Strengthening Digital Safety Education Through Cross-Sector Collaboration

Weihan Angela Ng
angeng@ethz.ch
Eidgenössische Technische Hochschule Zürich
Zürich, Switzerland

Abstract

The growing prevalence of digital threats, including cyberbullying, misinformation, and privacy breaches, underscores the urgent need for structured online trust and safety education. While various online safety curricula exist, gaps in accessibility, empirical validation, and cross-cultural adaptability persist. This position paper advocates for a collaborative framework involving youths, governments, and academia to enhance online trust and safety education. By actively engaging young people in the development of educational materials, integrating safety curricula into formal education, and ensuring academic research informs policy and practice, a more effective and inclusive digital literacy framework can be established. The paper outlines key challenges, evaluates existing educational approaches, and proposes a structured model that leverages cross-sector collaboration to create a resilient and adaptable online safety education ecosystem.

Keywords

Online Safety Education, Digital Trust and Safety, Cross-Sector Collaboration, Youth

1 Introduction

In an era where digital interactions have become deeply embedded in daily life, online trust and safety education is more critical than ever. The rapid evolution of digital threats, ranging from cyberbullying to misinformation and privacy breaches, necessitates a structured approach to equipping individuals—especially youths—with the knowledge and skills to navigate the digital landscape safely. Existing online safety curricula provide valuable resources for different stakeholders, but gaps remain in ensuring holistic, structured, and universally applicable education. Furthermore, current literature on online safety education underscores the need for greater empirical validation, youth involvement, and cross-cultural adaptability. This paper argues for the establishment of a more structured online trust and safety education framework, facilitated through partnerships between youths, governments, and academia. This paper begins by examining the roles of key stakeholders, followed by a critical review of existing educational efforts and research literature. It then presents a position argument and structured model to guide future development in this field.

2 Background

Effective online safety for children and youth requires multi-stakeholder collaboration involving governments, educators, digital firms, and international organizations to develop comprehensive, regulatory, and educational frameworks, as emphasized by Jang and Ko [10].

- **Public Sector**: Governments set policies and regulations that define the legal framework for online safety. They act as mediators, balancing the restrictiveness of policies with the rights and interests of companies and digital users [2]. Governments also provide public funding for initiatives that protect young users and enforce laws against threats in the cyber space [8].
- **Industry**: Tech companies develop and implement digital platforms, control user data, and influence digital experiences. They also have significant financial resources to support online safety initiatives and fund research and NGOs working in this space [7].
- **Academia**: Researchers contribute to knowledge discovery, especially in emerging digital safety concerns [1]. Universities play a crucial role in studying online risks, developing new technologies for safer interactions, and informing policies and industry practices [12].
- **NGOs**: NGOs operate at the grassroots level, serving as a bridge between governments, industry, and academia. They often provide community-based interventions, direct education, and advocacy for youth-friendly policies [11]. Youth-led NGOs, in particular, offer relatable and flexible approaches to digital safety.

These distinct yet interconnected roles highlight the necessity of coordinated collaboration to address the evolving digital risks faced by young people.

3 Literature Review

Ensuring a safe and trustworthy online environment requires a well-structured and comprehensive approach to online safety education. This section explores the current landscape of online safety curricula and education literature, categorising existing initiatives based on their intended audiences and focus areas. By reviewing these frameworks, we identify key strengths, limitations, and gaps in accessibility, implementation, and cultural relevance. Understanding these aspects is crucial in developing a more integrated and effective approach to digital safety education that meets the evolving needs of diverse stakeholders, including children, educators, policymakers, and aspiring professionals.

3.1 Review of Existing Online Safety Curricula

Online safety curricula can be classified into three primary categories: (1) programmes for children and educators, (2) frameworks for educational institutions and policymakers, and (3) resources for aspiring professionals in the trust and safety domain. While these categories provide a structured approach to identifying target audiences, they also reveal critical gaps in accessibility, applicability across different regions, and the integration of emerging digital threats.

The first category, targeting children and educators, includes initiatives such as Google's *Be Internet Awesome* [9], NetSmartz [13], and the eSafety Commissioner's Classroom Resources [5]. These programmes focus on fundamental digital safety concepts such as cyberbullying, privacy, and digital citizenship. However, many of these resources are developed within specific cultural contexts, limiting their universal applicability. Additionally, access to digital materials remains a barrier in regions with limited technological infrastructure.

The second category consists of curricula developed for educational institutions and policymakers, such as the UK Government's *Teaching Online Safety in Schools* [16] and the eSafety Commissioner's Best Practice Framework [4]. These frameworks integrate online safety into national education systems and provide professional development opportunities for educators. However, their effectiveness depends on local implementation, and they require adaptation to be relevant across diverse cultural and educational settings.

The third category comprises curricula designed for general audiences, including aspiring trust and safety professionals. Notable among these is the Trust & Safety Professional Association's (TSPA) curriculum [15], which provides in-depth insights into policy enforcement, transparency, and platform governance. Although accessible and comprehensive, some of its content is highly technical, making it less approachable for individuals without prior exposure to digital governance.

These classifications highlight the need for a more integrated approach that bridges these categories, ensuring that online safety education is accessible, inclusive, and relevant across cultural and professional contexts.

3.2 Review of Online Safety Education Literature

The literature on online safety education can be grouped into three key areas: preventive education for adolescents and youth, technological interventions for online safety, and cybersecurity education within school curricula. While these studies contribute significantly to the field, they also expose several critical gaps that must be addressed.

Preventive education research focuses on the risks faced by adolescents, including cyberbullying, privacy breaches, and online fraud [6, 17]. While these studies provide strong theoretical frameworks, they often lack empirical validation and cross-cultural applicability. Furthermore, emerging digital threats such as AI-generated misinformation are underexplored.

The second category of literature, focusing on technological interventions, explores the role of AI, content filters, and parental controls in enhancing online safety [12]. A key strength of the research by Mwijage and Ghosh is its emphasis on youth involvement in designing online safety technologies. However, challenges in scaling these solutions, particularly in under-resourced communities, remain a significant barrier to widespread adoption.

The third category, cybersecurity education within school curricula, highlights the importance of structured frameworks and teacher training (Iradat, 2024; Walsh et al., 2022). While these studies advocate for collaboration between educators, policymakers, and cybersecurity professionals, they often overlook the direct engagement of young people in shaping these initiatives. Additionally, existing research remains geographically limited, raising concerns about the global applicability of proposed frameworks.

While the literature offers valuable insights, real-world case studies provide concrete examples of how collaborative efforts can translate research into practice. The following initiatives illustrate how youth organisations, academia, and industry have worked together to advance online trust and safety.

3.3 Case Studies

This section highlights real-world examples of how diverse stakeholders, including youth organisations, academia, and industry, collaborate to advance digital safety through targeted initiatives and research-driven solutions.

3.3.1 Cyber Youth Singapore (CYS) and the Surf Safe Campaign. Cyber Youth Singapore (CYS) exemplifies the impact of NGOs in addressing online safety. As a youth-led national movement, CYS empowers young Singaporeans with digital skills and knowledge to navigate the digital world. The *Surf Safe Campaign*, a two-year outreach and education initiative, assisted secondary school students in navigating the rapid digitalisation accelerated by COVID-19. Covering topics such as cyberbullying, media literacy, and data protection, the campaign reached 58 schools and nearly 40,000 students by the end of 2023 [3]. CYS's effectiveness lies in its **collaborative approach**. It partners with both the public and private sectors to design and implement targeted initiatives, ensuring that programmes are directly relevant to young users. Through partnerships with government agencies and private industry players, CYS has been able to expand its reach and impact.

3.3.2 TUM Think Tank's Frontiers in Digital Child Safety. A strong example of academia-industry collaboration is the Frontiers in Digital Child Safety initiative led by the Technical University of Munich (TUM) in partnership with Apple. This project, funded by Apple with a commitment of $500,000, brings together scholars from TUM, Harvard University, and the University of Zurich to tackle emerging challenges in digital child safety [14]. This initiative highlights the importance of academia in researching cutting-edge safety solutions and the industry's role in providing resources and technical expertise. By fostering a global research community, the project explores innovative technological, educational, and policy-based solutions for digital child protection.

4 Research Motivation

Despite existing efforts in online safety education, there remain significant gaps in accessibility, implementation, and empirical validation. Current curricula often fail to adequately engage youths as active participants in shaping online safety strategies. Additionally, there is limited integration between educational institutions, policymakers, and industry stakeholders in the development of these frameworks.

This paper seeks to address the following research question: ***How can a structured, collaborative approach involving youths, governments, and academia enhance the effectiveness and inclusivity of online trust and safety education?*** This question is crucial because the digital landscape continues to evolve, presenting new risks that traditional online safety curricula may not adequately address. By identifying a more comprehensive and adaptable educational framework, this research aims to contribute to a safer and more resilient digital environment.

5 Argument

This section presents a position argument advocating for a comprehensive, multi-stakeholder approach to online safety education. It emphasises the importance of youth involvement, policy integration, and evidence-based research to ensure educational strategies are effective, inclusive, and sustainable.

5.1 Youth-Centric Development

A meaningful online safety framework must begin with active youth involvement in both design and evaluation processes. Their perspectives are crucial in ensuring that educational materials remain relevant and engaging. This includes incorporating participatory design approaches where youths contribute to content development, providing feedback on digital safety tools, and co-developing educational materials that reflect real-world online risks and experiences. Youth-driven initiatives, such as peer-led online safety campaigns, can also enhance engagement and relatability.

5.2 Government and Policy Integration

Governments have a critical role in mainstreaming digital safety into formal education and policy ecosystems. This requires cross-sector collaboration to develop standardised yet adaptable frameworks that address diverse cultural contexts. Policymakers must ensure that online safety education is mandated in schools at all levels, supported by adequate teacher training and continuous assessment of its effectiveness. Additionally, governments should collaborate with technology companies to develop responsible digital policies that align with education strategies, creating an ecosystem that reinforces trust and safety principles.

5.3 Academic Research & Empirical Validation

Academic research provides the backbone for evidence-based and scalable safety interventions. This includes longitudinal studies assessing the impact of online safety programmes on youth behaviour, research on the effectiveness of different pedagogical approaches, and comparative studies examining how various countries implement digital safety education. Academia should also work closely with governments and NGOs to translate research findings into actionable policies and scalable educational models.

6 Proposed Framework

A sustainable and effective online trust and safety education framework requires collaboration between three key stakeholders: youths, governments, and academia.

6.1 Understanding Youths' Online Behaviour

A foundational step in improving online safety education is gaining a much stronger understanding of the wide spectrum of youth behaviours in the online space. This requires academic research conducted in collaboration with schools on a wide scale. Research should explore youths' digital habits, their perceptions of risks, and the coping mechanisms they employ when encountering online threats. By capturing diverse experiences across different demographic groups, such research can provide valuable insights for policymakers and educators.

6.2 Research-Informed Gap Analysis

Building on insights from youth behavioural studies, it is essential to identify gaps that persist within current online safety curricula. Findings from academic studies should be leveraged to identify gaps in existing curricula and resources available for online trust and safety education. Once these gaps are identified, governments should collaborate with education ministries to develop curricula tailored to specific environments and age groups. This ensures that online safety education remains relevant, engaging, and appropriate for learners at different developmental stages.

6.3 Governmental Role in Enforcing Guardrails

In parallel with curricular reforms, there must be strong institutional safeguards and policies that protect young users on digital platforms. Beyond education, governments play a crucial role in ensuring that the necessary digital guardrails are in place to protect youths online. Regulation and enforcement must work hand-in-hand with education by setting clear policies on online platform accountability, youth privacy protections, and digital content moderation. Governments should also ensure that these guardrails evolve alongside emerging digital threats, such as AI-generated misinformation and sophisticated cyber threats.

6.4 Youths as Co-Creators of the Curricula

Rather than being passive recipients of online safety education, young people should be engaged as co-creators. Initiatives such as youth-led digital literacy campaigns and participatory design workshops can enhance engagement and ensure that educational materials reflect the realities of young people's online experiences.

6.5 Institutions' Role in Research Validation

Academic institutions play a crucial role in validating online safety education initiatives. Universities should collaborate with governments and youth organisations to conduct longitudinal studies assessing the effectiveness of different online safety education models. This will help refine existing strategies and ensure that interventions are grounded in empirical evidence.

7 Implications & Future Directions

The proposed framework has wide-reaching implications across policy, education, industry practices, and academic research.

7.1 Impact on Policy and Education

A structured and collaborative approach to online trust and safety education has the potential to influence national and international policy frameworks. Governments can leverage the proposed model to integrate online safety education into formal curricula across primary, secondary, and tertiary levels. Policymakers should work towards standardising online safety education guidelines while allowing flexibility for contextual adaptation across different regions and cultural settings. Furthermore, regulatory frameworks should be designed to ensure that digital platforms prioritise safety through transparency, responsible data practices, and robust content moderation strategies.

In addition, education ministries should collaborate with researchers to continuously refine and update digital literacy programmes, ensuring they remain relevant in response to emerging online threats. By embedding online safety into educational policy, governments can foster a generation of digital citizens who are better equipped to navigate the complexities of the online world.

7.2 Contributions to Industry and Technology Development

The digital safety landscape is continuously evolving, requiring proactive industry engagement. There is need for technology companies to play a more active role in online safety education, both in terms of platform design and corporate social responsibility initiatives. Companies should invest in developing tools that promote safer online interactions while collaborating with academia and governments to ensure their platforms are accessible, inclusive, and aligned with evidence-based educational strategies, striking a balance between business needs and responsibility.

7.3 Limitations of Study

While this paper proposes a structured and collaborative framework for online trust and safety education, it does not fully address the underlying power dynamics that may complicate cross-sector collaboration. In particular, structural inequalities between large technology firms and smaller youth-led NGOs may influence decision-making processes, resource distribution, and agenda-setting within collaborative efforts. These imbalances can hinder the co-creation of educational frameworks that genuinely reflect diverse youth perspectives. Future work should examine these power asymmetries more closely and explore mechanisms—such as participatory governance models or youth advisory councils—that can help mitigate unequal stakeholder influence and ensure more equitable collaboration.

7.4 Directions for Future Research

While this paper provides a proposed position for a more structured approach to online trust and safety education, several key areas remain open for further exploration:

(1) **Cultural and Regional Adaptations**. Given the diversity of digital experiences worldwide, further investigation is needed into how online safety education frameworks can be effectively adapted to different cultural, linguistic, and socio-economic contexts. Comparative studies examining international best practices could help refine universally applicable principles while maintaining local relevance.
(2) **The Role of Emerging Technologies**. With the rise of artificial intelligence, augmented reality, and decentralised digital ecosystems, research should explore how these technologies can be harnessed to enhance online trust and safety education. AI-driven personalised learning platforms, gamified safety modules, and immersive simulations could offer innovative solutions for more effective digital literacy training.

By addressing these research gaps, the field of online trust and safety education can continue to evolve, ensuring that individuals across all demographics are empowered with the necessary skills to navigate the digital landscape safely. Ultimately, fostering cross-sector collaboration between governments, academia, and industry will be critical in building a resilient and inclusive digital future.

8 Conclusion

The increasing complexity of digital threats necessitates a structured and collaborative approach to online trust and safety education. While existing curricula and research provide valuable insights, significant gaps remain in accessibility, empirical validation, and youth engagement. By fostering partnerships between youths, governments, and academia, a more effective, inclusive, and adaptable online safety education framework can be developed. Such an initiative will not only enhance digital resilience among young people but also contribute to a safer and more informed online environment for all.

Use of LLMs

The GPT-4o model had been utilised to proof-read and to summarise some sections of this work.

Acknowledgments

I would like to thank Cyber Youth Singapore for the provision of information relating to the Surf Safe Campaign and TUM Think Tank for the informal sharing of their programmes in the summer of 2024, which led me to the deeper reading of the research group's initiatives. Much thanks to the Centre for Advanced Technologies in Online Safety (CATOS), Singapore, for inspiring this paper.

References

[1] Rosanna Bellini, Emily Tseng, Noel Warford, Alaa Daffalla, Tara Matthews, Sunny Consolvo, Jill Palzkill Woelfer, Patrick Gage Kelley, Michelle L Mazurek, Dana Cuomo, et al. 2024. Sok: Safer digital-safety research involving at-risk users. In *2024 IEEE Symposium on Security and Privacy (SP)*. IEEE, 635–654.
[2] Moeller Ch. 2013. Safety in online media – freedom of the media; safety of media actors and media education. *Vestnik Chelyabinskogo Gosudarstvennogo Universiteta* 22 (313) (2013), 37–39. Note: Journal title and publisher translated from Russian.
[3] Cyber Youth Singapore. 2025. Cyber Youth Singapore Official Website. https://www.cyberyouth.sg
[4] eSafety Commissioner. n.d.. Best Practice Framework for Educators. https://www.esafety.gov.au/educators/best-practice-framework

[5] eSafety Commissioner. n.d.. Classroom Resources for Online Safety. https://www.esafety.gov.au/educators/classroom-resources
[6] David Finkelhor, Kerryann Walsh, Lisa Jones, Kimberly Mitchell, and Anne Collier. 2021. Youth internet safety education: Aligning programs with the evidence base. *Trauma, violence, & abuse* 22, 5 (2021), 1233–1247.
[7] Jake Goldenfein and Monique Mann. 2023. Tech money in civil society: Whose interests do digital rights organisations represent? *Cultural Studies* 37, 1 (2023), 88–122.
[8] Seymour E Goodman. 2014. Building the nation's cyber security workforce: Contributions from the CAE colleges and universities. *ACM Transactions on Management Information Systems (TMIS)* 5, 2 (2014), 1–9.
[9] Google. n.d.. Be Internet Awesome. https://beinternetawesome.withgoogle.com/en_us
[10] Yujin Jang and Bomin Ko. 2023. Online safety for children and youth under the 4Cs framework—A focus on digital policies in Australia, Canada, and the UK. *Children* 10, 8 (2023), 1415.
[11] Larry Minear. 1987. The other missions of NGOs: education and advocacy. *World Development* 15 (1987), 201–211.
[12] Tajuddin Mwijage and Arup Ghosh. 2024. The Role of Technology in Enhancing Adolescent Online Safety: Current Trends and Future Directions. *SoutheastCon 2024* (2024), 1610–1614.
[13] National Center for Missing and Exploited Children. n.d.. NetSmartz. https://www.missingkids.org/netsmartz/home
[14] Technical University of Munich. 2025. Frontiers in Digital Child Safety. https://tumthinktank.de/project/frontiers-in-digital-child-safety/.
[15] Trust and Safety Professional Association (TSPA). n.d.. Trust and Safety Curriculum. https://www.tspa.org/curriculum/ts-curriculum/
[16] UK Government. n.d.. Teaching Online Safety in Schools. https://www.gov.uk/government/publications/teaching-online-safety-in-schools/teaching-online-safety-in-schools
[17] Simona-Nicoleta Voicu and Ioan Crăciun. 2023. Preventive online safety education for teenagers. *International Journal of Legal and Social Order* 3, 1 (2023).