

Preventing Radicalisation to Violence Online: Algorithms, Evolving Technologies, and Resilience

Dr Narantuya Ganbat

Commissioner, NHRCM

29 May 2026, OSCE Security Committee, Vienna

Excellencies, distinguished colleagues,

It is an honour to join this important discussion on preventing radicalisation to violence online. I speak from the perspective of the National Human Rights Commission of Mongolia, and from the experience of having recently concluded our two yearlong public inquiry on human rights and digital technologies.

It was a major undertaking for the Commission. The inquiry covered a wide range of issues: personal data protection, artificial intelligence, digital public services, surveillance technologies, health data, children's rights, and social media regulation. One of its central lessons was clear: digital governance is no longer a technical matter alone. It is now a democratic governance issue, a human rights issue, and, increasingly, a security issue.

The theme of today's dialogue captures this reality very well. Radicalisation to violence online is not only about individual pieces of harmful content. It is also about the systems that organise attention, shape behaviour, and reinforce social identity. Engagement-driven design, algorithmic personalisation, cross-platform connectivity, private groups, encrypted spaces, and now generative and immersive technologies can expose users to harmful material, connect them to reinforcing communities, and normalise violence or hatred over time. The OSCE concept note rightly emphasises that these risks are shaped by platform design choices and cannot be addressed through content removal alone.

This is particularly important when we speak about children and young people. Online pathways to harm can begin with curiosity, loneliness, grievance, humour, gaming culture, conspiracy narratives, or the search for belonging. They may be ideological, but they may also be driven by nihilism, obsession with violence, or the pursuit of notoriety. This complexity requires careful, proportionate and evidence-based responses.

Scholarly work helps explain why this matters: moral and emotional language often travels further online, and expressions of outrage can be rewarded through likes, shares and comments. Over time, users may learn that anger receives attention, while platforms may unintentionally magnify perceptions of hostility between groups. Online environments may inflate perceptions of intergroup hostility, making people believe that “the other side” is more extreme or hostile than it actually is.

This matters for prevention. Radicalisation is often not a sudden event. It may develop through repeated exposure, emotional reinforcement, social validation, and gradual desensitisation. Where platform systems reward anger, humiliation, fear, and moral condemnation, they can deepen polarisation and make harmful narratives feel socially justified.

Mongolia’s experience is relevant, but I wish to frame it carefully. Mongolia is not facing a dominant online violent extremism problem. Rather, Mongolia’s experience shows how the broader conditions that weaken democratic resilience can emerge before violence occurs: disinformation, cyber harassment, unlawful disclosure of personal data, online threats, fraud, hate-based narratives, and polarising content.

This matters because Mongolia is already a highly connected society. The Commission’s research on social media regulation notes that, as of 2025, around 83 per cent of Mongolia’s population — approximately 2.9 million people — are internet users, and around 74.4 per cent are registered social media users. Facebook is especially significant: at the beginning of 2025, around 2.5 million users in Mongolia were connected to Facebook, making it one of the most influential public communication spaces in the country. These figures are important not because they suggest that social media is inherently harmful, but because they show the scale at which online design choices now affect public debate, trust, identity and democratic resilience in Mongolia.

As a democracy in Northeast Asia, Mongolia’s commitment to an open, plural and rights-respecting digital space has significance beyond our borders. It is also part of the wider conversation among democratic partners, including within the European Union and OECD, on how to protect trust, democratic resilience and human rights in the digital age.

For Mongolia, therefore, the question is not only how to regulate harmful online content. It is how to protect democratic space while addressing genuine online harms. This is a delicate balance. Overbroad regulation can chill expression, restrict journalism, or politicise content control. But the absence of regulation can leave citizens, especially children and vulnerable groups, exposed to manipulation, abuse and harm.

The Commission's inquiry therefore recommended a comprehensive but rights-based approach to social media regulation. This includes clearer responsibilities for platforms, stronger remedies for users, safeguards against unjustified account restriction or content removal, and an independent mechanism that is separate from the Commission and properly resourced to address harmful content. The recommendations also call for platforms to publish information in Mongolian on content removals, account restrictions and the reasons for moderation decisions affecting Mongolian users.

The inquiry also recommended that standards for identifying and restricting content that incites hatred or divides social groups should be aligned with the Rabat Plan of Action, and implemented without undermining freedom of expression. It further recommended specific programmes for vulnerable groups, older persons, persons with disabilities and rural communities.

Evolving technologies add urgency. Generative AI can produce persuasive synthetic text, images, audio and video at scale. In smaller-language environments, harmful content may spread quickly while detection and moderation capacity remains limited. This makes local language capacity, trusted reporting channels, safety-by-design, privacy-by-design, and cooperation with independent institutions essential.

Resilience is the other side of prevention. It cannot be built by law alone. It requires media and information literacy, independent journalism, fact-checking, mental health support, child protection, and trusted public communication. It also requires democratic trust. A society that trusts its institutions, its media and its neighbours is less vulnerable to manipulation and violent mobilisation.

Finally, no state can address these challenges alone. The platforms are global. The technologies are evolving rapidly. Harmful narratives, extremist subcultures and

synthetic media cross borders easily. The OSCE can add real value by supporting principled dialogue, practical cooperation, and rights-based standards.

Our shared task is to reduce pathways to violence without weakening the open societies we seek to protect. That means safer design, accountable platforms, capable institutions, empowered users and resilient communities.

Thank you.