or viewing what is presented to them, but they remain viewers who do not like, share, or comment.

Social media, through algorithmic recommendations and engagement-driven business models, has redefined how attention is captured and maintained. Unlike traditional forms of media, where sales were a primary target, social media is reliant on dynamic interactions between users that are heavily influenced by the platforms that they build. This self-sustaining system has allowed for the constant amplification of content, which has created both opportunities and challenges for organizations who seek to influence public perceptions.

Actors, Influence, and Agency

Throughout the research project, 5 key actor groups were identified. Each of these groups interacted with each other to form the social media landscape in a geographic region and were observed to have varying levels of influence on change, as well as agency to act on their own.

Platform

These are the social media platforms that were observed throughout the research project. All platforms considered were backed by medium or large companies: Meta,

Alphabet, ByteDance, and X Corporation, and all have a bigger impact on the information ecosystem (Alaimo et al., 2019). Because these companies are pivotal forces in the social media space, they have a high influence on both the information that populates their platforms, as well as on their competitors' algorithms. Their widespread adoption and larger revenue also give them more agency to act independently, and many of these companies can directly interact with and negotiate with governments (Godoy, 2025).

Policy

This group is made up of national governments. As policymakers, they hold a large influence over the nations they operate in, their control over the information space on social media is subject to the cooperation of social media platforms, unless platforms are completely restricted within their particular region (Mickoleit, 2014; Rochefort, 2020). They have high agency and can influence other actors through policy, fines, or state-controlled censorship.

Proponents

This group is made up of both local political actors, as well as non-state actors, who are the primary backers of political influence operations.

Political Parties

The key distinction between political parties and governments is the officiality in their positions. They are able to more easily take advantage of social media's engagement-driven algorithms to gain support and popularity, and have a considerable amount of influence over the information ecosystem (Dommet et al., 2021). Their agency can be more restricted due to their sensitivity to how they are perceived by the public (Graham & Schwanholz, 2020).

Non-state Actors

These are high-power organizations or governments outside of the target audience's nation that can run multinational influence operations (Dommet et al., 2021). They have a considerable amount of agency, but despite their considerable financial backing, they have a middling level of influence on the information space and mostly ride on existing state conflict.

Partners

The Partners group is made up of news media, as well as the third-party or non-government organizations who act as information partners in the system. They are larger

groups who represent the actions, agendas, or rights of other actor groups, and are deeply involved in the online information space.

News Media

News media are described as either private or state-backed news media sites, and the journalists who contribute to them. They are seen as partners because they are responsible for delivering information to people through their platforms. Their agendas heavily depend on the group that they represent, as state-backed media is more aligned with government aims, while independent media would be more aligned with social media users (Schulz et al., 2022). Though there are variations in agency that are represented in each nation's press freedom score, all news media are subject to the governments within the countries they are in, giving them a middling level of agency. Their influence is considerably high compared to other actors in the information space, but there is an overall decline of trust in news globally (Newman et al., 2024).

Third-Party or Non-Government Organizations

This group is composed of grassroots or non-government organizations that aim to address issues on the online information space, primarily through response-based actions, such as digital literacy or fact-checking initiatives (Vo, 2022). Members of this sub-group primarily align with the protection of social media users. Similar to news media, they are subject to changes implemented by both social media platforms and their local governments, limiting their agency. Their influence over information is lower than news media due to their limited audience reach (Juneja & Mitra, 2022).

People

These are the receivers of both information and regulation from previous actors. Though they are also able to contribute to the information space by creating their own content or actively engaging with others on social media, their influence is mostly limited because they act independently (Schulz et al., 2022; Vo, 2022). They have a considerable amount of agency, though this is still subject to any regulation from platforms or policies.

Together, these groups interact with each other to create a complex ecosystem where each actor's influence and agency shapes the flow of information online. These relationships determine what information is seen and shared, as well as what is represented as more believable. These mechanisms, while empowering individuals to engage with others and find a wider audience, allows for actors with malicious intentions

to spread harmful or misleading content.

Misinformation, Disinformation, and Malinformation

In the large and complex world of social media, where users are faced with copious amounts of information—real, fake, and sometimes a conflation of both—it is important to distinguish its different categories. Wardle & Derakshan created a conceptual framework to differentiate misinformation, disinformation, and malinformation (2017).

- **Misinformation** is information that is objectively false, but is not created with the intention of causing harm to other entities. Content that is misleading falls under this category.
- **Malinformation** is information that is based on reality, but is used to harm or slander an individual, organization, group, or country. Harassment and hate speech are examples of information that fits this category.
- Finally, **disinformation** is intentional or purposeful misinformation, where information is falsely created for the sake of harming another entity. This information that falls under this category is often fabricated or manipulated to better resonate with its audiences.

The interplay between the algorithmic mechanisms and actor motivations creates a system that rapidly disseminates information, regardless of whether it's accurate or not. As Wardle & Derakshan's framework reveals, users are unknowingly exposed to harmful content that utilizes social media's vulnerabilities to reach wider audiences (2017). These dynamics are what create the ideal setting for coordinated influence operations that can manipulate a user's behaviors and perceptions online.

Influence Operations

In the world of social media, influence can mean many things. Influence operations are coordinated efforts that involve the spread of information to achieve a specific effect among a target audience (Chi, 2023). These operations are driven by the interests of the various actors who run them, and they often aim to disrupt the flow of accurate information to their audiences.

If you're a social media influencer, it pertains to how much you can sell something to your audience: a product, a trend, or an idea. For platforms, influence can refer to the

control they have over who sees what. For corporations outside of platform ownership, it can refer to the partnerships they have within their networks that allow them to push some ideas more than others. Influence operations in theory are neutral, however it is worth noting that many of the recorded outcomes relating to them have been presented in a negative light (Pastor-Galindo et al., 2025).

Influence operations make up a significant part of the complexities in social media issues and social media ownership (Bateman et al., 2021). Operations can look like the active spread of disinformation, but can also be the amplification of specific topics to make them seem more important to people (Tozo, 2024). Influence operations can be carried out by governments, organizations, or individuals to shape narratives, deepen divides, or sway actions and behavior in the real world .

Influence operations also exist outside of social media, and are part of a large network of tactics to alter people's perception. There is history of longer-term influence operations being carried out through pre-internet media, like newspapers, radio, and television, and that exposure has led to better reception to some political candidates over others, or to reject news that relates to certain geopolitical issues (Bateman et al., 2021). On the other hand, short-term influence operations tend to be carried out on social media platforms, and have been shown to affect political beliefs, discrimination, and skepticism over medical information.

Influence Operation Tactics

Because influence operations are complex and are shaped by their actors, agendas, and audiences, it is difficult to come up with a concrete list of ways that they are carried out. However, there are several tactics that are more common than others. Figure 2 illustrates a common approach to influence operations that aim to reshape publicly perceived narrative, integrating several approaches to manufacture engagement and expand content reach.

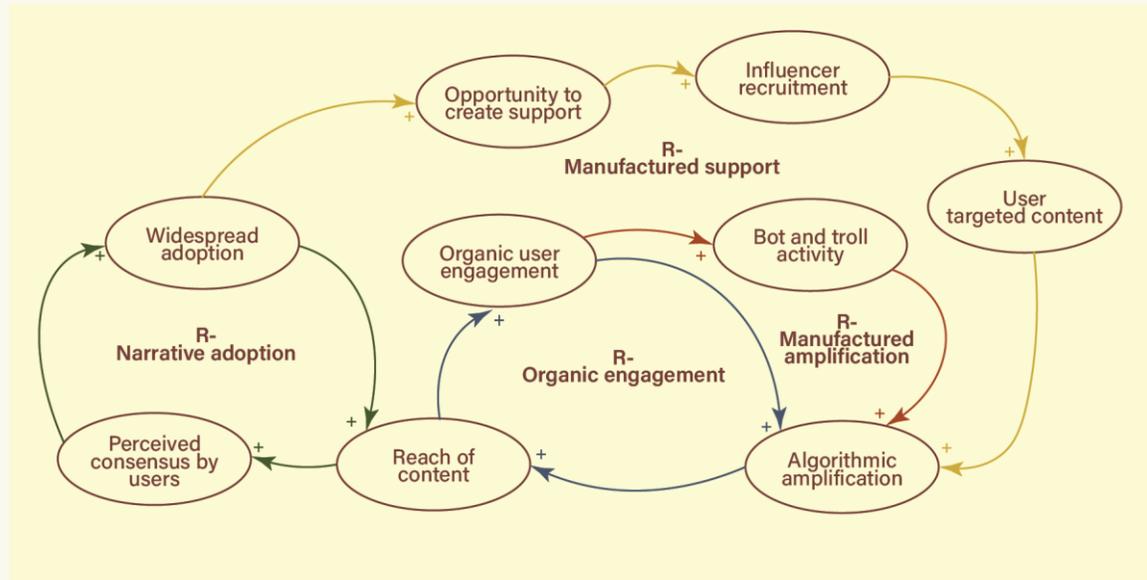


Figure 2. Integrated influence operation tactics.

Disinformation Campaigns

This tactic is characterized by the active spread of false information to negatively skew perceptions about an individual, organization, or country. On social media, these false narratives can be further amplified through the use of bots, fake accounts, and even influencers (Marom, 2024). This tactic was utilized by Russia during the 2016 U.S. Elections, when divisive ads that capitalized on political movements and social issues were widely disseminated on Facebook (Zandrozny et al, 2018).

Bot Networks and Troll Farms

This tactic involves crowding the online space with divisive messaging to manipulate users' perceptions regarding certain topics. While both are effective in spreading disinformation on social media platforms, they are fundamentally different (Connaughton & Pukallus, 2015). Bots primarily share spam and irrelevant information that makes it difficult for users to properly engage with social media spaces regarding an issue. On the other hand, trolls are real people who often engage in divide-and-conquer tactics to create disruptive environments by engaging with others or targeting specific individuals. In a 2020 report, the Oxford Internet Institute found that between January 2019 and November 2020, over 300,000 "cyber troops" accounts backed by political parties were removed from Twitter and Facebook (Bradshaw et al., 2021).

them to have platforms take down content in order to protect the integrity of the country. During the COVID-19 pandemic, this act was put into action against posts that criticized the government's management of the pandemic (Sombatpoonsiri & Mahapatra, 2024).

Though influence operations can be classified a list of types, it is important to note that these tactics are not mutually exclusive of each other. Similar to the causal loop illustrated in Figure 2, they can be used together in a coordinated attack to more effectively disrupt information online. Due to the interplay of these techniques resulting in highly complex and individualized approaches to influence operations, it has been increasingly difficult to mitigate the issues they bring to social media.

Types of Regulation and Mitigation

Despite the challenges brought about by social media's complex systems, there have been several attempts made by those in the Platform, Policy, and Partners actor groups to regulate misleading information online. For existing regulations, this paper will use the typology outlined by Rochefort's comparative social media policy analysis (2020), adding a fourth classification for more response-based third-party initiatives:

Industry Self-Regulation

This form of regulation is a minimalist approach, where private social media companies are encouraged to address their own shortcomings (Rochefort, 2020). This involves voluntary action from platforms, such as establishing new standards or policies that are not already prescribed by government policy. An example of this would be Meta's moderation of hate speech, or its partnerships with fact-checking initiatives.

Limited Government Regulation

This type of regulation involves the enforcement of narrowly defined standards, and these are often focused on one aspect of an organization or industry operation, not necessarily addressing the overall structure of the social media space, such as a focus on only data tracking and collection while not impacting overall privacy concerns (Rochefort, 2020).

Comprehensive Government Regulation

This form of regulation is meant to broaden the scope of public authority, targeting the root causes of issues. This targets every player in the industry to undermine existing

structures and business models (Rochefort, 2020). One example of this form of regulation is the growing idea of regulating social media platforms like they are public utilities to limit their ability to exploit user engagement.

Third-party Responsive Initiatives

This type is more responsive than previous classifications, which are all meant to restrict user behavior. These initiatives come from actors with limited influence, such as non-governmental or grassroots organizations who react to and correct information in the social media space (Courchesne, 2021). A good example of this type of response is the creation of external fact-checking initiatives.

The landscape for mitigating the issues surrounding information flow on social media is a reflection of the fragmented efforts to achieve balance between accountability, autonomy, and evolving technologies. Each of these mitigation types has its own strengths and weaknesses, with more targeted approaches accurately addressing a small subset of problems, while overarching policies fail to create lasting change. The consideration of comprehensive solutions and frameworks indicates that perceptions surrounding social media, and people's reliance on it, is now changing. Finally, third-party initiatives are highlighted as the group that fills in gaps in the system, to varying degrees of success. Altogether, these different approaches point to the shaky dynamics between mitigating harm and restructuring how social media platforms function, which shapes the research's analysis on accountability and responsibility on social media's evolving information spaces.

Chapter Summary:

Over time, social media has evolved from a space that connects personal networks to an entity that both bridges and divides communities. Social media platforms, which are now designed to maximize engagement, have unintentionally created echo chambers within themselves, which can lead users to fall deeper into their own biases. In this online ecosystem where truth is subjective, malicious motivations have led actors to weaponize social media's features to influence public perceptions for their own ideologies. Though these harms have come to light, regulation and mitigation have been fragmented, struggling to keep up with rapid technological developments. This first chapter outlines the evolution of early social media systems to what they are now.

Chapter II: Influence Operations Around the World

Because both social media and influence operations span across different geographies, this research project makes an effort to look at social media-based operations from a global perspective.

Global Comparison of Social Media Influence Operations

To analyze influence operations around the world, the research utilized information from the Reuters annual Digital News Report, Meltwater and We Are Social's Digital 2024 report, the Reporters Without Borders global index, peer-reviewed journals, and local news articles to evaluate 66 countries across 12 geographical regions. These regions were selected to provide a comprehensive global overview of influence operations. Within geographical regions, countries were selected based on availability of information, activity of the influence operation landscape, and existing policy variances in order to achieve a more heterogeneous representation of each geography.

This information is categorized and synthesized in Table 1, which outlines influence operations by geography, and highlights each region's key actors, common tactics, target audiences, and platform usage patterns. By structuring information by region, the research is able to comparatively analyze geographies, and identify both universal characteristics and distinct patterns of information manipulation. The table, which also outlines existing mitigations from Platforms, Policy, and Partners, provides an understanding of the interplay between different actor groups, and serves as a foundation for highlighting key issues and pinch points within the system.

Table 1
Influence Operations and Mitigations by Geographic Region

Region	Countries Evaluated	Major influence operations and tactics	Public response	Mitigations enacted	Challenges
North America	United States of America	Influencer campaigns backed by Kremlin organizations to alter perceptions about current geopolitical conflict.	Awareness and skepticism about astroturfing and political advertising	Partnerships with other nations to study disinformation tactics	Several instances of TikTok and Facebook systems failing to detect both misinformation and political advertising
	Canada	Bot networks backed by political parties to boost social media scores or manufacture support during the election period.	Significant number of people still share misleading content	Regular reporting on suspicious activity in the information space by state-backed organizations	Higher levels of polarization due to partisan culture
		Content restriction and post removal backed by platforms and the state to drive conversation away from local controversy.	Awareness of internet information space has led users to create things like	Platform-wide banning of larger known Kremlin networks	Perceived freedom is important to people in this region, making them more sensitive to change
		Troll armies spreading climate misinformation to promote right-wing practices and policies	“algospeak” to defy current measures	Efforts to limit use through institutional bans, platform bans, or age restrictions	Unstable political situations have led to counterproductive policy revisions, such as Meta’s relaxing of hate speech restrictions.
Latin America	Argentina	Trolls and bot networks employed by foreign actors (Russia, China, and Israel) to justify geopolitical conflict or to push certain political candidates	More susceptibility to influence operations, exhibited by higher trust in disinformation and influencer politicians	WhatsApp “fake news” detector launched by independent news media	Primarily responsive solutions enacted by third-party organizations or citizen-led initiatives
	Brazil				
	Chile				
	Colombia	State using both targeted censorship through platform-wide or internet restrictions, and broad media censorship through political connections to minimize dissent and control which narratives reach the public	Heavier restrictions on state narratives have led independent media to practice self-censorship	Restrictions on platforms, like X	Patterns of unethical relationships between media companies and government/politicians have led to homogeneity in media narratives
	Mexico	Use of personal brand and populist narratives by politicians to foster widespread support and assume celebrity-like status			Global shift to combatting misinformation and influence operations have started a migration to messaging platforms like WhatsApp, which has gained more widespread use than social media platforms in this particular region
	Peru				
	Venezuela				

MANUFACTURING PERCEPTION

Western Europe	Austria	Bot networks run by Russia and Iran to drive negative sentiment for Ukraine and Palestine	Good awareness and minimal lasting effect among general population, but unable to curb spread of misinformation.	Development of comprehensive, multi-level frameworks to counter both mis- and disinformation by several countries	Few frameworks have been put into action due to limited budgets or reliance on the cooperation of social media platforms
	Belgium				
	France	Disinformation campaigns carried out by bots and troll networks, backed by both foreign and local actors to create information noise during the election period.	Concerns around susceptibility are mostly for citizens who identify as right-wing.	National efforts starting to take down and curb disinformation while also maintaining freedom of speech, starting with an organization that reports on suspicious behavior	Measures to remove “illegal content” lack specificity regarding dissent or criticism, leading to potential censorship of individual voices.
	Germany				
	Netherlands				
Switzerland	Disinformation via redirection to messaging platforms, backed by right-wing extremists aiming to align targets with their political strategies.		State-prescribed partnerships with platforms to remove “illegal content”	Growing threat of right-wing extremist communities on platforms like Telegram or WhatsApp	
Northern Europe	Denmark	Troll farms and harassment campaigns linked to politicians, governments, and military aiming to manufacture harassment against opposing organizations	Minimal impact on outcomes of local elections, and operations mostly act as noise in social media spaces	Setting limits on foreign donations supporting political parties	Existing relationships between some political parties and social media giants
	Finland				
	Ireland				
	Norway	Bot and troll networks backed by foreign actors spreading terrorist content on social media to both intimidate citizens and also create a negative image of the EU.	Uptick of both hate speech and harassment on TikTok and X with minimal moderation	Holding platforms accountable for poor moderation	Ambiguity of the Digital Services Act are hard to effectively automate, and may also infringe free speech
	Sweden				
	United Kingdom				
			Raising minimum age limits for social media usage	Harassment and hate speech have remained unaddressed in the regulation of social media content	
			Sanctioning high-power nations like Russia for foreign influence operations		

MANUFACTURING PERCEPTION

Eastern Europe	Bulgaria	Troll farms and disinformation channels run through messaging apps have been backed by Russia to improve perceptions regarding the Russia-Ukraine conflict	Existing preference for populist/strong-arm approaches to leadership make them more likely to agree with Russian influence operations or extremist narratives	Governments requesting for platform checks during election periods	Some government legislations are pushing against independent media and dissenting voices
	Czech Republic			Banning of Russian state media accounts on Telegram and troll accounts on social media	Approaches, while there are many, are mostly one-sided, and have received little support from social media giants
	Hungary	Bot networks employed by political parties during election periods to manufacture widespread support and boost social media ratings		Suing of platforms for deceptive advertising	
	Poland			Grassroots efforts to create an international fact checking collecting	
	Romania			EU pushing for platform transparency for political advertising	
	Slovakia Ukraine	Use of populist narratives by influencer politicians to gain celebrity-like status and increase public support	Citizens doubt some forms of state media, and overall trust is declining		
Southern Europe	Croatia	Influencer campaigns from Russia to spread anti-NATO and anti-Ukraine sentiments	Population has less trust in media, politics, and social media, but aren't largely impacted by influence operations	Some government leaders calling for accountability and equitable power in both politics and social media platforms	Platform or messaging bans limit widespread use of entire population, having users bear the consequences
	Greece			Growing support for regulation of social media for young teenagers	Lack of active cooperation with platforms, leading to a one-sided approach to solving issues
	Italy	Policies and bans on criticism enacted by the state to minimize dissent or push media to self-censor			
	Portugal			Policies enacted that call for transparency of online services, including social media	
	Spain	Increased use in messaging apps for disinformation, spyware, and spamoflauge to drive further polarization	Growing doubt in both social media platforms and news media within more polarized countries Disagreement with governments who choose to ban entire platforms		

Sub-Saharan Africa	Ghana	Several troll farms found to be operated out of this region, backed by foreign actors, primarily due to the abundance of human labor.	A significant number of people engage with and follow bot and troll content.	No active mitigations through policies	Despite high susceptibility to influence, governments have yet to catch up with policy, only resorting to restriction in some countries.
	Kenya				
	Nigeria				
	South Africa	Bots, trolls, and influencers also utilized by local political parties during elections to malign political competitors' reputations.	Online hate campaigns have at times, driven real world election violence.	Some restrictions for certain social media and messaging applications	Due to lower overall internet penetration in this region, fact checking and digital literacy initiatives have had limited success
	Tanzania	Bots and trolls amplifying real posts from local pro-Kremlin supporters, pushing more believable pro-Russia sentiments.	In the few attempts of government to block applications or platforms, citizens responded negatively due to growing use and reliance.		Press freedom is also limited in this region, leading to homogenous media narratives
	Zimbabwe				
Middle East & North Africa	Algeria	Troll farms spreading misinformation during local elections, primarily backed by political parties	High susceptibility to inauthentic/spamoufl age pages	Heavy control through policies against disinformation, but also against dissent	Unilateral control from platforms and policies puts more pressure on users and limits freedom or exposure to different narratives
	Egypt				
	Iraq	Censorship through policy backed by the government to combat disinformation, extremism, and "threats to national security" but also to quell dissenting voices	Actively engage with redirections to messaging apps, where extremist rhetoric can spread		Strong-arm approach puts user privacy at risk, especially in pushing citizens to use local varieties of social media platforms
	Jordan				
	Morocco				
	Palestine	Keyword and content censorship by platforms to silence certain narratives and influence public perception	Citizens unable to speak out against government or policy		Platforms have alliances with high-conflict nations that may benefit from controlling narratives
	Saudi Arabia				
	United Arab Emirates				

North & Central Asia	Kazakhstan	State-backed surveillance through local versions of social media platforms to control media and local perceptions	High control has led to homogenous media narratives	Strong-arm control against dissent and critique against policy and government leaders	Limited power and agency granted to users and media
	Mongolia				
	Russia	Use of policies to criminalize dissent	Self-censorship done by both media and individual users		Wider variation in internet usage in this region, and overall lower press freedom, increasing susceptibility to influence operations
	Tajikistan	Increased spread of hate speech and discrimination through social media and news media to drive polarization and create information noise		Mixed approaches to media control in different countries, with some being multi plural but poorly managed, and others subject to heavy control	
	Turkmenistan		Citizens tolerate noise on social media and are subject to nationally prescribed platforms and applications		
	Uzbekistan				
East Asia	China	Influencer recruitment and disinformation spread backed by China to reshape perceptions about Covid-19 response or spread CCP rhetoric	Use of social networking sites for collecting news and information is lower in this region, indicating lower trust	Government partnership with messaging platforms to address criminal activity on apps	Power imbalances between nations in conflict leads to increased need for influence operations
	Hong Kong				
	Japan				
	South Korea	Keyword censorship by state and platforms of algospeak to quell dissent and discussion of sensitive topics			Increased censorship of dissenting voices in strong-arm nations, silencing dissent and criticism
	Taiwan		More active use of messaging apps or video apps		Growing concern of using messaging apps to skirt policies for criminal use (i.e. sextortion)

MANUFACTURING PERCEPTION

South & Southeast Asia	India	Active spread of emotionally charged content by extremists to spread radical content and create negative perceptions towards certain groups	Populations in this region are highly interactive and reactive on social media, and are susceptible to influence	Fact checking and digital literacy initiatives to respond to proliferation of mis- and disinformation	Lack of support from platforms and policy to combat amount of disinformation present
	Indonesia				
	Myanmar				
	Pakistan	Platform/internet blocking or red-tagging of individuals to villainize or silence dissenters			
	Philippines				
	Thailand	Use of populist narratives by politicians to reshape narratives and foster widespread support		ASEAN-prescribed guidelines for combatting fake news	Blocking of platforms or content and/or criminalizing "disinformation" with unclear framework leads to excessive censorship Proliferation of locally backed influence operations has not necessitated change ASEAN unable to establish proper guardrails like the EU, instead offering suggestions or recommendations that have yet to be implemented
Oceania	Australia	Bots and troll networks from China, suspected to be for espionage and surveillance	Good awareness of local influence operations, giving it a lower effectiveness Local parents support restrictions of social media for young teens due to known negative effects	Collaboration between Meta and government to curtail disinformation Pressuring Big Tech with fines for failure to crack down on misinformation Third-party framework for commitment to online safety and reporting with voluntary participation from platforms	Platform shifts pose a threat to healthy information space, exhibited by Facebook's switch to community notes model from fact-checking Slow shift to Telegram for use in influence operations due to efficacy in tackling influence ops on social media spaces Government fearful of AI acceleration of radicalization; no real recommendations yet Citizens are politically aligned, low demand for multi pluralism in media
	New Zealand	Disinformation spread by both local and foreign actors to sway political opinions during elections			

North America

North America is a highly active region for both foreign and local influence operations, particularly due to the multitude of international government and trade relations with other countries. This region has a distinct left-right partisan culture that is much stronger than other geographies, and this is reflected in the influencer campaigns, bot networks, and trolls that are active on the North American information space (Passantino & Lyngaas, 2024; Orr, 2019; Djuric, 2024). Many platforms have revised their policies in response to the US political situation, aligning content regulation to align with the current administration's value of free speech (Hendrix, 2025; Reinstein & Zahn, 2025).

The general public is aware of the existence of bots and trolls on social media, and despite widespread recognition of misinformation and disinformation on platforms, there is a significant number of people still sharing and engaging with misleading content (MrGraeme, 2023; Rogers, 2019). Populations are aligned with this region's prioritization of free speech, showing dissatisfaction with regulations or bans imposed on platforms (Ren, 2025; Di Placido, 2025).

Mitigations enacted have utilized international partnerships, though no comprehensive policies have been put in place (Skopeliti, 2024; Global Affairs Canada, 2025). This region has focused mostly on gathering information on foreign interference, or have focused on actively reporting on the online information space to correct large scale disinformation campaigns.

Latin America

Most large-scale disinformation campaigns in this region are backed by foreign actors, with local influence operations mostly shaped by either targeted or broad media censorship (Realuyo, 2024; Baroud, 2024; PEN International, 2024; Erb, 2019; Mitchell, 2023). In Latin America, more politicians have sizeable social media followings, which often impacts public perception and support (Caiero, 2021; Freeman, 2023; Grainger, 2024).

This region has shown higher levels of trust in misleading content than most other geographies evaluated, and they are also highly active on social media, making them an effective target for influence operations (Bell, 2022; Worley, 2018). News media in these countries either have affiliations with specific political parties, or they are pushed into

self-censorship by the state through supposed anti-terrorist policies that protect national security (Media Ownership Monitor, 2019, Soloff, 2024).

Most mitigations to influence operations and disinformation have been from non-government or third-party organizations, many of them focused on fact-checking or digital literacy (Serrano, 2017; Arzaba, 2023). Governments have taken some action against platform negligence, but cases like this have ended in complete platform bans, which gains negative reception from citizens (France-Passe, 2024).

Western Europe

In Western Europe, the online information space is most active around election periods, with disinformation campaigns backed by both foreign and local actors to create noise on social media through trolls and bots (Recorded Future Insikt Group, 2025; O'Carroll, 2024a). Outside of these active periods, there is a significant number of bot networks run by Russia and Iran to reshape perceptions about geopolitical conflicts surrounding Ukraine and Palestine (Antoniuk, 2024).

Users are largely unaffected by influence operations, and election outcomes do not seem to largely shift based on social media (Recorded Future Insikt Group, 2025). However, there is still an active spread of misinformation and misleading content (NL Times, 2025). Though more people remain unfazed by disinformation and emotionally charged content, there is some concern about susceptibility of those who identify with right-leaning ideologies (Törnberg & Chueri, 2025).

Mitigation in this region has been consistent across multiple countries, with active efforts to develop comprehensive frameworks that both combat misinformation and disinformation (Global Affairs Canada, 2024; Gilah, 2025). Though still in its early stages, there are state-backed efforts to curb the spread of misleading content by taking it down, though the long-term impacts on freedom of speech have yet to be evaluated.

Northern Europe

This region has had several troll farms and harassment campaigns backed by politicians, governments, and military targeting opposing organizations and parties (O'Sullivan, 2020; Bush & Kurzynski, 2020). Foreign interference primarily takes shape in bot and troll networks spreading terrorist content to create negative perceptions surrounding the European Union (O'Carroll, 2024b).

Though there is minimal impact on real-world politics and elections, the noise created by misleading content and disinformation is still prevalent (Dragomir et al., 2024). Moreover, there is an increase of hate speech and harassment on different platforms that hasn't been addressed by both platforms and policy (Yle News, 2021).

Governments are mostly focused on combatting foreign interference, with many countries in this region limiting foreign donations provided to political parties (Garcia & O'Carroll, 2024). Though still in early stages of discussion, nations have considered taking more action by holding platforms and high-power nations accountable for their contributions to the disruptions in the social media information space (Ravikumar, 2024; Bounds, 2024).

Eastern Europe

Eastern Europe is a target for several disinformation campaigns run by Russia to shift perceptions about their geopolitical conflict with Ukraine, and these have started to move out of mainstream social media into 1:1 channels like messaging platforms (Tejkalová, 2023; Demagog, 2022; The Kyiv Independent News Desk, 2024; Buziashvili & Mammadova, 2024). Similar to other European regions, election periods are the most active period for local operations, often carried out through bot networks to manufacture engagement and support (Szicherle & Krekó, 2021; Florkiewicz, 2025, Haeck et al., 2024).

A distinguishing factor for this region is the higher preference for strong-arm approaches and populist values, and this has been utilized by both foreign actors and local politicians to gain support (Nehring & Sittig, 2023). However, there is declining trust in both state media and social media (Than, 2024; Newman et al., 2024).

More governments are actively seeking partnerships with platforms to protect national elections, and this aligns with the European Union's push for transparency surrounding political advertising. Additionally, there have been efforts from platforms to block large Russian state media accounts (Bagwe, 2022; European Commission, 2024). However, there is growing concern for some legislations pushing against independent media and independent voices (Szicherle & Krekó, 2021).

Southern Europe

In this region, there have been several attempts by Russia to undermine NATO and Ukraine through influencer campaigns (Rossi, 2025; Martin, 2025). There are more bans

and policies against criticism of the state in this geography, which has led pushed many media channels into self-censorship (Mastracci, 2023; International Federation of Journalists, 2023).

The population remains largely unaffected by influence operations, but there is also lower trust in general media and politics, especially in more politically polarized countries (Cabrera, 2024; Newman et al., 2024). As demonstrated by the Telegram ban in Spain, many citizens were dissatisfied with the government's attempts to regulate the spread of information through restriction (Alpañes, 2024).

Governments in this region are actively calling for platform accountability and a more equitable distribution of power within politics and social media, indicating a step forward in discovering more comprehensive solutions (Davor, 2021; Wilson & Keaten, 2025). There is also growing support for policies controlling social media exposure for young teenagers (Gencturk, 2024).

Sub-Saharan Africa

This region is unique because it is a target of troll farm operations, as well as a purveyor of them, though many of these locally run troll farms operate activities that target other countries (The Zimbabwe Independent, 2024; Ward et al., 2020). Many of the operations that target these regions are locally backed and are concentrated around elections to malign political competitors (France, 2023).

Because the internet penetration rate is lower and more varied in this region, digital literacy is lower, and a significant number of users engage with and follow both bot and troll content (Jonathan, 2023). Some hate campaigns coordinated online have also led to real-world violence, indicating susceptibility to polarizing and disruptive narratives. Due to growing use and reliance on social media platforms, the few attempts to restrict applications have been met with negative response (Africa Center for Strategic Studies, 2024).

Possibly due to the large variation in internet penetration and social media adoption, policies and mitigations have yet to be proposed within Sub-Saharan Africa. However, there are active studies exploring and reporting on influence operations within the continent (Africa Center for Strategic Studies, 2024).

Middle East & North Africa (MENA)

This region is characterized by heavy control over the information space, with censorship through policy enacted to combat disinformation (Miller, 2021). However, this tends to quell dissenting voices and lead to self-censorship (Global Voices, 2021; Gulf Times, 2020). Despite the control over information, there are still numerous troll farms discovered to be spreading misinformation during elections, primarily backed by political parties (Jones, 2020; Kharroub, 2022; Shtaya, 2024).

This region is largely connected to the internet and is highly active on social media (Newman et al., 2024). However, unlike European regions, they have a high susceptibility to spoof accounts and pages (Lim et al., 2020). Despite this, there has been little to no negative response from users, though this may also be due to pre-existing policies against extremism and threats to national security (Fatafta, 2020).

North & Central Asia

In North & Central Asia, similar to the MENA region, many countries have a heavy influence on social media and the online information space (Tsedendamba, 2024; Stryker, 2020). Governments in this region have not only pushed for local versions of common social media platforms, but have also enacted policies that criminalize dissent (Novastan, 2023).

This increased control has led to self-censorship by both media and individual users, and news media has mostly taken on a single narrative (CABAR Asia, 2021). However, hate speech and discrimination that does not target the state is common on social media, and this is what primarily contributes to the noise on the online information space (Davaasharav, 2019).

The region's common tactic of overseeing information through policy has been effective at minimizing foreign interference, but this also infringes on free speech and privacy (CABAR Asia, 2021; ARTICLE 19, 2025). Despite most states also having heavy control over news media, there are a few outliers that through a lack of management, have created media multi pluralism.

East Asia

The most active influence operations in East Asia have been backed by China, targeting Taiwan and Hong Kong (Kao, 2022; Lee, 2025; Bodeen, 2025). These operations utilize

local influencers who aim to spread rhetoric aligned with the Chinese Communist Party. Within China, there have also been active efforts by the state to censor specific topics, despite attempts by citizens to codify keywords to avoid restrictions (Davidson, 2024).

The use of social media platforms as an information source is lower in this region than in others, with use mostly relegated to media-sharing platforms to follow local news media (Newman et al., 2024). Instead, there is heavy usage of different messaging applications, though this has also faced setbacks due to the growing criminal activity on these encrypted platforms (McCurry, 2024).

This region has a clear divide in approaches, with China exercising strong control and media censorship over its state and by extension, Hong Kong (Hawkins & Lin, 2024; Marvi, 2019). There is an active restriction of most international social media platforms in China, and a ban on TikTok placed in Hong Kong (Somani, 2024). On the other hand, Japan and South Korea have not needed to respond through policy due to the populations' lower reliance on social media for information (Newman et al., 2024).

South & Southeast Asia:

South & Southeast Asia is an active region for social media influence operations, though most are locally operated (Ong & Tapsell, 2022; Dash & Mitra, 2024; Internews, 2024; Wijayanto et al., 2024). Radical content shared by extremist groups has gone unchecked, and in some cases, has led to real world consequences (Amnesty International, 2022; Elliott, 2024). Many governments in this region have also attempted to control the information space by red-tagging, censoring, or criminalizing dissent, or restricting internet access in militarized regions of conflict (Amnesty International, 2024; Disinformation Social Media Alliance, 2024; AP News, 2025; Sombatpoonsiri & Mahapatra, 2024). A common influence tactic in this region also involves politicians' use of social media to gather a following and gain support by appealing to populist narratives (Digital Reach, 2024).

As social media and the internet become increasingly accessible to people in this region, their engagement is highly interactive and reactive (Newman et al., 2024). There is a considerable amount of trust that citizens put in what they see on online platforms, as well as a relatively low digital literacy rate, which makes them highly susceptible to influence operations (Martinus, 2023). Local actors also take advantage of financial

needs within some of these nations, utilizing influencers and trolls to help carry out operations (Paulo, 2022; Mundhra, 2022; Seah, 2024).

Though there have been no comprehensive policy-based approaches to solving for the proliferation of misleading content online, the Association for Southeast Asian Nations (ASEAN) has prescribed general guidelines to help combat misinformation and disinformation (Irwansyah, 2024). On the other hand, South Asia has been subjected to more restrictive policies, such as the revisions to the Indian IT Act, which allows the state to censor content through a government-backed fact-checking unit (Singla, 2024). Non-government and third-party organizations have attempted to respond to these issues by establishing fact-checking units and promoting digital literacy through “pre-bunking” (Cabato, 2023; Catellani, 2025).

Oceania

Though Oceania generally has a more robust response to influence operations, this region has had several disinformation campaigns backed by local and foreign actors around the election period (Coper, 2025). Governments have also suspected China of running operations for espionage and surveillance, and have since increased their focus on combatting foreign interference (Antoniuk, 2023; Martin, 2023).

Citizens have a good awareness of local and foreign disinformation campaigns, and they have not largely affected election results (Carson et al., 2024). They are also more receptive to age-based social media restrictions on young teens, indicating open mindedness to other proposed mitigations (Morning Report, 2024).

There has been active collaboration between governments and large platform owners like Meta to take down and curb disinformation, and they have also pressured big tech companies with fines for failing to uphold their responsibilities (Jose & Kaye, 2025; Speers & Truu, 2024). One unique approach, despite its mixed results, has been New Zealand’s Aotearoa Code of Practice for Online Safety and Harms, which is a voluntary agreement to uphold online safety and has been signed by most large platforms (Netsafe, 2023).

Key Patterns and Observations

Though each region has its own unique characteristics, approaches, and public response, there were several patterns that could be observed when evaluating them in a group.

- Use of social media for various operation types:** Different influence operation types have been adapted to work for various platform types. Troll farms are most active on X and Facebook, where replies and comments are a core form of interaction and engagement. Spamuouflage and redirections to spoof websites or messaging channels spread more easily on Facebook, which displays external links as their own type of media. X has been known for its large number of bot networks, and its randomized curation of short posts allows for easy account creation and post generation. Finally, TikTok and Instagram, which are both media-sharing platforms that focus on short-form content, have been the most common place for recruited influencers to operate, as these platforms allow users to more effectively inject humor and personality into their content, making them more trustworthy.
- Motivations behind local vs. foreign operations:** Though specific approaches may vary, there are common patterns between the intentions behind local and foreign influence operations. Local operations are primarily coordinated to gain or maintain power, either by pushing for a specific political candidate into an official government position, or to keep who is currently occupying the role in place. On the other hand, while foreign actors may have made some attempts to interfere with local elections, most of their efforts are meant to reshape international perceptions around geopolitical tensions.
- Social media skepticism:** There is increased awareness globally that what people see online may not be objectively true, and overall trust in social media is declining. However, this awareness varies significantly by geographies, and regions with higher internet penetration rates like Europe and Oceania being largely unaffected by influence operations when compared with developing regions with South & Southeast Asia and Sub-Saharan Africa.
- The importance of state control:** Despite the previous observation, one key factor that contributes to system outliers is the amount of control the state has over media spaces. With higher control, digital literacy and internet access have little impact on how people respond, as illustrated by the MENA and North &

Central Asia regions, which have relatively high and low internet penetration rates.

- **Common challenges:** Though mitigation approaches have been different in each region, they face many of the same challenges. The state often struggles to find balance between free speech and platform regulation. Oftentimes, these approaches are unilateral in nature, putting pressure on platform resources. This has led to a lack of cooperation from platforms, and by extension, the users whose online experiences begin to suffer.

Chapter Summary

This chapter dissects social media influence operations from a global perspective. Combining literature review findings with existing analytical data, the resulting table maps existing tactics, behaviors, and mitigations across geographic regions. Through comparative analysis, this chapter reveals the varied approaches of social media platforms, policymakers, and third-party organizations, and the gap these approaches may have with public response. Influence operations, despite variations in each region, capitalize on cultural gaps and existing polarization, and achieve a greater rate of success due to engagement-based content amplification, and users' susceptibility to agreeable, politically charged content.

Chapter III: Illustrating Core Systemic Issues

Using the macro-level patterns and observations from the previous chapter as a guide, the following chapter delves deeper into the dynamics within the social media influence operations landscape. To gain a better understanding of the fundamental cause-and-effect relationships between systemic elements, semi-structured interviews with both experts and active social media users were conducted and distilled into a set of key factors that shape influence operations on social media.

The core systemic issues can best be classified through five perspectives: Platform, Policy, Proponent, Partner, and People. These issues outline both the factors that make influence operations so effective, and the tensions that lead to inefficiency in present mitigations. Causal loop diagrams will also be used to further illustrate some of the tensions that contribute to how these issues persist within the system.

Platform Perspective

Platforms and the companies that run them are in a position where they can implement variations of their business models across different geographies, but they continue to aim for the same things: engagement and reach (Panda et al., 2017). Though they are subject to policies and regulations in the countries that they operate in, they also largely respond to the preferences of their users. Platforms are faced with the challenge of protecting users from misleading content online, protecting freedom of speech online, and also ensuring they meet their business goals through advertising and engagement, and this has led to an inconsistent approach that has led coordinated influence operations on the information space to go largely unchecked.

Algorithmic Amplification

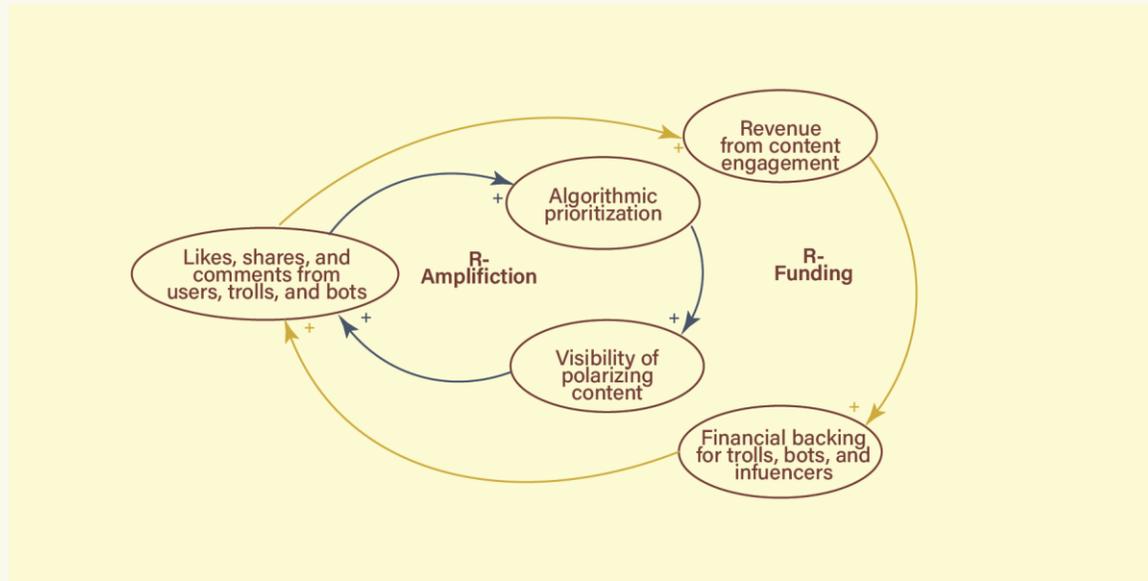


Figure 4. Algorithmic amplification boosting content engagement revenue.

As outlined in Figure 4, by using the algorithmic amplification present on social media, influence operations achieve greater success because they can use troll farms, bot networks, or recruited influencers to manufacture engagement and increase visibility. As engagement can also lead to revenue through paid content, actors can have an additional financial incentive for running operations. Users and policies have yet to fully understand how content amplification works due to algorithms being proprietary to each social media platform (Kim & Routledge, 2022). Though there are many speculations and general ideas behind how viral content achieves the reach that it does, the science that goes into platform success continues to be a black box. Though governments and third-party organizations have called for more transparency for political advertising, platforms have mostly provided the minimum amount of information required to maintain competitive advantage over other platforms.

Platform Polyamory

All participants, aligned with general user behavior, use multiple platforms (DataReportal, 2025). Those who engage with short-form video would be more likely to use similar platforms. Conversely, other types of platforms (micro-blogging, social network) would also be able to respond to any unmet user needs and fulfill other purposes. Most interview participants engaged with TikTok for entertainment, Instagram for entertainment and personal expression, Facebook to connect with family, and X to reach a wider community. This lack of loyalty increases competition between platforms,

and pushes them to replicate features to hopefully convert users to more heavily rely on one application.

Shift to Short-form

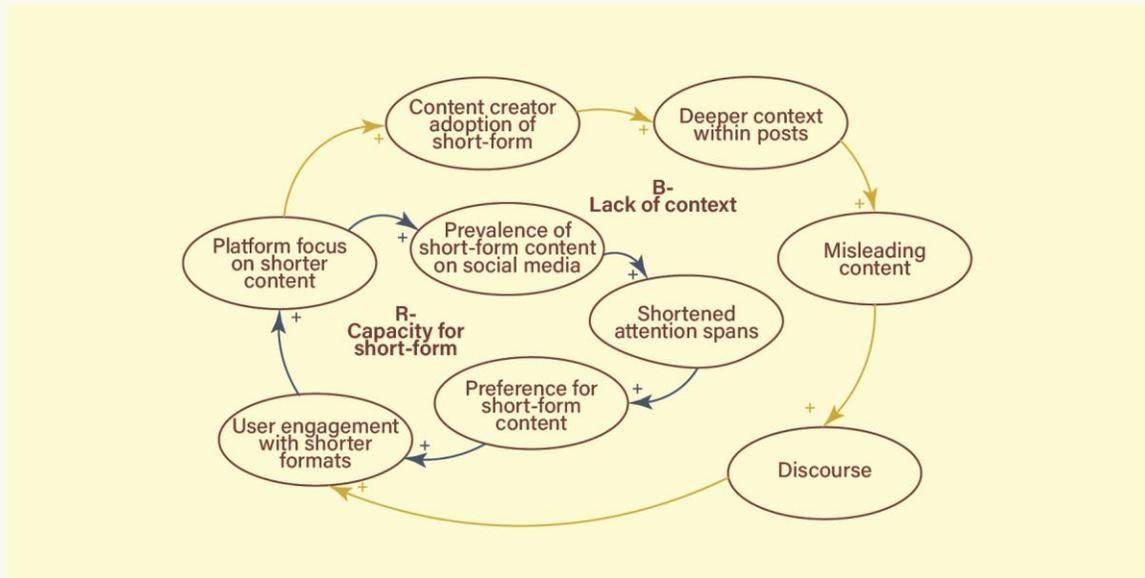


Figure 5. Shift to short-form leading to more misleading content online.

Users who were interviewed indicated their preference for short-form content because it was more convenient, simpler, and easier to focus on than lengthier content alternatives. Half of the interviewed users also expressed their avoidance of longer form content such as deep dives on YouTube or multi-part videos. Though there are many factors at play that have led to more limited capacities to focus—such as digitization leading to decreased reading, increased physical and mental stress, and the sheer amount of information that has become available to people—social media provides a singular online space for it to continue (Hari, 2022). Many social media and advertising guides have highlighted the limited capacity that users have for digesting content, pushing advertisers and influencers to ensure attention is grabbed within 3-5 seconds at the risk of a user scrolling away (Trufil, 2023). Furthermore, platforms are creating features that support this decrease in attention, such as the “tap and hold” feature to play a video at 2x speed, allowing users to interact with minute-long videos in half the time and consume more information (Torres, 2024). As illustrated in Figure 5, the multitude of factors that reinforce the prioritization of short-form has inadvertently led to less context within posts, increasing the amount of misleading content online.

Platform Moderation Over Time

Many of the interviewed experts discussed how platforms remain inconsistent in their approach to moderation because they primarily respond to news and real-world events. Because social media in its present form is still a relatively new technology, regulation and moderation only became a consideration in the last decade or so. Following both Trump's presidential win and the Rohingya massacres in Myanmar in 2016, social media platforms became increasingly aware of the lasting effects that online information on their users, and both Meta and Twitter (now X) responded with revised advertising and content policies (Su, 2018; Wagner, 2017). However, Trump's recent presidential campaign and eventual win have led to a rollback in moderation, with both X and Meta reducing fact-checking initiatives and relaxing hate speech moderation (O'Brian & Ortutay, 2022; Zilber, 2025). Though these recent shifts allow users to freely share content as they please, they also allow for the proliferation of misleading or harmful information, and create the ideal setting for disinformation-based influence operations.

Community Notes Models

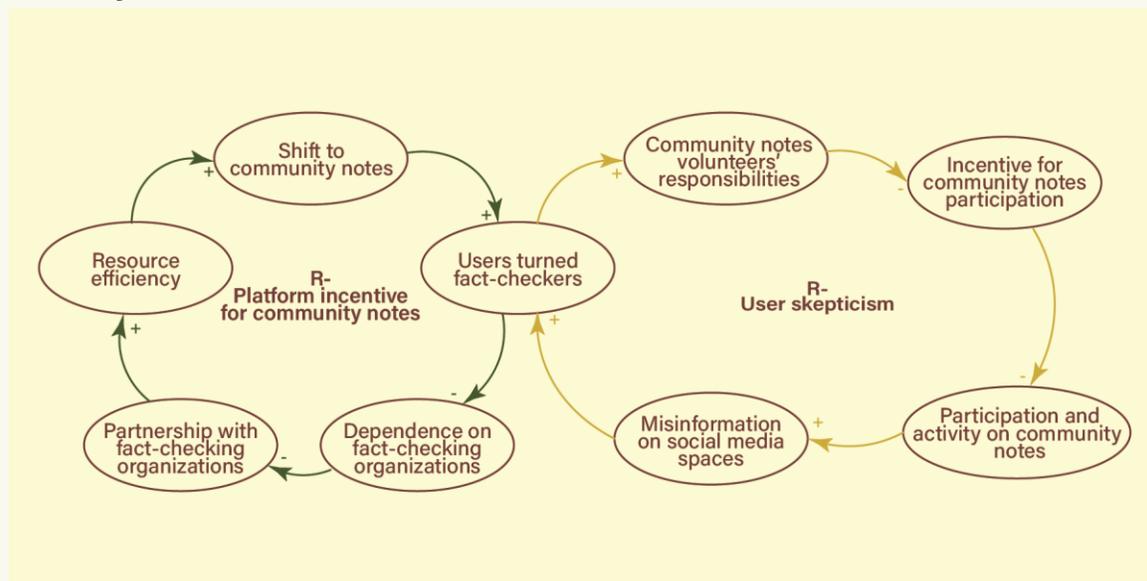


Figure 6. Resource-efficient community notes models heavily rely on users.

Several of the experts interviewed expressed concern over the growing popularity of Community Notes over formal fact-checking. In the past, large platforms like Meta relied heavily on fact-checking organizations to maintain information quality on social media (Zahn, 2025). However, it recently made a shift to a Community Notes model, similar to the example in Figure 6, mirroring X's approach to having users volunteer to verify information. While this is more resource-efficient than formal partnerships with fact-

checking organizations, and uses a majority vote to determine whether something is correct or not, it also shifts the responsibility over the information space to volunteering users, who are not incentivized or paid for their efforts to correct information online. While it has increased the speed of checking misinformation, cited sources have been found to be unreliable, and the number of notes actually published to the public have been steadily decreasing (Lee & Silverman, 2025).

Policy Perspective

Policy, which is implemented by the government, has the overarching aim of protecting the rights of its citizens. Due to policy's geographic boundaries, and social media platforms' ability to go past them, governments have had trouble implementing regulations that comprehensively address the issues brought about by influence operations online. Attempts at regulation have also been met with resistance from both platforms and users.

Responding to Yesterday's Problems

Several interviewed experts expressed their concerns about the failure of policy to keep up with existing issues surrounding influence operation tactics. While policy can take between several months to several years before being implemented, changes on social media can be implemented overnight so long as they do not infringe on any existing laws. This allows for the rapid development of new tactics and approaches to coordinated influence operations. One good example of this is covered in an evaluation across three electoral cycles in the Philippines, where influence operations had moved on to creating self-sustaining disinformation ecosystems, while policy was responding to the previous cycle's issues surrounding micro-influencers (Fallorina et al., 2023).

Unilateral Policy Response

In most of the observed regions, regulation through policy was primarily done without active collaboration from platforms. These attempts to mitigate the spread of harmful or misleading information online have mostly shifted financial burden onto social media platforms through fines and increased moderation, and this has led to decreased cooperation (Global Witness, 2023). In Oceania and some regions of Europe, there have been increased efforts to create collaborative approaches to content moderation, and this has led to more participation and proactiveness from larger social media platforms (Jose & Kaye, 2025).

Control as a Form of Influence

While policies aim to protect people through the regulation of the online information space, too much information can become its own influence operation. Though the level of control varies by nation, some strong-arm governments utilize their power over what citizens see to quell dissent. The gray area between censorship and regulation has yet to be clearly defined, and this feeds into negative user perceptions surrounding content moderation.

Proponent Perspective

Proponents, who can be local or non-state actors, are those who coordinate influence operations on social media. Though they are subject to local policies and platform changes, they are free to adapt and develop new approaches in response. Because these operations are also large in scale and sometimes outsource labor to other countries, they are hard to fully mitigate and address.

Flexibility in Approach to Narrative Control

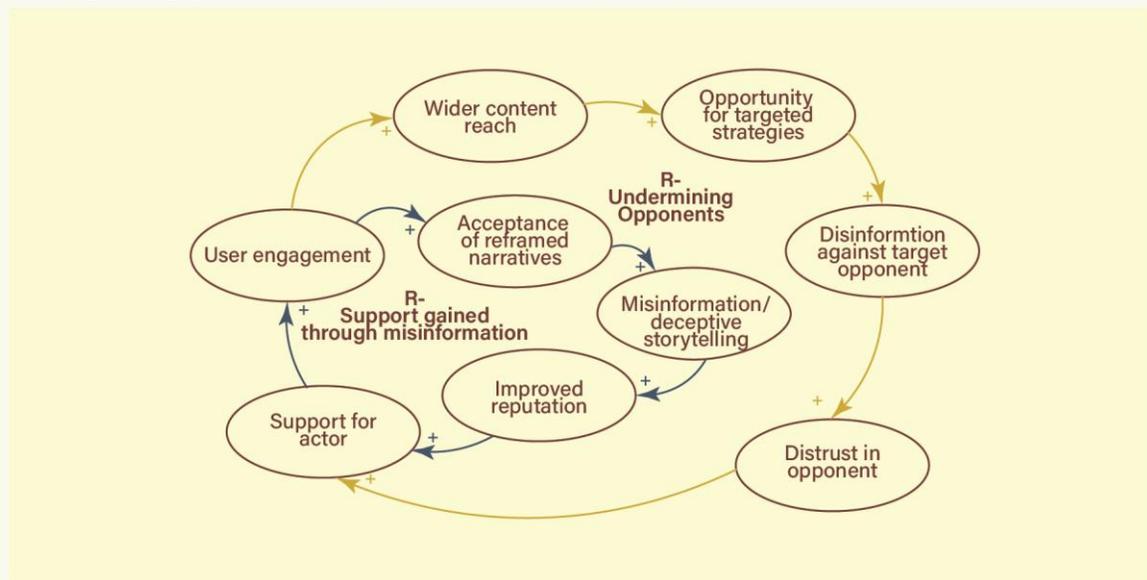


Figure 7. Common approach used in narrative control influence operations.

Within the regions observed, the most common agenda behind foreign actors and political parties' influence operations was narrative control, aiming to increase their own support while undermining the reputation opponents. The dynamics within this tactic are exhibited in Figure 7, where building distrust in an opponent through disinformation can inadvertently lead to more support for the actor. Because Proponents are more concerned about the end goals than they are how they achieve them, they are able to

have also established government-backed fact-checking units, which has raised suspicion of potential censorship (Funke & Benkelman, 2019; Funk et al., 2024). These actions often undermine the trust and credibility that users have towards these initiatives.

Ineffectiveness of Media Literacy Initiatives

Another concern that experts raised during their interviews was the minimal effectiveness of media literacy initiatives. Because of the complexity of influence operations and the agendas behind them, there is no reliable set of steps that can be followed to be able to critically evaluate information online. Because content is often created to evoke an emotional reaction, it is difficult to expect every user to be able to approach social media information objectively. Experts also mentioned that voluntary engagement with media literacy content is considerably low, and it rarely reaches their intended audiences.

People Perspective

This group is made up of the users driving engagement on social media. Though they are the biggest group, they are made to respond to changes and actions implemented by all previous groups. However, their interests are often represented by Platforms and Partners, who aim to either keep their attention as users or protect their rights.

Motivations Behind Social Media Use

Participating users mentioned the threefold purpose of social media in their lives: to stay connected, to remain informed, and to be entertained. Social media platforms have evolved to be able to answer all of those needs, and this has created a strong dependency on social media to solve many different things. These mixed motivations have blurred the lines for a platforms purpose: news has become a source of entertainment and vice versa.

Content and Engagement Preferences

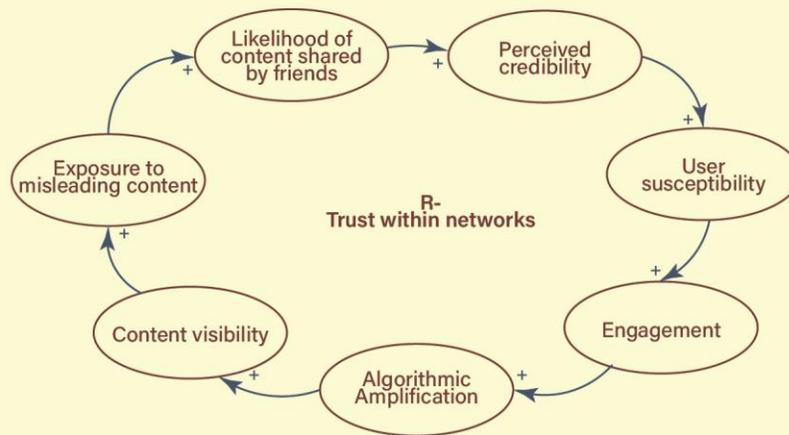


Figure 9. Increased trust in content shared with a user's social network.

When users were asked about content preferences, many trusted news media and journalists' posts more, though exposure to this content was minimal. Posts shared by a friend were positively received, and participants explained that this was mostly due to the higher likelihood of shared values with those within their networks. This is illustrated in Figure 9, where platforms utilize this inherent trust within a person's network to capitalize on engagement. When asked about online influencers, most participants expressed skepticism about the content shared, stating that influencers may take things out of context for the sake of engagement. Overall, participants' engagement was primarily driven by who is sharing the information, and what values they stand for.

Knowledge of Filter Bubbles

Interviewed users were all aware of how they were in filter bubbles, though they were unsure of how content is selected for them. Overall, there is general satisfaction in how content is curated on Facebook, Instagram, and TikTok. Most dissatisfaction came from content that users deemed "irrelevant" making its way to their platform pages. Though users are tolerant of information that is moderately disagreeable, they perceive extreme views as either trolls, bots, or a form of rage-bait. These reactions indicate a reluctance from users to leave their filter bubbles. When questioned regarding these sentiments, several participants expressed their view of social media being a source of comfort and reprieve from stress, and that they didn't see a reason for change.

Vigilance and Information Verification

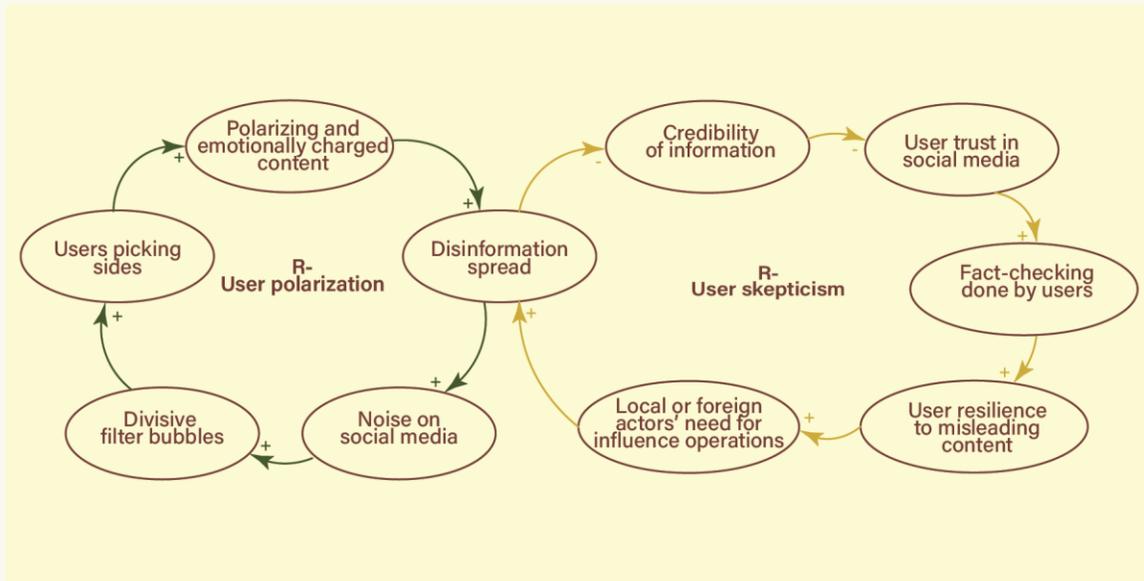


Figure 10. Disinformation leads to both user polarization but also to increased vigilance.

Figure 10 illustrates how while misleading content can drive further polarization through emotionally charged content, more users are now becoming aware of disinformation and are beginning to take initiative to fact-check on their own. While many of the interviewed users remain vigilant to more controversial or emotional content, methods for fact-checking vary. Though some go to news platforms or external sources to check information, most do secondary searches on social media platforms. Participants who said this was their preferred method cited that they viewed the opinions of other users as a method to combat potential bias.

Chapter Summary

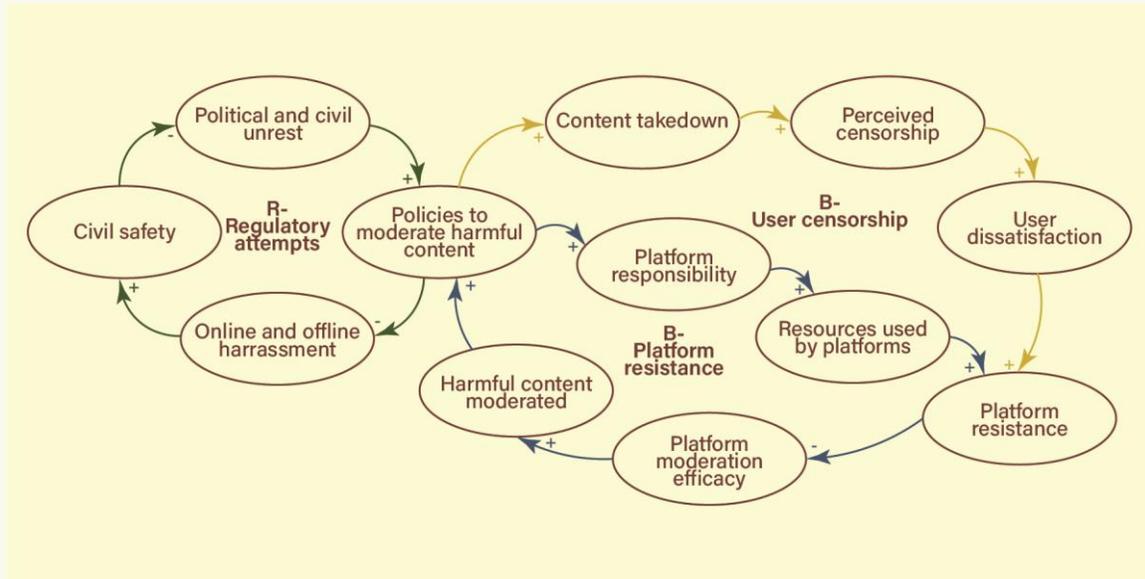


Figure 11. Conflicting agendas and approaches leading to fragmented solutions.

The Challenge of Perspective: Overall, grouping key factors and tensions within the system by perspective highlighted the importance of developing a multi-modal approach to protecting information on social media, and this is outlined in Figure 11. When regulatory measures promote civil safety and decrease political unrest, they are deemed effective and are built upon further. However, the efficacy of these policies is inhibited by users' perceived censorship and platforms' resistance to cooperate. Aside from the varying definitions of success in each group, these factors have highlighted how the shifting of burden and responsibility in solving issues has also driven more inefficiencies in facilitating positive change.

Chapter IV: Conclusion

The preceding chapters have detailed specific strategies and regional variations in social media influence operations, as well as the key patterns and issues with information integrity on social media spaces. Building on these findings, this chapter introduces a synthesis map (Figure 12) that outlines and summarizes these systemic elements, actors, and dynamics into a comprehensive overview. By mapping the relationships within the system, the map provides a deeper insight into how influence operations propagate and persist. Appendix A contains a short guide to navigating the synthesis map.

Key Actor Groups and System Relationships

Central to political influence operations on social media are the various stakeholders and their agendas, influence, and interconnections. Platforms design social media systems that drive engagement and shape what is seen by its users. Political and non-state actors utilize these systems to amplify their own narratives, and users are relegated the role of being consumers of content. Governments are caught between this interplay, and are divided in their motivations of regulate information online to either protect citizen safety or to quell dissenting voices.

Systemic Pinch Points

A key insight drawn over the course of the research has been the identification of systemic pinch points, which are areas in the information ecosystem that are most resistant to change. Among them are algorithmic amplification, individual incentives and profits, the large network of anonymous trolls and bots, and user susceptibility all work together to challenge existing mitigations.

- Algorithms, which prioritize engagement over accurate information, propagate emotionally charged and misleading content at a faster rate than current mitigations can manage.
- Government regulations act on a smaller geographic scale, while platforms span across these boundaries, giving them more flexibility and agility to skirt policies.
- Both local and non-state actors take advantage of existing divides, and capitalize on political tensions or lower digital literacy to maximize the effectiveness of influence operations.

Despite all existing approaches achieving varied success, the dynamics within the underlying social media information indicate resilience against most policies and solutions due to how large and complex platforms have become. Most power to create change lies in platforms and governments, while actor groups with lower influence are left to respond to change instead.

Issues in Creating Change

Throughout the systemic exploration, the uncovered systemic issues revealed three overarching barriers to change:

- Platforms are resistant to changing their algorithmic models, as lowering engagement could lead to financial repercussions and decreased social media usage.
- National policies and third-party initiatives cannot fully address globally operated and rapidly changing tactics used in influence operations.
- Human factors, such as preexisting biases and psychological comfort maintain existing structures that allow for the propagation of misleading content, despite all proposed and enacted interventions.

Due to the individual motivations of actor groups, limited reach of mitigations, and large network of platforms, the existing approaches have targeted individual symptoms, rather than overhauling how the system works. This directly challenges the underlying purposes that back both platforms and their users, who are largely happy with the way things on social media are, despite all the issues it may bring about.

The systemic dynamics that are outlined and mapped within this research illustrate why political influence operations are increasingly effective on spaces like social media. The tensions between engagement-driven platforms, fragmented approaches to regulation, and human factors creates an environment that rejects existing solutions. Though a truly comprehensive solution has yet to be implemented, future mitigations would benefit from using a systems-based approach that addresses all actor motivations to create meaningful change.

In an era where social media is deeply intertwined with human identity, and where it is able to shape discourse and politics, it is important to understand the processes by which influence operations are carried out. This synthesis map acts as a starting point which empowers individuals with knowledge of how information flows online, and can act as a foundation for research on future solutions.

Limitations of the Study

Though there were active efforts to approach this research project holistically, the following limitations are acknowledged following the completion of this report.

- **Large, Volatile Industry:** The social media and technology industries are constantly evolving, and the findings and insights drawn from this research may become outdated as platforms and policies change.
- **Platform Specificity:** As this research primarily focused on social networking, media-sharing, and microblogging platforms, it may not fully account for

operations and tactics carried out on other platform types, such as private messaging apps.

- **Primary Research:** Interviews were conducted across 3 of the outlined regions, which may not fully represent global perceptions surrounding social media. Additionally, interviews were reliant on participants' self-awareness, which may be subject to specific biases.
- **Causal Loop Diagrams:** Although causal loop diagrams are made to illustrate potential relationships and dynamics within systems, they are not made to represent and encapsulate every variable. Due to the complexity of social media's information ecosystems and user behaviors, identified feedback loops would benefit from further validation.

Further Research

This research project performed a systemic exploration to uncover the complex elements, influencer, relationships, and dynamics within political influence operations on social media. Utilizing the Systemic Design Toolkit as a guide, mixed methods were used to establish a comprehensive overview of political influence operations on the social media information space.

This research goes beyond surface-level observations to provide individuals with a deeper and more nuanced understanding of the existing factors that drive and support the success of influence operations online. By utilizing different sources of information, analyzing regional contexts, and identifying pinch points within the system, the project also offers a global overview from which future leverage points can be drawn.

This research can act as a foundation for future investigations for specific and more targeted mechanisms that are used to shape societal and cultural attitudes. Future research can also build upon these findings to design and test interventions and comprehensive mitigation strategies.

Bibliography

- Africa Center for Strategic Studies. (2024, March 13). *Mapping a surge of disinformation in Africa*. Africa Center. <https://africacenter.org/spotlight/mapping-a-surge-of-disinformation-in-africa/>
- Aïmeur, E., Amri, S. & Brassard, G. Fake news, disinformation and misinformation in social media: a review. *Social Network Analysis and Mining*. **13**, 30 (2023). <https://doi.org/10.1007/s13278-023-01028-5>
- Alaimo, C., Kallinikos, J., & Valderrama, E. (2019). Platforms as service ecosystems: Lessons from Social Media. *Journal of Information Technology*, 35(1), 25–48. <https://doi.org/10.1177/0268396219881462>
- Alpañés, E. (2024, March 24). Telegram suspension in Spain: "it's like closing a province because a robbery occurred there." *EL PAÍS English*. <https://english.elpais.com/technology/2024-03-24/telegram-suspension-in-spain-its-like-closing-a-province-because-a-robbery-occurred-there.html>
- Amnesty International. (2022, September 29). *Myanmar: Facebook's systems promoted violence against Rohingya; meta owes reparations – new report*. <https://www.amnesty.org/en/latest/news/2022/09/myanmar-facebooks-systems-promoted-violence-against-rohingya-meta-owes-reparations-new-report/>
- Amnesty International. (2024, October 14). Philippines: "I turned my fear into courage": Red-tagging and state violence against young human rights defenders in the Philippines. <https://www.amnesty.org/en/documents/asa35/8574/2024/en/>
- Antoniuk, D. (2023, July 25). China allegedly turns to transnational criminals to spread disinformation in Australia. *The Record*. <https://therecord.media/china-disinformation-cybercrime-groups-australia-aspi-report>
- Antoniuk, D. (2024, June 11). Switzerland notes increase in cyberattacks ahead of Ukraine Peace Summit. *The Record*. <https://therecord.media/ukraine-peace-summit-switzerland-cyberattacks-warning>
- AP News. (2025, January 3). *Myanmar's military rulers enact cybersecurity law with wide-ranging censorship provisions*. <https://apnews.com/article/internet-online-censorship-law-repression-8128ba7a2c02555217c6a64ab641eaf6>
- ARTICLE 19. (2025, January 21). *Central Asia: Freedom of expression online*. <https://www.article19.org/central-asia-freedom-of-expression-online/>
- Arzaba , A. (2023, December 11). *Narcofiles, political influencers, digital mercenaries: 2023's best investigative stories from Spain and Latin America*. Global Investigative Journalism Network. <https://gijn.org/stories/2023-best-spain-latin-america/>

- Bagwe, M. (2022, December 21). *Ukraine takes down domestic Pro-Kremlin Bot Farms*. Bank Information Security. <https://www.bankinfosecurity.com/ukraine-takes-down-domestic-pro-kremlin-bot-farms-a-20761>
- Barbrook-Johnson, P., Penn, A.S. (2022). Causal loop diagrams. In: *Systems Mapping*. Palgrave Macmillan, Cham. https://doi.org/10.1007/978-3-031-01919-7_4
- Baroud, R. (2024, December 20). *When censorship kills - social media and the war on the Palestinian narrative*. Palestine Chronicle. <https://www.palestinechronicle.com/when-censorship-kills-social-media-and-the-war-on-the-palestinian-narrative/>
- Bell, C. C. (2022, September 13). *Chile's progressive new constitution rejected by voters after campaign marred by misinformation*. The Conversation. <https://theconversation.com/chiles-progressive-new-constitution-rejected-by-voters-after-campaign-marred-by-misinformation-190371>
- Bing, C., Paul, K., & Satter, R. (2024, September 9). Russia focusing on American social media stars to covertly influence voters. *Reuters*. <https://www.reuters.com/world/russia-focusing-american-social-media-stars-covertly-influence-voters-2024-09-09/>
- Bodeen, C. (2025, January 4). Taiwan says China is redoubling efforts to undermine democracy with disinformation. *AP News*. <https://apnews.com/article/taiwan-china-disinformation-3f05dac36399bf672a702100147bf8fa>
- Bounds, A. (2024, December 3). Tiktok says it took down "cluster" of pro-Russian influencers in Romania. *Financial Times*. <https://www.ft.com/content/9c888f82-2e2d-44d4-ac1e-fcddb830cec5>
- Bradshaw, S., Bailey, H., Howard, P.N. (2020). *Industrialized disinformation: 2020 global inventory of organized social media manipulation*. Programme on Democracy & Technology. <https://demtech.oii.ox.ac.uk/wp-content/uploads/sites/12/2021/02/CyberTroop-Report20-Draft9.pdf>
- Braghieri, L., Levy, R., & Makarin, A. (2022). Social media and mental health. *American Economic Review*, 112(11), 3660–3693. <https://doi.org/10.1257/aer.20211218>
- Bush, D., & Kurzynski, M. (2020, January 28). *Poland: Presidential election 2020 scene-setter*. FSI. <https://cyber.fsi.stanford.edu/news/poland-scene-setter>
- Buziashvili, E., & Mammadova, S. (2024, April 9). *Kremlin-linked telegram channels seed anti-ukraine and anti-West narratives in Bulgaria*. DFRLab. <https://dfrlab.org/2024/04/03/kremlin-linked-telegram-channels-bulgaria/>

- CABAR Asia. (2021, November 23). *How are the authorities in Central Asia trying to control the internet? An interview with Hugh Williamson*. <https://cabar.asia/en/how-are-the-authorities-in-central-asia-trying-to-control-the-internet-interview-with-hugh-williamson>
- Cabato, R., (2023, April). *Journalist Fellowship Paper - How to cover malign influence campaigns*. Reuters Institute for the Study of Journalism. United Kingdom. https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2024-01/RISJ%20Fellows%20Paper_Regime%20Cabato_HT23_Final.pdf
- Cabrera, K. C. (2024, November 14). *Polarization in Portugal's media and social networks: A call for change*. LLYC. <https://llyc.global/en/ideas/polarization-in-portugals-media-and-social-networks-a-call-for-change/>
- Caeiro, C. (2021, September 29). *Bolsonaro's social media plan shows his election worry*. Chatham House. <https://www.chathamhouse.org/2021/09/bolsonaros-social-media-plan-shows-his-election-worry>
- Carson, A., Grömping, M., Gravelle, T. B., Jackman, S., & Phillips, J. B. (2024). Alert, but not alarmed: Electoral disinformation and trust during the 2023 Australian voice to Parliament referendum. *Policy & Internet*. <https://doi.org/10.1002/poi3.429>
- Catellani, P. (2025, January 22). Can prebunking outpace fake news?. *The Jakarta Post*. <https://www.thejakartapost.com/opinion/2025/01/23/can-prebunking-outpace-fake-news.html>
- Chan, J. (2024). Online astroturfing: A problem beyond disinformation. *Philosophy & Social Criticism*, 50(3), 507-528. <https://doi.org/10.1177/01914537221108467>
- Chi, C. (2023, November 9). *What you need to know about influence operations*. PhilSTAR. <https://www.philstar.com/headlines/2023/11/09/2309599/what-you-need-know-about-influence-operations>
- Connaughton, S. L., & Pukallus, S. (2025). *The Routledge Handbook of Conflict and Peace Communication*. Routledge.
- Coper, E. (2025, January 19). The new public square is fact-free social media – and it couldn't come at a worse time for Australia. *The Guardian*. <https://www.theguardian.com/commentisfree/2025/jan/20/the-new-public-square-is-fact-free-social-media-and-it-couldnt-come-at-a-worse-time-for-australia>
- Courchesne, L., Ilhardt, J., & Shapiro, J. N. (2021). Review of Social Science Research on the impact of countermeasures against influence operations. *Harvard Kennedy School Misinformation Review*. <https://doi.org/10.37016/mr-2020-79>

- Dash, S., & Mitra, T. (2024). Decoding the playbook: Multi-modal characterization of coordinated influence operations on Indian social media. *ACM Journal on Computing and Sustainable Societies*, 2(4), 1–19. <https://doi.org/10.1145/3675760>
- DataReportal. (2025). *Global social media statistics*. <https://datareportal.com/social-media-users>
- Davaasharav, M. (2019, January 26). *Mongolia to combat hate speech in online media*. Friedrich-Ebert-Stiftung in Asia. <https://asia.fes.de/news/mongolia-to-combat-hate-speech-in-online-media.html>
- Davidson, H. (2024, October 23). China cracks down on “uncivilised” online puns used to discuss sensitive topics. *The Guardian*. <https://www.theguardian.com/world/2024/oct/23/china-meme-online-pun-crackdown-rules>
- Davor, D. (2021). Social Media, freedom of expression, and the legal regulation of fake news in Croatia. *Studies of the Central European Professors’ Network*, 141–174. https://doi.org/10.54237/profnet.2021.mwsm_5
- Dekoninck, H., & Schmuck, D. (2022). The mobilizing power of influencers for pro-environmental behavior intentions and political participation. *Environmental Communication*, 16(4), 458–472. <https://doi.org/10.1080/17524032.2022.2027801>
- Demagog. (2022, January 4). *Ukrainian refugees and disinformation: Situation in Poland, Hungary, Slovakia and Romania*. https://demagog.org.pl/analizy_i_raporty/ukrainian-refugees-situation-in-poland-hungary-slovakia-and-romania/
- Dhingra, M., & Mudgal, R. K. (2019). Historical evolution of social media: An overview. International Conference on Advances in Engineering Science Management & Technology (ICAESMT) - 2019, Uttaranchal University, Dehradun, India. <https://doi.org/10.2139/ssrn.3395665>
- Di Placido, D. (2025, January 29). TikTok's 'cute winter boots' trend, explained. *Forbes*. <https://www.forbes.com/sites/danidiplacido/2025/01/29/tiktoks-cute-winter-boots-trend-explained/>
- DigitalReach. (2024, February 13). *Report: Operating the propaganda machine: Social media accountability in the 2022 Philippine election*. <https://digitalreach.asia/news/report-operating-the-propaganda-machine-social-media-accountability-in-the-2022-philippine-election-2/>
- Disinformation Social Media Alliance. (2024, December 23). *The influence of social media on global conflicts: Bias and impact on public discourse*. DISA. <https://disa.org/the-influence-of-social-media-on-global-conflicts-bias-and-impact-on-public-discourse/>
- Dixon, S.J. (2024). *Number of social media users worldwide from 2017 to 2028*. Statista. <https://www.statista.com/statistics/278414/number-of-worldwide-social-network-users/>

- Djuric, M. (2024, April 20). Liberal government turning to influencers to appeal to Canadian youth. *Global News*. <https://globalnews.ca/news/10439152/federal-liberal-government-influencers-young-canadians/>
- Dommett, K., Kefford, G., & Power, S. (2020). The digital ecosystem: The new politics of party organization in parliamentary democracies. *Party Politics*, 27(5), 847–857. <https://doi.org/10.1177/1354068820907667>
- Dragomir, M., Rúas-Araújo, J. & Horowitz, M (2024). Beyond online disinformation: assessing national information resilience in four European countries. *Humanities and Social Sciences Communications* 11(101). <https://doi.org/10.1057/s41599-024-02605-5>
- Draper, D. (2022, July 5). Summarizing the Section 230 Debate: Pro-Content Moderation vs Anti-Censorship. *Bipartisan Policy Center*. <https://bipartisanpolicy.org/blog/summarizing-the-section-230-debate-pro-content-moderation-vs-anti-censorship/>
- Elliott, V. (2024, November 19). “Cow vigilantes” in India are attacking Muslims and posting it on Instagram. *Wired*. <https://www.wired.com/story/cow-vigilantes-india-instagram/>
- Enke, N., & Borchers, N. S. (2019). Social media influencers in strategic communication: A conceptual framework for strategic social media influencer communication. *International Journal of Strategic Communication*, 13(4), 261–277. <https://doi.org/10.1080/1553118X.2019.1620234>
- Erb, S. (2019, March 18). *Highlighting issues the powerful want to ignore*. Deutsche Welle. <https://akademie.dw.com/en/new-digital-alternatives-are-highlighting-issues-the-powerful-would-like-to-ignore/a-47569081>
- European Commission. (2024, April 29). *Commission opens formal proceedings against Facebook and Instagram under the Digital Services Act*. https://ec.europa.eu/commission/presscorner/detail/en/ip_24_2373
- Fallorina, R., Lanuza, J.M., Sanchez, J.G., Ong, J.C., & Curato, N. (2023). *From disinformation to influence operations: The evolution of disinformation in three electoral cycles*. Internews. <https://internews.org/wp-content/uploads/2023/06/InternewsPH-June-2023-The-Evolution-of-Disinformation-in-Three-Electoral-Cycles.pdf>
- Fatafta, M. (2020, December 17). *From free space to a tool of oppression: What happened to the internet since the Arab Spring?*. The Tahrir Institute for Middle East Policy. <https://timep.org/2020/12/17/from-free-space-to-a-tool-of-oppression-what-happened-to-the-internet-since-the-arab-spring/>
- Fisher, M. (2022). *The Chaos Machine: The inside story of how social media rewired our minds and our world*. Little, Brown and Company.

- Fitzpatrick, N. (2018). Media manipulation 2.0: The impact of social media on news, competition, and accuracy. *Athens Journal of Mass Media and Communications*, 4(1), 45–62.
<https://doi.org/10.30958/ajmmc.4.1.3>
- Florkiewicz, P. (2025, January 10). Poland identifies Russian group that aims to sway elections, deputy PM says. *Reuters*. <https://www.reuters.com/world/europe/poland-identifies-russian-group-allegedly-aiming-sway-elections-deputy-pm-says-2025-01-10/>
- France 24. (2023, August 3). *Zimbabwe election disinformation spreads on WhatsApp*.
<https://www.france24.com/en/live-news/20230803-zimbabwe-election-disinformation-spreads-on-whatsapp-1>
- France-Presse, A. (2024, August 18). Social Platform X to cease local operations in Brazil, citing “censorship orders”. *Voice of America*. <https://www.voanews.com/a/social-media-platform-x-to-cease-local-operations-in-brazil-/7747017.html>
- Freeman, W. (2023, February 16). Nayib Bukele’s growing list of Latin American admirers. *Americas Quarterly*. <https://www.americasquarterly.org/article/nayib-bukeles-growing-list-of-latin-american-admirers/>
- Funk, A., Vesteinsson, K., & Baker, G. (2024, November 4). *Delegitimizing the messenger: The assault on fact-checkers*. Council on Foreign Relations.
<https://www.cfr.org/blog/delegitimizing-messenger-assault-fact-checkers>
- Funke, D., & Benkelman, S. (2019, July 11). Politicians are launching their own fact-checking projects. here’s why that’s problematic. *Poynter*. <https://www.poynter.org/fact-checking/2019/politicians-are-launching-their-own-fact-checking-projects-heres-why-thats-problematic/>
- García, C. A., & O’Carroll, L. (2024, May 30). Only seven EU countries require parties to reveal identity of all private donors. *The Guardian*.
<https://www.theguardian.com/world/article/2024/may/30/eu-countries-parties-private-donors-spain-france>
- Gencturk, A. (2024, December 12). Greece mulls curbing social media use for children under 15. *Anadolu Ajansi*. <https://www.aa.com.tr/en/europe/greece-mulls-curbing-social-media-use-for-children-under-15-/3422999>
- Gilah, R. (2025, January 2). *French official outlines government approach to countering foreign online operations*. SIPA | Institute of Global Politics.
<https://igp.sipa.columbia.edu/news/french-official-outlines-government-approach-countering-foreign-online-operations>
- Global Affairs Canada. (2024, October 29). *Government of Canada*.
https://www.international.gc.ca/world-monde/issues_development-

enjeux_developpement/peace_security-paix_securite/declaration_information_integrity-integrite.aspx?lang=eng

Global Affairs Canada. (2025, March 13). *Rapid Response Mechanism Canada: Global Affairs Canada*. GAC. <https://www.international.gc.ca/transparency-transparence/rapid-response-mechanism-mecanisme-reponse-rapide/index.aspx?lang=eng>

Global Voices. (2021, April 12). *Algerians' speech freedoms strained by media laws passed under covid-19 pretext*. <https://globalvoices.org/2021/04/08/algerians-speech-freedoms-strained-by-media-laws-passed-under-covid-19-pretext/>

Global Witness. (2023, September 15). *Mapping social media platforms' failure to tackle hate and disinformation*. <https://globalwitness.org/en/campaigns/digital-threats/a-world-of-online-hate-and-lies-mapping-our-investigations-into-social-media-platforms-failure-to-tackle-hate-and-disinformation/>

Godoy, J. (2025, April 14). Facebook owner Meta faces existential threat at trial over Instagram, WhatsApp. *Global Banking And Finance Review*. <https://www.globalbankingandfinance.com/US-TECH-ANTITRUST-META-PLATFORMS-fb73ba9a-b8f7-4c48-a814-1324737c397b>

Gorodnichenko, Y., Pham, T., & Talavera, O. (2021). Social media, sentiment and public opinions: Evidence from #Brexit and #USElection. *European Economic Review*, 136, 103772. <https://doi.org/10.1016/j.euroecorev.2021.103772>

Graham, T., & Schwanholz, J. (2020). Politicians and political parties' use of social media in-between elections. *Journal of Applied Journalism & Media Studies*, 9(2), 91-103. https://doi.org/10.1386/ajms_00017_1

Grainger, J. (2024, December 10). Tirades and trolls: How milei changed Argentina's political rhetoric. *Buenos Aires Times*. <https://www.batimes.com.ar/news/argentina/tirades-and-trolls-how-milei-changed-argentinas-political-rhetoric.phtml>

Gulf Times. (2020, November 30). *Expert discusses digital authoritarianism in the Middle East*. <https://www.gulf-times.com/story/679265/Expert-discusses-digital-authoritarianism-in-the-Middle-East>

Haeck, P., Paun, C., Cerulus, L., & Starcevic, S. (2024, November 29). Tiktok's Romanian reckoning. *POLITICO*. <https://www.politico.eu/article/tiktok-romania-reckoning-calin-georgescu-election-bytedance-china-russia/>

Hari, J. (2022). *Stolen focus: Why you can't pay attention--and how to think deeply again*. Crown.

Hawkins, A., & Lin, C. (2024, September 2). How China's internet police went from targeting bloggers to their followers. *The Guardian*.

<https://www.theguardian.com/world/article/2024/sep/02/how-chinas-internet-police-went-from-targeting-bloggers-to-their-followers>

Hendrix, J. (2025, January 7). *Transcript: Mark Zuckerberg announces major changes to Meta's content moderation policies and operations*. Tech Policy Press.

<https://www.techpolicy.press/transcript-mark-zuckerberg-announces-major-changes-to-metas-content-moderation-policies-and-operations/>

International Federation of Journalists. (2023, July 28). *Croatia: Bill on media is intolerable and must be fought*. <https://www.ifj.org/media-centre/news/detail/article/croatia-bill-on-media-is-intolerable-and-must-be-fought>

Internews. (2024). *Nexus of manipulation: Anatomy of influence operations in the Philippines*. <https://internews.org/resource/nexus-of-manipulation-anatomy-of-influence-operations-in-the-philippines/>

Irwansyah. (2024, March). *ASEAN guideline on management of government information in combating fake news and disinformation in the media*. Association of Southeast Asian Nations. <https://scholar.ui.ac.id/en/publications/asean-guideline-on-management-of-government-information-in-combat>

Jackson, J. (2023, July 27). *What are influence operations and why are we investigating them?*. The Bureau of Investigative Journalism. <https://www.thebureauinvestigates.com/stories/2023-07-27/what-are-influence-operations-and-why-are-we-investigating-them/>

Jonathan, S. (2023, August 12). *We tracked TikTok man glorifying coups, calling for Russian alliance: Guess who and where he hails from?*. Dubawa. <https://dubawa.org/we-tracked-tiktok-man-glorifying-coups-calling-for-russian-alliance-guess-who-and-where-he-hails-from/>

Jones, M. O. (2020, May 19). *Anatomy of a disinformation campaign: The coup that never was*. *Al Jazeera*. <https://www.aljazeera.com/features/2020/5/19/anatomy-of-a-disinformation-campaign-the-coup-that-never-was>

Jones, P., & Van Ael, K. (2023). *Design journeys through complex systems: Practice tools for systemic design*. BIS Publishers.

Jose, R., & Kaye, B. (2025, March 18). *Meta vows to curtail false content, deepfakes ahead of Australia election*. *Reuters*. <https://www.reuters.com/technology/meta-vows-curtail-false-content-deepfakes-ahead-australia-election-2025-03-18/>

Juneja, P., & Mitra, T. (2022). *Human and technological infrastructures of fact-checking*. *Proceedings of the ACM on Human-Computer Interaction*, 6(CSCW2), 1–36. <https://doi.org/10.1145/3555143>

- Kao, J. (2022, August 2). *Astroturfing: How hijacked accounts and dark public relations faked support for China's response to COVID-19*. Media Manipulation Casebook. <https://casebook-static.pages.dev/case-studies/astroturfing-how-hijacked-accounts-and-dark-public-relations-faked-support-chinas/>
- Kapos, S. (2024, October 2). Republicans are livid after the debate moderators fact-checked JD Vance. *POLITICO*. <https://www.politico.com/news/2024/10/02/vp-debate-moderators-fact-check-republican-reaction-00182132>
- Kaye, L. K. (2021). Exploring the "Socialness" of social media. *Computers in Human Behavior Reports*, 3, 100083. <https://doi.org/10.1016/j.chbr.2021.100083>
- Kharroub, T. (2022, February 3). *Mapping digital authoritarianism in the Arab world*. Arab Center Washington DC. <https://arabcenterdc.org/resource/mapping-digital-authoritarianism-in-the-arab-world/>
- Kim, T. W., & Routledge, B. R. (2021). Why a right to an explanation of algorithmic decision-making should exist: A trust-based approach. *Business Ethics Quarterly*, 32(1), 75–102. <https://doi.org/10.1017/beq.2021.3>
- Koukaras, P., Tjortjis, C., & Rousidis, D. (2019). Social media types: Introducing a data driven taxonomy. *Computing*, 102(1), 295–340. <https://doi.org/10.1007/s00607-019-00739-y>
- The Kyiv Independent News Desk. (2024, November 20). *Russia responsible for 80% of foreign influence operations in Europe, Czech FM says*. The Kyiv Independent. <https://kyivindependent.com/russia-responsible-for-80-of-foreign-influence-operations-in-the-world-czech-fm-says/>
- Kyle, P. (2024, November 20). *Draft statement of strategic priorities for online safety*. GOV.UK. <https://www.gov.uk/government/publications/draft-statement-of-strategic-priorities-for-online-safety/draft-statement-of-strategic-priorities-for-online-safety#technology-and-innovation>
- Lee, D., & Silverman, C. (2025, March 18). Community notes can't save social media from itself. *Bloomberg*. <https://www.bloomberg.com/opinion/features/2025-03-18/community-notes-on-x-and-meta-can-t-save-social-media-from-itself>
- Lee, Y. (2025, January 8). *Recruitment of online influencers reveals a new tactic of China's United Front*. Global Taiwan Institute. <https://globaltaiwan.org/2025/01/recruitment-of-online-influencers-reveals-a-new-tactic-of-chinas-united-front/>
- Li, J., & Wagner, M. W. (2024). How do social media users and journalists express concerns about social media misinformation? A computational analysis. *Harvard Kennedy School (HKS) Misinformation Review*. <https://doi.org/10.37016/mr-2020-147>

- Lim, G., Maynier, E., Scott-Railton, J., Fittarelli, A., Moran, N., & Deibert, R. (2020, May 4). *Burned after reading: Endless mayfly's ephemeral disinformation campaign*. The Citizen Lab. <https://citizenlab.ca/2019/05/burned-after-reading-endless-mayflys-ephemeral-disinformation-campaign/>
- Loveridge D. *The STEEPV acronym and process – a clarification*. "Ideas in Progress". PREST Policy Research in Engineering, Science and Technology. Manchester: The University of Manchester; 2002; 29.
- Marom, A. (2024, October 30). *Influence warfare: How misinformation and psychological operations shape global power*. Riskmind. <https://www.riskmind.org/post/influence-warfare-how-misinformation-and-psychological-operations-shape-global-power>
- Martin, A. (2023, August 11). New Zealand intelligence report accuses China of cyber-enabled interference. *The Record*. <https://therecord.media/new-zealand-report-china-interference-cyber-intelligence>
- Martin, A. (2025, January 10). Bots identified pushing anti-nato messages in Croatian presidential runoff. *The Record*. <https://therecord.media/bots-pushing-anti-nato-messages-croatian-election>
- Martinus, M. (2023, February 15). *Can ASEAN mitigate fake news in Southeast Asia?*. FULCRUM. <https://fulcrum.sg/aseanfocus/af-can-asean-mitigate-fake-news-in-southeast-asia/>
- Marvin, R. (2019, June 13). *Chinese ddos attack hits telegram during Hong Kong protests*. PCMAG. <https://www.pcmag.com/news/chinese-ddos-attack-hits-telegram-during-hong-kong-protests>
- Mastracci, M. (2023, August 12). *Media Freedom in south-east Europe faces escalating digital threats*. Balkan Insight. <https://balkaninsight.com/2023/08/08/media-freedom-in-south-east-europe-faces-escalating-digital-threats/>
- McCurry, J. (2024, August 28). South Korea battles surge of deepfake pornography after thousands found to be spreading images. *The Guardian*. <https://www.theguardian.com/world/article/2024/aug/28/south-korea-deepfake-porn-law-crackdown>
- Media Ownership Monitor. (2019, April 2). *Grupo Clarín*. <https://argentina.mom-gmr.org/en/owners/companies/detail/company//grupo-clarin-1/>
- Meltwater & We Are Social. (2024). *Digital 2024: Global overview report*. <https://datareportal.com/reports/digital-2024-global-overview-report>

- Mickoleit, A. (2014), Social media use by governments: A policy primer to discuss trends, identify policy opportunities and guide decision makers. *OECD Working Papers on Public Governance*, No. 26, OECD Publishing, Paris, <https://doi.org/10.1787/5jxrcmghmk0s-en>
- Miller, E. (2019, August 21). *Egypt leads the pack in internet censorship across the Middle East*. Atlantic Council. <https://www.atlanticcouncil.org/blogs/menasource/egypt-leads-the-pack-in-internet-censorship-across-the-middle-east/>
- Miranda, S., Trigo, I., Rodrigues, R., & Duarte, M. (2023). Addiction to social networking sites: Motivations, flow, and sense of belonging at the root of addiction. *Technological Forecasting and Social Change*, 188, 122280. <https://doi.org/10.1016/j.techfore.2022.122280>
- Mitchell, R. (2024, July 26). *Internet censorship verging on service blocking ahead of Venezuela elections*. Internet Society Pulse. <https://pulse.internetsociety.org/blog/internet-censorship-verging-on-service-blocking-ahead-of-venezuela-elections>
- Morning Report. (2024, November 29). *Kiwi parents would support similar social media ban to Australia's*. RNZ. <https://www.rnz.co.nz/news/political/535160/kiwi-parents-would-support-similar-social-media-ban-to-australia-s-expert>
- MrGraeme (2023, 14 August). *PSA: There is an active astroturfing campaign on this subreddit* [Online forum post]. Reddit. https://www.reddit.com/r/britishcolumbia/comments/15rd6xm/psa_there_is_an_active_astroturfing_campaign_on/
- Mundhra, L. (2022, November 10). *India undergoing Troll Farm Arms Race: European experts caution against information manipulation*. Inc42 Media. <https://inc42.com/buzz/india-undergoing-troll-farm-arms-race-european-experts-caution-against-information-manipulation/>
- Nehring, C., & Sittig, H. (2023). *Blurring the truth disinformation in Southeast Europe* edited by dr. Christopher Nehring, guest lecturer of the KAS media programme see at Sofia University, Hendrik Sittig, director media programme see of the konrad-adenauer-stiftung. Konrad-Adenauer-Stiftung e.V.
- Netsafe (2023). *Aotearoa New Zealand Code of Practice for Online Safety and Harms*. <https://thecode.org.nz/about-the-code-new-zealand/>
- Newman, N., Fletcher, R., Robertson, C. T., Arguedas, A. R., & Nielsen, R. K. (2024). *Reuters Institute Digital News Report 2024*, 61–162. <https://doi.org/10.60625/risj-vy6n-4v57>
- NL Times. (2025, March 5). *Majority of Dutch worried AI will increase misinformation, cybercrime*. <https://nltimes.nl/2025/03/05/majority-dutch-worried-ai-will-increase-misinformation-cybercrime>

- Novastan. (2023, March 30). *Turkmenistan launches national messaging services*.
<https://novastan.org/en/turkmenistan/turkmenistan-launches-national-messaging-services/>
- O'Brien, M., & Ortutay, B. (2022, December 13). Musk's Twitter disbands its trust and safety advisory group. *AP News*. <https://apnews.com/article/elon-musk-twitter-inc-technology-business-a9b795e8050de12319b82b5dd7118cd7>
- O'Carroll, L. (2024a, July 12). Disinformation networks "flooded" X before EU elections, report says. *The Guardian*. <https://www.theguardian.com/world/article/2024/jul/12/disinformation-networks-social-media-x-france-germany-italy-eu-elections>
- O'Carroll, L. (2024b, November 13). Ireland Orders X, TikTok and Instagram to curb terrorist content. *The Guardian*. <https://www.theguardian.com/world/2024/nov/13/ireland-orders-x-tiktok-and-instagram-to-curb-terrorist-content>
- O'Sullivan, D. (2020, December 16). Facebook accuses people tied to French military of running troll accounts. *CNN*. <https://edition.cnn.com/2020/12/15/tech/facebook-france-africa-troll/index.html>
- Olaniran, B., & Williams, I. (2020). Social media effects: hijacking democracy and civility in civic engagement. *Rhetoric, Politics and Society*, 77–94. https://doi.org/10.1007/978-3-030-36525-7_5
- Ong, J. C., & Tapsell, R. (2022). Demystifying disinformation shadow economies: Fake news work models in Indonesia and the Philippines. *Asian Journal of Communication*, 32(3), 251–267. <https://doi.org/10.1080/01292986.2021.1971270>
- Orr, C. (2019, July 18). Twitter bots boosted the trending #trudeaumustgo hashtag. *Canada's National Observer*. <https://www.nationalobserver.com/2019/07/18/news/twitter-bots-boosted-trending-trudeaumustgo-hashtag>
- Panda, R., Mehta, B., & Karani, A. (2017). Business Models in Social Media. *Ternational Journal of Marketing and Financial Management*, 5(2), 33–42. <https://doi.org/https://doi.org/10.5281/zenodo.495250>
- Passantino, J., & Lyngaas, S. (2024, September 13). Right-wing influencers say they were dupes in an alleged russian influence operation. they're keeping their millions, for now. *CNN*. <https://www.cnn.com/2024/09/13/media/right-wing-media-influencers-tenet-russian-money/index.html>
- Pastor-Galindo, J., Nespoli, P., Ruipérez-Valiente, J. A., & Camacho, D. (2025). *Influence Operations in Social Networks*. <https://doi.org/https://doi.org/10.48550/arXiv.2502.11827>

- Paulo, D. A. (2022, September 4). Trolls for hire in Philippines: The concealed political weapon used in a social media war. *Channel News Asia*. <https://www.channelnewsasia.com/cna-insider/paid-troll-army-hire-philippines-social-media-elections-influencers-2917556>
- PEN International. (2024, December 9). Argentina: Milei's first six months threaten freedom of expression and cultural rights - PEN international: Promoting literature & defending freedom of expression worldwide. <https://www.pen-international.org/news/argentina-mileis-first-six-months-threaten-freedom-of-expression-and-cultural-rights>
- Ravikumar, S. (2024, October 28). UK sanctions Russian "disinformation agencies". *Reuters*. <https://www.reuters.com/world/uk/uk-sanctions-russian-disinformation-agencies-2024-10-28/>
- Realuyo, C. (2024, February 12). *The competition for influence in the Americas is now online*. Atlantic Council. <https://www.atlanticcouncil.org/in-depth-research-reports/issue-brief/the-competition-for-influence-in-the-americas-is-now-online/>
- Recorded Future Insikt Group (2025, February 13). *Stimmen aus Moskau: Russian influence operations target German Elections*. <https://go.recordedfuture.com/hubfs/reports/ta-ru-2025-0213.pdf>
- Reinstein, J., & Zahn, M. (2025, January 19). TikTok thanks trump after it begins restoring service to us users. *ABC News*. <https://abcnews.go.com/Business/tiktok-restoring-service/story?id=117857636>
- Ren, E. (2025, January 19). TikTok was briefly banned in the U.S. What is the Chinese social media RedNote that TikTok exiles are flocking to?. *Toronto Star*. https://www.thestar.com/news/canada/tiktok-was-briefly-banned-in-the-u-s-what-is-the-chinese-social-media-rednote/article_1349e80e-d284-11ef-af46-df501c7d7696.html
- Rochefort, A. (2020). Regulating social media platforms: A comparative policy analysis. *Communication Law and Policy*, 25(2), 225–260. <https://doi.org/10.1080/10811680.2020.1735194>
- Rogers, K. (2019, October 11). Canadians' misperceptions about immigration reflect disinformation online: experts. *CBC News*. <https://www.cbc.ca/news/politics/immigration-disinformation-election-2019-poll-1.5316934>.
- Rossi, E. (2025, March 5). *From Russia to Italy: That's not Amore*. CEPA. <https://cepa.org/article/from-russia-to-italy-thats-not-amore/>
- Schulz, A., Fletcher, R., & Nielsen, R. K. (2022). The role of news media knowledge for how people use social media for news in five countries. *New Media & Society*, 26(7), 4056–4077. <https://doi.org/10.1177/14614448221108957>

- Seah, N. (2024, December 17). *Democracy in the digital age: How buzzer culture is stinging Indonesia's democracy: New perspectives on Asia*. CSIS. <https://www.csis.org/blogs/new-perspectives-asia/democracy-digital-age-how-buzzer-culture-stinging-indonesias-democracy>
- Serrano, C. (2017, March 20). *To slow the spread of false stories on WhatsApp, this Colombian news site is enlisting its own readers*. Nieman Lab. <https://www.niemanlab.org/2017/03/to-slow-the-spread-of-false-stories-on-whatsapp-this-colombian-news-site-is-enlisting-its-own-readers/>
- Shtaya, M. (2024). Tech Harms on People and Elections in the Middle East and North Africa Region. *Global Coalition for Tech Justice*. <https://digitalaction.co/tech-harms-on-people-and-elections-in-the-middle-east-and-north-africa-region/>
- Singla, P. (2024, December 25). *Navigating India's Digital Dilemmas: Balancing Free Speech and misinformation in the age of regulation*. RGNUL Financial & Mercantile Law Review. <https://www.rfmlr.com/post/navigating-india-s-digital-dilemmas-balancing-free-speech-and-misinformation-in-the-age-of-regulati>
- Skopeliti, C. (2024, February 26). First thing: us leading global alliance to fight foreign government disinformation. *The Guardian*. <https://www.theguardian.com/us-news/2024/feb/26/first-thing-us-leading-global-alliance-fight-foreign-government-disinformation>
- Soloff, A. K. (2024, July 27). Mexico's troll bots are threatening the lives of activists. *VICE*. <https://www.vice.com/en/article/mexicos-troll-bots-are-threatening-the-lives-of-activists/>
- Somani, D. (2024, September 8). Countries where X, TikTok, and Facebook are banned. *The Times of India*. <https://timesofindia.indiatimes.com/technology/tech-tips/countries-where-x-tiktok-and-facebook-are-banned-check-out-the-list/articleshow/113168064.cms>
- Sombatpoonsiri, J., & Mahapatra, S. (2024, July 31). *Regulation or repression? government influence on political content moderation in India and Thailand*. Carnegie Endowment for International Peace. <https://carnegieendowment.org/research/2024/07/india-thailand-social-media-moderation?lang=en¢er=china>
- Speers, D., & Truu, M. (2024, September 11). Social media companies could be fined for not cracking down on misinformation under new bill. *ABC News*. <https://www.abc.net.au/news/2024-09-11/acma-crackdown-social-media-disinformation-bill/104339144>
- Stryker, C. (2020, November 16). *Budding digital authoritarianism in Central Asia*. Davis Center. <https://daviscenter.fas.harvard.edu/insights/budding-digital-authoritarianism-central-asia>

- Su, S. (2018, August 15). *Update on Myanmar*. Meta. <https://about.fb.com/news/2018/08/update-on-myanmar/>
- Szicherle, P., & Krekó, P. (2021, June 7). *Disinformation in Hungary: From fabricated news to discriminatory legislation*. Heinrich Böll Stiftung. <https://eu.boell.org/en/2021/06/07/disinformation-hungary-fabricated-news-discriminatory-legislation>
- Tejkalová, A. N. (2023, August 31). *The battle against disinformation and Russian propaganda in Central and Eastern Europe*. European Journalism Observatory. <https://en.ejo.ch/media-politics/the-battle-against-disinformation-and-russian-propaganda-in-central-and-eastern-europe>
- Than, K. (2024, October 5). Thousands protest in Hungary demanding end to state media "propaganda." *Reuters*. <https://www.reuters.com/world/europe/thousands-protest-hungary-demanding-end-state-media-propaganda-2024-10-05/>
- Törnberg, P., & Chueri, J. (2025). When do parties lie? Misinformation and radical-right populism across 26 countries. *The International Journal of Press/Politics*, 0(0). <https://doi.org/10.1177/19401612241311886>
- Torres, M. (2024, November 29). People are watching tiktoks at 2x speed - this is what it's doing to their brains. *Yahoo!* <https://www.yahoo.com/lifestyle/people-watching-tiktoks-2x-speed-080019797.html>
- Tozo, J. (2024, September 13). *Distributed denial of truth (DDOT): The mechanics of influence operations and the weaponization of social media*. Leading Managed Detection and Response. <https://www.trustwave.com/en-us/resources/blogs/spiderlabs-blog/distributed-denial-of-truth-ddot-the-mechanics-of-influence-operations-and-the-weaponization-of-social-media/>
- Trufil, G. (2023, April 22). *Digital marketing strategy - the importance of the first few seconds in digital ads*. Split Second Research. <https://splitsecondresearch.co.uk/2023/04/22/the-importance-of-the-first-few-seconds-in-digital-ads/>
- Tsedendamba, B. (2024, June 25). *Caught between two giants: How mongolia's democracy can counter authoritarian propaganda*. Power 3.0: Understanding Modern Authoritarian Influence. <https://www.power3point0.org/2024/06/25/caught-between-two-giants-how-mongolias-democracy-can-counter-authoritarian-propaganda/>
- Vo, Samantha. (2022). *Grassroots organizations using social media as a lobbying technique*. CGU Theses & Dissertations, 461. https://scholarship.claremont.edu/cgu_etd/461
- Wagner, K. (2017, October 24). Twitter is changing its advertising policies following Russia's election interference. *Vox*. <https://www.vox.com/2017/10/24/16536934/twitter-change-policy-see-new-advertisements-russia-presidential-election>

- Ward, C., Polglase, K., Shukla, S., Mezzofiore, G., & Lister, T. (2020, April 11). Russian election meddling is back – via ghana and nigeria – and in your feeds. *CNN*.
<https://edition.cnn.com/2020/03/12/world/russia-ghana-troll-farms-2020-ward/index.html>
- We Are Social. (2014). *Digital 2014: Global social media users pass 2 billion mark*.
<https://datareportal.com/reports/digital-2014-global-social-media-users-pass-2-billion-mark>
- Wijayanto, Berenschot, W., Sastramidjaja, Y., & Ruijgrok, K. (2024). The infrastructure of domestic influence operations: Cyber troops and public opinion manipulation through social media in Indonesia. *The International Journal of Press/Politics*.
<https://doi.org/10.1177/19401612241297832>
- Wilson, J., & Keaten, J. (2025, January 23). Spain's leader wants the EU to "make social media great again." here's how. *AP News*. <https://apnews.com/article/europe-social-media-pedro-sanchez-8feda2d68fee23a2cb4517befd3fa>
- World Bank. (2025). *Individuals using the internet (% of population)*. International Telecommunication Union (ITU) World Telecommunication/ICT Indicators Database.
<https://data.worldbank.org/indicator/IT.NET.USER.ZS>
- Worley, W. (2018, June 13). *Misinformation is shaping the Colombian election*. Centre for International Governance Innovation. <https://www.cigionline.org/articles/misinformation-shaping-colombian-election/>
- Yle News. (2021, April 20). Finland ranks second in press freedom, but online hate speech a growing threat. <https://yle.fi/a/3-11892787>
- Zadrozny, B., Popken, B., Collins, B., & Lehren, A. W. (2018, May 11). Congress releases 3,000 Russian ads - including some that targeted fans of Fox News. *NBC News*.
<https://www.nbcnews.com/tech/tech-news/sean-hannity-black-lives-matter-among-targets-russian-influence-campaign-n872926>
- Zahn, M. (2025, January 7). Here's why Meta ended fact-checking, according to experts. *ABC News*. <https://abcnews.go.com/US/why-did-meta-remove-fact-checkers-experts-explain/story?id=117417445>
- Zilber, A. (2025, January 7). Zuckerberg announces end to Facebook's third-party fact-checking, admits model became a tool for censorship: "too many mistakes." *New York Post*.
<https://nypost.com/2025/01/07/business/meta-boss-mark-zuckerberg-says-facebook-will-end-fact-checking/>
- The Zimbabwe Independent. (2024, June 11). *Pro-Kremlin Disinformation in Zimbabwe: A mix of state and non-state actors' influence*. <https://www.newsday.co.zw/theindependent/opinion-analysis/article/200028249/pro-kremlin-disinformation-in-zimbabwe-a-mix-of-state-and-non-state-actors-influence>

Appendix A: Navigating the Synthesis Map

This synthesis map, also shown in Figure 12 of this document, acts as a foundation for contextualizing the social media information space, outlining systemic vulnerabilities, and identifying potential intervention points. It acts as an introduction to the landscape of global political influence operations that are carried out on social media, and readers who are interested in learning more about the concepts covered may do so through the supplementing written paper.

Overall, the system is likened to the Tower of Babel, as social media has become a complex and overwhelming information space where there is constant communication breakdown. Beginning at the top left of the map, readers are introduced to what influence operations are, as well as some of the common approaches used in coordinated attacks. These tactics come into play in the second section of the map, which offers a summary of global influence operations and mitigations around the world. Also outlined are the five actor groups, who they are, their goals, and their power within the social media space. These groups are used to frame the key issues identified within the system, and their conflicting interests lead the audience to the biggest barrier to creating positive change: fragmented approaches. This is further explained through the conclusion points that challenge the reader to reconsider the social media space through a systemic lens, considering each actor group's perspective when thinking about why issues continue to persist.

By providing readers with an overview of the system and its challenges, the synthesis map can be food for thought that allows the audience to explore a large and ever-expanding space like social media. Because social media is a force that isn't going away, and it will continue to be a space that many people interact with, it would be beneficial for users to better understand how information flows to more efficiently and wisely navigate different social media platforms.

A PDF copy of **Synthesis map on global political influence operations on social media** was also published on OCAD University's Open Research Repository on April 30, 2025.