



Public Inquiry Into Foreign Interference
in Federal Electoral Processes and
Democratic Institutions

Enquête publique sur l'ingérence étrangère
dans les processus électoraux et les
institutions démocratiques fédéraux

Overview Report:

Introduction to Social Media

Prepared by: Research Council and Commission Counsel

Summary of Report

This Overview Report summarizes the nature and functions of social media and how they have developed over time.

It also summarizes how and why social media are used, as well as how they create both benefits and vulnerabilities for democracy and healthy political discourse. Vulnerabilities include the potential for foreign interference and foreign influence.

Finally, the report provides a brief overview of the various means by which social media are, or can be, regulated to protect the integrity of democratic institutions such as the elections process.

Note to Reader

Pursuant to Rules 42-44 of the Commission's *Rules of Practice and Procedure*, the following Overview Report contains a summary of background facts and documents relating to the Commission's mandate.

Overview Reports allow facts to be placed in evidence without requiring the facts and related documents to be presented orally by a witness during the public hearings.

Overview Reports may be used to assist in identifying issues relevant to the Commission, make findings of fact and enable recommendations to be made by the Commission.

Parties have been provided an opportunity to comment on the accuracy of this Overview Report. Commission Counsel and the Parties may call evidence from witnesses at the Inquiry that casts doubt on the accuracy of the content of the documents underlying this Report. The Parties may also make submissions regarding what, if any, weight should be given to this Report and the cited documents.

Contents

Summary of Report.....	2
Note to Reader.....	2
1. Introduction	4
2. What is Social Media?.....	4
3. Why Social Media Can be Exploited by Foreign Actors.....	8
4. Effects of Foreign Influence and Foreign Interference on Social Media Platforms.....	13
5. Responses to Risks Presented by Social Media	14

1. Introduction

- [1] This Overview Report provides a high-level understanding of the nature of social media, and their use by private persons, corporations and governments.
- [2] The report also covers how social media create both benefits and vulnerabilities for democracy and healthy political discourse. Vulnerabilities include the potential for foreign interference and influence.
- [3] Finally, the report discusses various means by which social media are, or can be, regulated to protect the integrity of democratic institutions such as the elections process.

2. What is Social Media?

- [4] Social media platforms are online resources that allow users to build connections and network with one another. Users do this by producing, interacting with, and consuming content generated by other users.¹
- [5] Many of these platforms are, first and foremost, businesses. As users engage in producing, consuming and interacting with content, their actions generate three kinds of data that platform owners use to generate revenue:
 - a. **Public data** include posts and content that users generate.
 - b. **Private data** include user actions like searches and “likes.”
 - c. **Meta-data** include information gleaned by platforms such as user locations and usage patterns.²
- [6] Social media platforms use servers and databases to store and manage the public, private and meta-data users generate through their content, profiles, interactions and activities. These data serve as input to predict user behaviours both on the platform and

¹ Ji, Q. (2024). *Social Media and Society*. Routledge, **COM0000458**.

² Ji, Q. (2024). *Social Media and Society*. Routledge, **COM0000458**.

*Overview Report:
Introduction to Social Media*



Public Inquiry Into Foreign Interference
in Federal Electoral Processes and
Democratic Institutions

elsewhere. The data allow social media platforms to target content and, in some cases, advertising. The data also serve as a product that social media platforms offer for sale. Because data are the primary currency for social media platforms, the more users a platform has, and the more time users spend on the platform, the higher the value of the network. This effect snowballs when social media platforms are densely populated. The more a platform generates data for platform owners, the more value they offer their users, who are then able to reach other users more efficiently, and so on.³

- [7] Social media date back at least to the early 1980's when Bulletin Board Systems became a popular means for like-minded tech thinkers to discuss and share files. This expanded to the public at large in the late 1980's with the arrival of "**Compuserve**," whose interface made networking attractive to less computer-oriented citizens.
- [8] In the early 2000's, platforms like **Friendster** and **MySpace**, followed by **Facebook** in 2004, and **Twitter** in 2006, aimed to improve the user experience in connecting individuals online, introducing features like news feeds and profiles.
- [9] From the 2010's, smartphones began to saturate the population, accelerating social media usage, and preparing the ground for platforms like **Instagram** (2010) and **Snapchat** (2011), which specialized in sharing visual content. That same year, Chinese tech giant Tencent launched **WeChat**, which became the dominant social media platform for Chinese-language speakers. In 2021, WeChat was reorganized into closely related apps **Weixin** and WeChat. Weixin users have Chinese phone numbers, while those with overseas numbers use WeChat. The two platforms have different terms and conditions, including different kinds of speech limits. Whereas Weixin is primarily used in China, WeChat has become the dominant social media platform for Chinese-language speakers outside China. In 2023, company registration filings showed that the Chinese government had taken a "golden share" of 1% in a domestic subsidiary of

³ Tom Casey, Rethinking Engagement: Challenging the Financial Model of Social Media Platforms Harvard Data Science Review. 27 January 2022.

<https://hdsr.mitpress.mit.edu/pub/9lj3cuah/relase/3#nli3jdav1tm>, **COM0000447**; Philip Meier, "How do Digital Platforms Make their money?" Blogpost posted by Humboldt Institut für Internet und Gesselschaft, 29 July 2019. <https://www.hiig.de/en/how-do-digital-platforms-make-their-money/>, **COM0000463**.

*Overview Report:
Introduction to Social Media*



Public Inquiry Into Foreign Interference
in Federal Electoral Processes and
Democratic Institutions

Tencent holdings, which allows the government to gain access to online data, and in some cases, board seats, and other benefits of that nature.⁴

- [10] Originally conceived as a messaging application, WeChat/Weixin expanded to include features such as mobile payments, gaming and e-commerce, transforming into a comprehensive platform for daily life in China, but also among Chinese diaspora members. Weixin is so pervasive in its use and services that it can be difficult for people in China to live without it. For example, even public transit uses Weixin's platform as a payment medium in China. Because so many in China use Weixin, diaspora members use WeChat, which can access and engage with Weixin accounts, as it is the easiest way to keep in touch with friends and family who remain within China.⁵
- [11] For Russian speakers, the key social media platform is **VKontakte**, with about 85 million users worldwide.⁶ This platform was founded in 2006 by Pavel Durov, who went on to found the free speech absolutist platform **Telegram**. VKontakte allows users to post public and private messages, play games, share and like audio and video content, and follow news. In 2014, Durov resigned as CEO, claiming that allies of President Vladimir Putin had taken control of the platform, and were using it to gather information

⁴ Ye, J. "Beijing takes 'golden share' in a Tencent subsidiary, records show", Reuters, October 19, 2023, **COM0000472**; I. Deng "Tencent draws a line between WeChat and Weixin, telling users to choose as China's strict new data laws come into effect", South China Morning Post, September 7, 2021, **COM0000509**.

⁵ Bajaj, Shelly Ghai. Disinformation in a Diverse Digital Landscape: Ethnocultural Diaspora Experiences and Impacts, Commission Paper 2024, **COM0000445(EN)/COM0000445.FR**; Shah, Saqib (May 14, 2016). "The history of social networking," *Digital Trends*. <https://www.digitaltrends.com/computing/the-history-of-social-networking/>, **COM0000467**.

⁶ A. Melkadez, Number of monthly active mobile users of VKontakte (VK) from March 2016 to December 2023, *Statista*, July 8, 2024: <https://www.statista.com/statistics/425429/vkontakte-mobile-mau/>, **COM0000510**.

*Overview Report:
Introduction to Social Media*



Public Inquiry Into Foreign Interference
in Federal Electoral Processes and
Democratic Institutions

on dissidents, as well as on Ukrainians at the start of the Euromaidan conflict.⁷ As of 2021, Russian state entities own a majority share of the company.⁸

- [12] Social media use in Canada is extremely high, even in comparison to similar countries. Indeed, 94% of Canadians are regular Internet users with at least one social media account.⁹
- [13] Notably, messaging applications including WhatsApp, Telegram, **Signal** and, for Chinese-language users, WeChat, are especially popular with Canadians. Their popularity is particularly high with those between 18 and 34, as well as with diaspora communities who use these technologies to communicate easily with social circles in their countries of origin.¹⁰
- [14] One key difference between these diverse platforms is the extent of, and means through which, they maintain security and privacy. A social media platform's approach to security and privacy may be influenced by consumer demand: for example, users of Signal prioritize security and privacy. Or a platform's approach may be driven in part by conformity with government regulations around content. For example, WeChat is subject to a range of access requirements by Chinese government regulations.¹¹ On every platform, "back end" technological components work together with "front end" user-interfaces to create an experience for social media users while managing the

⁷ A. Soldatov and I. Borogan, "Durov in Detention: The End of Tech Titan Immunity?", August 27, 2024, Center for European Policy Analysis: <https://cepa.org/article/durov-in-detention-the-end-of-tech-titan-immunity/>, **COM0000517**.

⁸ A. Marrow, "CEO of Russia's VK resigns as state assumes control of internet firm", Reuters, December 3, 2021: <https://www.reuters.com/article/technology/ceo-of-russias-vk-resigns-as-state-assumes-control-of-internet-firm-idUSL8N2SO3IY/>, **COM0000498**.

⁹ Gruzd, Anatoliy & Mai, Philip (2022). The State of Social Media in Canada 2022. Social Media Lab Toronto Metropolitan University. DOI: 10.6084/m9.figshare.21002848, **COM000462**.

¹⁰ Gruzd, Anatoliy & Mai, Philip (2020). The State of Social Media in Canada 2020. Ryerson University Social Media Lab. Version 5. DOI: 10.5683/SP2/XIW8EW, **COM0000453**. For a detailed breakdown of Canadians social media platform use, please see: <https://socialmedialab.ca/2022/09/14/survey-finds-canadians-are-spending-less-time-on-social-media-but-tiktok-is-the-exception/>

¹¹ Bajaj, Shelly Ghai. Disinformation in a Diverse Digital Landscape: Ethnocultural Diaspora Experiences and Impacts, Commission Paper 2024, **COM0000445(EN)/COM0000445.FR**.

complexities of data storage and analysis, conforming to government regulations and providing diverse levels of privacy and security.

- [15] Canadians use social media for many purposes, but key among them is news consumption. A sizable proportion of Canadians regularly get most of their news from social media (24%) or the Internet (33%).¹² In one survey, 61% of respondents reported that, even when they rely on news from traditional sources, they use a mobile device for access.¹³
- [16] Not everyone who uses social media posts content, i.e., produces public data. Factors like privacy concerns may impact how active social media users are. Researchers have suggested a “1–9–90 rule,” which means that, in any given online community, 90% of users consume content passively, 9% of users contribute to content editing and merely 1% of users actively create new content.¹⁴

3. Why Social Media Can be Exploited by Foreign Actors

- [17] Social media can have a strong positive effect on democracy: social networking can bring political news to the attention of more people, rallying citizens around issues, helping them to organize and thus helping to build civil society. It can thus increase political literacy and political engagement.¹⁵ But ample research demonstrates social

¹² Statistics Canada. Canadian Social Survey – Quality of Life, Virtual Health Care and Trust, 2023-11-10. <https://www150.statcan.gc.ca/n1/daily-quotidien/231110/dq231110b-eng.htm>, **COM0000468**.

¹³ Newman, Nic, Journalism, Media, and Technology Trends and Predictions 2022. Reuters Institute for the Study of Journalism Report. DOI: 10.60625/risj-ahx9-vm24, **COM0000464**.

¹⁴ Reed, C. (2020). How to leverage the 1–9–90 rule and become a leader on LinkedIn. *Forbes.com*. <http://www.forbes.com/sites/forbesbusinesscouncil/2020/07/10/how-to-leverage-the-1-9-90-rule-and-become-a-leader-on-linkedin/?sh=69f6f6eb7d32>, **COM0000466**.

¹⁵ Lorenz-Spreen, P., Oswald, L., Lewandowsky, S., & Hertwig, R. (2023). A systematic review of worldwide causal and correlational evidence on digital media and democracy. *Nature Human Behaviour*, 7(1), 74–101. <https://doi.org/10.1038/s41562-022-01460-1>, **COM0000461**; Jha, C. K., & Kodila-Tedika, O. (2020). Does social media promote democracy? Some empirical evidence. *Journal of Policy Modeling*, 42(2), 271–290. <https://doi.org/10.1016/j.jpolmod.2019.05.010>, **COM0000457**.

media also have characteristics that can be exploited to undermine democracy, and one way this can happen is through foreign interference.¹⁶

- [18] Several characteristics of social media create opportunities for foreign influence or interference. The first factor is **exposure**: since so many Canadians use social media, including to get news,¹⁷ it follows that this provides a path to saturate the population with messaging.
- [19] The second factor is the phenomenon researchers call “**echo-chambers**”: as Jamieson and Cappella define them, echo chambers are “a bounded, enclosed media space that has the potential to both magnify the messages delivered within it and insulate them from rebuttal.”¹⁸ Social media users can be exposed primarily to content that reinforces existing political leanings and, unlike in-person political discourse, people may have fewer occasions to engage civilly with people who disagree.
- [20] It is worth noting that not all researchers agree on the role of echo chambers. A growing number of studies suggest echo chambers have less impact than previously assumed. Furthermore, some researchers have questioned whether empirical evidence directly supports the claim that exposure to certain kinds of social media content *causes* rather than *aligns with*, political polarization. For example, recent studies, conducted across various election cycles, support the claim that concentrated messaging in social media feeds have no significant impact on people’s beliefs and ultimate voting intentions.¹⁹

¹⁶ Bajaj, Shelly Ghai. Disinformation in a Diverse Digital Landscape: Ethnocultural Diaspora Experiences and Impacts, Commission Paper 2024, **COM0000445(EN)/COM0000445.FR**; Olaniran, B., & Williams, I. (2020). Social media effects: Hijacking democracy and civility in civic engagement. Platforms, protests, and the challenge of networked democracy, 77 – 94. https://doi.org/10.1007/978-3-030-36525-7_5, **COM0000465**.

¹⁷ Statistics Canada. Canadian Social Survey – Quality of Life, Virtual Health Care and Trust, 2023-11-10. <https://www150.statcan.gc.ca/n1/daily-quotidien/231110/dq231110b-eng.htm>, **COM0000468**.

¹⁸ Jamieson, K. H., & Cappella, J. N. (2008). *Echo chamber: Rush Limbaugh and the conservative media establishment*. Oxford University Press. <https://doi.org/10.1002/j.1538-165x.2009.tb01921.x>, p. 76, **COM0000456**.

¹⁹ Guess, A., Malhotra, N., Pan, J. Barbera, P., et al. (2023) How do social media feed algorithms affect attitudes and behavior in an election campaign? *Science*. 381 (6656), pp. 398–404. <https://doi.org/10.1126/science.abp9364>, **COM0000455**; Groshek, J. & Koc-

- [21] The third factor is **sensationalism**: as Deibert notes, people are, on average, attracted to content that has sensational, extreme or shocking elements.²⁰ This is important because social media platforms use **algorithms** to determine what content users see: the more a user looks at content, the more of that kind of content the platform will show them, and the more likely the platform will be to show the content to others.²¹ In short, algorithms amplify sensational or extreme content.
- [22] Algorithms operate throughout the software architecture of a social media platform: they are mathematical models that use rules and calculations to process data in ways that generate useful information for social media platform owners. An algorithm processes the data produced by each individual user, and then applies the rules and calculations to determine what content social media users see. Algorithms use public, private and meta-data, including users' "likes," searches, the content they produce, the content they spend time on, whom they follow, where they are located, etc. Based on these data, the algorithm predicts which posts a user might be most interested in and prioritizes those in their feed.²² The effect is, in theory, to amplify echo chambers, and especially the fringe ends of the chamber. When algorithms amplify echo-chambers, suggesting content to users that is in line with their existing preferences, researchers – some of whom are skeptical of the phenomenon - may refer to this as a "**filter bubble**."²³
- [23] The exact rules and calculations that a social media platform uses, that is, the exact mathematical models it employs, is proprietary information not available to the public. The goal of these algorithms is to keep users engaged by showing them content they

Michalska, K. (2017). Helping populism win? Social media use, filter bubbles, and support for populist presidential candidates in the 2016 US election campaign. *Information, Communication & Society*, 20(9), 1389– 1407.

<https://doi.org/10.1080/1369118x.2017.1329334>, **COM0000452**.

²⁰ Deibert, R. (2020). *Reset*. House of Anansi Press, **COM0000450**.

²¹ Warnke, Lina, Maier, Anna-Lena, Gilbert, Dirk Ulrich (2024) Social media platforms' responses to COVID-19-related mis- and disinformation: the insufficiency of self-governance, *Journal of Management and Governance*. <https://doi.org/10.1007/s10997-023-09694-5>, **COM0000471**.

²² Ji, Q. (2024). *Social Media and Society*. Routledge, **COM0000458**.

²³ Borgesius, et al. "Should we worry about filter bubbles?" *Internet Policy Review*, Volume 5, Issue 1, 31 March 2016. DOI: 10.14763/2016.1.401, **COM0000446**.

are likely to enjoy. This encourages them to spend more time on the platform.

Sometimes called the attention economy, platforms are motivated to encourage users to stay, because that in turn generates more data and potential revenue for the platform owners.²⁴

- [24] Those who may be engaged in foreign interference or foreign influence can capitalize on these features of algorithms by using bots and troll farms. **Bots** are software robots that, like other robots, are created to save human labour. Social bots play this role on social media by mimicking human behaviour in an attempt to influence both algorithms and individual people. For example, bots can artificially make certain content, notably low credibility content, appear more popular. To do this, bots:

[...] amplify interactions with content as soon as it is created to make it look legitimate and to facilitate its spread across social networks. Next, [bots] try to increase public exposure to the created content and thus boost its perceived credibility by targeting influential users that are more likely to believe disinformation in the hope of getting them to “repost” the fabricated content.²⁵

- [25] Bots can also create content, and imitate humans in initiating friend requests or follows, as well as liking or even replying to posts. Humans may be unaware that they are engaging with a robot. Bots are also capable of scraping data from social media platforms, for example by following key words.²⁶ This is important because it allows not just companies but states to search for dissident material online.²⁷ These capabilities are expected to be greatly enhanced in the near future with the use of AI.²⁸

²⁴ Ji, Q. (2024). *Social Media and Society*. Routledge, COM0000458; Deibert, R. (2020). *Reset*. House of Anansi Press, **COM0000450**.

²⁵ Aïmeur, E., Amri, S. & Brassard, G. (2023) Fake news, disinformation and misinformation in social media: a review. *Soc. Netw. Anal. Min.* 13, 30. <https://doi.org/10.1007/s13278-023-01028-5>, **COM0000444**.

²⁶ Ji, Q. (2024). *Social Media and Society*. Routledge, **COM0000458**.

²⁷ Bajaj, Shelly Ghai. Disinformation in a Diverse Digital Landscape: Ethnocultural Diaspora Experiences and Impacts, Commission Paper 2024, **COM0000445(EN)/COM0000445.FR**.

²⁸ B. Wampler, “AI among us: Social media users struggle to identify AI bots during political discourse”, Notre Dame News, February 27, 2024: <https://news.nd.edu/news/ai-among-us-social-media-users-struggle-to-identify-ai-bots-during-political-discourse/>, **COM0000518**; Cybersecurity and Infrastructure Security Agency (United States), “Social Media Bots Infographic Set”, **COM0000508**.

*Overview Report:
Introduction to Social Media*



- [26] During the 2016 United States presidential election, Twitter identified over 50,000 bots that collectively reached millions of American users, attempting to impact public opinion.²⁹
- [27] A **troll**, like a bot, aims to artificially produce and promote content. Trolls, unlike bots however, are human, though trolls may in some cases make use of bots. An organized group of trolls is sometimes called a **troll farm**. The Russian “Internet Research Agency” (IRA) is perhaps the most famous example of a troll farm. A U.S. Department of Justice report found that the IRA had “used social media accounts and interest groups to sow discord in the U.S. political system through what it termed ‘information warfare’.” According to the same report, “the campaign evolved from a generalized program designed in 2014 and 2015 to undermine the U.S. electoral system, to a targeted operation that by early 2016 favored candidate Trump and disparaged candidate Clinton.”³⁰
- [28] Canadians’ active presence online means a broad, easily accessible audience for those who might attempt foreign interference or foreign influence. Trolls, troll farms and bots are then used to attempt to shape both the content and tone of online discourse, which social media platform algorithms may amplify. Together, these elements offer a prime environment for foreign actors who may wish to influence, or interfere in, our democracy. They can use these tools to harm democracy in at least three ways: subverting democratic norms, intimidating political opponents and poisoning the information environment.

²⁹ Ji, Q. (2024). *Social Media and Society*. Routledge, p163, **COM0000458**.

³⁰ *Report on the Investigation Into Russian Interference in the 2016 Presidential Election*, Special Counsel Robert S. Mueller III, U.S. Department of Justice, Washington D.C., March 2019. <https://www.justice.gov/archives/sco/file/1373816/dl>, **COM0000576**.

4. Effects of Foreign Influence and Foreign Interference on Social Media Platforms

- [29] Foreign actors may attempt to use social media to **subvert democratic norms**. For a democracy to function, it is necessary that citizens have some justified trust in the fairness and effectiveness of institutions, that they can engage civilly with fellow citizens who hold divergent viewpoints and that they treat each other, in political contexts, with respect even when they disagree.
- [30] But foreign state actors, using social media tools, can attempt to seed and foster doubt in institutions, normalize hateful discourse and thus increase polarization. In addition, by continuously painting everyday policy disputes as existential threats, they attempt to fray the fabric of democratic society.³¹ Their efforts are supported by the fact that extreme or sensational content tends to get more user engagement on social media platforms.³² By posting this kind of content, those intent on foreign influence and interference can count on algorithms to amplify it, because these algorithms amplify whatever initially garners more attention. This can make extreme views appear more common than they are, potentially normalizing polarization, while also decreasing political civility and mutual engagement.
- [31] Foreign state actors use social media platforms as a means to engage in **intimidation and political repression**, both of which may decrease targets' political participation.
- [32] There are two key ways social media platforms can be used to repress political participation. The first is to create reputational threats to deter people from running for office or campaigning for an issue. Social media platforms can be used to spread rumours or engage in **doxing**, a practice whereby a person's personal information is released on the Internet, enabling others to harass a person offline, including at the person's home. The possibility of such campaigns can deter political participation.

³¹ Council of Canadian Academies (2023). *Fault Lines: Expert Report on the Socioeconomic Impacts of Science and Health Misinformation*, **COM0000448**.

³² Deibert, R. (2020). *Reset*. House of Anansi Press, **COM0000450**.

- [33] Second, foreign states can monitor social media activity of diaspora members abroad, using that information to track their activities, connections and communications. Evidence suggests that information can then be used to threaten dissidents and their families. This can also deter political participation.³³ When citizens do not feel safe exercising their political rights in Canada, this harms our democracy.
- [34] For citizens to assess policy positions and make voting decisions, a rich and varied, but sound and trustworthy information environment is critical. Disinformation, which involves the intentional and malicious spread of false material (i.e. of misinformation), poisons that environment. The same three mechanisms are ripe for spreading mis- and disinformation on social media platforms: in addition to individuals who share, “like,” or otherwise promote it either wittingly or unwittingly, bots designed to act as pseudo-people can spread disinformation; trolls can both produce and spread disinformation; and algorithms can drive mis- and disinformation content into people’s feeds.³⁴

5. Responses to Risks Presented by Social Media

- [35] Societies have taken a range of different approaches to respond to the risks to democratic health that exist on social media platforms. These approaches may include legislative measures, as well as standards and guidelines that invite platforms to self-govern. In such cases, social media platforms, whether individually or collectively, draw up codes and commit to adhere to them.³⁵ The latter is sometimes called “**industry self-governance.**”
- [36] Managing the risks social media present to democratic society can be challenging. There are multiple reasons for this. First, democracies place a high value on free

³³ Bajaj, Disinformation in a Diverse Digital Landscape: Ethnocultural Diaspora Experiences and Impacts, Commission Paper 2024, **COM0000445(EN)/COM0000445.FR.**

³⁴ Warnke, Lina, Maier, Anna-Lena, Gilbert, Dirk Ulrich (2024) Social media platforms’ responses to COVID-19-related mis- and disinformation: the insufficiency of self-governance, *Journal of Management and Governance* <https://doi.org/10.1007/s10997-023-09694-5>, **COM0000471.**

³⁵ Cusumano, M. A., Gawer, A., & Yoffie, D. B. (2021). Can self-regulation save digital platforms? *Industrial and Corporate Change*, 30(5), 1259–1285. <https://doi.org/10.1093/icc/dtab052>, **COM0000449.**

expression. While freedom of expression in Canada is not an absolute right, there are both political and legal limits on what actions may be taken on social media that restrict expression. This makes determining which content ought to be limited on social media complex, particularly in cases where there is a dispute over the falsity of content.

- [37] Second, social media platforms' business model depends on maximizing user engagement. Since users engage more with inflammatory content, these businesses may have incentives to resist regulation of false or hateful content.
- [38] Third, the sheer volume of content, and the capacity of both humans and robots to learn how to evade safeguards makes the task of protecting democracy in the online environment challenging.
- [39] Bearing these challenges in mind, social media platforms in Canada do operate within regulatory frameworks. Some social media platforms take “self-governance” measures, and in addition, there are legal frameworks with which they must conform.
- [40] Social media platforms employ “flagging” and “tagging” approaches to self-governance. The **flagging** approach aims to identify and remove mis- and disinformation before users are exposed to it. The **tagging** approach leaves the false material in place but provides corrective information and links to sources so that users can engage with the false material more critically.³⁶
- [41] Material that users and platform administrators come across and identify as false or hateful can be fact-checked before it is either removed or tagged. “[B]y and large, fact-checking [and then tagging posts] can be effective in helping to fight fake news, with average news users becoming more attuned to accurate information after a single exposure to a fact-checking message”.³⁷

³⁶ Lazer, D. et al. (2018). The science of fake news. *Science*, 359(6380), 1094–1096. <https://doi.org/10.1126/science.Aao2998>, COM0000460.

³⁷ Walter, N., Cohen, J., Holbert, R. L., & Morag, Y. (2020). Fact-checking: A meta-analysis of what works and for whom. *Political Communication*, 37(3), 350–375. <https://doi.org/10.1080/10584609.2019.1668894>, COM0000469; Ji, Q. (2024). *Social Media and Society*. Routledge, p.141. COM0000458.

[42] Flagging and tagging by individuals can be laborious, particularly given the presence of bots and troll farms, which may rapidly and continuously flood social media platforms with false information. For this reason, many social media platforms use **algorithmic moderation**, a process where computer models are deployed to either classify or match content likely to violate guidelines. That material, like hand-flagged material, can then be checked and either tagged or removed. Where the algorithm suggests a bot may be involved, fake accounts can be removed as well.³⁸

[Algorithmic m]atching involves transforming a piece of content into a hash—a string of unique identifiers that is essentially a “digest” of the message. The system then compares that hash with a known library of hashes that are deemed problematic to see if the new content contains identical elements.³⁹

[43] The advantage of this approach is that it prevents individuals, bots, or trolls from simply re-posting material that has been removed. Classification algorithms use statistical patterns to assess the likelihood that any given piece of content violates the social media platform’s norms and guidelines.

[44] In Canada, social media companies that moderate online content by such means largely do so on their own initiative. Though Canadian law does regulate some aspects of online content – such as privacy protection, intellectual property, defamation or the dissemination of content in contravention of the *Criminal Code* (e.g. child pornography, incitement of hatred) –so far, Canada has not adopted any general legislative measures directed at prohibiting online content solely because it is untrue or inflammatory.⁴⁰

³⁸ Gorwa, R., Binns, R., & Katzenbach, C. (2020). Algorithmic content moderation: Technical and political challenges in the automation of platform governance. *Big Data & Society*, 7(1), <https://doi.org/2053951719897945>, **COM0000451**.

³⁹ Ji, Q. (2024). *Social Media and Society*. Routledge, p. 53, **COM0000458**.

⁴⁰ Emily Laidlaw, “Mis- Dis- and Mal-Information and the Convoy: An Examination of the Role and Responsibilities of Social Media,” Paper produced for the Public Order Emergency Commission, 2022: <https://publicorderemergencycommission.ca/documents/policy-papers/>, **COM0000459**.