THE EVOLVING STATE OF ONLINE SAFETY POLICIES

A SNAPSHOT IN TIME – SPRING 2026

Digital policy and online safety regulations have expanded rapidly but not evenly over the past decade. Governments are expected to protect citizens in digital spaces that are transnational, data-intensive, and dominated by private platforms whose business models and practices evolve faster than public policy. These platforms possess greater technical expertise, operational capacity, and access to data than the public authorities tasked with regulating them. This power imbalance makes "duty of care" a challenge to operationalize.

Child online safety is the most developed area of online safety regulation globally, supported by strong consensus, legal definitions, and established global cooperation mechanisms. It is increasingly reinforced through tech product design measures such as age-verification and parental controls.

Online safety for those over 18 is more complicated. Universally, responsibility for prevention of and response to cybercrimes, such as online fraud, versus technology-facilitated offences, such as sharing of non-consensual intimate images, is fragmented

across multiple agencies, creating complex pathways and inconsistent outcomes for those who experience them. People may be required to navigate confusing and overlapping reporting pathways across digital platforms and public institutions. These realities reflect deeper assumptions about whose safety is prioritized, what constitutes harm, and who should bear responsibility for harms that occur online or are facilitated through digital technologies.

Increasingly, both civil society and regulators agree that the efficacy of online safety measures should be evaluated not only by content removal statistics, but by outcomes such as timely responses, meaningful remedies, reduced risk of recidivism, and access to support services for survivors of harm. Recognizing the limits of reactive enforcement, regulators are increasingly turning to upstream approaches, including outcome-based regulation, co-regulation, safety-by-design incentives, and digital citizenship investments (public awareness, education) as prevention strategies.

Global Trend: Data protection has emerged as a foundational concept in digital governance. General Data Protection Regulation (GDPR) style frameworks, which center user rights and the responsibilities of duty-bearers, are becoming more universal. Many of today's data protection laws help address risks such as algorithmic discrimination or surveillance.

Functioning as an upstream counterweight to digital harm, data protection frameworks target the conditions which can enable harm proliferation; for example, by requiring that default settings are set to minimal data collection upon account creation. Because data protection rules apply across platforms and products, they often operate as a cross-cutting safety architecture. This matters in an environment where political will, legal clarity, and enforcement capacity vary significantly depending on the type of harm and the population affected.

Regional Models and the Fast-Moving Policy Space: Strong regional models are also shaping global thinking about online safety governance in ways that increasingly integrate data protection, platform accountability, and AI considerations. The European Union has emerged as a reference point through the Digital Services Act, comprehensive data protection frameworks, and emerging AI regulation, emphasizing systemic risk assessments, transparency obligations, and accountability for platform design and governance choices. National frameworks such as Australia's eSafety model and the United Kingdom's Online Safety Act also reflect efforts to consolidate fragmented responsibilities and impose clearer "duty of care" roles.

New Challenges: At the same time, generative AI, algorithmic content curation, automated moderation, and synthetic media are reshaping the scale, speed, and nature of digital experiences

including through deepfakes, automated harassment, and opaque decision-making systems that challenge existing legal categories, evidentiary standards, and enforcement tools. As new products and risks emerge more quickly than legislation can be drafted, implemented, or evaluated, regulators increasingly lean towards future-proof, principle-based approaches and adaptive governance.

In summary:

1. Regional cooperation can effectively rebalance the dynamic between regulators and platforms and address cross-border harms.

2. Data protection and broader data governance can be used as effective regulatory levers because they shape upstream conditions for risk, accountability, and transparency across platforms and products.

3. Prevention efforts such as safety-by-design and digital citizenship can leverage the strength of commercial tech sector and educational institutions.

4. Setting standards for AI adoption and use and delineating roles and responsibilities in handing AI-related harms, can ensure human agency and rights remain central to digital development.