

Digital Policy Hub – Working Paper

Platform Governance of Misogyny: Mainstream and Alternative Technology

Karmvir Kaur Padma

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.



Copyright © 2026 by Karmvir Kaur Padda

The opinions expressed in this publication are those of the author and do not necessarily reflect the views of the Centre for International Governance Innovation or its Board of Directors.

Centre for International Governance Innovation and CIGI are registered trademarks.

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

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

Karmvir Kaur Padda is a Digital Policy Hub postdoctoral fellow with a Ph.D. in sociology from the University of Waterloo. During her fellowship, she will examine how platform affordances shape the spread of gendered misinformation and extremist discourse across moderated and unmoderated online spaces.

Platform Governance of Misogyny: Mainstream and Alternative Technology

Karmvir Kaur Padda

Bottom Line Up Front

Misogynistic extremism persists across the digital ecosystem because current governance models focus too narrowly on removing harmful content from individual platforms. These approaches overlook the emotional dynamics, infrastructural features and cross-platform migrations that allow extremist communities to adapt and proliferate. Mainstream platforms limit visibility but rarely disrupt underlying grievances, while alternative-technology (alt-tech) environments provide ideological refuge and continuity. Effective intervention requires policy frameworks that operate at the ecosystem level, incorporate democratic oversight and address both the emotional and technological conditions that sustain misogynistic extremism. Without a shift toward coordinated, relational and infrastructure-focused governance, harmful narratives will continue to outpace existing regulatory tools.

Key Points

- Misogynistic extremism thrives through emotional bonds, grievance narratives and pseudo-scientific rationalizations that often evade traditional moderation.
- Content takedowns on mainstream platforms frequently push extremist communities toward alt-tech ecosystems with weak or nonexistent governance.
- Extremist users maintain multi-platform identities, circulating memes, narratives and coded language across both mainstream and fringe environments.
- Platform infrastructures, such as recommendation systems, search tools, monetization models and hosting arrangements, shape the visibility and persistence of harmful content.
- Democratic, ecosystem-level governance is needed to address the emotional, relational and technical factors that allow misogynistic extremism to persist.

Recommendations

- **Build ecosystem-level governance through cross-platform coordination mechanisms:** Establish systems that allow platforms to share non-identifying indicators of emerging harms, enabling earlier detection of misogynistic extremist activity without compromising user privacy. Support the development of shared governance standards that extend beyond terrorism designations and include gender-based extremist threats.
- **Require interpretive moderation capacity:** Equip platform moderation teams with feminist linguistic expertise, emotional-context awareness and access to gender-based violence specialists. Artificial intelligence systems should be designed to flag rhetorical patterns and not just slurs, so that nuanced forms of misogynistic extremism are routed for human review.
- **Develop grievance-aware governance frameworks:** Integrate emotional support, counter-narratives and healthier community pathways into platform governance. Provide optional mental-health referrals, elevate content emphasizing relational ethics and healthy masculinity, and develop soft off-ramps that offer credible alternatives to extremist spaces.
- **Regulate platform infrastructures alongside content:** Implement requirements for algorithmic transparency, gender-focused risk assessments, baseline moderation standards for services integrated into national infrastructure, and support for ethical hosting providers. Addressing visibility and engagement architectures is essential for disrupting the spread of misogynistic extremism.
- **Establish democratic, independent governance institutions:** Establish rights-based bodies that oversee platform transparency, assess emerging harms, coordinate cross-platform intervention and ensure procedural fairness. These institutions must remain independent from both governments and corporations to maintain legitimacy and public trust.

Introduction

Debates over how to govern social media platforms often assume that harmful content can be controlled through better rules, faster takedowns or more independent oversight. Yet research consistently shows that extremist communities are not eliminated by moderation; they adapt to it. When platforms impose bans or algorithmic restrictions, harmful groups frequently migrate, reorganize and reframe their narratives in ways that outpace regulatory reforms (Horne et al. 2025; Fielitz, Hitziger and Schwarz 2021; Rogers 2020; Urman and Katz 2022). Previous work on misogynistic extremism demonstrates the same pattern: Reddit’s removal of the subreddits r/Incels and r/Braincels did not neutralize the community but instead catalyzed its reconstitution on Incels.is, an alternative-technology (alt-tech) forum designed explicitly to circumvent the governance constraints of mainstream platforms (Padda Forthcoming 2026).

This migration produced what Naomi Klein (2023) calls “mirror worlds,” alternative ecosystems that preserve the technical architecture of mainstream platforms while inverting their ethical foundations. Within these mirrored environments, misogyny is not merely reproduced but re-engineered. Feelings of loneliness and rejection are reframed as evidence of systemic injustice, recasting incel identity as morally righteous and positioning feminism as a source of social decline (Padda Forthcoming 2026). This moral inversion — transforming grievance into virtue and hate into reason — creates a type of extremist discourse that evades traditional moderation. Because it often avoids explicit slurs, relying instead on pseudo-scientific rationalism, rights-based language and the emotional grammar of victimhood, much of this content slips past automated systems and inconsistent human moderation (Gillespie 2018).

These dynamics reflect deeper governance gaps across the digital ecosystem. Rather than focusing on individual platforms, this working paper examines how mainstream environments, including Reddit, YouTube, Meta and X/Twitter, interact with alt-tech infrastructures such as Gab, Incels.is, 4chan/8kun and fringe Telegram channels. Alt-tech spaces market themselves as “uncensored” zones, positioning themselves in opposition to big tech governance and attracting communities seeking both ideological refuge and operational freedom (Dunne 2024; Buckley and Schafer 2022). Studies show that these platforms become tightly coupled with extremist speech, often relying on inflammatory content to sustain engagement and user growth (Dehghan and Nagappa 2022). In this context, deplatforming from mainstream sites can inadvertently strengthen extremist communities by reinforcing their sense of persecution and pushing them toward systems with dramatically weaker oversight.

This paper argues that content-centric regulatory approaches, which dominate Canada’s online harms discussions and earlier proposals such as Bill C-63, address symptoms rather than causes. They focus on what platforms remove, not why harmful narratives persist or how they mutate across ecosystems. As a result, extremist ideologies such as misogynistic grievance-based extremism continue to circulate across the broader digital landscape, threatening democratic resilience, public safety and gender equality. These concerns sit squarely within the Centre for International Governance Innovation’s research priorities on digitalization, security and democracy, and data, economy and society, highlighting how platform design, data governance and algorithmic systems shape the conditions under which online harms evolve.

Mainstream and Alt-Tech Platform Ecosystems

Understanding the persistence of misogynistic extremism requires shifting from a single-platform perspective to an ecosystem approach. Mainstream and alt-tech platforms operate as interconnected but asymmetrical environments, each offering distinct governance regimes, technical affordances and ideological possibilities (Gillespie 2018; de Winkel 2023). Extremist communities navigate this ecosystem strategically. They use mainstream platforms such as Reddit, YouTube, X and Meta's services for visibility, recruitment and mainstream cultural influence, while turning to alt-tech spaces such as Gab, 4chan/8kun, Incels.is and fringe Telegram channels for ideological refuge, consolidation and escalation (Padda Forthcoming 2026; de Winkel 2023; Copland 2020; Davis and Chouinard 2016; Ribeiro et al. 2021; Massanari 2017). The durability of incel ideology illustrates how these platforms work together to sustain harmful movements even in the face of content moderation.

Mainstream platforms present the most formal governance structures, relying on combinations of content policies, algorithmic moderation, human review and, particularly in Reddit's case, volunteer moderators who shoulder substantial governance responsibility. Yet these systems consistently struggle with nuance, context and the emotional and rhetorical complexities of misogynistic extremism. Automated moderation systems, despite their scale, remain ill-equipped to interpret irony, coded language or the subtle moral inversions through which misogynistic narratives present themselves as rational or even justice-oriented (He, Hong and Raghu 2025; Caplan 2018). Human moderation offers greater nuance, but as Tarleton Gillespie (2018) and Nicolas P. Suzor (2019) emphasize, moderation is infrastructural labour — unevenly supported, inconsistently trained and frequently invisible. Reddit's model in particular, where corporate policies interact with thousands of independent volunteer teams, produces a patchwork of enforcement. As a result, even as formal policies against hate speech or gender-based harassment tighten, misogynistic content often persists in softened, quasi-academic or emotionally coded forms that evade both automated filters and inconsistent human oversight.

These limitations are compounded by the content-centric orientation of most platform policies. Because moderation is designed to remove posts rather than address the broader narratives and grievances animating extremist communities, reductions in visible content do not translate into ideological disruption. Deplatforming a subreddit or suspending an account reduces immediate visibility but does little to challenge the underlying worldview. Users adapt by reframing their language, adopting pseudo-scientific or self-help framing, or migrating to more permissive platforms (Padda Forthcoming 2026).

Alt-tech environments provide the ideal destination for this migration. Platforms such as Gab, 4chan/8kun and Incels.is explicitly position themselves as alternatives to “censorious” big tech, marketing themselves as spaces for absolute free expression and ideological authenticity (Buckley and Schafer 2022; Dunne 2024). Their governance frameworks are minimal by design: terms of service are either weakly enforced or structured around maximalist free speech rhetoric; hosting arrangements rely on foreign

jurisdictions or decentralized infrastructures; and moderation systems are nominal, symbolic or entirely absent. These features create ideal conditions for extremist content to flourish, and empirical studies show that such platforms become structurally dependent on inflammatory hate speech to maintain engagement and identity cohesion (Dehghan and Nagappa 2022). This intensification underscores that these platforms are not neutral containers for ideology, but what Adrienne Massanari (2017, 330) calls “toxic technocultures,” where specific affordances, anonymity, persistence and community self-policing actively shape the production and circulation of extremist worldviews. In this context, the formation of Incels.is was not simply a relocation of a banned community — it was an ideological intensification. The site reproduced Reddit’s core technical features while abandoning normative constraints. It presented itself as a “home” for those victimized by feminism and big tech alike, framing deplatforming as proof of persecution and ideological righteousness (Copland 2020; Davis and Chouinard 2016; Nashrulla 2019; Ribeiro et al. 2021).

Users moving between these spaces do not sever ties with mainstream platforms. Instead, they maintain multi-platform identities, shifting memes, arguments and narratives across contexts. These cross-ecosystem feedback loops are further reinforced by emotional reward structures: alt-tech spaces offer validation of grievances and humiliation, which users then carry back into mainstream spaces as supposedly legitimate concerns (Thorburn 2023; Ging 2019; Urman and Katz 2022; Rogers 2020).

These dynamics highlight why focusing on individual platforms is insufficient. Extremist ideologies survive because the digital ecosystem provides complementary functions: mainstream platforms amplify reach and cultural influence, while alt-tech platforms provide ideological refuge and operational continuity. As Joan Donovan, Becca Lewis and Brian Friedberg (2019) argue, alt-tech services are not isolated alternatives but parallel infrastructures that depend on mainstream platforms for visibility, recruitment and cultural traction. Thus, content moderation — no matter how well-intentioned — cannot meaningfully curb extremist misogynistic movements unless it accounts for the mobility of harmful narratives across governance boundaries.

This ecosystem logic sits at the core of contemporary challenges in digital governance. It demonstrates that current regulatory approaches, which focus on removing harmful content from individual platforms, address only the visible symptoms rather than the structural conditions that allow misogynistic extremism to persist. These dynamics provide the foundation for the governance failures explored in the next section and underscore why a holistic, democratic and cross-platform governance framework is now urgently required.

Diagnosing Governance Failures Across Mainstream and Alt-Tech Platforms

The persistence and evolution of misogynistic extremism across the digital ecosystem reveal deep structural failures in contemporary platform governance. Although mainstream and alt-tech platforms differ dramatically in scale, resources

and institutional legitimacy, they share one critical feature: neither governance model successfully disrupts the ideological, affective or social conditions that allow misogynistic extremist narratives to take root. These failures are not simply a matter of insufficient rules or inconsistent enforcement. Rather, they reflect fundamental limitations in how digital governance has conceptualized the nature of harm, the structure of online communities and the relational dynamics that bind users to extremist worldviews.

Mainstream approaches to content moderation remain anchored in content-centrism, the assumption that harmful material can be managed through better policies, improved detection systems or faster takedowns. This logic underpins everything from corporate trust and safety programs to legislative initiatives such as Canada's former Bill C-63 and its 2025 revival attempts, which focused primarily on defining categories of "harmful content" and accelerating response timelines (Bronskill 2025). Yet content-focused governance consistently fails to anticipate how communities adapt to new rules. Extremist subcultures, including incels, do not dissipate when their explicit content is removed; they reconfigure their rhetoric, adopt coded language or move to spaces where enforcement is weak. The emotional and ideological coherence of these communities — which is built on shared narratives of injustice, humiliation and resentment — remains intact, allowing them to reconstitute elsewhere (Padda Forthcoming 2026; Fielitz, Hitziger and Schwarz 2021).

The first governance failure, therefore, lies in treating extremism as content rather than as an affective and relational structure. Misogynistic extremism operates through what Kate Manne (2018) describes as the moral logic of entitlement: the belief that men are owed women's attention, affection and submission, and that feminist gains represent an unjust deprivation. This narrative framework does not require overt slurs or explicit calls for violence to remain operational. Instead, it manifests as pseudo-scientific reasoning about "male disadvantage," therapeutic-style accounts of loneliness or appeals to "fairness" framed through distorted uses of rights language. Automated moderation systems — optimized to detect surface-level toxicity — struggle to identify these subtle forms of misogyny, while human moderators often lack the feminist linguistic literacy needed to recognize moral inversion. This creates blind spots in enforcement, allowing misogynistic ideologies to survive even under stricter policy regimes (He, Hong and Raghu 2025; Caplan 2018).

A second governance failure stems from infrastructural gaps between platform design and regulatory expectations. As Gillespie (2018) argues, moderation is infrastructural labour — continuous, complex and deeply entangled with the architecture of platforms themselves. Reddit's volunteer-based moderation model illustrates the consequences of outsourcing governance to unpaid community labour (Guo and Caine 2025; He, Hong and Raghu 2025; Suzor 2019; Gorwa 2024). Corporate policies may prohibit gender-based hate, but moderators vary in capacity, expertise and willingness to enforce rules, producing uneven outcomes across subreddits. Similarly, algorithmic moderation on YouTube or TikTok relies on opaque systems that disproportionately flag marginalized users, amplify borderline content or miss harmful rhetoric disguised as self-improvement advice. These inconsistencies reveal that governance failures are often rooted not in the absence of rules but in platform architectures optimized for engagement rather than safety or democratic accountability (Suzor 2019; Myers West 2018).

A third governance failure emerges from deplatforming practices that unintentionally strengthen alt-tech ecosystems. When mainstream platforms ban extremist communities, they often do so without considering where those communities will go or how their narratives will mutate. Richard Rogers (2020) documents how high-profile deplatformings consistently result in migrations to alt-tech platforms such as Gab, 8kun and Telegram, where ideological norms shift from “subcultural” to explicit extremism. These spaces function as ideological pressure cookers: minimal moderation, permissive terms of service and a shared narrative of persecution intensify harmful content rather than disperse it. Donovan, Lewis and Friedberg (2019) conceptualize this phenomenon as the emergence of parallel platforms — alternative infrastructures that depend on mainstream platforms for visibility and recruitment but operate outside their governance reach. Within this context, Incels.is serves as a paradigmatic case: a site born from deplatforming that reconstructed Reddit’s affordances while rejecting all normative constraints, becoming a hub for more explicit, violent misogyny.

The failure to anticipate migration underscores the fourth major governance gap: cross-platform fragmentation. Regulatory and corporate frameworks typically treat platforms as discrete units responsible only for what happens within their own boundaries. Yet extremist ideologies move fluidly across the ecosystem, adapting their tone and format based on each site’s affordances. A user may post ironic misogynistic memes on Reddit, consume longform ideological content on YouTube and participate in explicit hate speech on Incels.is or Telegram — all under different usernames or pseudonyms. Current governance models are not designed to track these trajectories or intervene in the relational ties that keep communities intact despite deplatforming. Research on cross-platform migration demonstrates that following deplatforming, extremist communities rapidly recreate their network structures on alt-tech platforms (Padda Forthcoming 2026), operating across multiple governance environments simultaneously in ways that current regulatory frameworks are not designed to track.

A fifth governance failure lies in democratic accountability. Scholars argue that platform governance has become privatized by default, with corporate decisions about speech, moderation and algorithmic design effectively functioning as public policy without meaningful democratic oversight (Owen 2019; Gorwa 2019). Meanwhile, alt-tech platforms explicitly reject normative accountability and use anti-censorship rhetoric to cultivate ideological loyalty. Under these conditions, the governance of misogynistic extremism becomes trapped between two insufficient models: mainstream platforms that cannot fully regulate harmful narratives without undermining their own business models, and alt-tech sites that refuse regulation entirely. As feminist governance scholars argue, meaningful accountability requires attention not only to content but to the social conditions, emotional dynamics and power structures that shape digital participation (Gurumurthy and Chami 2021).

Finally, governance frameworks fail because they ignore the grievances and emotional drivers at the core of extremist misogyny. As feminist and critical extremism research shows, incel ideology is sustained not only by ideological claims but also by affective infrastructures — shared feelings of humiliation, loneliness, resentment and distrust (Padda Forthcoming 2026; Ging 2019). Content moderation can remove hateful posts but cannot resolve these affective bonds or the narratives that transform private frustration into political grievance. Without interventions that address these deeper emotional and

relational dynamics, extremist communities simply regroup elsewhere, armed with stronger narratives of persecution and authenticity.

These failures reveal that platform governance is not merely a technical challenge but a democratic one. Current systems are focused on reactive content removal, platform-specific enforcement and privatized oversight, which are structurally incapable of disrupting misogynistic extremist ecosystems. A new governance paradigm is required: one that is ecosystem-wide, grounded in democratic accountability, and responsive to the emotional and relational dynamics that drive extremist participation. The next section outlines policy interventions that operationalize this shift, offering practical pathways for democratic institutions, platforms and civil society actors to address the persistence of gender-based extremism online.

Policy Recommendations for Democratic, Ecosystem-Level Governance

The failures outlined in previous sections reveal that neither mainstream nor alt-tech platforms can meaningfully disrupt misogynistic extremism under current governance models. Content-focused regulation, inconsistent moderation and platform-by-platform enforcement address only the visible symptoms of ideological harm while neglecting its emotional, relational and infrastructural roots. Effective governance requires an ecosystem approach — one that integrates democratic oversight, interpretive expertise, socio-emotional interventions and cross-platform coordination. This section offers policy recommendations designed to complement one another and translate research insights into actionable governance strategies.

Build Ecosystem-Level Governance Through Cross-Platform Coordination Mechanisms

While platform governance is often described as siloed, major platforms have long employed signal-based and risk-oriented approaches in domains such as fraud, spam and, increasingly, child sexual exploitation and abuse (Gorwa and Thakur 2024). In the child safety space, cross-platform coordination initiatives such as the Lantern Project and collaborations facilitated by the Tech Coalition demonstrate that non-identifying behavioural signals, metadata and real-time threat analysis can be shared across services to detect and disrupt harmful activity at scale. These models show that ecosystem-level coordination is technically feasible and already operationalized in high-priority safety domains (ibid.). However, they remain narrowly scoped, largely opaque and under-examined in independent research, with limited democratic oversight or rights-based accountability.

In contrast to these signal-based governance systems in fraud, spam and child safety, approaches to misogynistic extremism remain fragmented, with each platform largely responsible only for activity within its own boundaries despite the cross-platform nature. To address this gap, governments and independent oversight bodies

should support coordinated, privacy-preserving mechanisms such as non-identifying “ideological risk signature” sharing, enabling early detection of emerging harms without centralized surveillance. While tools such as the Global Inter Forum to Counter Terrorism hash database demonstrate that cross-platform coordination is feasible, their narrow focus on recognized terrorist organizations excludes most gender-based extremism. Expanding governance frameworks to address misogynistic grievance-driven violence, supported by policy coordination and safeguards for transparency and due process, would fill a critical gap in the current regulatory landscape.

Require Interpretive Moderation Capacity: Feminist Linguistic Literacy and Emotional Context

Misogynistic extremism often presents itself through rational, therapeutic or pseudo-scientific frames, allowing it to bypass moderation systems optimized for explicit slurs or surface-level “toxicity.” Because its harms are embedded in narrative structure, affect and gendered entitlement rather than overt hate speech, conventional filters routinely miss it. Addressing this gap requires interpretive moderation capacity that integrates feminist linguistic literacy, emotional-context awareness and cultural expertise in both human and algorithmic systems. Moderators need training to recognize moral inversion and grievance reframing; artificial intelligence systems should flag discursive patterns associated with “scientific” misogyny or victimhood narratives; and hybrid teams that combine technical staff with gender-based violence experts are essential for accurately interpreting nuanced forms of harm.

Develop Grievance-Aware Governance Frameworks to Address Emotional Drivers of Extremism

Misogynistic extremism is sustained through affective infrastructures — shared emotions of humiliation, loneliness and resentment that are collectively reframed as political injury — so governance systems that ignore these emotional drivers inadvertently strengthen extremist communities. When bans occur without addressing the underlying relational dynamics, users often interpret moderation as persecution, reinforcing grievance identities and accelerating migration to alt-tech ecosystems. A grievance-aware framework must therefore engage not only with harmful content but with the narratives that make it meaningful, integrating emotional support, counter-narratives and pathways into healthier communities as part of a broader public-health model of digital harm reduction. This includes optional mental health referrals, the promotion of relational ethics and healthy masculinity, and soft off-ramps into alternative forms of belonging.

Regulate Platform Infrastructure, Not Just Content

Most regulatory frameworks, including Canada’s attempted online harms legislation, focus narrowly on removing illegal or harmful content while overlooking the infrastructural conditions that allow misogynistic extremism to persist. Effective governance must address the systems that structure visibility and engagement:

- recommendation algorithms that funnel users toward extremist communities through “self-improvement” or pseudo-scientific gender content;

- search functionalities that make misogynistic material easily discoverable;
- monetization systems that reward inflammatory creators; and
- the broader technical stack on which platforms depend, including cloud hosting, domain name system, app-store distribution, payment processors and limited internet service provider access.

These dependencies do not function as simple points of removal, but as governance leverage points through which baseline standards, transparency obligations and conditional access to essential services can be imposed.

At the same time, infrastructural governance faces clear legal and jurisdictional limits. Platforms hosted abroad, supported by decentralized or offshore infrastructure, and accessed through virtual private networks cannot be fully punted off the web by any single state (Laidlaw and Young 2018; Svantesson 2024). Experience from assertive regulatory regimes such as Australia's eSafety framework demonstrates that even tools such as court orders and local access blocking have limited reach against non-cooperative actors (Associated Press 2025). These constraints make it essential to distinguish between fully malicious platforms designed to evade regulation and mixed-use services — such as Telegram, Discord or WhatsApp — where harmful subcommunities coexist with legitimate uses and where regulatory pressure, intermediary liability obligations and reputational incentives can more plausibly reshape platform behaviour (Gorwa 2024). Infrastructure regulation should therefore be understood as harm reduction and friction creation.

Establish Democratic, Independent Governance Institutions

Finally, platform governance cannot remain privatized in the hands of corporations that lack democratic accountability, nor can it rely on alt-tech platforms that intentionally evade oversight; instead, democratic states must create independent, rights-based institutions capable of governing the digital public sphere without directly controlling speech. Such bodies would conduct public risk assessments on emerging harms, set minimum standards for transparency, oversee cross-platform coordination, audit algorithmic impacts and ensure procedural fairness. Crucially, these institutions must be structurally independent from both governments and corporations, drawing on civil society, researchers and digital rights advocates to ensure legitimacy. Their role is not censorship but democratic accountability, anchoring digital governance in international human rights frameworks and ensuring that the infrastructures shaping online life operate transparently, equitably and in the public interest.

Conclusion

Misogynistic extremism persists not because platforms fail to remove enough harmful content, but because the digital ecosystem enables its adaptation, migration and emotional resonance. This paper has shown that content-centric governance cannot address the infrastructural, affective and cross-platform dynamics that sustain grievance-based ideologies. Democratic, ecosystem-level governance grounded in

interpretive expertise, socio-emotional interventions, infrastructural regulation and independent oversight is therefore essential. By treating misogynistic extremism as a relational, technological and political problem rather than a purely content-based one, policy makers can build governance systems capable of disrupting harmful narratives while strengthening democratic accountability and digital public health.

Acknowledgements

I would like to thank my CIGI fellowship supervisor, Dr. Lai-Tze Fan, for her encouragement and support, which made the completion of this work possible. I am also deeply grateful to my CIGI expert, Dr. Robert Gorwa, and to Digital Policy Hub peer Clarence Sokolambe Lakpini for their generous engagement with this paper and their thoughtful feedback that helped refine my analysis. I also thank Dianna English, Meaghan Dietrich and the CIGI publications team for their support throughout the writing and submission process.

Works Cited

- Associated Press. 2025. "Telegram's CEO Durov in Dubai as France Continues to Investigate Criminal Activity on Messaging App." AP News, March 17. <https://apnews.com/article/durov-telegram-france-messaging-dubai-421a69e62ca419ff50d48a11fb944187>.
- Bronskill, Jim. 2025. "Liberals Taking 'Fresh' Look at Online Harms Bill, Says Justice Minister Sean Fraser." CBC News, June 29. www.cbc.ca/news/politics/liberals-taking-fresh-look-at-online-harms-bill-says-justice-minister-sean-fraser-1.7573791.
- Buckley, Nicole and Joseph S. Schafer. 2022. "'Censorship-Free' Platforms: Evaluating Content Moderation Policies and Practices of Alternative Social Media." Preprint, *arXiv*. <https://doi.org/10.31235/osf.io/yf9qz>.
- Caplan, Robyn. 2018. *Content or Context Moderation? Artisanal, Community-Reliant, and Industrial Approaches*. Report, Data & Society. https://datasociety.net/wp-content/uploads/2018/11/DS_Content_or_Context_Moderation.pdf.
- Copland, Simon. 2020. "Reddit Quarantined: Can Changing Platform Affordances Reduce Hateful Material Online?" *Internet Policy Review* 9 (4). <https://doi.org/10.14763/2020.4.1516>.
- Davis, Jenny L. and James B. Chouinard. 2016. "Theorizing Affordances: From Request to Refuse." *Bulletin of Science, Technology & Society* 36 (4): 241–48. <https://doi.org/10.1177/0270467617714944>.
- de Winkel, Tim. 2023. "Fringe Platforms: An Analysis of Contesting Alternatives to the Mainstream Social Media Platforms in a Platformized Public Sphere." Ph.D. dissertation, Utrecht University. <https://doi.org/10.33540/1921>.
- Dehghan, Ehsan and Ashwin Nagappa. 2022. "Politicization and Radicalization of Discourses in the Alt-Tech Ecosystem: A Case Study on Gab Social." *Social Media + Society* 8 (3). <https://doi.org/10.1177/20563051221113075>.
- Donovan, Joan, Becca Lewis and Brian Friedberg. 2018. "Parallel Ports: Sociotechnical Change from the Alt-Right to Alt-Tech." In *Post-Digital Cultures of the Far Right: Online Actions and Offline Consequences in Europe and the US*, edited by Maik Fielitz and Nic Thurston, 49–65. Bielefeld, Germany: transcript Verlag. <https://doi.org/10.25969/mediarep/12374>.

- Dunne, Sarah Anne. 2024. "Too Far Right? The Role of Alt-Tech and Mainstream Social Media Platforms in Far-Right Political Violence Dr Sarah Anne Dunne." UCD Centre for Digital Policy, September 5. <https://digitalpolicy.ie/too-far-right-the-role-of-alt-tech-and-mainstream-social-media-platforms-in-far-right-political-violence-dr-sarah-anne-dunne/>.
- Fielitz, Maik, Jana Hitziger and Karolin Schwarz. 2021. "Tech vs. Hate: Patterns and Dilemmas of Deplatforming German-Speaking Hate Actors." Knowledge creates democracy series 9: 196–207. Institute for Democracy and Civil Society. <https://doi.org/10.19222/202102/18>.
- Gillespie, Tarleton. 2018. *Custodians of the Internet: Platforms, Content Moderation, and the Hidden Decisions That Shape Social Media*. New Haven, CT: Yale University Press.
- Ging, Debbie. 2019. "Alphas, Betas, and Incels: Theorizing the Masculinities of the Manosphere." *Men and Masculinities* 22 (4): 638–57. <https://doi.org/10.1177/1097184X17706401>.
- Gorwa, Robert. 2019. "Regulating Them Softly." Centre for International Governance Innovation, October 28. www.cigionline.org/articles/regulating-them-softly/.
- — —. 2024. *The Politics of Platform Regulation: How Governments Shape Online Content Moderation*. Oxford Studies in Digital Politics. Oxford, UK: Oxford University Press.
- Gorwa, Robert and Ghanaraj Thakur. 2024. *Real Time Threats: Analysis of Trust and Safety Practices for Child Sexual Exploitation and Abuse (CSEA) Prevention on Livestreaming Platforms*. Report, Center for Democracy & Technology. November. <https://cdt.org/wp-content/uploads/2024/11/CDT-Research-Real-Time-Threats-hqp-final.pdf>.
- Guo, Cheng and Kelly Caine. 2025. "Throwaway Accounts and Moderation on Reddit." Preprint, *arXiv*. <https://doi.org/10.48550/arXiv.2501.17430>.
- Gurumurthy, Anita and Nandini Chami. 2021. *The Deal We Always Wanted: A Feminist Action Framework for the Digital Economy*. Friedrich-Ebert-Stiftung Working Group on Feminist Visions of the Future of Work. <https://collections.fes.de/publikationen/ident/fes/17008>.
- He, Qinglai, Yili Hong and T. S. Raghu. 2025. "Platform Governance with Algorithm-Based Content Moderation: An Empirical Study on Reddit." *Information Systems Research* 36 (2): 1078–95. <https://doi.org/10.1287/isre.2021.0036>.
- Horne, Benjamin D., Myles Bowman, Milo Z. Trujillo, Mauricio Gruppi and Cody Buntain. 2025. "Supply and Demand on Alt-Tech Social Media: A Case Study of BitChute." *Computational Communication Research* 7 (1). <https://doi.org/10.5117/CCR2025.1.4.HORN>.
- Klein, Naomi. 2023. *Doppelgänger: A Trip into the Mirror World*. Toronto, ON: Vintage Canada.
- Laidlaw, Emily B. and Hilary Young. 2018. "Internet Intermediary Liability in Defamation." *Osgoode Hall Law Journal* 56 (1): 112–61. <https://doi.org/10.60082/2817-5069.3389>.
- Manne, Kate. 2018. *Down Girl: The Logic of Misogyny*. Oxford, UK: Oxford University Press.
- Massanari, Adrienne. 2017. "#Gamergate and The Fapping: How Reddit's Algorithm, Governance, and Culture Support Toxic Technocultures." *New Media & Society* 19 (3): 329–46. <https://doi.org/10.1177/1461444815608807>.
- Myers West, Sarah. 2018. "Censored, Suspended, Shadowbanned: User Interpretations of Content Moderation on Social Media Platforms." *New Media & Society* 20 (11): 4366–83. <https://journals-sagepub-com.proxy.lib.uwaterloo.ca/doi/full/10.1177/1461444818773059>.
- Nashrulla, Tasneem. 2019. "Incels Are Running an Online Suicide Forum That Was Blamed For A Young Woman's Death." BuzzFeed News, June 6. www.buzzfeednews.com/article/tasneemnashrulla/incels-suicide-forum-woman-killed-herself.

- Owen, Taylor. 2019. "Introduction: Why Platform Governance?" In *Models for Platform Governance*, 3–6. Waterloo, ON: CIGI. www.cigionline.org/static/documents/documents/Platform-gov-WEB_VERSION.pdf.
- Padda, Karmvir K. Forthcoming 2026. "Mirror Worlds and Misogyny: A Mixed-Methods Analysis of Incel Discourse Across Mainstream and Alt-Tech Platforms." *New Media & Society*.
- Ribeiro, Manoel Horta, Jeremy Blackburn, Barry Bradlyn, Emiliano De Cristofaro, Gianluca Stringhini, Summer Long, Stephanie Greenberg and Savvas Zannettou. 2021. "The Evolution of the Manosphere across the Web." *Proceedings of the International AAAI Conference on Web and Social Media* 15 (1): 196–207. <https://doi.org/10.1609/icwsm.v15i1.18053>.
- Rogers, Richard. 2020. "Deplatforming: Following Extreme Internet Celebrities to Telegram and Alternative Social Media." *European Journal of Communication* 35 (3): 213–29. <https://doi.org/10.1177/0267323120922066>.
- Suzor, Nicolas P. 2019. *Lawless: The Secret Rules That Govern Our Digital Lives*. 1st ed. Cambridge, UK: Cambridge University Press. <https://doi.org/10.1017/9781108666428>.
- Svantesson, Dan Jerker B. 2024. "Global Content Orders: A Legal Analysis of eSafety Commissioner." *Centre for Space, Cyberspace & Data Law* 1: 17. https://bond.edu.au/sites/default/files/2024-06/LAW_Centre%20paper_Global%20content%20orders%20A%20legal%20analysis%20of%20eSafety%20Commissioner.pdf.
- Thorburn, Joshua. 2023. "Exiting the Manosphere. A Gendered Analysis of Radicalization, Diversion and Deradicalization Narratives from *r/IncelExit* and *r/ExRedPill*." *Studies in Conflict & Terrorism*: 1–25. <https://doi.org/10.1080/1057610X.2023.2244192>.
- Urman, Aleksandra and Stefan Katz. 2022. "What They Do in the Shadows: Examining the Far-Right Networks on Telegram." *Information, Communication & Society* 25 (7): 904–23. <https://doi.org/10.1080/1369118X.2020.1803946>.