Submission of the Citizen Lab at the Munk School of Global Affairs & Public Policy, University of Toronto, to the UN High Commissioner for Human Rights' report on 'Protecting human rights defenders in the digital age'

March 9, 2026

Authors (in alphabetical order):
Noura Aljizawi
Siena Anstis
Dr. Marcus Michaelsen
Claire Posno
Dr. Ron Deibert

For all inquiries regarding this submission, please contact:
Dr. Ronald J. Deibert, Director, The Citizen Lab, Munk School of Global Affairs, Professor of Political Science, University of Toronto, r.deibert@utoronto.ca

About the Citizen Lab

- The Citizen Lab is an interdisciplinary laboratory based at the Munk School of Global Affairs & Public Policy, University of Toronto, focusing on research, development, and high-level strategic policy and legal engagement at the intersection of information and communication technologies, human rights, and global security. The Citizen Lab's research, which combines methods from political science, law, computer science, and area studies, includes investigating the digital targeting of human rights defenders.

Introduction

- The Citizen Lab's submission focuses on ***digital transnational repression*** which entails the use of digital technologies by governments to surveil, intimidate, and silence human rights defenders, among other groups, living in exile or the diaspora. It is part of the broader practice of transnational repression which arises when states extend domestic political controls and coercion into the territory of other states using methods like physical assaults, enforced disappearances, extraditions, and killings.
- To address the global proliferation of these practices, the human rights violations they entail, and the obligation of host states to prevent and respond to them, it is critical that bodies like the UN High Commissioner for Human Rights systematically identify and document acts of (digital) transnational repression, characterize them explicitly as human rights violations, and assess the responsibility of state and private actors involved.
- This submission draws on the Citizen Lab's extensive research on digital transnational repression, including more than 100 interviews with human rights defenders from 24 countries of origin residing in 23 host countries, examining both the methods used and impacts on those targeted.

Methods of Digital Transnational Repression

Digital technologies offer governments engaging in transnational repression a low-cost and effective means to expand the scope and scale of cross-border threats against human rights defenders, reaching deep into other territories and the personal lives of targets. Digital transnational repression facilitates other forms of transnational repression as digital attacks typically set the stage for an escalation into physical attacks and other threats. Human rights defenders in exile and in the diaspora are targeted with the following methods of digital transnational repression:

- **Monitoring and surveillance:** Monitoring social media feeds and online communications to gather, analyze, and exploit information on the activities, daily habits, location, and social networks of targets in order to expose country of origin contacts or facilitate physical surveillance and further threats. Human rights defenders participating in our research reported that they were constantly feeling monitored by authorities from their origin country and often, the presumed data collection already had a chilling effect in and of itself. Some participants described being monitored and surveilled even offline; when they travelled for advocacy events, they were followed, photographed, and filmed by individuals they believed were affiliated with their country of origin.
- **Interception and targeted surveillance:** Hacking of electronic devices, email, and social media accounts to access private information, communications, and contacts. These forms of targeted, invasive surveillance can rely on phishing attacks, physical access to a device, or the remote use of spyware. Increasingly, states use commercial spyware against human rights defenders, which they can purchase on a thriving but highly non-transparent market of surveillance technologies. The companies working in this field exploit vulnerabilities in widely used operating systems and applications to provide their customers with access to phone calls, personal files, emails, chats, and geolocation data of targets, regardless of their location.
- **Intimidation and stigmatization:** Using private, false, and distorted information, as well as online harassment, to silence and discredit human rights defenders. Defamation campaigns take interviews, statements, and other public activities of human rights defenders out of context to misrepresent them, isolate them from home country audiences, or exacerbate fissures within diaspora communities. Women human rights defenders, in particular, face sexualized and gender-based abuse, direct messages with detailed fantasies of sexual assault and rape threats, or their photoshopped fake nude pictures being circulated online. Private information, at times obtained in surveillance operations, is exploited to shame women human rights defenders publicly. Such harassment can be government-coordinated, relying on artificial accounts, paid trolls or regime-affiliated social media influencers. Often, it is also amplified by users loyal to the government and other groups who attack publicly outspoken women human rights defenders based on shared misogynist views and patriarchal norms around women's bodies, sexuality, behaviour, or notions of family honour. The weaponization of gender, sexual orientation, and other characteristics of the targets' intersecting identities for the purposes of digital transnational repression potentially leads to further forms of violence and discrimination.
- **Disruption and censorship:** Curtailing expression on blogs, news/organizations' websites, and social media profiles through distributed denial-of-service (DDoS) attacks, false reports, spam comments, content filtering, and information manipulation. Online publications of human rights and news organizations operating

from exile are typically blocked for audiences inside the country. The content moderation mechanisms of social media platforms are also abused to manipulate new feeds and followers or to mass report the profiles of human rights defenders for allegedly spreading hate speech or pornography. Reported profiles are blocked or shut down by the platform and users often struggle to have them reinstated.

- **Advancements in artificial intelligence (AI)** are enhancing state actors' capabilities in digital transnational repression, particularly with respect to surveillance and online information manipulation. AI-powered surveillance systems process and analyse data on a large scale, automating the inference of social ties and political opinions, which can then be used to map out and prepare attacks against human rights defenders and their networks. AI tools are also used to identify vulnerabilities in software which can be exploited for attacks, write malicious code, and create highly realistic personalized messages for tailored scenarios of social engineering that manipulate targets into opening compromised documents or links. Generative AI further scales up disinformation production by automating the generation of increasingly hard-to-detect false and misleading media, including deepfake videos, images, audio and text, which can be used to fabricate false allegations, spread false information or statements, or sexually harass women human rights defenders online. The proliferation of these tactics due to increasingly accessible AI capabilities risks further enhancing the scope and impact of digital transnational repression, leading to greater harm for human rights defenders targeted with digital repression. The impacts can be particularly severe for vulnerable groups such as exiled women and queer human rights defenders, who are already disproportionately targeted by gendered digital violence and hostility.

Impacts On Human Rights Defenders in Exile

Because digital technologies play a central role in the professional activities and personal lives of human rights defenders, the different tactics and techniques of digital transnational repression significantly impact their safety, security, and well-being. Digital tools enable authoritarian governments to easily instill fear and uncertainty among exiled human rights defenders and members of civil society more broadly, undermine the social relationships within exile and diaspora communities and their countries of origin, and foster self-censorship and withdrawal from activism. As such, digital transnational repression severely impedes civil society's continued ability to use digital tools to exchange, organize and mobilize freely across borders. The impacts caused by digital transnational repression, include but are not limited to:

- **Harms to mental health and well-being**: Exiled human rights defenders interviewed for our research mention experiencing feelings of exhaustion, stress, anxiety, burnout, sleeplessness, and depression as a result of different threats. Constant fear and hypervigilance can foster paranoia and psychological trauma.

- **Deterioration of social relations:** Relationships with family and colleagues fracture under the stress and uncertainty caused by online harassment, digital attacks, and surveillance. Targeted human rights defenders reduce or avoid contact with home country family members and colleagues to protect them against government reprisals. Disinformation and surveillance can breed mistrust, pushing individuals into social isolation and withdrawal. Individuals at risk are also ostracized online and offline as other diaspora members avoid being associated with them for fear of government retaliation.
- **Chilling effects and self-censorship:** Human rights defenders at risk of digital transnational repression adapt their activities, self-censor, or opt for a low profile. They may stop participating in public events and protests, shun media interviews, or cease promoting their work online and in public spaces. Some entirely withdraw from activism to protect themselves and their families. The targeting of selected human rights defenders also has broader dampening effects on entire diaspora communities, as others witness the consequences of publicly speaking out against the home regime.
- **Professional impacts:** Mental harm caused by online harassment and other forms of intimidation negatively affects productivity, work and study routines, forcing some targets to give up or change their profession. Smear campaigns, defamation, and disinformation can negatively impact targets' professional careers and opportunities, with risk-averse employers distancing themselves, while smaller, under-funded organizations cannot guarantee protection.
- **Resource diversion:** Targeted individuals and organisations need to spend significant resources on navigating their security rather than continuing human rights work. This can include training and equipment for digital security. Reporting and blocking social media accounts involved in online harassment and smear campaigns can sap the time and energy needed for human rights work.
- **Risks to physical safety:** Digital threats are often a stepping stone for other types of offline threats, such as stalking, the disruption of public events, in-person harassment, blackmail at embassies and consulates, and threats against family members in the country of origin or in exile. Such tactics of "everyday repression" can lead to severe physical threats such as assault, kidnapping, and even killing attempts.
- **Human rights impacts:** All of the above leads to a series of extraterritorial human rights violations, including violations of the right to freedom of expression, peaceful assembly, freedom of association, the right to privacy and the right to life.

Recommendations

To the UN

- Develop and implement mechanisms to document cases of (digital) transnational repression and integrate reporting into existing UN human rights mechanisms such as the Universal Periodic Review and the Special Procedures.
- Include (digital) transnational repression in the work of human rights treaty bodies by reviewing such practices during country visits and inquiries, in thematic discussions, and in *Concluding Observations*. The Committee on Enforced Disappearances (along with the Working Group) has taken a first step in this direction by issuing a call for inputs on Enforced Disappearances in the context of transnational repression.
- Clarify in guidance by human rights treaty bodies (such as in *General Comments*) that acts of (digital) transnational repression constitute extraterritorial human rights violations for which perpetrating states can be held liable.
- Ensure that UN monitoring, reporting, and protection mechanisms address the gendered dimensions of digital transnational repression, including harassment, disinformation campaigns, and other forms of sexualized and gendered violence targeting exiled women and queer human rights defenders, and integrate gender-sensitive and intersectional approaches into relevant UN mechanisms.
- Encourage member states to develop holistic protection mechanisms for individuals targeted by (digital) transnational repression and develop guidance for host states on preventing and responding to (digital) transnational repression.
- Raise awareness on the human rights risks posed by (digital) transnational repression, and publicly address incidents as well as the resulting human rights violations.
- Facilitate the exchange of best practices among states and relevant stakeholders on addressing transnational repression.
- Strengthen international human rights law and the enforcement of existing treaties and conventions to address the challenges posed by (digital) transnational repression.

To Host States

- Adopt a comprehensive definition of (digital) transnational repression and its gendered dimensions that is used to coordinate countermeasures across government agencies and policy sectors.
- Review state immunity laws and implement necessary changes to ensure victims of digital transnational repression can pursue legal remedies.
- Review existing criminal laws to ensure thorough investigation of and accountability for acts of digital transnational repression.

- Task national cybersecurity agencies to conduct systematic investigations into threats and raise awareness on emerging tactics and techniques of (digital) transnational repression.
- Restrict licenses and export of commercial spyware and strengthen mechanisms of oversight and transparency for governmental use of surveillance technologies.
- Support holistic protection mechanisms grounded in trauma-informed and intersectional approaches for individuals targeted by digital transnational repression, including digital security, psychosocial wellbeing, and legal support.

To Technology Companies

- Explicitly recognize and address digital transnational repression, including its gendered dimensions, in policies, enforcement frameworks, and user guidelines.
- Conduct human rights due diligence across digital infrastructure. Companies whose services or infrastructure – including social media platforms, hosting services, and telecommunication providers – may facilitate digital transnational repression should assess and mitigate these risks in compliance with the UN Guiding Principles on Business and Human Rights.
- Invest sufficient resources to identify and mitigate digital transnational repression, with particular attention to its gendered dimensions. Platforms should ensure adequate regional expertise, language capacity, and contextual knowledge to detect and respond to these threats.
- Develop accessible, trauma-informed reporting channels for digital transnational repression that minimize re-traumatization, provide human-based support, and ensure that incidents are quickly identified, investigated, and addressed.
- Notify human rights defenders when they are identified as targets of digital transnational repression and provide clear and actionable guidance on protection measures, developed in partnership with civil society and exiled and diaspora communities.
- Increase transparency and accountability regarding state-sponsored threats targeting human rights defenders in exile.