Netsafe response to the UN Human Rights Office Call for Inputs on Child Rights and Safety in the Digital Environment

https://www.ohchr.org/en/calls-for-input/2026/call-inputs-child-rights-and-safety-digital-environment

Areas of expertise or focus

Netsafe is Aotearoa New Zealand's independent, non-profit online-safety organisation. It began in 1998 as a grassroots community education movement known as the Internet Safety Group and has evolved alongside the digital environment from early education and digital citizenship to frontline harm resolution, platform accountability, artificial intelligence and cross-border cooperation.

Since 2016, Netsafe has served as the Approved Agency under New Zealand's Harmful Digital Communications Act 2015, with reappointments in 2021 and 2026. This statutory role requires Netsafe to receive and assess complaints, investigate harm, use advice, negotiation, mediation and persuasion to resolve complaints, engage domestic and international service providers, and provide online-safety education. Netsafe therefore brings experience across prevention, individual remedy, education, research, technology, platform engagement and policy.

In FY26, Netsafe received 27,482 reports of online harm, including 7,792 complaints engaging the HDCA thresholds. It also supports schools and families, operates New Zealand's INHOPE-connected hotline for suspected child sexual abuse material, conducts youth-centred research, and works alongside partners across the Pacific to strengthen locally led online-safety capability.

Risks to children's rights and safety in the digital environment

What are the most significant risks to children's safety, physical or mental health, and rights in the digital environment in your country or region, and how are they shaped by regional context?

One of the most significant risks reported to Netsafe is online bullying. Young people express to us the negative emotional impact online bullying has on their mental health, relationships and confidence, both at home and at school.

Netsafe's evidence suggests that risks should not be treated as an undifferentiated list of harmful

content. They arise at several levels including interpersonal harms such as bullying, coercion and image-based abuse; commercial harms such as scams and exploitative advertising; exposure to violent, sexual or hateful content; and systemic risks created by product design, data practices, recommender systems and AI-mediated services. These risks overlap and a young person may experience synthetic intimate imagery, peer harassment, algorithmic amplification and an ineffective reporting process as parts of the same incident.

Generative AI illustrates this convergence. Netsafe's 2026 research with 1,202 people aged 13–24 found that nine in ten had used generative AI in the previous six months, 30% had used it for advice or emotional support, one in nine had encountered deepfakes, and only one-third said they would seek adult help if something went wrong. One-third never checked what information AI services collected, adjusted privacy settings or checked whether a tool was age appropriate. These findings support obligations for child-rights impact assessment, age-appropriate defaults, privacy protection, meaningful reporting and safeguards addressing relational dependency, synthetic sexual content and unsafe advice—not simply additional user education.

In Aotearoa New Zealand, these risks are shaped by children's increasing exposure to digital environments from a young age and by the growing role digital technologies play in education and social connection. Online services have become embedded in many aspects of children's daily lives, increasing both the opportunities and risks they encounter online (https://resource.netsafe.org.nz/Netsafe-Submission-to-Parliamentary-Inquiry-into-online-harms-encountered-by-young-NZers.pdf).

How do digital product design and business models drive or mitigate risks to children's rights online, including in AI-mediated products and services?

Product design should be evaluated by foreseeable outcomes, including whether it amplifies risk, creates compulsive or dependent use, facilitates unwanted adult contact, enables synthetic abuse, or makes remedy inaccessible. Businesses should be required to assess these risks before deployment and after material product changes, publish meaningful findings, and demonstrate whether mitigations work in practice.

The way social media platforms and other digital services collect data and profile users can also affect children's rights, especially privacy. Platforms can reduce these risks through default privacy settings, limits on data collection and stronger protections for young users. Responsibility should not sit solely with children and families as online platforms should be expected to identify and reduce foreseeable risks.

How are children's own views being sought and given due weight in the design, governance and regulation of digital environments in your region, by either companies or States?

Netsafe's experience distinguishes between consulting children and sharing influence with them. Its Youth Action Squad provides an ongoing mechanism through which young people shape campaigns, research and service priorities (https://netsafe.org.nz/our-work/youth-action-squad). During FY26, ten secondary-school students helped design a national survey that received more than 800 responses. Netsafe is also developing regional Young Digital Leaders groups in Otago, creating sustained local participation rather than relying only on national one-off consultation.

Participatory research undertaken with Auckland University of Technologies TOROA Centre treated students as contributing researchers. It found that young people often reject generic, moralising advice that underestimates their knowledge or places responsibility for systemic harm on individual behaviour. Effective participation therefore requires safe and accessible engagement, recognition of young people's expertise, transparent feedback about how their views influenced decisions, and continued involvement during design, implementation and evaluation.

How do digital environments, and the measures taken to make them safer, affect different groups of children differently — and how can equality and non-discrimination in access and protection be ensured?

Children do not experience digital environments in the same way. Their age and stage of development shape how they engage online. Their circumstances can also affect the risks they face and the support they can access. Netsafe's research with Save the Children New Zealand found that children's online activities and safety needs change significantly as they grow older ((Children and Youth Online Safety in Aotearoa New Zealand, https://cdn.sanity.io/files/8y8wsx0z/production/cdaa8f37aa4ca375042375f1ab74d5e9dd36403a.pdf).

Equal protection does not mean identical intervention. A child's culture or language may affect whether support is relevant and accessible. Geography can affect access to specialist help. Family circumstances and access to trusted adults can also affect a child's ability to seek support. Safety measures themselves can create inequality. For example, age assurance may depend on formal identification or require additional personal information. Reporting systems may also be difficult to use if they require high digital literacy or are not available in a child's language.

These impacts should be assessed before safety measures are introduced. Children should have alternative ways to access support where documentation, language or location creates a barrier. Global standards should protect children without assuming that every child has the same resources or support available to them. A child's access to support should not depend on whether they live in a large, well-resourced jurisdiction or a smaller one with more limited capacity.

How can children's digital literacy, skills and empowerment be supported, so that they can benefit from the opportunities of the digital environment and participate in it safely?

Children need the skills and confidence to participate in the digital world safely. Digital literacy must strengthen children's agency without shifting responsibility for unsafe systems onto them. It should start early and continue throughout a child's education. Education should help children understand online risks and build critical thinking. It should also help them understand how digital platforms shape their online experiences.

Netsafe's participatory research with Auckland University of Technology's TOROA Centre for Communication Research shows why young people should help shape this education. The research found that young people experience many online risks through peer networks, group chats and shared spaces. It also found that generic or moralising safety messages can fail to reflect the reality of their online lives. Digital resilience is stronger when young people are treated as knowledgeable participants rather than passive audiences. This means digital literacy should be developed with children, not simply delivered to them.

Netsafe puts this approach into practice through its education work. During FY26, Netsafe's Kete education platform reached 1,503 schools, representing 58.6% of New Zealand schools. A total of 10,210 users completed 17,076 microlearning sessions, while 31,717 resources were downloaded. Kete provides free online safety resources for schools and educators (https://education.netsafe.org.nz/).

Reach alone is not enough. Education should also be assessed for whether it builds children's confidence and improves how they respond when harm occurs. It should support help-seeking and strengthen children's ability to participate safely online (https://netsafe.org.nz/schools-and-kura).

Families and educators also need support. Research by Netsafe and Save the Children New Zealand found that children see parents as one of their most important sources of guidance online and want adults to better understand their online experiences (Children and Youth Online Safety in Aotearoa New Zealand, https://cdn.sanity.io/files/8y8wsx0z/production/cdaa8f37aa4ca375042375f1ab74d5e9dd36403a.pdf). Children and Youth Online Safety in Aotearoa New Zealand also examined what children wanted adults and technology companies to understand about their experiences online. Adults therefore need the knowledge and confidence to support children without replacing their agency or treating them as passive recipients of safety advice.

State obligations and regulatory responses

What noteworthy legislative, policy and regulatory measures are being taken in your region to

address risks to children's rights and well-being in the digital environment, and which are proving effective?

New Zealand's Harmful Digital Communications Act 2015 (HDCA) provides an accessible pathway to remedy when online harm occurs. As the Approved Agency, Netsafe can receive and assess complaints, explain options and use advice, negotiation or mediation to seek resolution before court action is considered. In FY26, 92.13% of HDCA complaints were resolved through Netsafe processes without District Court orders. Netsafe also completed 96.67% within 15 business days, despite a 22% increase in complaints meeting the HDCA thresholds.

This model shows that effective remedy does not need to begin with court proceedings. However, the HDCA largely operates after harm has occurred and focuses on harm to identifiable individuals. New Zealand is now considering measures that move further towards prevention. These include child safety risk assessments for platforms and reforms addressing AI-generated intimate imagery (https://www.beehive.govt.nz/release/government-moves-ban-u16s-social-media).

What broader, more systemic interventions are needed in your country/region to address online harms?

New Zealand needs to move further upstream. The HDCA provides an effective pathway once harm has occurred, but online safety cannot rely on responding after the event. Greater investment is needed in prevention. Platforms should assess foreseeable risks before harm occurs and build safety into their products from the outset. This should extend to platform design and recommender systems. Emerging risks from AI also need to be addressed (https://resource.netsafe.org.nz/Netsafe-Submission-to-Parliamentary-Inquiry-into-online-harms-encountered-by-young-NZers.pdf).

Prevention also requires sustained investment in digital safety education. Children need practical skills before harm occurs. Parents and educators need support too. When prevention fails, children need an accessible pathway to support and recovery. Netsafe supports a "No Wrong Door" approach so children and families have a trusted place to start. They should not have to navigate multiple systems or understand which legal pathway applies before asking for help.

Frontline evidence should then feed back into prevention and regulation. This allows policy and education to respond to the harms children are actually experiencing rather than continually reacting after harm has occurred.

How are States addressing their child rights obligations through legislative and regulatory approaches – for example, to ensure that digital tools and products used by children are safe and age-appropriate, and to ensure consistency with the human rights requirements of legality,

necessity, and proportionality where rights are restricted?

New Zealand's proposed Online Safety (Minimum Age and Child Safety Risk Assessment) Bill illustrates two different regulatory approaches. It would restrict children under 16 from holding accounts on covered platforms. It would also require platforms to assess risks to children. The risk assessment approach places responsibility on platforms to identify foreseeable harm before it occurs (https://www.legislation.govt.nz/bill/government/2026/339/en/latest/).

Netsafe supports stronger protections for children, but regulation must remain proportionate. Restrictions on children's access may affect their privacy or ability to participate online. They may also shift activity to other digital spaces. Measures should therefore be supported by evidence and assessed for unintended consequences. General Comment No. 25 requires restrictions on children's rights to be lawful, necessary and proportionate (https://www.ohchr.org/en/documents/general-comments-and-recommendations/general-comment-no-25-2021-childrens-rights-relation).

Netsafe's experience supports putting greater preventative responsibility on the services that create and shape digital environments. Age restrictions may form part of the response, but they should not replace safer design or meaningful platform accountability (https://resource.netsafe.org.nz/24-August-2026-Netsafe-statement-on-Governments-Social-Media-Age-Restriction-Bill.pdf).

How are States regulating age-inappropriate content, algorithmic amplification, and product liability in ways that protect children while respecting the full range of their rights?

New Zealand's existing framework is largely reactive. The HDCA provides an important remedy after a harmful digital communication has affected an identifiable person. This remains essential, but it does not directly address systemic risks created by platform design.

Regulation should also focus on preventing foreseeable harm before it occurs. This means placing greater responsibility on platforms for the environments they create. Recommender systems are particularly important because platforms influence not only whether content is available, but how widely and repeatedly it is shown. The European Union's Digital Services Act provides one model by requiring very large platforms to assess and mitigate systemic risks arising from their services (https://digital-strategy.ec.europa.eu/en/policies/digital-services-act-package).

Preventative regulation still needs to protect children's access to information and freedom of expression. Greater transparency is therefore essential. Regulators and researchers need to understand how recommender systems affect what children see and whether safeguards are working. This allows regulation to address risks at a systemic level rather than relying solely on removing content after harm has occurred.

Have States adopted or proposed measures that would disproportionately impact human rights, such as online freedom of expression or privacy, with the stated aim – however well intentioned – of protecting children in the digital environment? If so, please provide specific examples.

New Zealand's proposed Online Safety (Minimum Age and Child Safety Risk Assessment) Bill combines two different approaches. It would restrict children under 16 from holding accounts on covered platforms. It would also require platforms to undertake child safety risk assessments (https://www.legislation.govt.nz/bill/government/2026/339/en/latest/). The risk assessment provisions shift responsibility towards platforms and prevention. The age restriction places the restriction directly on children, so it requires a different level of rights scrutiny.

Age assurance also raises privacy concerns. New Zealand's Privacy Commissioner has warned that age-checking systems can create risks through data breaches and inaccurate age estimates. Personal information may also be reused to infer age (https://www.privacy.org.nz/tuhono-connect/statements-media-releases/privacy-commissioner-statement-on-under-16s-social-media-ban/).

General Comment No. 25 requires restrictions on children's rights to be lawful, necessary and proportionate. It also calls for child rights impact assessments of legislation affecting children (https://www.ohchr.org/en/documents/general-comments-and-recommendations/general-comment-no-25-2021-childrens-rights-relation). A platform risk assessment is not the same as assessing the impact of the restriction itself. That assessment should consider whether the restriction will reduce harm. It should also consider whether the same protection could be achieved through less restrictive measures.

Netsafe supports moving further towards prevention and safety by design. However, preventative regulation should put greater responsibility on the systems creating and shaping risk. It should not unnecessarily restrict children's rights in the process.

Responsibilities of business enterprises

What concrete measures have companies (or your company) implemented to meet their responsibility to respect children's rights under the UN Guiding Principles on Business and Human Rights and the Children's Rights and Business Principles, including with regard to AI products and services?

From a civil-society perspective, a key gap is the distance between companies' global policies and what children experience in practice at a country level. Companies should undertake child-rights impact assessments before launching high-risk products or major features. They should repeat

those assessments after material changes, particularly for AI-enabled products and recommender systems.

Companies should also report on outcomes at a country level. They should show whether measures introduced to reduce harm are actually working and allow independent scrutiny of those findings. Without this information, regulators and civil society cannot readily assess whether global commitments are protecting children in practice.

What accountability and remedy mechanisms provided by companies are proving most promising, and where are the greatest gaps, including across companies of different sizes and those operating across multiple countries and regions?

Netsafe's trusted escalation experience shows the value of specialist civil-society intermediaries. Netsafe has trusted relationships with a number of major platforms that enable serious reports to be escalated directly for review when standard reporting channels have not resolved the harm. This can help children and families who may struggle to navigate platform systems alone. However, escalation arrangements differ between companies and can depend on informal relationships. A child's access to remedy should not depend on geography, market size or individual platform contacts.

Netsafe led the development and initial establishment of the Aotearoa New Zealand Code of Practice for Online Safety and Harms in 2022. The Code established common commitments and public reporting. It also provided a complaints process and external scrutiny. This demonstrated the value of a shared voluntary framework, but also its limits. Voluntary commitments should not replace enforceable duties, independently verifiable information or effective individual remedy.

Digital safety, privacy, and the rule of law

How can children's protection from sexual exploitation and abuse be pursued in ways that are consistent with guaranteeing the right to privacy and access to secure communications for all?

Children's protection from sexual exploitation and abuse and their rights to privacy and secure communications should not be treated as competing objectives. Netsafe's INHOPE-connected work demonstrates the value of specialist reporting and trained assessment. It also enables referrals so that reports can reach the agency best placed to respond. In FY26, Netsafe received and processed 1,047 reports of suspected or confirmed child sexual abuse material through public and international reporting channels (INHOPE, https://inhope.org/EN/our-story).

Privacy-preserving tools also show that action can be taken without requiring access to the underlying image. Take It Down, developed by the National Center for Missing and Exploited Children, creates a digital fingerprint of an intimate image on the user's device. The image itself is

not uploaded. Participating platforms can then use the fingerprint to help detect and remove copies (https://takeitdown.ncmec.org/).

Any interference with private or secure communications should still be lawful, necessary and proportionate. It should also be subject to appropriate safeguards and oversight. Effective child protection should target the harm while preserving privacy wherever possible.

How can children be protected from abuse without over-criminalizing their own conduct, consistent with due process protections and the principle of evolving capacities?

Children can be protected from abuse without treating every harmful interaction between young people as a criminal matter. Responses should distinguish exploitation, coercion and predatory conduct from developmentally typical or consensual peer behaviour. Where appropriate, children should have access to independent advice and education before a matter enters the criminal justice system. Restorative intervention can also provide another pathway.

New Zealand's Harmful Digital Communications Act 2015 (HDCA) provides a practical example. As the Approved Agency, Netsafe can use advice, negotiation and mediation to address harm. This can protect the affected child while helping another young person understand the impact of their conduct and stop the harmful behaviour (https://netsafe.org.nz/our-work/helpline-services/the-harmful-digital-communications-act).

Serious offending may still require a criminal justice response. However, criminal processes should not be the automatic first response to every peer-to-peer incident. The response should reflect the child's age and evolving capacities. It should also protect the best interests and due process rights of the children involved.

International cooperation and options for a framework

What are the most critical gaps in addressing children's rights in the digital environment through international cooperation, and what forms of cooperation are most urgently needed?

Netsafe's experience working alongside locally led partners across the Pacific shows that international standards alone are not enough. Countries need the local capability to put them into practice. Netsafe has supported locally led initiatives in the Cook Islands, Fiji, Tonga, Vanuatu and Papua New Guinea. This work has reinforced the importance of building trusted local services and capability rather than importing a single model from larger jurisdictions.

International cooperation should therefore help smaller and lower-resourced jurisdictions build sustainable local capability. Global platforms also need escalation pathways that work

consistently across markets.

What options for a framework, including at the global level, should be considered to advance children's rights and safety in the digital environment? How might they relate to existing instruments and processes?

There are established international children's rights instruments, but there is no sufficiently consistent regional or global agreement on the minimum standards that companies should meet to protect children online. This leaves individual States to regulate global technology companies with vastly different levels of regulatory capacity and influence. For smaller countries, there is a significant power imbalance. They may be seeking changes from companies operating at a global scale without the regulatory or market leverage of larger jurisdictions.

A global framework should therefore build on the Convention on the Rights of the Child, General Comment No. 25 and the UN Guiding Principles rather than create competing rights. Greater regional and global cooperation could translate those existing principles into agreed minimum regulatory standards. These could cover child-rights-by-design, risk assessment, accessible reporting and remedy, transparency, and effective escalation pathways.

Any framework should also recognise that implementation will differ between countries. Netsafe's experience working alongside locally led partners across the Pacific shows the importance of building local capability rather than simply exporting a regulatory model developed in a larger jurisdiction. Smaller and lower-resourced countries need access to technical expertise, sustainable funding and regional cooperation. This would help address the imbalance they face when engaging with large global platforms.

Global companies should provide substantively equivalent safety and remedy regardless of the size of a market. A child should be able to access a trusted reporting pathway and receive age-appropriate information. They should also be able to escalate serious concerns and seek review of platform decisions. States should be able to assess whether these protections work in practice, not simply whether legislation or company policies exist.

Greater regional or global agreement on minimum standards could provide smaller States with more collective leverage and give global companies clearer and more consistent expectations across markets. The value of a global framework would therefore be to turn existing rights into more consistent minimum protections and practical accountability, while still allowing implementation to reflect local needs and capacity.