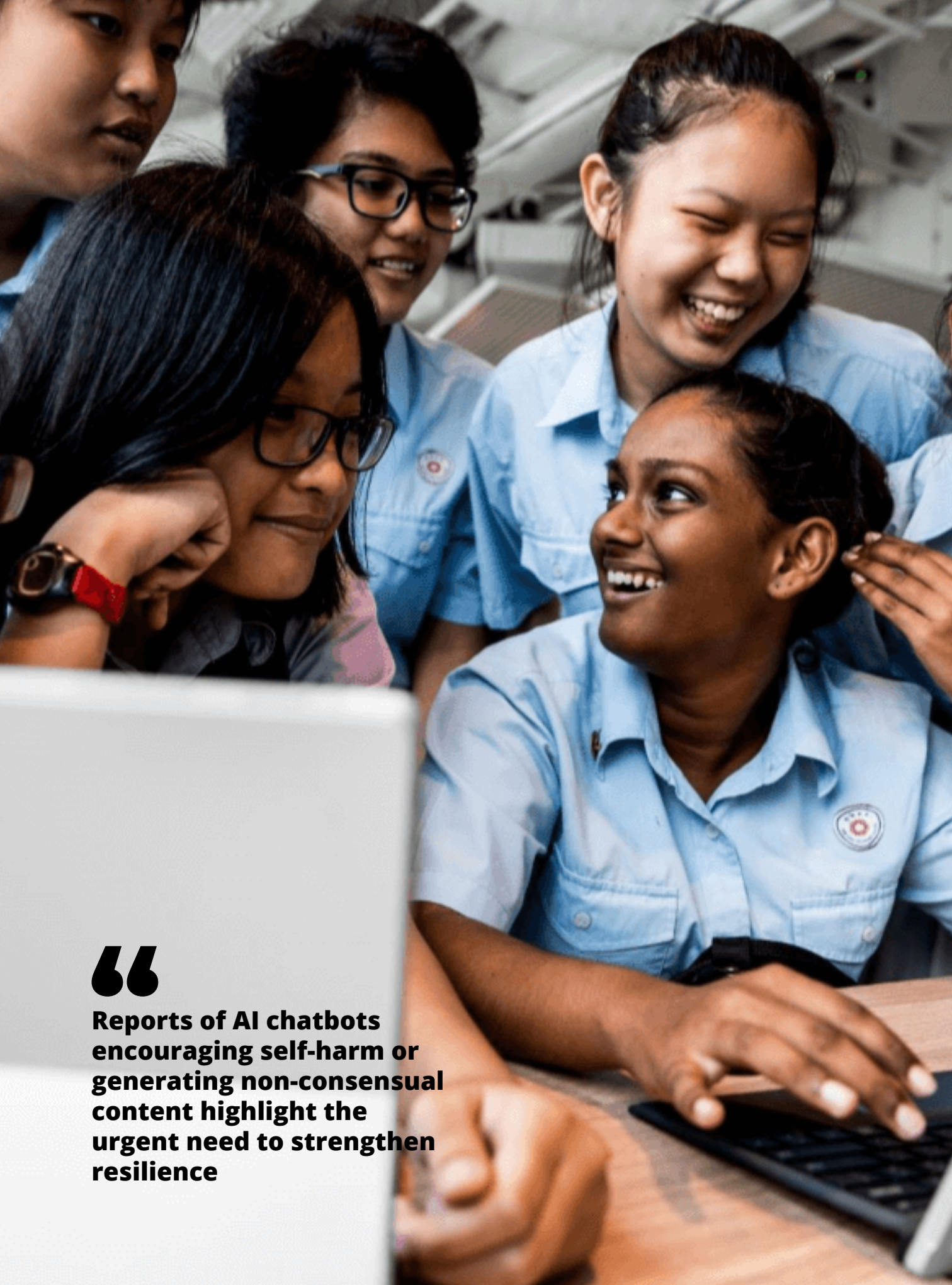

Reports of AI chatbots encouraging self-harm or generating non-consensual content highlight the urgent need to strengthen resilience

Digital guardians of the AI era: building youth cybersecurity resilience

by **Ziarla Mae Malabanan**

Artificial intelligence (AI) is reshaping how young people communicate, learn and express themselves online. At the same time, it exposes them to unprecedented risks – from algorithmic bias and privacy violations to harassment, exploitation and AI-driven scams. These threats are not hypothetical: reports of AI chatbots encouraging self-harm or generating non-consensual content highlight the urgent need to strengthen resilience.

This article argues that youth-led cybersecurity capacity building is an effective defence against AI-enabled threats.

> **"By combining technical training with policy literacy, ethical awareness and hands-on experience, young people can become proactive digital guardians able to detect, respond to and prevent cyber harm."**

Drawing on case studies from Southeast Asia, Europe and Africa, the article demonstrates how inclusive programmes, particularly those targeting underrepresented or vulnerable communities, empower youth to navigate digital risks confidently and responsibly. Investing in youth resilience not only strengthens individual safety but also creates a new generation capable of strengthening digital ecosystems.

Introduction

Young people are increasingly at risk from artificial intelligence, the United Nations Children's Fund[1] and the United Nations Educational, Scientific and Cultural Organization[2] warn. The risks include algorithmic bias, privacy violations and exposure to manipulated content. Tools like chatbots, AI-driven teaching applications, and content generators are transforming learning, social connection, and self-expression. While these tools can entertain or educate, they are often used without

[1] United Nations Children's Fund (UNICEF), Generative AI: Risks and Opportunities for Children (Florence: UNICEF Innocenti, 2023)

[2] UNESCO, Ethics of Artificial Intelligence – The Recommendation.

supervision, leaving youth vulnerable to data collection, manipulation, and exploitation.

An investigation by Triple J Hack, an Australian current affairs programme, revealed allegations of AI chatbots sexually harassing and encouraging self-harm of young people, as well as allegations of ChatGPT reinforcing delusions that led an individual to hospitalization. Young people connect with chatbots or AI companions (digital characters powered by AI) as a form of social outreach or emotional support, said a youth counsellor interviewed for the programme. In some cases, they were told they had "no chance of making friends", that they were "ugly or disgusting", or that they should "kill themselves".[3] Other similar incidents – including harassment during language learning – show these risks are neither isolated nor hypothetical.

"Beyond chatbots and similar tools, cybercriminals leverage AI to automate data collection, profiling, cyberbullying, and the creation of deepfake child sexual abuse material."

Even digitally fluent youth are vulnerable: LinkedIn-based employment scams, for example, steal sensitive personal or financial information by targeting college students and early-career professionals with convincing but fake job offers. Deepfake technology, once niche, is now widely accessible, as shown in a 2020 investigation, which enabled users to generate over 100,000 non-consensual images of minors generated by bots on Telegram.[4]

[3] A. McLennan, AI Chatbots Accused of Encouraging Teen Suicide as Experts Sound Alarm, ABC News, 12 August 2025.

[4] R. Karim and M. Seera, Digital Child Abuse: Deepfakes and the Rising Danger of AI-Generated Exploitation, Monash Lens, 25 February 2025.

These examples, showing that AI increases the online risks youth face, also highlight the importance of cybersecurity resilience. One of the most effective long-term solutions is youth-led capacity building in cybersecurity. By participating in or leading these programmes, young people develop digital literacy, preparedness, and resilience, enabling them to navigate AI-enabled threats safely. This type of resilience not only shields individuals but also strengthens the overall integrity of digital ecosystems.

Framing cybersecurity education and AI awareness around youth – defined as persons aged 15–24[5] – empowers digital natives to become proactive, informed, and resilient participants in the online world.

Youth as digital guardians

Because young people already live in digital ecosystems, they learn and adopt new technologies quickly. However, tech fluency alone is not enough. They also need training in policy literacy, including how laws, regulations, and data governance operate in cyberspace. This objective can be achieved by participation in campaigns, workshops, and opportunities to engage directly with policymakers.

> **“With appropriate training, young people can turn their technological fluency into a defensive advantage, spotting AI-driven manipulation or suspicious patterns more naturally than less tech-immersed groups.”**

[5] United Nations, Frequently Asked Questions United Nations for Youth, United Nations.

While early exposure to cybersecurity creates a culture of safe behaviour by default, it also equips youth to shape future policies, products, and education. Early engagement in cybersecurity and AI ethics through a bottom-up approach helps reinforce safer, more responsible digital environments.

Case studies in youth-led resilience

Across the globe, non-profit groups are investing in the next generation of cybersecurity leaders.

In 2022, the Association for Southeast Asian Nations (ASEAN), in partnership with the ASEAN Foundation and Microsoft Asia Pacific, launched a regional programme to raise awareness, expand knowledge, and strengthen cybersecurity skills among young people. What distinguished this programme was its focus not on elite universities, but on underprivileged and often overlooked communities, including unemployed or underemployed youth and female-only groups breaking into information technology.

> **"Seven ASEAN countries, together with the FutureReadyASEAN.org initiative, delivered free, localized training reaching tens of thousands, proving cybersecurity can empower youth."**

A representative from CyberGuardiansPH, one of the partners of the programme, said: "The ASEAN Cybersecurity Skilling Programme has allowed the youth to appreciate cybersecurity as a tool to protect themselves online. We have participants with no IT or STEM backgrounds who realised that cybersecurity is not as intimidating as it sounds".[6] Beyond participants, even some trainers have since pursued careers in cybersecurity, a trend echoed by participants across all seven countries.

In Europe, the CyberPeace Builders, run by the Geneva-based CyberPeace Institute, mobilize over 1,300 volunteer professionals to safeguard non-governmental organizations. Since 2021, they have completed over 1,150 missions. In one example, participants helped a humanitarian organization in Eastern Europe recover from a ransomware attack, regaining access to critical donor databases and restoring operations within 48 hours.

South Africa's CyberM8 initiative trained thousands of learners, helped small information and communications technology businesses adopt advanced digital skills, and sparked a nationwide culture of cyber awareness. For example, this initiative guided a local tech start-up to implement robust cybersecurity protocols to protect against a phishing attack on their client database. In another example, CyberM8 hosted workshops that reached over 500 students in a single week. These are two examples of youth-led initiatives that have contributed to strengthening cybersecurity.

From Southeast Asia to Europe and Africa, these initiatives prove that when youth are given inclusive, hands-on training and real-world opportunities, they become the first line of defence against the world's pressing digital threats. At their core, these initiatives follow a clear strategy: They combine education, practical support, and community-focused empowerment in ways that can scale and adapt to local cultures.

[6] Z. Malabanan, "A safe cyberspace for the ASEAN community,". ASEAN Magazine, 2022,

Pathways to youth-led cyber resilience

As AI transforms daily life, technical fluency alone is insufficient to address an increasingly complex threat landscape. Young people must pair technical skills with ethical, social, and legal awareness. AI-driven threats – from deepfakes to automated phishing – outpace legislation and traditional security tools, demanding human defenders capable of adapting to this ever-changing digital reality.

Investing in youth-led capacity building achieves exactly this objective. Providing inclusive, hands-on programmes that blend technical skills with ethics, policy literacy, and community engagement, will cultivate a generation of cybersecurity-savvy youth who do more than simply react. They will innovate, develop new tools and defensive AI systems, and challenge predictable patterns of attack with creative problem-solving. This approach will ensure a sustained movement toward responsible digital users and resilient cyber ecosystems.

About the Author

Ziarla Mae Malabanan is a certified Project Management Professional building a full-time career in cybersecurity. She has led youth empowerment and capacity-building programmes across ASEAN and Europe, including leading the 2017 ASEAN Youth Engagement Summit, helping establish the Student Youth Think Tank for Europe-Asia Relations in 2021, and directing a regional ASEAN cybersecurity programme in 2022. Her work equips young people with the skills, knowledge and leadership to navigate digital risks responsibly.

JOIN OUR SUMMER SCHOOLS

Specialized programs in Migration and Human Rights, Artificial Intelligence (AI) and Ethics

Migration and Human Rights

Gain practical knowledge on global migration challenges, human rights protection, and policy development through expert-led sessions, case studies, and interactive exercises.

Artificial Intelligence, Ethics and Human Rights

Explore the intersection of artificial intelligence, ethics, and human rights through practical exercises, expert insights, and real-world case studies.

unicri
United Nations
Interregional Crime and Justice
Research Institute