Summary Report

Author: Dr. Shelly Ghai Bajaj, Postdoctoral Fellow, University of Waterloo & the Balsillie School of International Affairs

Panel Theme: Disinformation, Digital Space & Democratic Processes

Key Issues: This report will address several of the questions identified in the "Schedule and description of policy roundtables" including:

i) the context of foreign interference, attribution, and the challenges associated with attribution (Q2, Q3);
ii) the spread and experiences of MDM by diaspora communities and distinct considerations in developing mitigation strategies for Canada's diverse communities affected by MDM (Q8, Q4, Q5), and;
iii) approaches to confronting the challenge posed by MDM to democratic political institutions, strategies for diverse communities affected by MDM, and strategies to detect, deter, and counter MDM (Q1, Q8, Q9).

Assessment:

Foreign interference during democratic elections, including the use of disinformation, poses unique threats and challenges for liberal democracies. While research and analysis tend to focus on disinformation and electoral interference, the spread of disinformation has impacts beyond elections. Disinformation can increase polarization, sow discord and erode social cohesion, undermine democratic trust, and/or increase 'information pollution.' This report underscores how the growing complexity of the disinformation in the digital information space means that disinformation spread is a continuous and constant process that occurs between and beyond election cycles.[1]

To effectively address disinformation threats and harms in the digital information space, strategies must align with Canada's commitment to pluralism and diversity, a liberal media and information environment, and the protection of individual rights and civil liberties. The disinformation threats to national security and public safety of today and tomorrow require comprehensive whole-of-government and whole-of-society approaches to help produce a robust form of capacity and resiliency to disinformation.

[1] Although there are conceptual distinctions between mis-, dis-, and mal-information (MDM) this report opts to use the term 'disinformation' to refer to false, inaccurate, and misleading information that causes individual or collective harm. For a more robust discussion on why the distinctions between MDM is difficult, please see Bajaj, S.G. & Momani, B. "Introduction: Digital Disinformation and Democracy in Canada" in S.G. Bajaj & B. Momani (eds.), *Disinformation and Democracy in the Digital Age: A Canadian Perspective*. University of Toronto (forthcoming).

The growing complexity also means that understanding the differential experiences and impacts of disinformation cannot be a one-size-fits all approach. Mitigation strategies need to be designed to consider the short, intermediate, and long-term objectives that build trust. Approaches to building capacity and resiliency must account for: i) the increasingly variegated technological landscape of platforms and digital spaces as well as digital tools and tactics for the spread of disinformation; ii) a shifting geopolitical context, and; iii) the differential experiences and impacts of disinformation for certain subpopulations like ethnocultural diasporas.

The Evolving Landscape of Digital Disinformation

i) Technological landscape:

The rapid development of digital information communication technologies (ICTs) means that disinformation can flow at an unprecedented scale, scope, and speed. The mechanisms by which it spreads is also shaped by the digital information landscape. Digital automation, data harvesting and mining, predictive analytics, algorithmic amplification, bot and troll networks, and generative AI in the form of audio, text-based, and visual deepfakes allows for disinformation content to be produced and disseminated at scale. These technologies can be combined with dissemination strategies including online astroturfing[2] and microtargeting.[3] The technological drivers of disinformation, on both the content creation side and the dissemination side, are easy to access, inexpensive, and increasingly sophisticated.

The platforms on which disinformation can spread is also evolving with unique digital infrastructure for many of the platforms. Much of our understanding of disinformation is based on open and more public digital spaces like Facebook, X, Instagram, and YouTube. The platform landscape, however, is increasingly differentiated and fragmented. The exponential growth of the micro vlogging platform, TikTok, from 2020 onwards demonstrates how quickly platforms can become key sources of information.[4] The digital information space also includes a wide range of end-to-end encrypted (E2EE) or partially encrypted private chat and direct messaging apps like Facebook/Instagram Direct Messenger, WhatsApp, Telegram, Signal, and Discord. The Chinese 'super-app,' WeChat, offers a distinct mix of functions that combine elements of other social media platforms.

These digital spaces are also inherently transnational. These digital spaces defy our understanding of state boundaries. This means that there are spillover effects from disinformation campaigns providing indirect pathways for disinformation transfer. In

---

[2] Chan, J. (2024). Online astroturfing: A problem beyond disinformation. *Philosophy & Social Criticism*, *50*(3), 507-528. https://doi.org/10.1177/01914537221108467

[3] Dawson, J. (2021). Microtargeting as Information Warfare. *The Cyber Defense Review*, *6*(1), 63–80. https://www.jstor.org/stable/26994113.

[4] Newman, N., et al (2024) Reuters Institute digital news report 2024. Reuters Institute for the study of Journalism, p.11. https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2024-06/RISJ_DNR_2024_Digital_v10%20lr.pdf.

addition, the very nature of digital information spaces means that disinformation content is constant and continuous, rather than a single, isolated episode. Digital disinformation also travels across platform boundaries. This creates increasingly complex digital information environments with disinformation spreading beyond and between election cycles.

i) Geopolitical context:

The broader geopolitical context is also in a state of flux. Multilateralism and the rules-based order that has structured the post-war international institutional and security system is increasingly strained. The global arena is increasingly characterized by multipolarity, great power competition, populist-nationalist governments coupled with democratic backsliding and creeping authoritarianism, and alternative multilateral arrangements like the BRICS+.

Against these shifting geopolitical dynamics, the digital information space is increasingly used as an operational domain for geopolitical tensions and engagement. With the low-barrier-to-entry, there is a diversification of state and non-state actors in the digital information space. There are well known state actors but also emerging states that are well-institutionalized and organized in their ability to operate strategically. There are also intermediary states involved in the supply chain of disinformation production and dissemination including content farms operating in the Global South. In addition to state actors, a diverse range of non-state actors such as automated bot and troll networks, hacking collectives, cyber troops, and lone wolves operate to disseminate disinformation.

ii) Disinformation within Ethnocultural Diasporas

Disinformation is not experienced universally and interacts with identity in complex ways, producing differential experiences and impacts for many of Canada's ethnocultural diasporas. First, these communities have distinct digital information environments. Ethnocultural diasporas belong to transnational digital information environments and use private encrypted chat and direct messaging apps at rates that are higher than Canadian averages. In an original survey conducted among Arab, Chinese, and South Asian Canadians, 81.5 percent reported belonging to chat groups on chat and direct messaging apps with international membership.[5] These chat groups are transnational, connecting information sources across hard borders. In the same survey, 40 percent of respondents reported observing someone being called out for sharing mis/disinformation.

The private and closed nature of these spaces makes these spaces vulnerable to possible exploitation by threat actors. Our research also suggests that a significant portion of disinformation circulating in these spaces is third language content, providing distinct

[5] Key findings are highlighted in this report from an original survey conducted among Arab, Chinese, and South Asian communities. The survey was designed, administered and analyzed by a research team, including the author of the report, in 2023. The survey focused on digital information sharing practices and behaviours in digital spaces, including private and encrypted chat apps to understand how and why disinformation circulates in these spaces within these communities. The research was funded through the Digital Citizenship Contribution Program through Heritage Canada.

mechanisms for disinformation to be microtargeted to ethnocultural diasporas. This is consistent with findings of election-related disinformation campaigns in the United States targeting the LatinX communities.[6]

These communities also face differential disinformation impacts and harms. At an individual-level, users often feel overwhelmed by disinformation circulating within their online and offline information networks. Moreover, individual users report negotiating challenging online encounters surrounding information sharing, including calculations of when it is safe or unsafe to share or engage with content. There are very real concerns around censorship, surveillance, and reprisals from home countries that can alter the everyday digital behaviours and practices of individual users.[7]

At the group-level, ethnocultural diasporas face a double burden of disinformation labour around disinformation circulating from within their communities and disinformation narratives that target their communities.[8] Ethnocultural diasporas often report experiencing hate as a result of disinformation (46 percent of survey respondents). A majority (51 percent) reported feeling at least some marginalization against their communities because of online disinformation.[9]

Increasing Complexity of Disinformation and Digital Spaces: Policy Considerations & Implications

There are several policy-related implications of the technological and geopolitical contextual factors shaping foreign interference and the differential spread and impacts of digital disinformation. First, the diversification of the landscape obfuscates the origins and intent of disinformation content. The challenge of attribution is further complicated by the blurring distinctions between domestic and foreign digital information environments as many of these spaces are transnational and global in scope. Even if attribution of content is possible and a pattern of foreign interference is established, disinformation, once available in the digital information space, can take on a life of its own, moving across platform boundaries.

However, while attribution is difficult and challenging, it is an important data point and can help with the strategic foresight of future threats. Coordinated disinformation operations, computational propaganda, and foreign information manipulation techniques may reveal identifiable patterns over time and across liberal democracies. As an antidote to the shifting geopolitical context, deepening and extending cooperation with allies is important in the digital information space. Identifying these patterns, especially if done through

---

[6] For a more comprehensive discussion on disinformation threats and harms for diasporas from a comparative perspective, see Bajaj, S.G. 2024. Digital Disinformation Threats and Ethnocultural Diasporas. In Adlakha-Hutcheon, G. and Kelshall, C. (eds.). *(In)Security: Identifying the Invisible Disruptors of Security*. Pp. 56-66. Switzerland: Springer Nature. https://doi.org/10.1007/978-3-031-67608-6.

[7] These insights are gleaned from semi-structured focus group discussions.

[8] Bajaj, S. G., & Momani, B.. Forthcoming. Digital Diasporas, COVID Disinformation and Chat Apps: The Spread and Implications of COVID Disinformation. In *Disinformation and Democracy in a Digital Age: A Canadian Perspective.* Toronto, CA: University of Toronto Press.

[9] This data is based on our survey of Arab, Chinese, and South Asian Canadians.

leveraging multilateral arrangements like the G7 Rapid Response Mechanism, can help better anticipate disinformation threats over the longer term.

Other government responses include investing in public education curriculum to include digital literacy, but also a comprehensive curriculum with an emphasis on critical thinking skills, civics, history, social science, and humanities education to bolster information literacy. Of course, this suggests the need for greater intergovernmental cooperation based on a shared commitment to public education about disinformation. Increased communication and coordination at all levels of government can help to build trust in democratic institutions and processes.

Government support for research and collaborative research networks and partnerships across sectors including academia, government and policy, civil society, and industry is also required. Innovative solutions that leverage digital technologies to detect and counter disinformation in real-time can be developed by such partnerships. Examples from the United States like the Atlantic Council's Foreign Interference Attribution Tracker that uses open-source intelligence (OSINT), key metrics, and an ethical and transparency framework can provide insights into how to build Canadian capacity.

While legal-regulatory approaches to platform governance and reducing online harms and technological interventions like content moderation and false, misleading, inaccurate, and altered content warnings have a place, for private encrypted chat apps these interventions are limited and objectionable. The protection of privacy afforded by E2EE helps to differentiate these spaces from other platforms. Increased surveillance, censorship, or securitizing these digital spaces can do more harm than good, eroding trust with ethnocultural diasporas who are already disproportionately targeted by disinformation and subject to distinct harms from disinformation. Over the long term, top-down approaches in these digital spaces can diminish confidence in a liberal media and information environment.

The diversity of the digital information space, including private encrypted chat apps, requires responses that are bottom up rather than top-down. In these spaces, the government and government agencies are best considered to play a supporting role through funding, facilitating outreach, and creating opportunities for cross-sector collaboration. In India, for example, a research project involving WhatsApp and civil society organizations developed a voluntary crowd-sourced tipline for disinformation content shared on WhatsApp in the lead-up to the 2019 general election. Researchers found that the tipline was it effectively capturing disinformation content quickly and often before such content could appear in more public digital spaces.[10]

Mitigation strategies for private encrypted chat apps requires the prioritization of digital agency of individual users and the experience and expertise of civil society organizations

[10] Other jurisdictions have tried similar tiplines. See, Kazemi, A., Garimella, K., Shahi, G. K., Gaffney, D., & Hale, S. A. (2022). Research note: Tiplines to uncover misinformation on encrypted platforms: A case study of the 2019 Indian general election on WhatsApp. *Harvard Kennedy School (HKS) Misinformation Review*. https://doi.org/10.37016/mr-2020-91.

as connective and trusted intermediaries for these communities. Working with more vulnerable communities that deepens and extends trust must be grounded in a firm commitment upholding high ethical standards regarding the protection of individual privacy, anonymization, and transparency. Individual users and civil society organizations are aware of the dynamics of disinformation spread and distinct impacts of disinformation *within* and *on* their communities. Seeing individual users and civil society organizations as equal partners in building resiliency and capacity must be a guiding principle.

Policy responses to dealing with disinformation threats, harms, and impacts should be grounded in comprehensive frameworks of strengthening human security[11] and building of trust, especially among more vulnerable communities. This requires government action, but also must go beyond government action. The idea that the 'best defence is a strong offence' requires strategies to cultivate robust and durable societal resilience and capacity to digital disinformation threats and foreign interference.

[11] For evolving conceptions of human security that considers the intersection of identity, human security, and emergent and digital technologies, see Adlakha-Hutcheon, G. & Kelshall, C. 2024. *(In)Security: Identifying the Invisible Disruptors of Security*. Switzerland: Springer Nature. https://doi.org/10.1007/978-3-031-67608-6.