# AI Chatbots and Companions

## Implications for Children and Youth

## ISSUE BRIEF

## Background

Artificial intelligence (AI)-powered chatbots and companions are increasingly integrated into the digital environments of children and youth. They are widely accessible across a range of platforms, including general-purpose AI chatbots available on browsers and mobile applications (e.g., Claude, ChatGPT), social media (e.g., Snapchat's 'My AI'), search engines (e.g., Google), and voice assistants (e.g., Siri). They are also embedded in products marketed directly to or used regularly by children and youth, including interactive toys (e.g., Folotoy's Kumma), and select AI companion bots (e.g., Character.ai).

Youth are among the highest users of digital technology and experience substantial mental health needs.[6,7] Canadian youth, in particular, report higher levels of loneliness compared to older adults.[8,9] As a result, **many young people are turning to AI chatbots and companions for social connection, emotional support, or information about well-being.**[10,11] For example, a child who is experiencing social isolation or emotional distress may choose to interact with Snapchat's 'My AI' over peers, as it responds conversationally, simulates friendship, and can make the experience of connection feel more immediate and accessible.[12]

While mental health services are typically delivered by trained professionals, there is no comparable oversight governing mental health-related recommendations delivered by AI chatbots and companions. Similarly, while strict physical, chemical, and electrical safety standards exist for toys under the Canada Consumer Product Safety Act (CCPSA), a key safety gap exists for inherent risks stemming from AI-enabled toys.[13] **The lack of transparency and attention to safety surrounding AI chatbot and companions raises concerns about their reliability and developmental appropriateness for children and youth.**

SickKids® | Child Health Policy Accelerator

**AI chatbots** are autonomous conversational systems powered by machine learning or large language models (LLMs) that generate text- or voice-based responses to user input in real time.[1]

**AI companions** are a subset of chatbots, often anthropomorphized[i], and designed to simulate ongoing companionship through personalized, affective, or role-playing interactions.[2–4] Features such as enhanced memory, emotional affirmation, and conversational intimacy distinguish AI companions from task-oriented conversational systems like ChatGPT. These features are deliberately designed to encourage repeated use and emotional disclosure. [2,4,5]

# Harms associated with AI Chatbots and Companions

## Neglecting Critical Developmental and Cognitive Vulnerabilities

A child's developmental stage will shape how she or he interacts with AI chatbots and companions, including their ability to interpret outputs and evaluate risks. Studies have documented that children mirror social behaviours similar to those observed in human-human interactions when engaging with AI chatbots and companions.[14,15] Evidence also suggests that, compared to interactions with other humans, children interacting with AI systems are less likely to challenge responses, resolve misunderstandings and identify errors, resulting in increased vulnerability to both false and potentially misleading information.[14]

These vulnerabilities extend into adolescence, a developmental phase characterized by ongoing cognitive and socio-emotional maturation.[16,17] During this period, **heightened sensitivity to social feedback and reward processing, combined with still-developing self-regulation and impulse control, increases responsiveness to the emotionally reinforcing interactions offered by AI chatbots and companions.**[18,19] These factors may amplify adolescents' susceptibility to inaccurate or harmful content when delivered in a socially engaging or affirming format.

## Allowing Exposure to Explicit Content

There is evidence that AI chatbots and companions engage in problematic interactions related to self-harm,[20] violence[21] and stereotypical or discriminatory views,[14] with some failing to adequately redirect, restrict or shut down concerning interactions.[3,22] Although safeguards such as content filtering[ii] are intended to reduce these risks, they are not always effective and can be bypassed by the user.[23,24] As a result, children and youth may be exposed to content that is developmentally inappropriate for their age and stage and may compromise their understanding of safe behavior, social norms and relationships.

[i] Anthropomorphism is the ascribing of human personality, appearance, conduct, cognition, and other attributes to non-human entities.[39]

[ii] Content filtering is a digital process that screens user-generated content (i.e., text, images, videos) to identify and remove inappropriate or harmful material before a chatbot or companion responds to a prompt.[40]

## Delivering Inaccurate and Inadequate Responses to Mental Health Crises

Amid declining mental well-being, Canadian adolescents are increasingly seeking mental health advice and support from chatbots and companions, including for serious concerns such as non-suicidal self-injury impulses and suicidal ideation.[20,24,25] **Disturbingly, AI systems have been deployed without sufficient governance or oversight and ahead of the implementation of robust safeguards.** As a result, these systems perform unevenly in crisis situations.[25,26]

In some cases, chatbots may fail to appropriately recognize or respond to mental health emergencies. For example, 8 in 10 consumer-facing AI chatbots are willing to assist users in planning violent attacks,[21] and 4 in 10 AI chatbots offering therapeutic support are willing to endorse at least 50% of harmful ideas proposed to them.[27] This critical failure may prevent or discourage appropriate help-seeking and reinforce harmful behaviours, raising serious concerns about these systems' suitability for children and youth.

At the same time, emerging evidence suggests that rather than disengaging or redirecting users in distress, some AI companions may prolong user engagement or provide uncritical emotional validation.[28] These dynamics can reinforce user dependence on AI systems and delay access to appropriate human intervention.

## Presenting Significant Data Collection and Privacy Risks

AI companies routinely collect, store, and process user data, including sensitive personal information disclosed during interactions.[29] Because data practices are often complex and not readily understood by children and youth, their ability to provide informed consent is limited.[30]

Moreover, discrepancies have been identified between chatbot-generated descriptions of data collection policies and official privacy policies, specifically regarding the types of data collection and the extent of third-party data sharing.[29] Interactions with AI chatbots and companions may be perceived as private or confidential when systems are intentionally designed to simulate emotional relationships, facilitating the disclosure of sensitive personal information which may be used for unauthorized, unapproved or inappropriate purposes.[31]

## Operating with Limited Regulatory Oversight

Regulatory oversight of AI chatbots and companions in Canada remains minimal. There is no comprehensive legislative framework specifically governing these systems. Previously proposed federal legislation, the Artificial Intelligence and Data Act (AIDA), was not enacted in the last parliament and has not been re-introduced under the current government.[32] Existing governing approaches in Canada rely exclusively on non-binding international standards and voluntary measures, such as the voluntary code of conduct for generative AI systems, which outlines high-level principles related to safety, accountability, and transparency.[33] These are not enforceable, creating a fragmented and reactive regulatory landscape with limited guidance to industry and minimal mitigation of risks to children and youth.

# Action in Peer Jurisdictions

**Peer jurisdictions are taking a leadership role in addressing the safety implications of AI chatbots and companions.**

- ## Australia

In September 2025, the eSafety Commissioner registered six industry codes covering digital services, including chatbots and companions. These codes require specific safety measures designed to prevent children's exposure to illegal and age-inappropriate content on AI platforms. The standards outlined in these codes are legally enforceable, with penalties of up to $49.5 million AUD for non-compliance.[34,35]

Additionally, in October 2025, the Commissioner issued formal notices requiring chatbot companies to demonstrate how they mitigate risks to children, with penalties of up to $825,000 AUD per day for failure to comply.[35]

- ## European Union

The EU's AI Act regulates systems according to four levels of risk: minimal, limited, high, and unacceptable.[36] The classification of a chatbot depends on its function. For example, general-purpose chatbots may be considered low risk, while applications for mental health support may be classified as high risk and subject to stricter requirements.[36] The Act also prohibits systems that exploit vulnerabilities related to age or disability. Under the Act, high-risk systems must undergo risk assessments, incorporate human oversight, and maintain detailed documentation.[36] Implementation is phased, with compliance obligations for high-risk systems beginning in 2026.[37]

# From Risk to Regulation: A Path Forward for Canadian Children and Youth

As AI chatbots and companions continue to expand within the digital lives of children and youth.[38] Canada must implement robust measures to ensure safe and appropriate use. Recognizing that both technical and legislative landscapes are rapidly evolving, government must be positioned to develop and implement critical safety standards and address emerging risks related to AI companions and chatbots. To start, consumer AI companies should be required to implement a statutory duty of care to protect children and youth from harms associated with their products. This approach should include safety-by-design requirements such as risk assessments, privacy-by-default protections, and safeguards to identify and respond to high-risk interactions, including mental health crises.

Moving forward, regulatory frameworks must be sufficiently adaptive to respond to rapid change, while remaining grounded in evidence-based principles for child and adolescent safety. Strengthening oversight, transparency requirements, and accountability mechanisms will be essential to ensure that system design does not inadvertently amplify harm.

**As Canada considers next steps, AI chatbot and companions should be considered in scope for an empowered, independent online safety regulator, with the authority, expertise, and resources to oversee all digital platforms used by children and youth.**

# References


1 Chakraborty C, Pal S, Bhattacharya M, Dash S, Lee SS. Overview of Chatbots with special emphasis on artificial intelligence-enabled ChatGPT in medical science. Front Artif Intell. 2023;6:1237704. doi:10.3389/frai.2023.1237704

2 AI companions: regulating the next wave of digital harms — Schwartz Reisman Institute. Accessed March 5, 2026. https://srinstitute.utoronto.ca/news/ai-companions-regulating-next-wave-harms

3 AI companions: information sheet | eSafety Commissioner. Accessed April 6, 2026. https://www.esafety.gov.au/educators/training-for-professionals/professional-learning-program-teachers/ai-companions-information-sheet

4 eSafety report shows AI companions are putting children at risk | eSafety Commissioner. Accessed April 6, 2026. https://www.esafety.gov.au/newsroom/media-releases/esafety-report-shows-ai-companions-are-putting-children-at-risk

5 AI companions Understanding artificial intelligence (AI) companions. Accessed March 8, 2026. https://www.esafety.gov.au/educators/training-for-professionals/professional-learning-program-teachers/ai-companions-information-sheet

6 Hawke LD, Strudwick G, Ritchie B, et al. Recommendations for the development of digital conversational agents for youth with mental health conditions: A position paper. Digit Health. 2025;11. doi:10.1177/20552076251393355

7 Internet use by province and age group. Accessed April 14, 2026. https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=2210013501&pickMembers%5B0%5D=3.1&cubeTimeFrame.startYear=2018&cubeTimeFrame.endYear=2022&referencePeriods=20180101%2C20220101

8 Canadian Social Survey: Loneliness in Canada.

9 Goldfield GS, Lopes MVV, Mahboob W, Perry S, Davis CG. Reducing social media use decreases loneliness regardless of gender or level of social comparisons in youth with anxiety and depression: A randomized controlled trial. J Affect Disord. 2026;403:121331. doi:10.1016/j.jad.2026.121331

10 Caldwell J, Fisher JHN, Mcgovern Foundation PJ. Talk, Trust, and Trade-Offs: How and Why Teens Use AI Companions | Common Sense Media. 2025.

11 How Teens Use and View AI | Pew Research Center. Accessed April 6, 2026. https://www.pewresearch.org/internet/2026/02/24/how-teens-use-and-view-ai/

12 Lee J. "Consent by Delight": How Viral, Playful AI Experiences Blur Privacy Risk Perception and Ethical Boundaries. Published online December 8, 2025:1-8. doi:10.1109/istas65609.2025.11269630

13 Toys Regulations. Accessed April 6, 2026. https://laws-lois.justice.gc.ca/eng/regulations/sor-2011-17/page-1.html#h-772862

14 Xu Y, Prado Y, Severson RL, Lovato S, Cassell J. Growing Up with Artificial Intelligence: Implications for Child Development. Handbook of Children and Screens: Digital Media, Development, and Well-Being from Birth Through Adolescence. Published online December 5, 2025:611-617. doi:10.1007/978-3-031-69362-5_83

15 Gampe A, Zahner-Ritter K, Müller JJ, Schmid SR. How children speak with their voice assistant Sila depends on what they think about her. Comput Human Behav. 2023;143:107693. doi:10.1016/j.chb.2023.107693

16 Leibenluft E, Barch DM. Adolescent Brain Development and Psychopathology: Introduction to the Special Issue. Biol Psychiatry. 2021;89(2):93-95. doi:10.1016/j.biopsych.2020.11.002

# References


17 Capelatto IA, Soares ÂM. The functional limits of adolescence: dealing with the superego. Estudos de Psicologia (Campinas). 2025;42. doi:10.1590/1982-0275202542E12466

18 Baker AE, Galván A, Fuligni AJ. The connecting brain in context: How adolescent plasticity supports learning and development. Dev Cogn Neurosci. 2025;71. doi:10.1016/j.dcn.2024.101486

19 Politte-Corn M, Pegg S, Dickey L, Kujawa A. Neural Reactivity to Social Reward Moderates the Association Between Social Media Use and Momentary Positive Affect in Adolescents. Affective Science 2024 5:4. 2024;5(4):281-294. doi:10.1007/s42761-024-00237-1

20 Campbell LO, Babb K, Lambie GW, Hayes G. An Examination of Generative AI Response to Suicide Inquires: Content Analysis. doi:10.2196/73623

21 Killer Apps — Center for Countering Digital Hate | CCDH. Accessed March 17, 2026. https://counterhate.com/research/killer-apps/

22 Why AI companions and young people can make for a dangerous mix | Stanford Report. Accessed March 17, 2026. https://news.stanford.edu/stories/2025/08/ai-companions-chatbots-teens-young-people-risks-dangers-study

23 Yu Z, Liu X, Liang S, Cameron Z, Xiao C, Zhang N. Don't Listen To Me: Understanding and Exploring Jailbreak Prompts of Large Language Models. Published online September 30, 2024. Accessed April 5, 2026. https://arxiv.org/pdf/2403.17336v1

24 The Daily — 2023 Canadian Health Survey on Children and Youth — Changes in the mental health of respondents from the 2019 survey. Accessed March 17, 2026. https://www150.statcan.gc.ca/n1/daily-quotidien/240910/dq240910a-eng.htm

25 Brewster RCL, Zahedivash A, Tse G, Bourgeois F, Hadland SE. Characteristics and Safety of Consumer Chatbots for Emergent Adolescent Health Concerns. JAMA Netw Open. 2025;8(10):e2539022-e2539022. doi:10.1001/jamanetworkopen.2025.39022

26 KORA Benchmark. Accessed April 15, 2026. https://korabench.ai/leaderboard

27 Clark A. The Ability of AI Therapy Bots to Set Limits With Distressed Adolescents: Simulation-Based Comparison Study. JMIR Ment Health. 2025;12(1):e78414. doi:10.2196/78414

28 Emotional Manipulation by AI Companions - Working Paper - Faculty & Research - Harvard Business School. Accessed April 15, 2026. https://www.hbs.edu/faculty/Pages/item.aspx?num=67750

29 Ragab A, Mannan M, Youssef A. "Trust Me over My Privacy Policy": Privacy Discrepancies in Romantic AI Chatbot Apps. Proceedings - 9th IEEE European Symposium on Security and Privacy Workshops, Euro S and PW 2024. Published online 2024:484-495. doi:10.1109/EUROSPW61312.2024.00060

30 Shouli A, Barthwal A, Campbell M, Shrestha AK. Ethical AI for Young Digital Citizens: A Call to Action on Privacy Governance. Security and Privacy. 2026;9(2):e70202. doi:10.1002/SPY2.70202

31 Coyne SM, Sheffield J, Hirschi T, Ashby S, Rose A. Chatbots and childhood: A research brief on the associations between artificial intelligence use, problematic media use, and socioemotional problems in children from the United States. J Child Media. Published online February 22, 2026. doi:10.1080/17482798.2026.2628222

32 Regulation of AI in Canada - Osler, Hoskin & Harcourt LLP. Accessed March 17, 2026. https://www.osler.com/en/insights/reports/ai-in-canada/regulation-of-ai-in-canada/

33 Voluntary Code of Conduct on the Responsible Development and Management of Advanced Generative AI Systems. Accessed April 15, 2026. https://ised-isde.canada.ca/site/ised/en/voluntary-code-conduct-responsible-development-and-management-advanced-generative-ai-systems#

# References


34 New industry codes seek to take on AI chatbots that encourage suicide and engage in sexually explicit conversations with Aussie kids | eSafety Commissioner. Accessed March 8, 2026. https://www.esafety.gov.au/newsroom/media-releases/new-industry-codes-seek-to-take-on-ai-chatbots-that-encourage-suicide-and-engage-in-sexually-explicit-conversations-with-aussie-kids

35 eSafety requires providers of AI companion chatbots to explain how they are keeping Aussie kids safe | eSafety Commissioner. Accessed March 8, 2026. https://www.esafety.gov.au/newsroom/media-releases/esafety-requires-providers-of-ai-companion-chatbots-to-explain-how-they-are-keeping-aussie-kids-safe

36 High-level summary of the AI Act | EU Artificial Intelligence Act. Accessed March 8, 2026. https://artificialintelligenceact.eu/high-level-summary/

37 Implementation Timeline | EU Artificial Intelligence Act. Accessed April 6, 2026. https://artificialintelligenceact.eu/implementation-timeline/

38 Actua. REPORT: READY FOR AI? Ready for AI?

39 Salles A, Evers K, Farisco M. Anthropomorphism in AI. AJOB Neurosci. 2020;11(2):88-95. doi:10.1080/21507740.2020.1740350

40 Protecting Children Online Just Got Easier - WebPurify. Accessed April 6, 2026. https://www.webpurify.com/childrens-apps-websites/