

Policy Brief No. 1 – January 2024

New Technologies Challenge Freedom of Thought: Cases and Directions for Research

Mai Mavinkurve and Maroussia Lévesque

Key Points

- New technologies ranging from social media to brain implants that interface with computers pose unprecedented risks to our cognitive autonomy.
- As yet, there is no consensus on where the line falls between legitimate influence and unlawful manipulation.
- Emerging technologies strain current regulations' ability to protect freedom of thought.
- Policy makers and other stakeholders need to prioritize freedom of thought in governing future technological deployments.

Introduction

Until recently, the ability to keep one's thoughts private was taken for granted. But recent developments call into question this assumption, as increasingly subtle, powerful and invasive technologies are becoming more pervasive. This policy brief offers case studies illustrating how technology threatens our freedom of thought and points to some directions for research.

Technology-Enabled Interference with Freedom of Thought

Technology has seeped into every facet of our daily life. We have embraced it as a means to enhance our decision making for daily activities ranging from our means of transportation to our place of work, our choices and locations for our meals, the extent of physical activity we need, and the type of news or information we consume. Initially, these tools appeared harmless, providing us with new efficiencies and productivity. However, over time a subtle, perhaps unconscious shift transpired. Modern and emerging technologies have now assumed a commanding role by establishing a direct conduit to the deepest recesses of our minds and thoughts.

About the Authors

Mai Mavinkurve is a CIGI senior fellow and a serial entrepreneur. Mai has founded and run companies in the emerging technology space building distributed data, artificial intelligence (AI) and blockchain solutions for industry.

Mai is an industry expert and adviser in applied AI and data governance. As a member of Innovation, Science and Economic Development Canada's Economic Strategy Table on digital industries, she led the subgroup that developed recommendations on national data strategy and intellectual property. She has also represented Canada at the G7 ministerial meetings on AI and the future of work. Mai has advised on Ontario's Digital and Data Task Force, as well as co-chaired a data governance initiative with the Standards Council of Canada. Mai is an experienced engineer and executive leader with a focus on the practical enterprise applications and implications of AI and data. Mai has been acknowledged as one of the 30 Most Influential Women in AI in Canada.

Maroussia Lévesque is a CIGI senior fellow, a doctoral candidate at Harvard Law School, an affiliate at the Berkman Klein Research Center, and a member of the Indigenous Protocol and Artificial Intelligence working group. Building on previous work on AI and human rights at Global Affairs Canada, she currently researches AI governance.

She previously worked for Quebec's public inquiry commission on electronic surveillance and clerked for the Chief Justice at the Quebec Court of Appeal, and holds degrees from Harvard, McGill and Concordia Universities.

These developments call into question our ability to keep our thoughts private, to shield them from manipulation, and ultimately to think for ourselves.

Concepts such as influence, persuasion, manipulation and brainwashing are not new. International human rights and domestic constitutional law place the inner sanctum of the mind off limits from external interference (for details, see Alegre 2021, 2022; Laidlaw, forthcoming 2024). However, with recent advances in certain technologies the exertion of influence has assumed a nuanced yet potent form that strains analog-era legal regimes. This influence operates on a large scale, driven by mass swaths of data and automated algorithms, while targeting individuals at a micro-level. It is gradually eroding our capacity to exercise agency over our thoughts and potentially compromising critical thinking faculties (see Table 1). A crucial question arises: Where should the line be drawn between permissible influence and outright infringement upon our inherent right to freedom of thought? Who has oversight and where do responsibilities lie?

Each technology has a varying degree of urgency and impact. However, they all demonstrate the potentially detrimental effect that technology can have on our freedom of thought, our inner sanctuary, the "*forum internum*" (Alegre 2022).

Social Media and Mental Health

Mental health issues appear to be unintended consequences of the skewed incentive model built into social media applications, which seek to keep users addicted. As American statistician and political science professor Edward Tufte said in the Netflix documentary *The Social Dilemma*, "There are only two industries that call their customers 'users': illegal drugs and software" (Orlowski 2020). Some of society's youngest and most vulnerable users are being particularly affected.

As Samantha Bradshaw and Tracy Vaillancourt write, "the increased impact of social media on the way we think, feel and behave provides a new opportunity to evaluate challenges around teen mental health and well-being on digital platforms through

Table 1: Technologies Impacting Freedom of Thought

Now	Short Term	Medium and Long Term
<ul style="list-style-type: none"> Social media 	<ul style="list-style-type: none"> Generative AI Workplace surveillance 	<ul style="list-style-type: none"> AR and VR BCIs

Source: Authors.

Note: AI = artificial intelligence; AR = augmented reality; BCIs = brain-computer interfaces; VR = virtual reality.

the lens of freedom of thought” (Bradshaw and Vaillancourt, forthcoming 2024). A case in point is the story of Molly Russell, a 14-year-old from the United Kingdom, who tragically took her own life in 2017. She was an avid Instagram user. A British coroner recently concluded that the internet “affected her mental health in a negative way and contributed to her death in a more than minimal way” (Walker 2022; Satariano 2022). Her case has sparked considerable controversy, highlighting the profound influence that social media can exert on an individual’s behaviour. In 2021, leaked internal research documents from Facebook revealed that its app Instagram worsens body image issues for one in three teenage girls. The conclusions of its own corporate documents make clear that social media has contributed to mental health issues (Gayle 2021).

The impact of TikTok on freedom of thought is also of concern. According to a report issued by the Center for Countering Digital Hate (CCDH), TikTok pushes content about eating disorders, self-harm and suicide to some teens within a matter of minutes of their joining the platform. Rather than mitigating the dissemination of harmful material, TikTok appears to target the most vulnerable teens with a larger quantity of harmful content. The report further reveals the alarming frequency with which the platform bombards these users with body image and mental health material: every 39 seconds (CCDH 2022, 5). According to researcher Maen Hammad, “the algorithm forces you into a spiral of depression, hopelessness, and self harm, and it’s terribly difficult to get out of that spiral once the algorithm thinks it knows what you want to see. And it’s extremely alarming to see how easy it is for children to fall into this spiral” (quoted in Gilbert 2023). A report co-authored by Hammad found that TikTok’s recommendation algorithm is so sophisticated that within minutes of a new user joining its platform, its youngest users can be inundated with violent, extremist and misogynistic incel content (Ekō 2023).

The US surgeon general, Dr. Vivek Murthy, has raised a grave concern regarding what he views as the United States’ pre-eminent health issue: the pervasive mental health crisis affecting numerous people across different facets of life (Bloomberg 2022). Murthy has emphasized that technology platforms can have adverse effects for many people, and that excessive social media use is a significant contributory factor to mental health problems affecting children, adolescents and young adults (US Department of Health and Human Services 2021, 2023). Indeed, overuse of social media emerges as a noteworthy risk for suicide among young adults.¹ The Social Media Victims Law Center reports that teens and youth are being heavily influenced by negative content on social media platforms, which can lead to mental health issues, and that social media is linked to suicide, depression and anxiety.² The impact of social media on mental health is an ongoing concern, with many experts warning of its detrimental effects.³

These examples unveil a distressing reality. In the absence of protective mechanisms, the most vulnerable segments of our society are profoundly impacted. The destructive consequences arising from online content propagated by automated algorithms that exploit our thoughts, preferences, aversions and actions cannot be disregarded.

Social Media and Politics

Social media technology has been wielded as a tool to fuel polarization, electoral interference and the destabilization of democracy. The notorious Cambridge Analytica scandal during the

1 See <https://socialmediavictims.org/effects-of-social-media/>.

2 Ibid.

3 Ibid.

2010s stands as a prominent illustration of this phenomenon. The company managed to acquire access to the personal data of millions of Facebook users, without their consent. This harvested data was subsequently utilized to construct detailed profiles of users and to deploy targeted political ads during the 2016 US presidential election (Raymond 2022). The scale of the issue went far beyond the United States, and incidents of this nature were revealed across Southeast Asia (Rahn 2018).

While political influence and propaganda are not new, the Cambridge Analytica scandal revealed the effectiveness of social media as a platform through which to manipulate unsuspecting voters. Consequently, a new paradigm emerged, characterized by the systematic dissemination of misinformation, disinformation and fake news campaigns by politicians. This calculated strategy applies the principle of the “liar’s dividend” (Chesney and Citron 2019, 1758). It consists of casting doubt on democratic institutions to create a sense of confusion and mistrust. This dangerous encroachment on freedom of thought jeopardizes people’s self-determination and autonomy and is harmful to democracy itself. Recently, translated Chinese military documents have revealed that the military sees the nature of warfare as “shifting from destroying bodies to paralyzing and controlling the opponent’s mind” (Farahany 2023). TikTok and its parent company, the Chinese company ByteDance, have also faced scrutiny for national security risks (Vergun 2023).

Social media provides these and many other citable examples of where and how technology has trespassed on our ability to think and act freely. While other emerging technologies may not yet exhibit such overt manifestations, their underlying operational framework remains fundamentally aligned with the problematic manipulation and alteration of thought processes to maximize the probability of desired future actions. The transgressions of social media have transcended the boundaries of benign influence, having hijacked our cognitive autonomy in multiple domains. However, the quandary of accountability looms large, sparking fervent debates that have yet to yield any definitive answers.

Generative Artificial Intelligence

The age of artificial intelligence (AI) has begun (Gates 2023). In the past decade alone, the advancement of AI has been remarkable. In 2013, only a handful of companies possessed the capability to incorporate AI or machine learning into their products. Today, however, the emergence of generative AI and innovative tools such as ChatGPT have democratized access to the transformative power of AI. It is now accessible to organizations of all kinds.

When it comes to the *forum internum*, generative AI triggers apprehension due to its capacity to erode trust, undermine our perception of reality and challenge notions of truth. Generative AI can produce novel outputs across various domains. It can generate images, music, audio, text, code and even entire websites. It is a powerful field within AI to create new content without direct human intervention.

Generative AI is improving at a rapid pace. In February 2023, a peculiar video surfaced on YouTube, purportedly capturing an unlikely conversation between Canadian Prime Minister Justin Trudeau and American podcast host Joe Rogan (Vega 2022). The outlandish nature of such an interview occurring was alone sufficient to raise suspicions regarding its authenticity and suggest that it was fabricated. In March 2023, AI-generated images of former US president Donald Trump being arrested circulated on Twitter (Belanger 2023). These images were considerably more challenging to recognize as artificially generated. Then, in April 2023, a malicious individual deployed an AI-generated voice of a young girl as part of an elaborate scheme to extort \$1 million from her frightened and concerned mother, who was deceived into believing her daughter had been kidnapped (Paul 2023). These examples illustrate that in the span of a few months, generative AI went from being a curiosity to a persistent disruptor of our digital information landscape.

Combining the capabilities of social media and generative AI will only compound the risks to freedom of thought. Social media platforms provide a means of mass distribution and behavioural micro-targeting. Generative AI provides a means

of mass content creation. When combined, these two enablers could supercharge computational propaganda. Computational propaganda refers to the use of automated tools, algorithms and computational techniques to manipulate public opinion and spread propaganda on social media platforms.⁴ Leveraging computational tools to amplify certain narratives or spread false or misleading information, computational propaganda manipulates online conversations. One of the most concerning aspects of generative AI in the context of freedom of thought is the increasing difficulty for users to discern between real and fake information, leading to the potential for mass manipulation of public opinion. It is expected that the sophistication of generative AI tools will only continue to grow, bringing with it the need for guidelines and guardrails around its use.

Workplace Surveillance

A *nudge* is “any aspect of the choice architecture that alters people’s behavior in a predictable way without forbidding any options or significantly changing their economic incentives. To count as a mere nudge, the intervention must be easy and cheap to avoid” (Thaler and Sunstein 2008, 6). Digital nudging has found applications within the user experience designs of diverse consumer digital applications. Capitalizing on the inherent biases that influence all our decision making (Weng 2020), companies can subtly shape, encourage or manipulate users to take certain desired actions. This practice initially appeared innocuous, but increasingly sophisticated algorithms have transformed it from mere enhancements to disconcerting surveillance and control systems, raising concerns regarding privacy and the extent of manipulation involved. These are all ways in which our freedom of thought is systematically eroded.

Digital nudging is becoming increasingly prevalent within workplace tools. With the rise of remote work, one particular facet experiencing growth is worker surveillance. The expanded collection of data on employees through technology has paved the way for algorithmic management — an approach that leverages data and algorithms to

monitor, influence and ultimately manage worker behaviour (Ajunwa 2023). Various software and hardware solutions are now available to monitor and track employees, empowering companies to make determinations regarding matters such as worker remuneration and productivity processes.

Several examples highlight the expanding landscape of technological surveillance and its implications:

- Canon Information Technology’s Beijing office employs a system that mandates employees to smile into a camera for building entry. Company representatives state that this approach promotes a positive atmosphere and encourages employees to create an uplifting environment (Reuter 2021).
- The contemporary workplace offers a wide array of productivity tools and office digital assistants, each claiming to enhance workers’ efficiency throughout the day. These tools monitor employees’ work patterns and provide recommendations to boost productivity. The extent to which this data is processed and utilized poses a direct encroachment on workers’ established work methods, particularly when implemented without their full knowledge and consent (Kantor and Sundaram 2022).
- Certain workplaces provide employees with physical devices capable of tracking their movements and locations. For instance, Amazon asked delivery drivers to sign biometric data consent forms so the company could collect data around the movements and facial expressions of their drivers (Hautala 2021).
- Smart buildings serve as another illustration. These structures possess technology that makes decisions on behalf of occupants, raising questions about the preservation of free will when technology assumes decision-making roles (Lecomte 2022).

These instances underscore the expanding reach of surveillance technologies and raise concerns regarding privacy, consent and the potential erosion of individual autonomy within various domains, be it the workplace, academic settings or our everyday environments.

The scope and scale of workplace monitoring is increasing. Employing an assortment of tools and technologies, workplace surveillance

4 See www.oii.ox.ac.uk/research/projects/computational-propaganda/.

is progressively encroaching upon personal boundaries, verging on intrusiveness. Companies assert that such monitoring endeavours enhance productivity. Presently, a multitude of worker “engagement” tools exist, constituting a feedback loop that guides individuals in modifying their behaviour (Indeed Editorial Team 2022). Nevertheless, as these tools persist over time, there arises the potential for them to erode employee autonomy and undermine their agency.

Augmented and Virtual Reality

Augmented reality (AR) and virtual reality (VR) technologies are rapidly advancing, and their potential to change how we interact with the world and each other is enormous. Products such as the Oculus Rift and Microsoft’s VR and mixed reality headsets like the HoloLens, among other devices, are becoming increasingly popular. However, these technologies also pose risks to freedom of thought, privacy and other human rights:

- **Control of individuals within the virtual world.** It’s possible that in the future, people may spend a lot of time in a VR environment interacting with others and in virtual places created by a small group of companies. As with social media, these virtual worlds could create environments where people’s thoughts, behaviour and opinions could be influenced or controlled. In such a scenario, an entire surveillance environment might result where all aspects of life are monitored and controlled.
- **Blurred lines between real and virtual.** The immersive nature of VR technologies makes them particularly addictive. These technologies might result in scenarios where VR becomes increasingly realistic, so that in the future it will be harder to distinguish between what is real and what is virtual. These blurred lines could have significant consequences for our mental health and well-being.
- **Infringement on what we feel and think.** The devices that enable AR/VR experiences collect a lot of data about individuals. These data points include biometric information such as retinal scans, pupil tracking, heart rate, body

temperature, fingerprint scans and so forth. In a paper titled “Watching Androids Dream of Electric Sheep: Immersive Technology, Biometric Psychography, and the Law,” Brittan Heller (2020) introduced the term *biometric psychography* to describe individuals’ involuntary sharing of data regarding their physical reactions. For example, our pupils may dilate when we see something we like. AR/VR technologies routinely collect such biometric data, which is particularly distressing given users’ incomplete understanding of these practices. Despite nominal consent, such data harvesting can infringe privacy when the data provides the basis for making decisions about individuals (Carbonneau 2022). In a virtual world where sensory biometric data is constantly collected, our reactions can result in our rights to feel, or think what we want in private, to be infringed. Some researchers believe that privacy in the metaverse will not be possible (Rosenberg 2023).

These technologies have the potential to disrupt how we interact with the world. Consequently, we need policies and laws to ensure that we can mitigate the risks that they present (Madiega et al. 2022).

Brain-Computer Interfaces

Neuralink, the brainchild of Elon Musk, obtained US Food and Drug Administration (FDA) approval in May 2023 to initiate human testing (Levy, Taylor and Sharma 2023). This venture aims to develop brain implants capable of interfacing with computers — thereby creating a brain-computer interface, or BCI — to enable individuals with disabilities to gain visual or motor functions. These implanted chips receive sensory inputs from the body that are then processed by computers to generate signals that facilitate physical activities. Neuralink is not the only player in the BCI industry; other companies have received FDA approval and been testing their implants (US Government Accountability Office 2022). One noteworthy example is a device called the “brain-spine interface” developed by a team from the Swiss Federal Institute of Technology in Lausanne (Lewis 2023). Variations of this system have demonstrated the capacity to translate thought signals from a skull implant into electrical signals for the spine.

Remarkably, a paralyzed test subject successfully controlled movements by simply thinking about actions like walking or climbing stairs.

The BCI field has grown in recent years and shown potential for rehabilitating specific physical disabilities. The prospect of a computer working in partnership with the human brain to control physical body functions, as well as external devices, is both exciting and daunting. The US Department of Defense has invested in BCI research for military applications, such as controlling drones via thoughts and monitoring a soldier's cognitive workload (Binnendijk, Marler and Bartels 2020).

Some critics of BCI fear that it could override a human's free will and self-control (Saha et al. 2021). There is apprehension regarding the extent to which computers may make decisions on behalf of humans, potentially impeding their ability to intervene. While BCI technology is still in its early stages and considered experimental, numerous ethical considerations such as privacy, safety, security, accountability and human autonomy are already clear (ibid.). It is imperative that necessary guardrails be put in place as the field advances and the technology becomes more sophisticated.

Conclusion

In an increasingly interconnected world, the convergence of social media, generative AI, workplace surveillance, AR/VR technologies and BCI has profound implications for our capacity to maintain control over our thoughts and opinions. The case studies above are but a few examples of the myriad ways in which emerging technologies challenge the traditional guardrails that were established in the past century to protect our freedom of thought. Historian and philosopher Yuval Noah Harari (2023) has warned that these new tools have the potential to "hack" human beings, compromising our agency and autonomy. He emphasizes the importance of maintaining our ability to retain control over our thoughts and decisions in the face of advancing technologies, and encourages society to critically examine and navigate the ethical implications of these developments (ibid.).

Most importantly, though, to combat these attacks on our mind, we must apply the rule of law and put in place the appropriate legal and technological boundaries to ensure we create a future we want to live in. It is high time to consider freedom of thought a top-of-mind technology policy concern requiring dedicated resources, greater expertise and decisive action.

Works Cited

- Ajunwa, Ifeoma. 2023. *The Quantified Worker: Law and Technology in the Modern Workplace*. Cambridge, UK: Cambridge University Press. <https://doi.org/10.1017/9781316888681>.
- Alegre, Susie. 2021. *Protecting Freedom of Thought in the Digital Age*. CIGI Policy Brief No. 165. Waterloo, ON: CIGI. www.cigionline.org/publications/protecting-freedom-of-thought-in-the-digital-age/.
- . 2022. *Toward an International Agenda to Protect the Forum Internum*. CIGI Policy Brief No. 176. Waterloo, ON: CIGI. www.cigionline.org/publications/toward-an-international-agenda-to-protect-the-forum-internum/.
- Belanger, Ashley. 2023. "AI-faked images of Donald Trump's imagined arrest swirl on Twitter." *Ars Technica*, March 21. <https://arstechnica.com/tech-policy/2023/03/fake-ai-generated-images-imagining-donald-trumps-arrest-circulate-on-twitter/>.
- Binnendijk, Anika, Timothy Marler and Elizabeth M. Bartels. 2020. "Brain-Computer Interfaces: U.S. Military Applications and Implications, An Initial Assessment." Santa Monica, CA: RAND Corporation. www.rand.org/pubs/research_reports/RR2996.html.
- Bloomberg. 2022. "Mental Health Is Biggest Health Concern in US Says Surgeon General." November 2. www.bloomberg.com/news/videos/2022-11-02/mental-health-is-biggest-health-concern-in-us-says-surgeon-general.
- Bradshaw, Samantha and Tracy Vaillancourt. Forthcoming 2024. *Freedom of Thought, Social Media and the Teen Brain*. Legitimate Influence or Unlawful Manipulation? Policy Brief No. 9. Waterloo, ON: CIGI.
- Carbonneau, Sara. 2022. "Biometric Psychography and its Implications on Personal Autonomy." *The Extended Mind*, August 9. www.extendedmind.io/the-extended-mind-blog/biometric-psychography-and-its-implications-on-personal-autonomy.
- CCDH. 2022. *Deadly by Design*. Washington, DC: CCDH. <https://counterhate.com/research/deadly-by-Design>.
- Chesney, Bobby and Danielle Citron. 2019. "Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security." *California Law Review* 107 (6): 1753–1820. <https://doi.org/10.15779/Z38RVOD15J>.
- Ekö. 2023. *Suicide, Incels, and Drugs: How TikTok's deadly algorithm harms kids*. March. San Francisco, CA: Ekö. https://s3.amazonaws.com/s3.sumofus.org/images/eko_Tiktok-Report_FINAL.pdf.
- Farahany, Nita. 2023. "TikTok is part of China's cognitive warfare campaign." *The Guardian*, March 25. www.theguardian.com/commentisfree/2023/mar/25/tiktok-china-cognitive-warfare-us-ban.
- Gates, Bill. 2023. "The Age of AI Has Begun." *GatesNotes* (blog), March 21. www.gatesnotes.com/The-Age-of-AI-Has-Begun?WT.mc_id=20230321100000_Artificial-Intelligence_BG-LI_&WT.tsrc=BGLI.
- Gayle, Damien. 2021. "Facebook aware of Instagram's harmful effect on teenage girls, leak reveals." *The Guardian*, September 14. www.theguardian.com/technology/2021/sep/14/facebook-aware-instagram-harmful-effect-teenage-girls-leak-reveals.
- Gilbert, David. 2023. "TikTok Is Pushing Incel and Suicide Videos to 13-Year-Olds." *Vice News*, March 21. www.vice.com/en/article/qjv4jw/tiktok-incels-targeting-young-users.
- Harari, Yuval Noah. 2023. "Yuval Noah Harari argues that AI has hacked the operating system of human civilisation." *The Economist*, April 28. www.economist.com/by-invitation/2023/04/28/yuval-noah-harari-argues-that-ai-has-hacked-the-operating-system-of-human-civilisation.
- Hautala, Laura. 2021. "Amazon drivers must consent to biometric monitoring or lose jobs, reports say." *CNET*, March 23. www.cnet.com/tech/tech-industry/amazon-drivers-must-consent-to-biometric-monitoring-or-lose-jobs-reports-say.
- Heller, Brittan. 2020. "Watching Androids Dream of Electric Sheep: Immersive Technology, Biometric Psychography, and the Law." *Vanderbilt Journal of Entertainment and Technology Law* 23 (1): 1–51. <https://scholarship.law.vanderbilt.edu/jetlaw/vol23/iss1/1/>.
- Indeed Editorial Team. 2022. "16 Employee Engagement Tools That Companies Can Use." *Indeed*, December 9. www.indeed.com/career-advice/career-development/employee-engagement-tools.
- Kantor, Jodi and Arya Sundaram. 2022. "The Rise of the Worker Productivity Score." *The New York Times*, August 14. www.nytimes.com/interactive/2022/08/14/business/worker-productivity-tracking.html.

- Laidlaw, Emily. Forthcoming 2024. *Technology-Facilitated Mind Hacking: Protection of Inner Freedoms in Canadian Law*. Legitimate Influence or Unlawful Manipulation? Policy Brief No. 5. Waterloo, ON: CIGI.
- Lecomte, Patrick. 2022. "Smart buildings: What happens to our free will when tech makes choices for us?" *The Conversation*, December 20. <https://theconversation.com/smart-buildings-what-happens-to-our-free-will-when-tech-makes-choices-for-us-189285>.
- Levy, Rachel, Marisa Taylor and Akriti Sharma. 2023. "Elon Musk's Neuralink wins FDA approval for human study of brain implants." *Reuters*, May 26. www.reuters.com/science/elon-musks-neuralink-gets-us-fda-approval-human-clinical-study-brain-implants-2023-05-25.
- Lewis, Dyani. 2023. "Brain–spine interface allows paralysed man to walk using his thoughts." *Nature*, May 24. www.nature.com/articles/d41586-023-01728-0.
- Madiega, Tambiana, Polona Car, Maria Niestadt and Louise Van de Pol. 2022. "Metaverse: Opportunities, risks and policy implications." PE 733.557. European Parliamentary Research Service, June. [www.europarl.europa.eu/RegData/etudes/BRIE/2022/733557/EPRS_BRI\(2022\)733557_EN.pdf](http://www.europarl.europa.eu/RegData/etudes/BRIE/2022/733557/EPRS_BRI(2022)733557_EN.pdf).
- Orlowski, Jeff, dir. 2020. *The Social Dilemma*. Boulder, CO: Exposure Labs.
- Paul, Andrew. 2023. "A mom thought her daughter had been kidnapped — it was just AI mimicking her voice." *Popular Science*, April 14. www.popsci.com/technology/ai-vocal-clone-kidnapping/.
- Rahn, Wesley. 2018. "Facebook's data woes in Southeast Asia." *Deutsche Welle*, April 6. www.dw.com/en/cambridge-analytica-causing-trouble-for-facebook-in-southeast-asia/a-43286109.
- Raymond, Nate. 2022. "Facebook parent Meta to settle Cambridge Analytica scandal case for \$725 million." *Reuters*, December 23. www.reuters.com/legal/facebook-parent-meta-pay-725-mln-settle-lawsuit-relating-cambridge-analytica-2022-12-23/.
- Reuter, Dominick. 2021. "Employees at a Beijing office have to smile for an AI camera to get through the front doors, change the temperature, or print documents." *Business Insider*, June 17. www.businessinsider.com/workers-at-chinese-office-have-to-smile-at-ai-camera-2021-6.
- Rosenberg, Louis. 2023. "New research suggests that privacy in the metaverse might be impossible." *VentureBeat*, February 20. <https://venturebeat.com/cdn.ampproject.org/c/s/venturebeat.com/virtual/new-research-suggests-that-privacy-in-the-metaverse-might-be-impossible/>.
- Saha, Simanto, Khondaker A. Mamun, Khawza Ahmed, Raqibul Mostafa, Ganesh R. Naik, Sam Darvishi, Ahsan H. Khandoker and Mathias Baumert. 2021. "Progress in Brain Computer Interface: Challenges and Opportunities." *Frontiers in Systems Neuroscience* 15: 578875. www.frontiersin.org/articles/10.3389/fnsys.2021.578875/full.
- Satariano, Adam. 2022. "British Ruling Pins Blame on Social Media for Teenager's Suicide." *The New York Times*, October 1. www.nytimes.com/2022/10/01/business/instagram-suicide-ruling-britain.html.
- Thaler, Richard H. and Cass R. Sunstein. 2008. *Nudge: Improving Decisions about Health, Wealth, and Happiness*. New Haven, CT: Yale University Press.
- US Government Accountability Office. 2022. "Science & Tech Spotlight: Brain-Computer Interfaces." GAO-22-106118, September 8. www.gao.gov/products/gao-22-106118.
- US Department of Health and Human Services. 2021. *Protecting Youth Mental Health. The U.S. Surgeon General's advisory*. Washington, DC: US Department of Health and Human Services. www.hhs.gov/sites/default/files/surgeon-general-youth-mental-health-advisory.pdf.
- . 2023. *Social Media and Youth Mental Health: The U.S. Surgeon General's Advisory*. Washington, DC: US Department of Health and Human Services. www.hhs.gov/sites/default/files/sg-youth-mental-health-social-media-advisory.pdf.
- Vega, Manuela. 2023. "Viral video of Joe Rogan interviewing Justin Trudeau was generated by AI and highlights danger of deep fakes." *Toronto Star*, February 22. www.thestar.com/news/canada/2023/02/22/viral-video-of-joe-rogan-interviewing-justin-trudeau-was-generated-by-ai-and-highlights-danger-of-deep-fakes.html.
- Vergun, David. 2023. "Leaders Say TikTok Is Potential Cybersecurity Risk to U.S." *US Department of Defense News*, April 6. www.defense.gov/News/News-Stories/Article/Article/3354874/leaders-say-tiktok-is-potential-cybersecurity-risk-to-us/.

Walker, Andrew. 2022. *Molly Russell: Prevention of future deaths report*. Regulation 28 report to prevent future deaths. October 13. London, UK: North London Coroner's Service. www.judiciary.uk/prevention-of-future-death-reports/molly-russell-prevention-of-future-deaths-report/.

Weng, Sicheng. 2020. "The digital nudge: intersecting behavioral economics and UX Design." Medium, July 1. <https://bootcamp.uxdesign.cc/behavioral-economics-nudging-and-user-experience-design-610940b41d75>.



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.

À propos du CIGI

Le Centre pour l'innovation dans la gouvernance internationale (CIGI) est un groupe de réflexion indépendant et non partisan dont les recherches évaluées par des pairs et les analyses fiables incitent les décideurs à innover. Grâce à son réseau mondial de chercheurs pluridisciplinaires et de partenariats stratégiques, le CIGI offre des solutions politiques adaptées à l'ère numérique dans le seul but d'améliorer la vie des gens du monde entier. Le CIGI, dont le siège se trouve à Waterloo, au Canada, bénéficie du soutien du gouvernement du Canada, du gouvernement de l'Ontario et de son fondateur, Jim Balsillie.

Credits

Managing Director and General Counsel [Aaron Shull](#)
CIGI Senior Fellow and Project Co-Leader [Susie Alegre](#)
Director, Program Management [Dianna English](#)
Program Manager [Jenny Thiel](#)
Publications Editor [Lynn Schellenberg](#)
Senior Publications Editor [Jennifer Goyder](#)
Graphic Designer [Abhilasha Dewan](#)



This policy brief was made possible thanks to the financial support of the Konrad-Adenauer-Stiftung (KAS) Canada.

Copyright © 2024 by the Centre for International Governance Innovation

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

For publications enquiries, please contact publications@cigionline.org.



This work is licensed under the Creative Commons Attribution 4.0 International License. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>. For re-use or distribution, please include this copyright notice.

Centre for International Governance Innovation and CIGI are registered trademarks.

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