Centre for International Governance Innovation

Digital Policy Hub

Digital Policy Hub — Working Paper

Governing Decentralized Social Media in a Centralized Policy World

Alexander Martin

Fall term — 2025–2026 cohort

About the Hub

The Digital Policy Hub at CIGI is a collaborative space for emerging scholars and innovative thinkers from the social, natural and applied sciences. It provides opportunities for undergraduate and graduate students and post-doctoral and visiting fellows to share and develop research on the rapid evolution and governance of transformative technologies. The Hub is founded on transdisciplinary approaches that seek to increase understanding of the socio-economic and technological impacts of digitalization and improve the quality and relevance of related research. Core research areas include data, economy and society; artificial intelligence; outer space; digitalization, security and democracy; and the environment and natural resources.

The Digital Policy Hub working papers are the product of research related to the Hub's identified themes prepared by participants during their fellowship.

Partners

Thank you to Mitacs for its partnership and support of Digital Policy Hub fellows through the Accelerate program. We would also like to acknowledge the many universities, governments and private sector partners for their involvement allowing CIGI to offer this holistic research environment.

About CIGI

The Centre for International Governance Innovation (CIGI) is an independent, non-partisan think tank whose peer-reviewed research and trusted analysis influence policy makers to innovate. Our global network of multidisciplinary researchers and strategic partnerships provide policy solutions for the digital era with one goal: to improve people's lives everywhere. Headquartered in Waterloo, Canada, CIGI has received support from the Government of Canada, the Government of Ontario and founder Jim Balsillie.

About the Author

Alexander Martin is a Digital Policy Hub doctoral fellow and a Ph.D. candidate in science and technology studies at York University. His research focuses on information integrity, digital governance and digital communication infrastructure in Canada. He is a deputy editor at Open Canada, where he works with authors on Canadian foreign policy and technology policy issues. He previously held policy and strategic communications roles at the North Atlantic Treaty Organization's Office of the Chief Scientist and the United Nations Information Integrity Unit, focusing on mis- and disinformation and hate speech in peace operations. His work examines how public-interest digital infrastructure and governance can strengthen democratic resilience and digital sovereignty in Canada.

67 Erb Street West
Waterloo, ON, Canada N2L 6C2
cigionline.org

Policy Memo

Governing Decentralized Social Media in a Centralized Policy World

Alexander Martin

Bottom Line Up Front

Decentralized social media platforms such as Mastodon and Bluesky offer community-governed, interoperable alternatives to centralized social media. Yet, in Canada, current and emerging platform regulations do not reflect how decentralized platforms operate. These rules could overburden volunteer-run communities, discourage experimentation and push users back to dominant, foreign-owned platforms, which could undermine Canada's digital sovereignty goals. Regulators should instead adopt participatory, interoperability-based frameworks that protect these users, while supporting community-run platforms and long-term investment in Canadian digital infrastructure.

Key Points

- **Decentralized and community-based moderation reshapes governance.** Federated social media platforms such as Mastodon and Bluesky rely on local rules and community norms rather than centralized control, giving communities greater discretion over speech, safety and participation.
- **Centralized regulation is unfit for decentralized systems.** Canadian platform policies, including the proposed Online Harms Act and the Online News Act, were designed for large corporate platforms and do not clearly recognize volunteer-run and federated communities.
- **Participatory governance builds legitimacy and resilience.** Taiwan's civic technology ecosystem offers case studies that could be useful reference points for Canada in considering how to include participatory and decentralized approaches in future digital policy frameworks.
- **Proportionate policy is essential to digital sovereignty.** Regulations that reflect the realities of decentralized and open-source platforms can protect users while supporting community control, interoperability and a more democratic digital ecosystem.

Recommendations

- **Adopt a tiered, proportionate regulatory model:** The Canadian Radio-television and Telecommunications Commission (CRTC), alongside the Department of Justice, should recognize the public interest and domestic potential of the fediverse and apply proportionate, flexible oversight scaled by platform size, governance model and risk.
- **Promote cooperative compliance and transparency:** The Department of Canadian Heritage (Canadian Heritage) and the CRTC should work with alternative social media communities to support open-source moderation tools.
- **Leverage interoperability and portability for digital sovereignty:** Innovation, Science and Economic Development Canada (ISED), Canadian Heritage and the Strategic Innovation Fund should support open standards, account portability and Canadian-based federated platforms through targeted funding and infrastructure, to expand user choice, reduce reliance on foreign platforms and help Canada build self-governed digital infrastructure.
- **Invest in research and public-interest infrastructure:** The Social Sciences and Humanities Research Council, ISED and Canadian Heritage should fund research on decentralized moderation, misinformation and civic trust alongside investment in public-interest digital infrastructure.
- **Embed participatory policy co-design:** Canadian Heritage and ISED should involve civil society, platform administrators, developers and users in regulatory design, so policy reflects how decentralized systems operate. Canada should treat federated platforms as public-interest digital infrastructure when shaping future regulation.

From Centralized Platforms to Decentralized Governance

Canada and most Western policy makers continue to regulate social media platforms through frameworks built for large, centralized companies. Yet decentralized and federated social media platforms such as Mastodon and Bluesky operate through fundamentally different governance structures based on community stewardship and distributed control that have not been taken into account in most policy discussions. This working paper presents the argument that applying centralized regulatory models to decentralized platforms risks overburdening volunteer-run administrators, narrowing digital diversity and ultimately undermining Canada's broader goals for digital sovereignty.

This policy challenge is reminiscent of historical attempts to govern the internet. Early internet ideals framed digital space as open and decentralized, most notably in John Perry Barlow's "Declaration of the Independence of Cyberspace," in which he imagined the internet as a space free from centralized control (Barlow 1996). However, as social media companies such as Facebook emerged, participation expanded while power became increasingly centralized, embedding user interaction within surveillance-based business models (Srnicek 2016, 31–32; Zuboff 2019). This era grew out of the Californian Ideology, where rhetoric of openness and freedom increasingly masked the concentration of corporate power (Barbrook and Cameron 1996, 45; O'Mara 2019, 109).

Today, people are increasingly trying to take back control over how social media platforms are designed and governed. One prominent example is Mastodon, which operates within the "fediverse" or "federated universe," a network of independently run servers, or "instances," that interoperate through open protocols, allowing people to join one community while still interacting with others if they choose (Woloshyn and Fraser 2023, para. 5). Instances are essentially small, self-governed social media communities, each with its own rules and moderation practices. This model redistributes platform governance across communities rather than concentrating it in a single corporate authority.

Most fediverse platforms use the ActivityPub protocol, which enables this cross-platform communication, in contrast to the closed protocols used by platforms such as Facebook and X. ActivityPub works like email, where a user on one platform can communicate with someone on another, much like how a Gmail user can send messages to someone using Microsoft Outlook (LSE Library 2025, 17:11). While Bluesky relies on a different technical architecture called the Authenticated Transfer, or AT, Protocol, it follows the same principle of interoperability, illustrating a broader move away from centralized platform control toward distributed, community-based governance (*AP News* 2024, para. 6).

Launched in 2016, Mastodon is one of the largest and most prominent platforms in the fediverse. It comprises roughly 27,000 instances worldwide with between 10 and 14 million users (Gehl 2025, 6). Registrations to Mastodon surged after Elon Musk's 2022 takeover of Twitter, rising from 60 to 80 per hour to more than 3,500 per hour (Escritt 2022, para. 11). Each instance is independently run, with its own rules, moderation practices and community norms, which together form a federated network (Gehl 2025, 65).

Only two months after Musk's takeover, many technology journalists declared Mastodon a failure because it did not scale in the same way as X (Hoover 2023, para. 3). However, there are an increasing numbers of users who are moving from platforms like X to decentralized social media such as Bluesky, as they look for alternatives that promise responsible content moderation and community accountability (Di Placido 2024, paras. 13–14). The fediverse is now larger than at any point in its history, with growth further accelerated by Meta's decision to make its new platform, Threads, interoperable with the fediverse through the ActivityPub protocol. Usership has reached approximately 400 million monthly users (Perez 2026, para. 7).

The analysis presented in this working paper situates the fediverse in relation to digital sovereignty, understood as the ability of communities or governments to control their own digital infrastructure, data and online spaces rather than leaving them in the hands of foreign corporations (Government of Canada 2025). Mastodon is used as the primary case study because it has a longer history, having been founded in 2016, as well as a more established user base and a stronger body of academic research. This is in contrast to newer platforms, such as Bluesky, for example, which was founded in 2019 and formally launched in 2023 (Knibbs 2025, para. 3). Drawing on scholarship on decentralized social media and content moderation, this paper examines the fediverse through Mastodon, identifies gaps in Canadian regulation and proposes proportionate policy pathways that support community-governed platforms.

How the Fediverse Governs Itself: Covenants

In the fediverse, trust and accountability come from shared agreements between communities and the people who run them rather than from a single central authority. In the fediverse, the Mastodon Server Covenant is the clearest example of how this works in practice. To be listed in Mastodon's public directory, where users can search for and join an instance, a server must maintain four conditions: "active moderation against racism, sexism, homophobia and transphobia"; "daily backups"; "at least one other person with emergency access to the server infrastructure"; and a "commitment to give users at least three months of advance warning in case of shutting down."[1] Through the covenant, Mastodon turns community norms into enforceable expectations, giving users a clear basis for trust and making it possible for decentralized communities to govern themselves without a central authority.

Robert W. Gehl and Diana Zulli describe this process as "covenantal federalism," where small units consent to band together while following a shared ethical code (Gehl and Zulli 2023, 3276). A reciprocal relationship is created in which administrators must maintain trust by stewarding their instances or users can withdraw and migrate elsewhere. At the same time, users are also accountable since they can be removed from communities that they violate and must then find another community (ibid., 3282–83). Tarleton Gillespie's (2018) concept of "custodians of the internet" can be borrowed here to understand this. Actors in social media networks are responsible together for the

1 See https://joinmastodon.org/covenant.

ethical, social and informational health of their communities (ibid., 140). Applying this concept to the fediverse, they govern themselves through shared rules and user choice, where communities keep power by deciding who they connect to and how their spaces are run (Rozenshtein 2024, 187).

Why Top-Down Moderation Falls Short and Why Bottom-Up Matters

Community-based moderation on platforms like Mastodon has roots in earlier forms of decentralized online governance. Systems such as Usenet in the 1980s were made up of small communities that set their own rules and maintained discussion without a central authority (Friedl and Morgan 2024, 1). Kevin Driscoll reinforces this history by showing how early hobbyist networks relied on volunteer administrators who shaped culture, enforced rules and built trust long before corporate platforms existed (Driscoll 2022, 3).

Mastodon continues this approach: each instance defines its own rules, but while following the standards laid out by the Mastodon Server Covenant. As Alan Z. Rozenshtein (2024, 182) describes it, "no instance can control the behaviour of another, and there is no central authority." This structure allows people to speak while giving others the ability to limit their exposure to unwanted content, by simply not following them. In doing so, it puts into practice the principle that freedom of expression does not imply an entitlement to unlimited reach, as users are not necessarily exposed to everybody's content and they construct their own algorithms (ibid., 186). These dynamics reflect what Rozenshtein calls the "moderator's trilemma" (ibid., 181). Centralized platforms have huge and very different user bases; they rely on blanket, top-down rules to manage them; and, at the same time, they want to avoid upsetting large groups of users, as well as the politicians who represent them (ibid.).

By distributing authority across many instances and relying on community-driven moderation, the fediverse gives up uniformity in favour of context-sensitive approaches (ibid., 186–87). This structure supports softer forms of moderation that reduce the visibility of harmful content rather than remove it outright. Erin Kissane and Darius Kazemi (2024, 20) note that "medium-sized Fediverse servers can offer high-touch, context-sensitive, moderation that differs sharply from that of central platforms." This model allows moderators to operate closer to their communities and with far more manageable user-to-moderator ratios, making it better suited to applying local norms, detecting harmful behaviour early, and avoiding the large-scale moderation failures seen on global platforms.

Public debates about harmful online content often focus on tone and outrage rather than on how moderation systems actually work in practice (Rossini 2022, 418). In this context, Mastodon treats information integrity not as a centralized enforcement problem but as a shared responsibility, where communities collectively govern the spaces they create and maintain. Content moderation itself is one of the main ways decentralized and centralized platforms differ, a distinction that has become more important as the reach of social media has enabled harmful content to spread quickly and widely (United Nations 2024, 3).

Research shows that Mastodon communities are most concerned with terms of service (ToS) risks related to toxic content (Cramer, Maxam and Davis 2025, 2). A large-scale analysis of thousands of instance ToS found that the most common rules explicitly prohibit racism, sexism, homophobia, transphobia, xenophobia and casteism

(Gehl 2025, 77). While Mastodon's ToS are generally short and straightforward, the average ToS for centralized social media platforms would take 15 to 17 minutes to read at a normal pace. Yet 93 percent of users agreed to the ToS after spending only 51 seconds reading them (Obar and Oeldorf-Hirsch 2020, 141). This contrast shows that Mastodon uses its ToS to clearly communicate community values and expectations rather than to obscure power through dense and unreadable legal language.

Such deception by centralized platforms is reinforced by the underlying business model, which depends on algorithmic amplification to drive advertising revenue and therefore reward engagement over safety (Srnicek 2016, 36–38). As former Facebook employee and whistle-blower Frances Haugen explained, making algorithms safer would reduce time spent on the platform, ad clicks and profits (de Guzman 2021, para. 4). Following the change in ownership of Twitter (now X), researchers observed an immediate rise in hateful content containing racist, homophobic or transphobic slurs, which increased by roughly 50 percent after the takeover (Hickey et al. 2025, 3). Although X's Community Notes content moderation system allows users to add annotations to misleading posts, more than 90 percent of submitted notes are never published, leaving most proposed corrections invisible to the wider user base (Braga, Tardáguila and Soares 2025, 17). Additionally, participation remains extremely low, with fewer than one percent of users contributing, while posts from non-English speakers or the Global South are disproportionately excluded. This exposes communities to misinformation as Community Notes fail to account for local context and situated knowledge (Bassett 2025, 6).

One thing Mastodon and other decentralized social media platforms get right is maintaining a manageable ratio of moderators to users for content moderation. Take, for example, the instance Indieweb.social, which has about 11,000 registered users and 1,500 active users, with three moderators, that is, roughly one per 500 active users (Flipboard 2024, 9:00). Kissane and Kazemi (2024, 20) found in their research that "the most lightly staffed instance provides one moderator per approximately 1,800 members, and several provide at least one moderator per 100 members." At first glance this may seem low, but in contrast X employed only 2,294 moderators to serve all of the countries of the European Union, as revealed under the Digital Services Act (Chee and Mukherjee 2023, para. 4).

Decentralized Moderation in Practice: Tools, Enforcement and Limits

Local context, community and the decentralized structure of Mastodon are core strengths of the fediverse. In 2019, the far-right white supremacist social network Gab joined Mastodon, quickly becoming one of the platform's largest instances (Makuch 2019, para. 3). Gab was already linked to real-world violence, including the Pittsburgh synagogue attack carried out by Robert Gregory Bowers, who had been active on the site (ibid., para. 3).

In the year that followed, the fediverse demonstrated why its design is more resilient to bad actors. One key tool is defederation, which allows an instance to disconnect from another instance that violates its norms, preventing users from seeing content from those instances. This is one of the most powerful tools available to instance administrators, enabling communities to protect themselves without relying on centralized enforcement or mass user removals (Colglazier, TeBlunthuis and

Shaw 2024, 10). Alongside defederation, instance administrators relied on a combination of human moderation, silencing tools that limit the visibility of users or entire instances unless explicitly followed, instance-level codes of conduct, shared blocklists of harmful servers and hashtags such as #Fediblock, used to identify bad actors (ibid.). Together, these practices limited Gab's reach and reduced its ability to spread across the network.

Notably, almost one year after joining the fediverse, Gab's CTO Rob Colbert announced that the instance would withdraw from the network on May 29, 2020, citing its growing isolation and stating that "the implementation will change and we won't be compatible with this Fediverse" (quoted in Caelin 2022, 162). By that point, Gab had become 99.3 percent insular, meaning that only 0.7 percent of the wider fediverse could still interact with it, effectively confining Gab's content to its corner (ibid.). In effect, the fediverse's decentralized moderation system had succeeded in containing Gab without requiring a centralized ban, demonstrating how network-level governance can limit the reach of harmful actors.

Beyond Mastodon, Bluesky offers another view into how moderation is developing on decentralized platforms. Bluesky allows users to create and share moderation lists that block or flag large groups of accounts, giving people collective ways to shape what they see rather than relying solely on platform-wide enforcement. Users can apply their own moderation labels or subscribe to labels created by others, allowing moderation to operate through shared filters and user choice (Lai et al. 2025, para. 3).

However, as Samantha Lai and Yoel Roth note, "Currently, moderation tooling for federated platforms is developed and maintained by volunteers either at their own cost or through short-term grants" (Lai and Roth 2024, para. 16). As a result, many of the systems responsible for keeping these platforms safe depend on unpaid labour and unstable funding. One example is Blacksky, a moderation service built on the Bluesky network that helps users curate and filter content, according to shared norms and safety goals. Although it is one of the most developed moderation services on the platform, it operates with only 10 paid staff and relies on a mix of grants, seed funding and user donations (Hallinan, Rothenstein and John 2025, 18; Hendrix 2025, 27:00).

Bluesky's moderation tools have expanded rapidly, but they remain uneven in how they are used. The platform hosts 14,000 moderation lists, which are user-created lists of accounts that can be used to mute or block accounts altogether. Many of the largest lists contain tens or even hundreds of thousands of accounts, suggesting they rely heavily on automated tools to stay up to date (Hallinan, Rothenstein and John 2025, 12). For users who want more fine-grained control, Bluesky also supports labelling frameworks that tag content or accounts for hate speech, explicit material or graphic imagery, allowing moderation to operate through shared filters rather than centralized takedowns (ibid.; Hogg et al. 2024, 11).

Canada's Platform Governance Gap

Policy frameworks designed for centralized systems often clash with the decentralized logic of the fediverse. Notably, roughly two-thirds of Canadians think the government has a role to play in forcing platforms to take responsibility for harmful content

(Lockhart 2025, 49). Earlier policy approaches already revealed the structural mismatch between centralized regulatory assumptions and decentralized, volunteer-run platform governance, including the proposed Online Harms Act (Bill C-63). Introduced in February 2024, the bill was later tabled and ultimately died on the order paper when Parliament was prorogued in January 2025 (Zimonjic 2024, para. 1; Lau 2025, paras. 1–2). Although framed as a major reform, the bill reflected centralized platform assumptions, requiring safety plans, risk mitigation and regulator-mandated takedowns within 24 hours for non-compliance (Tenove and Tworek 2024, paras. 5–6). Although the bill required regulators to consider platform size and resources, its vague language could still impose heavy burdens on small, volunteer-run fediverse instances without clear exemptions or tailored obligations (Gehl 2024, para. 8).

Decentralized content moderation often struggles to scale effectively and lacks stable funding, particularly when compared to centralized commercial platforms (Friedl and Morgan 2024, 7). In practice, even a single flagged post could exceed an instance's technical or financial capacity. Administrators, who are everyday users, lack the compliance systems, funding and automated tools available to centralized social media platforms, making it difficult to scale moderation (Gehl 2024, paras. 50–54). Going forward, legislation should take these decentralized spaces seriously and recognize their ability to connect people without locking them into a single platform.

This issue extends beyond individual users and affects how large Canadian institutions become dependent on centralized platforms. Notably, prior to the Online News Act, the CBC had tied much of its public reach to Facebook. As Fenwick McKelvey and Robert W. Gehl argue, organizations such as the CBC, and governments more broadly, should invest in decentralized alternatives, so they are not forced to operate at the mercy of companies like Meta or X (McKelvey and Gehl 2022, para. 13). The priorities of these companies can change rapidly, as illustrated by Musk's acquisition of Twitter, and by the time those changes occur, institutions are often already locked in. Corey Doctorow's concept of "enshittification" explains how platforms entrench dependence and redirect power and value toward themselves (Doctorow 2023, para. 2).

Canada's Online News Act shows this dependence clearly: Google agreed to pay $100 million to keep news in search results, while Meta blocked news links entirely, exposing Canada's reliance on foreign-owned platforms (Aiello 2023, para. 1). As a result, Canadians now miss roughly 11 million news views per day on Facebook and Instagram, even though most users still rely on those platforms for news (Media Ecosystem Observatory 2024, 3). This highlights an opportunity for policy to support more resilient, Canadian-based alternatives rather than reinforcing reliance on global tech firms. Notably, the European Commission has created a Mastodon instance to reduce its dependence on corporate platforms (Oldereide 2022, para. 10). Canada has made similar efforts, including the creation of non-government Gander Social, a federated platform built on the same open protocol as Bluesky (Drescher 2025, paras. 3–5). The Canadian government could build on this momentum by funding federated platforms or developing its own decentralized public digital infrastructure.

International Lessons from Decentralized Governance

Internationally, Taiwan offers a clear example of how decentralized digital tools can strengthen democratic governance. Following the Sunflower Movement, a student-led occupation of Taiwan's national legislature protesting opaque economic policy making, Taiwan institutionalized platforms such as vTaiwan and Join for public decision making (Siddarth, n.d., 4, 6). These tools are designed to structure participation differently from traditional comment threads. Rather than amplifying the loudest voices, they allow participants to register agreement or disagreement with specific proposals, making areas of consensus and division more visible while limiting domination by a small number of actors.

These platforms give people a direct way to shape public policy through open and structured participation. At its peak, vTaiwan involved roughly 200,000 participants, about one percent of Taiwan's population, and addressed 28 policy questions, with approximately 80 percent leading to a legislative outcome. Many of these focused on technology-related issues such as ride sharing, non-consensual intimate images, financial technology and artificial intelligence (Weyl, Tang and the Plurality Community 2024, 68). As Divya Siddarth notes, democratic governance requires constant iteration, flexibility and the reworking of power relations to succeed (Siddarth, n.d., 6). A similar ethos underpins Mastodon and the fediverse, where users exercise meaningful control over the systems that govern them. Taken together, these cases offer useful reference points for Canada as it considers how participatory and decentralized approaches might inform future digital governance initiatives.

Recommendations

- **Adopt a tiered, proportionate regulatory model:** The Canadian Radio-television and Telecommunications Commission (CRTC), in coordination with the Department of Justice, should recognize the public interest and domestic potential of the fediverse and other community-run social media and apply proportionate, flexible oversight, scaled by platform size, governance model and risk, that protects these communities while ensuring accountability.

- **Promote cooperative compliance and transparency:** The Department of Canadian Heritage (Canadian Heritage) and the CRTC should work with alternative social media communities to support open-source moderation tools, Canadian-focused shared blocklists managed by communities, rather than impose centralized enforcement models that do not fit decentralized networks.

- **Leverage interoperability and portability for digital sovereignty:** Innovation, Science and Economic Development Canada (ISED), Canadian Heritage and the Strategic Innovation Fund should support open standards, account portability and Canadian-based federated platforms through targeted funding and infrastructure. Strengthening open protocols and domestic alternatives would expand user choice, reduce reliance on foreign platforms and help Canada build self-governed digital

infrastructure, ensuring that decisions about platforms are guided by Canadian laws and values rather than foreign interests.

- **Invest in research and public-interest infrastructure:** The Social Sciences and Humanities Research Council, ISED and Canadian Heritage should fund research on decentralized moderation, misinformation and civic trust alongside investment in public-interest digital infrastructure.
- **Embed participatory policy co-design:** Canadian Heritage and ISED should involve civil society, platform administrators, developers and users in regulatory design so that policy reflects how decentralized systems actually operate. Drawing on models such as Taiwan's participatory digital governance, Canada should treat federated platforms as public-interest digital infrastructure when shaping future technology policy regulation.

Acknowledgements

I would like to thank Dianna English for believing in me and for her guidance and help throughout the construction of this paper. Second, my thanks to Dr. Ian Stedman for his constant support and understanding. Lastly, I would like to thank Meaghan Dietrich and the Digital Policy Hub program at large, as well as my fantastic reviewers.

Works Cited

Aiello, Rachel. 2023. "'No Concessions' St-Onge says in $100M a year news deal with Google." *CTV News*, November 29. www.ctvnews.ca/politics/no-concessions-st-onge-says-in-100m-a-year-news-deal-with-google-1.6665565.

AP News. 2024. "What is Bluesky, the fast-growing social platform welcoming fleeing X users?" *AP News*, November 15. https://apnews.com/article/bluesky-x-twitter-dcc4f92f9d0386dbacc3f30376b38dbc.

Barbrook, Richard and Andy Cameron. 1996. "The Californian Ideology." *Science as Culture* 6 (1): 44–72. https://doi.org/10.1080/09505439609526455.

Barlow, John Perry. 1996. "A Declaration of the Independence of Cyberspace." February 8. www.eff.org/cyberspace-independence.

Bassett, Kayla. 2025. *X's Community Notes and the South Asian Misinformation Crisis*. Center for the Study of Organized Hate, June 30. www.csohate.org/2025/06/30/x-community-notes-south-asia/.

Braga, Roberta, Cristina Tardáguila and Marcelo Soares. 2025. *A Deep Dive into X's Community Notes: An Analysis of English and Spanish Contributions Between 2021 and 2025.* Digital Democracy Institute of the Americas. https://ddia.org/en/a-deep-dive-into-xs-community-notes-report.

Caelin, Derek. 2022. "Decentralized Networks vs. the Trolls." In *Fundamental Challenges to Global Peace and Security*, edited by H. Mahmoudi, M. H. Allen and K. Seaman, 143–68. Cham, Switzerland: Palgrave Macmillan. https://doi.org/10.1007/978-3-030-79072-1_8.

Chee, Foo Yun and Supantha Mukherjee. 2023. "Musk's X has a fraction of rivals' content moderators, EU says." Reuters, November 10. www.reuters.com/technology/musks-x-has-fraction-rivals-content-moderators-eu-says-2023-11-10/.

Colglazier, Carl, Nathan TeBlunthuis and Aaron Shaw. 2024. “The Effects of Group Sanctions on Participation and Toxicity: Quasi-experimental Evidence from the Fediverse.” Preprint, arXiv, April 2. https://arxiv.org/abs/2404.02109v1.

Cramer, Geoffrey, William P. Maxam III and James C. Davis. 2025. “Engineering patterns for Trust and Safety on social media platforms: A case study of Mastodon and Diaspora.” *Journal of Systems and Software* 222: 112322. https://doi.org/10.1016/j.jss.2024.112322.

de Guzman, Chad. 2021. “The Facebook Whistleblower Revealed Herself on *60 Minutes*. Here’s What You Need to Know.” *TIME*, October 4. https://time.com/6103645/facebook-whistleblower-frances-haugen/.

Di Placido, Dani. 2024. “The X (Twitter) Exodus to Bluesky, Explained.” *Forbes*, November 19. www.forbes.com/sites/danidiplacido/2024/11/19/the-x-twitter-exodus-to-bluesky-explained/.

Doctorow, Cory. 2023. “The ‘Enshittification’ of TikTok.” *WIRED*, January 23. www.wired.com/story/tiktok-platforms-cory-doctorow/.

Drescher, Nathan. 2025. “Canada is building its own social network.” *The Logic*, May 23. https://thelogic.co/news/gander-canadian-social-network/.

Driscoll, Kevin. 2022. *The Modem World: A Prehistory of Social Media*. New Haven, CT: Yale University Press.

Escritt, Thomas. 2022. “Mastodon: What is the social network hailed as a Twitter alternative?” Reuters, November 7. www.reuters.com/technology/mastodon-what-is-social-network-hailed-twitter-alternative-2022-11-07/.

Flipboard. 2024. “Moderation and Migration for a Better Social Web, with Fediverse Leader Tim Chambers.” January 9. YouTube video, 51:12. www.youtube.com/watch?v=Sa5dULONq_M.

Friedl, Paul and Julian Morgan. 2024. “Decentralised content moderation.” *Internet Policy Review* 13 (2). https://doi.org/10.14763/2024.2.1754.

Gehl, Robert W. 2024. “Reading the Online Harms Act with my Fediverse Admin Hat On.” *FOSS Academic* (blog), April 10. https://fossacademic.tech/2024/04/10/Online%20Harms%20Act.html.

– – –. 2025. *Move Slowly and Build Bridges: Mastodon, the Fediverse, and the Struggle for Democratic Social Media*. Oxford, UK: Oxford University Press.

Gehl, Robert W. and Diana Zulli. 2023. “The digital covenant: non-centralized platform governance on the Mastodon social network.” *Information, Communication & Society* 26 (16): 3275–91. https://doi.org/10.1080/1369118X.2022.2147400.

Gillespie, Tarleton. 2018. *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media*. New Haven, CT: Yale University Press.

Government of Canada. 2025. “Digital Sovereignty: A Framework to improve digital readiness of the Government of Canada.” Policy paper, November 12. www.canada.ca/en/government/system/digital-government/digital-government-innovations/cloud-services/digital-sovereignty/digital-sovereignty-framework-improve-digital-readiness.html.

Hallinan, Blake, Orly Rothenstein and Nicholas A. John. 2025. “Stuck in the middleware with you: the challenges of capitalizing a market-oriented approach to platform governance.” *Communication and Change* 1: 19. https://doi.org/10.1007/s44382-025-00018-9.

Hendrix, Justin. 2025. “Rudy Fraser on Building Blacksky and the Future of Middleware,” March 3, in *Tech Policy Press Podcast*, produced by *Tech Policy Press*, podcast, 37:58. www.techpolicy.press/podcast-rudy-fraser-on-building-blacksky-and-the-future-of-middleware.

Hickey, Daniel, Daniel M. T. Fessler, Kristina Lerman and Keith Burghardt. 2025. “X under Musk’s Leadership: Substantial hate and no reduction in inauthentic activity.” *PLOS One* 20 (2): e0313293. https://doi.org/10.1371/journal.pone.0313293.

Hogg, Luke, Renée DiResta, Francis Fukuyama, Richard Reisman, Daphne Keller, Aviv Ovadya, Luke Thorburn, Jonathan Stray and Shubhi Mathur. 2024. “Shaping the Future of Social Media with Middleware.” Preprint, arXiv, December 13. www.arxiv.org/pdf/2412.10283.

Hoover, Amanda. 2023. “The Mastodon Bump Is Now a Slump.” *WIRED*, February 7. www.wired.com/story/the-mastodon-bump-is-now-a-slump/.

Kissane, Erin and Darius Kazemi. 2024. *Findings Report: Governance on Fediverse Microblogging Servers*. August 20. Infrastructure Insights Fund. https://infrastructureinsights.fund/wp-content/uploads/2023/11/fediverse-governance.pdf.

Knibbs, Kate. 2025. “Bluesky Is Plotting a Total Takeover of the Social Internet.” *WIRED*, May 19. www.wired.com/story/big-interview-jay-graber-bluesky/.

Lai, Samantha and Yoel Roth. 2024. “Online Safety and the ‘Great Decentralization’: The Perils and Promises of Federated Social Media.” *Tech Policy Press*, October 17. www.techpolicy.press/online-safety-and-the-great-decentralization-the-perils-and-promises-of-federated-social-media/.

Lai, Samantha, Yoel Roth, Renée DiResta, Kate Klonick, Mallory Knodel, Evan Prodromou and Aaron Rodericks. 2025. “New Paradigms in Trust and Safety: Navigating Defederation on Decentralized Social Media Platforms.” Carnegie Endowment for International Peace, March 25. https://carnegieendowment.org/research/2025/03/fediverse-social-media-internet-defederation?lang=en.

Lau, Mandy. 2025. “Canada’s Online Harms Bill Is Dead (Again): Three Questions to Consider for the Next Round.” *Tech Policy Press*, April 28. www.techpolicy.press/canadas-online-harms-bill-is-dead-again-three-questions-to-consider-for-the-next-round/.

Lockhart, Angus. 2025. *Survey of Online Harms in Canada 2025*. May. Toronto, ON: The Dais. https://dais.ca/wp-content/uploads/2025/05/OnlineHarms2025.pdf.

LSE Library. 2025. “Big tech is broken: what can we do about it now?” May 27. YouTube video, 1:24:51. www.youtube.com/watch?v=wSSj-kEXxyM.

Makuch, Ben. 2019. “The Nazi-Free Alternative to Twitter Is Now Home to the Biggest Far Right Social Network.” *VICE*, July 11. www.vice.com/en/article/the-nazi-free-alternative-to-twitter-is-now-home-to-the-biggest-far-right-social-network/.

McKelvey, Fenwick and Robert W. Gehl. 2022. “Canada’s public broadcaster should use Mastodon to provide a social media service.” *The Conversation*, November 8. https://doi.org/10.64628/AAM.n53cydfqg.

Media Ecosystem Observatory. 2024. “Old News, New Reality: A Year of Meta’s News Ban in Canada.” Information Ecosystem Brief, August. https://static1.squarespace.com/static/6450265301129e5dbabfe8a2/t/66c36dba3066f124adb4000c/1724083643261/Meta%2BNews%2BBan%2BReport-08-19.pdf.

Obar, Jonathan A. and Anne Oeldorf-Hirsch. 2020. “The biggest lie on the Internet: ignoring the privacy policies and terms of service policies of social networking services.” *Information, Communication & Society* 23 (1): 128–47. https://doi.org/10.1080/1369118X.2018.1486870.

Oldereide, Andréa. 2022. “EU debuts on Mastodon, Twitter exodus fits policy goals.” *Luxembourg Times*, November 22. www.luxtimes.lu/europeanunion/eu-debuts-on-mastodon-twitter-exodus-fits-policy-goals/1341634.html.

O’Mara, Margaret. 2019. *The Code: Silicon Valley and the Remaking of America.* New York, NY: Penguin Press.

Perez, Sarah. 2026. “Threads edges out X in daily mobile users, new data shows.” *TechCrunch*, January 18. https://techcrunch.com/2026/01/18/threads-edges-out-x-in-daily-mobile-users-new-data-shows/.

Rossini, Patrícia. 2022. “Beyond Incivility: Understanding Patterns of Uncivil and Intolerant Discourse in Online Political Talk.” *Communication Research* 49 (3): 399–425. https://doi.org/10.1177/0093650220921314.

Rozenshtein, Alan Z. 2024. “Moderating the Fediverse: Content Moderation on Distributed Social Media.” In *Media and Society After Technological Disruption*, edited by Kyle Langvardt and Justin (Gus) Hurwitz, 177–92. Cambridge, UK: Cambridge University Press. www.cambridge.org/core/books/media-and-society-after-technological-disruption/moderating-the-fediverse/A9018A2868DBE98ED59BA66BC6C25913.

Siddarth, Divya. n.d. “Taiwan: Grassroots Digital Democracy That Works.” New York, NY: RadicalxChange Foundation. www.radicalxchange.org/updates/papers/Taiwan_Grassroots_Digital_Democracy_That_Works_V1_DIGITAL_.pdf.

Srnicek, Nick. 2016. *Platform Capitalism*. Hoboken, NJ: Wiley.

Tenove, Chris and Heidi Tworek. 2024. “What Lessons Did Canada Learn Before Creating Its Online Harms Bill?” *Tech Policy Press*, March 12. https://techpolicy.press/what-lessons-did-canada-learn-before-creating-its-online-harms-bill.

United Nations. 2024. *United Nations Global Principles for Information Integrity: Recommendations for Multi-stakeholder Action*. August. https://doi.org/10.18356/9789211065756.

Weyl, E. Glen, Audrey Tang and the Plurality Community. 2024. *Plurality: The Future of Collaborative Technology and Democracy*. www.plurality.net/.

Woloshyn, Roxanna and Ashley Fraser. 2023. “What is the fediverse and why does Threads want to join?” *CBC News*, July 19. www.cbc.ca/news/business/fediverse-explainer-1.6905837.

Zimonjic, Peter. 2024. “The Liberals are set to unveil an online harms bill today. Here’s what you need to know.” *CBC News*, February 26. www.cbc.ca/news/politics/online-harms-bill-explainer-1.7124140.

Zuboff, Shoshana. 2019. *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. First edition. New York, NY: PublicAffairs.