# (Gen)eration AI:

## Safeguarding youth privacy in the age of generative artificial intelligence

Tiffany Kwok and Christelle Tessono | March 2025

Toronto Metropolitan University

the dais

# Acknowledgements

The Dais is a public policy and leadership think tank at Toronto Metropolitan University, working at the intersection of technology, education and democracy to build shared prosperity and citizenship for Canada.

For more information, visit dais.ca
20 Dundas St. W, Suite 921, Toronto, ON M5G 2C2

@daisTMU

**How to Cite this Report**
Kwok, Tiffany and Christelle Tessono. *(Gen)eration AI: Safeguarding youth privacy in the age of generative artificial intelligence*. The Dais. 2025. https://dais.ca





This project has been funded by the Office of the Privacy Commissioner of Canada (OPC); the views expressed herein are those of the author(s) and do not necessarily reflect those of the OPC.

**Design and Illustration**
Mariana Rodrigues

**Copy Editor**
Suzanne Bowness, CodeWord Communications

**Contributors**
Nina Rafeek Dow
André Côté

**Acknowledgments:**
This project could not have been made possible without the participation and insights provided by all the interviewees:

Howie Bender, Rubicon Publishing
Paulla Bennett, York Region District School Board
Andrew Bieronski, Quizizz
Jovine Chan, York Region District School Board
Angela Chen, Stanford Accelerator for Learning
Tom D'Amico, Ottawa Catholic School Board
Remo George Joseph, Quizizz
Indra Kubicek, Digital Moment
Jacques Marcoux, Canadian Centre for Child Protection (C3P)
Brenda McPhail, McMaster University
David Porter, David Porter and Associates
Valerie Steeves, University of Ottawa

The Dais proudly engages a diverse group of funders to support and catalyze our work, consistent with our values, and subject to a thorough internal review. As a non-partisan, public-interest institute, we only accept funds from organizations that support our mission and enable us to undertake work independently, with full editorial control. The names of all of our financial supporters are publicly and transparently displayed on all online and printed material for each project or initiative.

# Authors

**Tiffany Kwok**
Policy Analyst

Tiffany Kwok (she/her) is a policy analyst at the Dais. She is passionate about tech and urban policy, service design, and research, and has experience working in the public sector and various non-profit organizations both in Canada and the UK. Tiffany holds a Bachelor's degree in Political Science and Urban Studies from the University of Toronto and a Master's degree in Digital Technologies and Policy from University College London.

**Christelle Tessono**
Policy and Research Assistant

Christelle Tessono (she/her) conducts research at the intersections of digital technology, human rights, and governance. This has led her to work on a variety of projects related to political advertising on social media platforms, gig work, facial recognition technology, privacy, and AI governance. Christelle holds a Bachelor of Arts in political science from McGill University and is currently pursuing graduate studies at the University of Toronto's Faculty of Information.

# Table of Contents

# 1 Executive Summary

**BOLD IDEA:** **As generative AI tools become commonplace for children and teenagers, so must governments, school administrators, and technologists introduce new practical privacy policy interventions.**

Chatbots as digital companions, AI-driven teaching tools, and content-generation tools are all examples of how generative AI (genAI) is transforming how young people learn, connect, and express themselves.

However, the rapid rise of genAI use has exposed youth to unprecedented privacy risks. Unlike adults, children and teenagers often trust technology— and now AI systems—without fully understanding how they work, making young people more vulnerable to data collection, manipulation, and exploitation.

From unauthorized data collection and profiling to the spread of deepfake child sexual abuse material (CSAM) to AI-generated cyberbullying, genAI is already amplifying the risks youth face online. Policymakers, technologists, educators, and parents must act now to implement concrete safeguards and mitigate harm.

This report examines the unique privacy risks children and teenagers face when using genAI tools, particularly in educational and social media settings. Using a mixed-method research approach, including a literature review, expert interviews, and case studies, the report assesses existing policy frameworks, identifies gaps in protections, and recommends best practices for policymakers, technologists, educators, and parents to mitigate risks and enhance youth privacy.

**The rapid rise of genAI use has exposed youth to unprecedented privacy risks.**

## Findings: Privacy and associated risks of generative AI for youth

1. **Unconsented data collection and processing:** Children may unknowingly provide personal information to generative AI platforms, which can be collected and used for purposes they do not fully understand. AI systems often retain inputted data, making it difficult to control or delete information once shared.

2. **Profiling and targeted advertising:** AI-powered platforms can create persistent data profiles of children, leading to targeted advertisements and automated profiling, even as privacy preferences change over time.

3. **Data breaches and cybersecurity threats:** Educational institutions and AI providers have been subject to cyber attacks, exposing sensitive student information such as names, health records, and geolocation data.

4. **Surveillance and overreach:** The adoption of genAI in educational settings has led to increased data collection, tracking students' learning patterns, behaviours, and even personal interests, raising concerns about surveillance and lack of autonomy.

5. **Bias and misinformation:** GenAI tools trained on biased datasets can reinforce harmful stereotypes, misrepresent historical facts, or generate misleading information, which may impact a child's learning and perception of the world.

6. **Harmful content and AI companionship risks:** AI-generated content can be manipulated to create harmful material, including cyberbullying content, misinformation, and deepfake child sexual abuse material (CSAM). Furthermore, AI chatbots designed for companionship may lead children to form emotional bonds with non-human entities, which can impact their mental health and social development.

7. **Dark patterns and manipulative design:** Many AI platforms employ "dark patterns" —design tactics that subtly encourage users to share more data or stay engaged longer than intended. This can be particularly exploitative for children.

## Policy recommendations

### **For policymakers:** implement stronger privacy laws and governance

- **Mandate youth-specific privacy protections:** Future privacy laws must explicitly account for youth vulnerabilities, requiring stricter consent mechanisms and "privacy-by-design" principles.
- **Ban data collection without explicit consent:** Similar to the European Union's GDPR protections for minors, establish regulations that prohibit AI systems from collecting youth data unless explicit, informed consent is obtained.
- **Enforce accountability for tech companies:** Introduce legal consequences for companies that fail to secure youth data, mismanage AI-generated content, or use deceptive consent practices.
- **Create specific guidelines for AI in education:** Develop clear policies for how genAI should be used in classrooms, ensuring that AI-driven tools do not compromise student privacy or academic integrity.

### **For technologists and AI developers:** design AI with privacy first

- **Adopt Privacy-Enhancing Technologies (PETs):** Implement differential privacy, homomorphic encryption, and anonymization techniques to minimize risk.
- **Limit data collection and retention:** Set and enforce strict data minimization policies to ensure AI models only store essential information for the shortest duration possible.
- **Enhance transparency and youth-friendly design:** Develop clear, age-appropriate privacy dashboards and real-time data usage alerts for young users and their guardians.
- **Strengthen AI safety and cybersecurity**: Regularly audit AI models to identify and patch vulnerabilities that could lead to data leaks or adversarial attacks.

### **For educators and school boards:** strengthen AI literacy and digital safeguards

- **Conduct Privacy Impact Assessments (PIAs):** Require school boards to evaluate AI tools before approving them for student use.
- **Promote AI and digital literacy education:** Integrate AI literacy into K-12 curricula to equip students with critical thinking skills for assessing AI-generated content.
- **Develop clear guidelines for AI use in schools:** Establish policies that govern how AI can be used in the classroom, ensuring that it complements—rather than replaces—traditional learning methods.

### **For parents and guardians:** advocate for safe AI use at home

- **Monitor AI interactions and privacy settings:** Regularly review the AI-powered tools children use and adjust privacy settings accordingly.
- **Educate children on AI risks:** Teach children about AI's capabilities and limitations, emphasizing responsible engagement and data protection.
- **Encourage critical thinking:** Foster skepticism about AI-generated content and encourage children and youth to verify information through multiple sources.

By embedding youth privacy protections into AI design, governance, and educational policies today, we can ensure that the next generation experiences the benefits of AI without compromising their safety, security, or autonomy.

# 1 Introduction

Podcasts produced from PDF documents in a matter of minutes, imaginary but realistic images of a vacation that never happened, and outputs brainstormed in seconds from a user prompt—these are all the product of generative artificial intelligence (genAI). Since the public release of ChatGPT in late 2022, this rapidly evolving technology has changed how workflows and processes are optimized and how ideas and content are generated. From applications that can create photorealistic images and videos from a few well-crafted sentences to chatbots that can participate in human-like discussions and relationship-building, excitement for opportunities to come are undeniable.

And yet, beyond the optimization, customization, and quick generation of new content, harms and risks of the technology have emerged just as quickly.

The terms "youth," "children and teenagers," and "minors" are used interchangeably throughout the report to refer to school-aged individuals. For children and teenagers, user behaviour during interactions with genAI tools often differs from adults, as these young users have a greater tendency to trust the tool's outputs and responses. Despite knowing that an AI system is a machine, youth may continue to treat the system as human, confiding in and providing personal information that one might not disclose otherwise.[1] This relationship with technology comes with new privacy risks. Without fully understanding how genAI systems work before using them, or how inputted data might be used, youth are at risk for many harms, including unconsented data collection and processing, the inability to correct or control data, and ultimately a violation of their human right to autonomy and privacy. This can have long-lasting implications—creating a data profile of a user despite changing privacy preferences, making users privy to tailored advertisements, profiling from automated systems, and even to having their data compromised in cyber attacks.

Other concerns also exist. They include biased data and outputs, "dark patterns"[2], (elements that draw on behavioural science, user data, and predictive algorithms to manipulate users into doing things they don't want to do) embedded in the design of the user interface, and content created for malicious use. Some of the harmful uses of the technology has led to concerns around children and teenagers being at risk for bullying from genAI content created by peers, dealing with adverse impacts on mental health from growing relationships formed with AI companions, and facing temptation to over-rely on genAI tools for school assignments. Other harms have been exacerbated by the availability and ease with which child sexual abuse material (CSAM) can be created.[3]

In general use cases, the core of genAI technology is powered on training data— an aspect that enables genAI tools to provide customized outputs, which are further refined as more data is provided. This has been seen as beneficial in some cases; for example, where applications used to optimize lesson planning for teachers have saved time and burdensome administration. In some cases, users sharing personal or confidential information with a chatbot led to companies like Samsung, J.P. Morgan, and Apple to limit or ban the use of ChatGPT due to privacy and corporate confidentiality risks associated with this information being entered into the tool's training data, and potentially reshared through other outputs.[4]

This report focuses specifically on privacy risks and implications surrounding children's and teenagers' use of genAI tools, and what can be done to protect them from these harms. We employed a mix of methods to explore this evolving field: a literature review, a scan of policy initiatives addressing genAI technology, privacy, and data protection, and 12 expert interviews with individuals from school boards, non-profits, an education technology company, the startup space, and academia.

Our study focuses on applications in K-12 education and on social media (which are used ubiquitously by many youth). The report begins with a brief overview of the main genAI tools being used by children and teenagers at the time of writing. This is followed by an examination of the state of policy and technical interventions in Canada and other key jurisdictions. The final section outlines privacy design principles, and best practices for technologists, educators, parents, and policymakers.

**Despite knowing that an AI system is a machine, youth may continue to treat the system as human, confiding in and providing personal information that one might not disclose otherwise.**

# 2 What is Generative AI?

## The whole wide world of generative AI

The growing use of generative AI tools follows a decades-old trend in digitization, whereby an increasing number of activities are mediated through digital platforms.

According to the Organisation for Economic Cooperation and Development (OECD), generative AI can be defined as AI models that "create new content in response to prompts, based on their training data."[5] GenAI applications can be used to generate text, code, images, audio, and video. There are an increasing number of companies that are developing or adopting generative AI tools in their products. These technological developments can be experienced across multiple sectors, for example in healthcare (e.g. patient case notes, clinical diagnostics, etc.), in the banking sector (e.g. as client-facing chatbots), in the financial sector (e.g. as virtual assistants for market research), and in education (e.g. for curriculum development and student assessment). However, for the purposes of this report, we will focus on genAI tools that are directly present in everyday lives of Canadian children and teenagers through K-12 education as well as on social media platforms.

In recent years, digitization has led to the use of big data analytics to optimize student performance at school.[6] Technology is usually adopted for the following purposes: information search (e.g. digital libraries), school management (e.g. communications with parents, teachers, and students), and teacher educational tools (e.g. quizzes, automated proctoring, or grading). Right now, the use of genAI in schools is growing. An increasing number of companies have developed technologies for learning purposes, or are now adopting genAI applications as a part of their existing products. Generative AI companies are also adapting their technologies to operate in educational contexts.

In terms of uses of genAI in social media, we know it is currently being to train and serve as chatbots,[7] to create content on platforms (video, images, text, audio, etc.), as well as to train algorithms from metadata and other types of user-generated content. Another related context that we believe is important to note, but are not within scope of this report, is in video game development.[8] Furthermore, generative AI tools are increasingly being used in the provision of mental health support such as crisis help-lines.[9] Companies will develop chatbots using the transcripts from help-line phone calls to train their systems.

**Table 1:** A non-exhaustive list of popular uses of generative AI in education and on social media platforms

| Application | Name | Utility for school boards and educators | Utility for youth | Risks to youth |
|---|---|---|---|---|
| Conversational Chatbot / Companion AI | • Character. AI<br>• Replika<br>• Snapchat's My AI,<br>• Meta AI | Teacher's assistant | Emotional support on social media | Develop strong emotional bonds<br><br>Receive false, problematic, harmful or misleading information<br><br>Sensitive and personal information is used by malicious actors against youth |
| Image, audio and video generation and editing | • Dall-E<br>• Stable Diffusion<br>• Adobe<br>• Google's Veo 2<br>• Canva<br>• Suno | Create visual and auditory content for pedagogical purposes | Visual and audio content for assignments<br><br>Create memes and humorous content for social media | Pictures, video and audio from the youth could be edited to (un) intentionally harm them.<br><br>CSAM (child sexual abuse material) developed by malicious actors<br><br>Media developed could be unrepresentative and perpetuate harmful biases |
| Translation | • HeyGen | Translate to promote accessibility and comprehension | Support reading comprehension, translation to native language etc. | Poor and inaccurate translation of concepts may further marginalize youth, which may lead to poor academic performances and sense of belonging |
| Text editor, text generation | • Grammarly<br>• OpenAI's ChatGPT<br>• Microsoft Copilot<br>• Google's Gemini | Curriculum development, automated grading, assessment development | Essay drafting and editing, coding development, editing, problem solving, summarize text, support reading comprehension | Inability to apply and understand academic integrity norms<br><br>Overreliance on tools may lead to undeveloped critical thinking and writing skills |
| Academic assessment, Teaching Assistant, AI-based tutoring | • Quizlet<br>• Khanmigo<br>• Brisk<br>• SchoolAI<br>• Quizizz | Develop grading rubrics, create quizzes, tests, exams and other assessment methods | Practice quizzes, tests, exams and other assessment methods | Assessments and AI-based tutoring may not accurately grasp students' learning style and needs |
| Literature scanning, academic search engines | • Perplexity<br>• R3<br>• Elicit | Find and analyze scholarship as well as primary and secondary sources | Find and analyze scholarship as well as primary and secondary sources | Degradation of reading comprehension and research skills |
| Learning management system | • Google Classroom<br>• Canvas* | Organize learning materials and student-teacher communications in a central platform | Access learning materials and student-teacher communications in a central platform | Sensitive and personal information may be exposed due to cyber attack |
| *The above learning management system examples have integrated features (E.g. Google Workspace for Education has Gemini) | | | | |

## GenAI and privacy harms

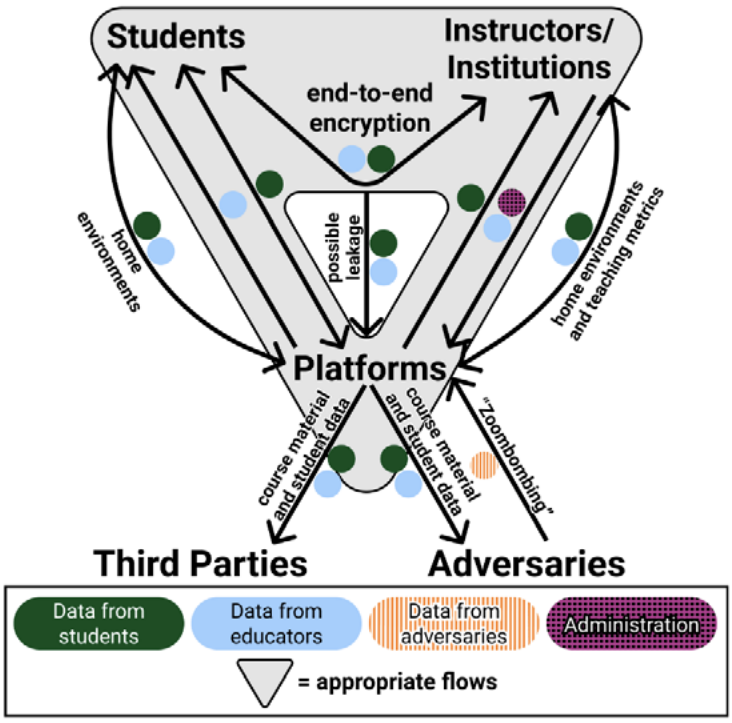
Recognizing privacy harms specific to children and teenagers is vital as genAI tools continue to become more widely used. Virtual learning environments currently being adopted in education feature information technology infrastructures that enable continuous data collection—which include, but are not limited to, student records, grades, geolocation, browsing habits, and personal information.[10] In the absence of technical and policy guardrails preventing how data is collected and subsequently used, student information can be sold to a third-party and used for purposes that schools, teachers, and students did not consent to or are unaware of.[11] Even if anonymized, student data collected by genAI tech providers can be re-identified and used for other purposes (malicious or commercial) by a third-party or the provider themselves as re-identification techniques have increasingly become more sophisticated.[12] For instance, a study conducted by Na et al. (2018) showed that a machine-learning algorithm was able to re-identify approximately 80 percent of children and 95 percent of adults based on de-identified physical activity data.[13] This indicates the ease with which sensitive data can be exposed to advertisers, governments, and strangers.[14]

**In the absence of technical and policy guardrails preventing how data is collected and subsequently used, student information can be sold to a third-party and used for purposes that schools, teachers, and students did not consent to or are unaware of.**

In recent years, a growing number of institutions have experienced data breaches from cybersecurity attacks. For instance, the Toronto District School Board uses PowerSchool, a cloud-based program to store student and staff information. In January 2024, PowerSchool was hacked, which led to sensitive personal data from 1985 to 2024 to be compromised, including medical information, home addresses, health card numbers, student names, and dates of birth among others.[15] GenAI technology providers are not immune to this problem. Other cases indicate that the risks are present, such as Character.ai's data breach of usernames, voices, and chats due to internal error,[16] and now defunct educational technology company AllHere's mismanagement of sensitive student data, as called out by a whistleblower.[17]

The adoption of digital technologies in the classroom undermines traditional norms of intellectual privacy and safety.[18] Prior to digital technology adoption, classroom discussions, assignments, performance indicators, and other relevant metrics, didn't leave the classroom or the school. Students were free to discuss without concern that what they contribute to the classroom would not be used in other contexts. Now, the classroom extends to wherever the student accesses it, with their laptop, tablet, and even smartphone. Digital technologies are constantly tracking their performance, as well as having access to information outside the classroom such as their geolocation and lifestyle habits (e.g. extracurricular interests, nutrition, browsing history, etc). The large amounts of data collected about students enables unprecedented forms of surveillance. As illustrated in Figure 1, through the adoption of digital technology in the classroom, data flows between not only students and their teachers, but also between the educational technology platforms, providing data to third parties, and potential adversarial actors looking to use data in harmful ways. This raises a range of risks to the individual—from having their data sold to marketers for targeted advertising,[19] to having data "profiles" created that have the potential to follow students into adulthood, despite changing privacy preferences.

**Figure 1:** This model depicts data flows between students, instructors, schools, platforms, third parties, and adversarial actors. This model was developed by Cohney et al. (2021).[20]




Moreover, social media platforms have a wide range of privacy implications that may affect the children and youth who use them.[21] With the growing adoption of genAI, the risks of privacy violation and data breaches increase. For instance, platforms such as Snapchat have launched conversational chatbots that act as companions to their users. The information that is disclosed by youth through these conversations can be used by platforms for a variety of undisclosed purposes.[22] For example, although there are mechanisms in place to "collect consent," the process by which it is collected is often unclear, so user data is not collected via informed consent. Further, we do not know whether this data is de-identified and anonymized.

**Digital technologies are constantly tracking their performance, as well as having access to information outside the classroom such as their geolocation and lifestyle habits (e.g. extracurricular interests, nutrition, browsing history, etc). The large amounts of data collected about students enables unprecedented forms of surveillance.**

## Harms beyond privacy

Other types of harms beyond privacy are also important to mention. For instance, researchers have raised concerns about the degradation of pedagogical conditions. This concern builds on the decades-long digitization of education, whereby a teacher's autonomy is compromised through the displacement of grading and curriculum delivery onto automated decision-making tools. Furthermore, since a majority of these tools are developed by for-profit technology providers, there is also a concern that pedagogical decisions are being shaped by these corporate interests instead of the public interest.[23] In this case, genAI would accelerate this trend due to the technology's advanced capabilities.

Experts we spoke with raised concerns that genAI is eroding educational norms and values such as academic integrity, attribution, and originality. Other concerns include the degradation of critical thinking skills of students due to heavy reliance on generative AI tools to interpret texts.[24] Regarding its application in school settings for skills identification and career counselling, there are concerns that students won't be adequately evaluated, which could be detrimental to their professional development.

Beyond AI's impact on learning and academic performance, experts also raised concerns about the hyper-personalization of education afforded by genAI tools. Hyper-personalization refers to the set of technical approaches that analyze student learning patterns and capabilities to develop exercises and interfaces that ensure better engagement and performance outcomes. Specifically, experts argued that these technologies will facilitate the isolation of youth and increase challenges in developing strong social bonds with their peers. This could render youth even more vulnerable to phishing and catfishing scams as well.

Finally, and this is an issue that predates genAI tools, there are risks of ingrained biases and potentially discriminatory outputs, as genAI tools may not capture all lived realities, and histories of marginalization due to an overrepresentation of dominant Western epistemologies. For instance, this may look like adopting derogatory language such as slurs, creating contested socio-political framings (e.g. not recognizing a particular Indigenous nation), or assuming gendered representations of work (e.g. only presenting men as doctors) to name a few examples.[25] Moreover, "hallucinations", the term for chatbot responses that contain false or misleading information, are also significant problems with genAI technologies.[26] This may lead to not only misunderstanding issues and academic material, but may also expose students to harmful practices (e.g. if a user asks about nutritional or health advice for instance and gets a non-standard answer or advice).

On social media platforms, genAI can harm youth when bad actors use it to create child sexual abuse material (CSAM) deepfakes of minors. There are growing cases nationally and internationally, where bad actors will use tools like open-source image platform Stable Diffusion to create this harmful imagery.[27] It is also important to note that both youth (e.g. classmates, teammates, peers, etc), as well as malicious adult actors, can engage in this harmful behaviour.

There has also been an increase in the popularity and use of conversational chatbots. Examples of companies include Character.AI and Replika. These tools are often designed to optimize engagement, and as a result may lead to users developing deep emotional connections that open the door to a variety of mental health challenges. For instance, in October 2024, a 14-year-old American tragically died by suicide after developing an intense emotional attachment with a companion AI.[28]

# 3 The Policy Landscape: GenAI Privacy and Data Protections for Children

In this section, we provide a brief overview and analysis of privacy laws and how they apply to generative AI technologies that directly interact with children and teenagers.

## Canadian policy landscape

In Canada, there are two privacy laws at the federal level. The *Personal Information Protection and Electronic Documents Act* (PIPEDA) outlines how federally regulated private sector companies should collect, use, and disclose personal information when conducting their commercial activities in Canada.[29] PIPEDA applies to businesses like banks, airlines, and technology providers such as social media platforms, as well as educational technology companies that develop genAI, or use genAI in their student-facing products. Second, there is the *Privacy Act*, which outlines how federal government institutions should collect, use and disclose personal information in their activities. This could include federal departments and agencies that deliver employment insurance, pensions, and tax refunds.[30]

PIPEDA adopts a consent-based framework that requires companies to seek the consent of their clients before processing their personal information. Personal information consists of information about an identifiable individual, which can include their name, age, identification numbers, blood type, ethnic origin, opinion, evaluations, comments, disciplinary actions, medical records, etc.[31] PIPEDA enforces the following 10 principles: accountability, consent, accuracy, safeguards, openness, individual access, compliance, limiting collection, accuracy, identifying purposes, and limiting use, disclosure, and retention. This means that businesses have to apply those principles when undergoing their commercial activities.

## Do we have up-to-date privacy legislation?

It is important to note that there have been efforts to modernize PIPEDA and regulate AI systems in recent years. Most recently, the government tabled *Bill C-27: The Digital Charter Implementation Act* in June 2022. Bill C-27 consisted of the following three Acts: the *Consumer Privacy Protection Act* (CPPA), the *Personal Information and Data Protection Tribunal Act* (PIDPTA), and the *Artificial Intelligence and Data Act* (AIDA). The CPPA sought to recognize privacy as a fundamental human right and prioritizes the importance of protecting children from privacy harms. The proposed PIDPTA would serve as an administrative tribunal to hear appeals and administer penalties for contraventions to the CPPA. AIDA sought to regulate AI systems by requiring companies to adopt risk mitigation measures, and prohibit the development of systems built with illegally obtained personal information. In early 2025, the prorogation of parliament ended the legislative session, with C-27 dying on the order paper.[32]

The Office of the Privacy Commissioner of Canada (OPC) is an agent of parliament that oversees compliance with PIPEDA and the *Privacy Act*. The OPC is empowered to investigate complaints, perform audits, pursue court action, undertake research, and promote public awareness about privacy issues.[33]

The OPC's analysis of PIPEDA suggests that there are series of no-go zones in the processing of personal information:

1. collecting, using or disclosing personal information in ways that are otherwise unlawful;
2. profiling or categorizing individuals in a way that leads to unfair, unethical or discriminatory treatment contrary to human rights law;
3. collecting, using or disclosing personal information for purposes that are known or likely to cause significant harm to the individual;
4. publishing personal information with the intent of charging people for its removal;
5. requiring passwords to social media accounts for the purpose of employee screening; and
6. conducting surveillance on an individual using their own device's audio or video functions.[34]

The OPC recently outlined a series of nine principles for how developers and providers of generative AI technologies should apply PIPEDA in their commercial activities. These consist of practical ways businesses can operationalize PIPEDA's fair information principles. They include, but are not limited to, acknowledging the legal authority of consent in the use of personal information during the training, development, and use of generative AI systems.

## Zooming in: provinces and school boards

Some provinces (Quebec, Alberta and British Columbia) have developed their own privacy laws for the private sector. Given that education is a provincial responsibility, privacy laws are subsequently operationalized in their respective education statutes. Ontario is the first to pass legislation that specifically addresses AI in the education system.

The Ontario government passed the first provincial law that specifically governs the use of AI in the education system. In November 2024, Ontario introduced Bill 194: *Strengthening Cyber Security and Building Trust in the Public Sector Act*. The Bill amends the province's *Freedom of Information and Protection of Privacy Act* (FIPPA) and introduces the *Enhancing Digital Security and Trust Act* (EDSTA). EDSTA establishes requirements for the responsible use of AI systems by public-sector institutions, which include school boards. Moreover, the legislation acknowledges the uniquely sensitive nature of data about minors, as well as systems that directly interact with them. As a result, the law will establish provisions for the creation of technical standards, directives, and disclosure mechanisms that school boards must abide by when processing digital information about individuals under 18.[35]

The provinces of Quebec, Alberta, and British Columbia have developed privacy legislation that applies to the private sector. In 2022, the Quebec government enacted the *Act respecting the protection of personal information in the private sector*. This Act was developed to ensure that privacy law in Quebec was aligned with the *General Data Protection Regulation* (GDPR) in Europe. As a result, the law builds on the existing consent framework by adding special provisions relating to the processing of biometric data, heightened notices of data breaches, and mandated privacy risk assessments.[36]

Beyond legal frameworks, a growing number of provincial education ministries have published guidelines on the use of AI systems. For instance, in British Columbia, the Ministry of Education and Child Care developed the Considerations for *Using AI Tools in K-12 Schools*, which outlines the possible applications, accessibility, data and security, as well as tailored guidance for school boards, district leaders, and teachers.[37] In Quebec, the minister of Education released the *L'utilisation pédagogique, éthique et légale de l'intelligence artificielle générative* (The educational, ethical and legal use of generative artificial intelligence), which outlines criteria for when and how generative AI should be used in pedagogical contexts, and includes ethical considerations, as well as discussion on the legal obligations per Quebec's privacy laws.[38] In New Brunswick, *Recommended Approaches to Generative Artificial Intelligence* was developed by education stakeholders and partners, and advises school communities on how to view AI as a collaboration tool, and to use it responsibly and effectively.[39]

Certain school boards have taken the initiative to create their guidelines for the use of AI (particularly generative AI tools like ChatGPT). For instance, the Waterloo Catholic District School Board in Ontario developed guidelines for AI and generative AI tools for its students and educators.[40] The Ottawa Catholic District School Board has also established guidelines, which include grade-based guiding principles, information for parents/custodians/guardians, and access to a free one-hour online primer course on generative AI.[41]

In summary, private sector technology companies developing and providing generative AI products must abide by federal privacy legislation, except in Quebec, Alberta and British Columbia where their respective provincial laws are considered default. Furthermore, school boards must abide by provincial public sector privacy legislation. Ontario is the first province to have passed legislation relating to the use of AI systems in educational contexts. While the province has carved out special protections for youth, implementation has just begun. Many provincial ministries of Education have developed guidelines for AI use within school boards, with some boards introducing their own policies.

## Lessons from other jurisdictions

In recent years, many countries have introduced a wide range of regulatory and technical approaches to address youth privacy and data protection. The advent, quick evolution, and widespread access to genAI technology and tools has amplified calls for safeguards to be put in place. Yet policy interventions exist in various forms—from legislation both presently existing and under consideration, "design codes" for children (sets out standards online services should follow), and voluntary guidelines and frameworks. This section highlights three key lessons drawn from other jurisdictions in their efforts to govern and enforce privacy and data protection of youth with genAI products and services.

### Privacy and data protection regulations that establish specific provisions for youth

Many other jurisdictions have acknowledged that special provisions are required for the privacy protection of youth. This has been recognized through different policy interventions—including but not limited to the European Union's General Data Protection Regulation (GDPR), *AI Act*, and *Digital Services Act* (DSA); the United Kingdom's *Age-Appropriate Design Code* (AADC), and the *United States' Children's Online Privacy Protection Act* (COPPA). Beyond interventions at the nation state-level, international agreements and jurisdictional tools establish specific protections. Examples include the United Nations Convention on the Rights of the Child, and the non-profit organization, California IT in Education (CITE), which established the California Student Data Privacy Agreement.

***European Union***

The EU's legal framework is most comprehensive. Three key pieces of legislation—the GDPR, AI Act, and DSA—address different aspects of AI and data governance. As a broad overview, GDPR harmonizes data privacy laws across Europe, the AI Act classifies different risk levels for AI systems, and states obligations for developers, providers, and users, and the DSA introduces responsibilities for online platforms and new rights for users like data protection, transparency, and accountability. Core features of genAI technology are believed to be inherently in violation of the GDPR. The following rights are difficult to promise with the nature of training data used to power an AI system:

**Right to access:** the right to obtain a copy of personal data that has been collected (with genAI, you can't just "access" your data because it's been inputted into the black box system - an AI system in which it's internal workings are a mystery to its users)

**Right to rectify:** the right to correct inaccurate personal data that's been collected (with genAI, you can't just "access" and correct your data because it's been inputted into the black box system)[42]

**Right to object to data processing:** the right to object to the processing of personal data (with genAI, inputs are generally inputted into the training data); this right is only possible for some (e.g. EU, UK with strict data protection regimes) due to the option to opt out of training

There is also the issue of consent, which was already overridden due to the fact that the majority of data used to initially power these systems was obtained by "scraping" (extracting existing information, i.e., text, images, from the internet.)[43] Italy initially banned OpenAI's ChatGPT, and later levied a €15 million fine on OpenAI, due to the lack of legal basis around unauthorized collection and processing of personal data to train the algorithms, among other concerns.[44] In other cases, data deletion requests under GDPR have been unsuccessful—with Meta requiring those who request deletion to submit evidence that their personal information appears in its outputs, and genAI companies OpenAI and Midjourney failing to respond to an image deletion request.[45]

However, focusing on child-specific protections, the GDPR does indicate that the interests of children as data subjects should have high priority, very often outweighing any legitimate interest to process an individual's data,[46] and that children under 13 can only give consent with permission from a parent.[47] The AI Act specifies that AI systems that exploit any vulnerabilities of an individual due to their age are prohibited.[48] And the DSA offers specific protections to children, including the right to protection of personal data and privacy, and requiring online services to provide child-friendly terms and conditions.[49]

***United Kingdom***

The UK's *Age-Appropriate Design Code* (AADC) of the *Data Protection Act* 2018 provides standards for companies to develop online services in a way that uphold children's rights and best interests. From minimizing data collection and retention and switching off geolocation settings, to avoiding "nudge" techniques and providing transparency by sharing privacy information and published terms in language suited to the age of a child, these standards support compliance with the GDPR. Similar age-appropriate "design codes" have also been passed by the states of California and Maryland in the US. In 2021 (the year the AADC came into full effect), 42 platform changes were logged across four platforms (Meta, Google TikTok, and Snap)—including "by default" design, ensuring that personal data is processed with the highest privacy protection by default.[50]

***United States***

The US *Children's Online Privacy Protection Act* (COPPA), which took effect in 2000, outlines specific safeguards for children under the age of 13 and is applicable to any service operator and company collecting and using children's personal information, focusing on websites directed to children. Besides requiring parental consent to collect and use children's personal information, and a clear and comprehensive online privacy policy, some elements of the law are likewise challenged by the inherent nature of genAI technology. Whereas the law requires that operators must provide parents with the opportunity to access their child's personal information to review and/or delete, this is made impossible due to the black box nature of training data and algorithms. The US Senate also passed the *Kids Online Safety and Privacy Act* (KOSPA) at the end of 2024, which was meant to update COPPA—however, the Act did not pass prior to the switch in government in January 2025.

Recognizing that the process to pass legally-binding policy interventions is lengthy, organizations like CITE in the state of California have introduced tools like the California Student Data Privacy Agreement.[51] Established to support the implementation of the California Education Code, this agreement acts as a practical tool for school districts looking to adopt new technologies with digital providers. It also aligns with the *Family Educational Rights and Privacy Act* (FERPA), COPPA, and California student privacy laws including the California Education Code and the *Student Online Personal Information Protection Act* (SOPIPA).[52] The creation of such agreements for practical implementation speaks to the need for youth-specific privacy and data protection regulations and tools.

## Global legal frameworks place specific obligations on companies and organizations with respect to data about children and teenagers

Organizations have a responsibility to release products that are designed with children's privacy in mind, from the ground up. Specific to genAI technologies, a company must carefully consider what types of user data to collect, use, and store, and efforts should be made to minimize the data collected. Article 5(1) of the EU GDPR states that "personal data shall be: c) adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed ('data minimisation')."[53] Recital 38 of the GDPR goes further to say that children merit special protection, and this in particular should apply to cases including the use of personal data of children to create user profiles and when using services offered directly to a child.[54] Standard 8 in the UK's AADC likewise states the need to collect and retain only the minimum amount of personal data.[55]

Beyond data collection and processing, embedding transparency and accessibility into the design of the tool and its user-facing interactions are crucial to protecting user privacy. In developing privacy policies and terms of service, companies should use language and layout of terms that is accessible and conducive to children and teenagers' understanding. Article 12(1) of the EU's GDPR states that any communication related to data processing should be communicated in a concise, transparent, intelligible, and easily accessible form, using clear and plain language, in particular for any information addressed specifically to a child."[56] Standard 4 in the UK's AADC similarly states that privacy information provided "must be concise, prominent and in clear language suited to the age of the child." The standard suggests providing "bite-sized" explanations at different points when the use of the data occurs.[57]

Companies can also take other explicit actions to mitigate privacy harms towards youth—including bans on the use of "nudge techniques" (design features which lead users to follow the designer's preferred paths in the user's decision making) and "dark patterns" (each designed to influence behaviour and decision-making), as stated in Article 25 of the EU's DSA[58], and requiring high privacy by default, as stated in Article 25 of the GDPR.[59]

**Beyond data collection and processing, embedding transparency and accessibility into the design of the tool and its user-facing interactions are crucial to protecting user privacy.**

## Existing interventions for children's safety and privacy have had mixed success

Due to the longer timeframes required to pass laws, a range of technical product design measures offer another option for intervention. For example, setting limits on user ages and establishing age verification features have been used by different social media platforms and digital services. Following an initial ban of OpenAI's ChatGPT in Italy due to privacy violations, OpenAI introduced age verification mechanisms to prohibit access for users under 13 years of age.[60] However, not only have age limitations been somewhat ineffective (due to the ease with which users can "trick" the age requirement), on the flip side, some mechanisms have been privacy-invasive in asking for biometric data and personal documentation.[61] Although age-assurance mechanisms remain a contested intervention, existing efforts include the UK's ICO Opinion on Age Assurance, and existing international standards bodies like International Organization for Standardization (ISO) continue to explore other approaches.[62]

Other potential privacy and data protection interventions include the use of timely pop-up messages or other "friction" points to warn users when their data is being collected and used, and incorporating child-friendly ways (e.g. diagrams, cartoons) to present privacy information.[63] Publicly available templates for child-friendly privacy notices have been developed by organizations like the educational consultancy Data Protection Education, offering animations and graphics for schools to use with their students.[64] Sector-specific guidelines are also necessary to appropriately support youth in different settings. Australia's eSafety Commissioner developed a National Safer Technologies 4 Schools (ST4S) Initiative, offering a standardized risk-assessment framework and toolkit for schools to evaluate digital tools and services.[65]

## Summary

- GenAI privacy and data protections for youth are protected by PIPEDA at the federal level. Provinces including Quebec, Alberta, and British Columbia also have their own privacy laws for public-sector institutions. The province of Ontario recently passed Bill 194, which outlines responsibilities of school boards when engaging in the processing of digital information about individuals under 18.

- Guidelines are being published on the use of AI systems at both the ministerial and school board level. These act as more ad-hoc policy interventions in addition to traditional privacy laws.

- A wide range of regulatory approaches are being implemented in key jurisdictions including the EU, UK, and the US. These interventions address different stakeholder responsibilities, the development of genAI tools at all stages of the product life cycle, and responsible use of genAI tools in the context of children and teenagers.

- Industry-led technical interventions to enact privacy and data protection for youth include age-assurance methods like age-gating and age verification, timely pop-up messages or friction points, and incorporating child-friendly ways to present privacy information. Practical guidance frameworks for education settings support school boards and educators looking to incorporate genAI technology in their schools.

# 4 Best Practices for Safeguarding Children and Teen Privacy

## Privacy design principles as a common foundation

Privacy design principles can offer a common foundation for incorporating privacy considerations into the design and use of technologies like genAI. While privacy design principles are non-binding and less detailed than design codes and laws and regulation, design principles offer a systematic, proactive approach to ensuring that privacy, data protection, and user rights are embedded in policies and technology products.

"Privacy by Design" was a framework developed by former Ontario Information and Privacy Commissioner Ann Cavoukian in the 1990s, with the aim to embed privacy into the design specifications of technology.[66] The framework is based on seven foundational principles:

1. Proactive not reactive; preventative not remedial
2. Privacy as the default setting
3. Privacy embedded into design
4. Full functionality - positive-sum, not zero-sum
5. End-to-end security - full lifecycle protection
6. Visibility and transparency - keep it open
7. Respect for user privacy - keep it user-centric

Privacy by Design has been recognized across Canada and internationally as one of the most widely used privacy frameworks both in the public and private sectors, and has been applied to information technologies, organization practices, and information architectures. Adherence to principles is often up to interpretation, and not legally binding. The unequal adoption of privacy design principles across sectors and users calls for firmer action and standards in best practices to implement said principles.

In addition to Cavoukian's principles, a variety of organizations have also released privacy design principles to guide their work, many of which are values-based. From prioritizing fairness, accountability, security, and transparency, companies committed to these principles have taken action to increase clarity of privacy policies and their terms of service, and school boards have issued guidelines for safe and ethical AI use in institutions.

The Privacy by Design principles offer an intuitive, common foundation for understanding the contributors to privacy as it relates to AI systems used by youth. They can more directly guide the development of AI governance models and AI products for communities of policymakers, technologists, and privacy practitioners. The following sections will explore best practices for protecting youth privacy geared to each of these groups.

## For policymakers: Youth interests as a priority for genAI governance

With the exception of Ontario's nascent Bill 194, there is no legal framework for governing youth privacy and genAI specifically in Canada today. For policymakers, genAI's rapid deployment—and the vulnerabilities for specific populations like children and youth—create urgency in establishing governance frameworks, with important lessons for Canada from other jurisdictions that are further ahead.

**Youth privacy should be specifically recognized in AI governance.** Privacy-by-design should be mandated for companies and the development of all tools and features. Youth understanding of privacy differs from that of adults, requiring specific measures to be put in place to truly obtain consent and to understand what is being agreed to.[67] Whether this is established through simplified privacy dashboards and prompts,[68] or a complete ban on the use of dark patterns in the development of a tool,[69] youth privacy requires additional attention and design measures beyond foundational privacy-by-design for adults.

**Future iterations of privacy laws should set standards for clear consent mechanisms, transparency, and data collection, use, and processing.**
These standards should be put in place with accountability mechanisms and appropriate penalties. Regulators should be timely and consistent in establishing a standard of acceptable risk levels. Establishing codes of practice that are specific to sectors that develop and deploy genAI tools can be one avenue for timely, actionable support for technologists and teams innovating new technologies. Specific to youth, limits to data collection should be established and enforced, akin to the EU's GDPR limitations on appropriate data collection and processing. These measures will ensure that companies and technologies cannot indiscriminately collect data for training purposes.

**Policymakers and governments must take an informed, participatory approach by working alongside educators, child development experts, and youth to develop future governance and safeguards.** This will ensure that aspects like standards for clear wording of privacy policies and terms of service, and mechanisms for obtaining consent are youth-friendly and conducive to protecting youth privacy. User testing for different types of privacy protection features with youth focus groups can help inform design decisions such as where, when, and how privacy notices should be displayed and written. Without involving youth in the development of policy interventions, any safeguards or policies will not adequately reflect the needs and lived experiences of young people. Regulatory sandboxes are a potential avenue for organizations to work alongside regulators in both the development of their technology and the testing of a policy intervention.

**For practical tips, see the accompanying tool:** Best Practices for Safeguarding Children's Privacy with GenAI Tools

## For technologists and privacy practitioners: Privacy considerations in the development and maintenance of genAI tools

Generative AI technology poses a new challenge for the preservation of privacy. For technologists and privacy practitioners, there are risks to understand and manage, but also proactive steps that can be taken.

**Privacy can be unknowingly compromised by sharing personal or sensitive information with a chatbot.** Australia's Office of the Victorian Information Commissioner (OVIC) published a report on the use of ChatGPT by a child protection worker, who had entered a significant amount of personal information relating to the child.[70] Concerns revolve around the "black box" of how inputted data is used to train generative AI models, the potential data leakage risks,[71] and the opportunity for cybersecurity attacks to retrieve inputted data.

**Cyber attacks pose a significant threat, with the potential to compromise children's privacy.** With 90 percent of successful cyber attacks on generative AI systems resulting in sensitive data leakage, action must be taken at every stage of development and maintenance of systems to mitigate privacy risks.[72] Our research found several privacy-related cyber attacks documented in literature and media, with the aim to leak data from the model, to bypass safety guardrails, or to poison the model's training data. While there are a wide range of attacks, the cyber threats can be grouped by two target vulnerabilities: attacks executed through inputs from users, and through training data sets. Both are described briefly below.

Attacks executed through inputs from users have the potential to cause generative AI systems to produce outputs that go against safety guardrails, leaking personal data or hateful/illegal material. For example, "jailbreaking" and "prompt injection" attacks use specific, intentionally-crafted prompts to bypass existing safety and moderation guardrails or content filters. These generate malicious outputs. They can include distribution of misinformation in its outputs, disclosure of sensitive information, and the execution of malicious code.[73] Model inversion attacks can extract or infer individuals' data used in the training data of the models. This inherent vulnerability created through user input has serious implications for personal data that may be accidentally entered into the systems.

Attacks executed by targeting the training data directly are akin to traditional concepts of cyber attacks. By polluting the model's training data, all subsequent outputs of the system are poisoned—potentially spreading false or offensive information, or revealing restricted data. "Data poisoning" attacks allow malicious actors to change the behaviour of a generative AI system—with backdoor data poisoning attacks injecting mislabelled training examples into a training set, resulting in the model learning the attacker's intended output.[74] For example, hundreds of poisoned datasets were discovered to have been uploaded onto the Hugging Face platform—a platform used for machine learning model collaboration –that have the potential to enable attackers to inject malicious code and impact many other models if downloaded and used.[75]

**Developers of generative AI tools can take proactive steps to protect the privacy of users.** Developers can make use of privacy-enhancing technologies (PETs) like data synthesis algorithms, differential privacy, and homomorphic encryption.[76] Although these technologies are not foolproof,[77] they are mitigating measures to implement as foundational protections. Data should be pre-trained and fine-tuned, should receive regular evaluations of its output filter, and should be minimized, anonymized, and de-identified.[78] For technologies that are youth-facing, efforts should be made to use synthetic data (artificially generated data rather than data produced by real-world events). If youth data is collected, easily comprehensible notices should be made at the initial access point of the tool, and data processing and time limits for retention should be transparent, accessible, and held accountable by a dedicated governing body.

Throughout the development of the technology, feedback and reporting mechanisms should be put in place for testing users to provide input to developers' further improvements of the model and safety assurance.[79] Feedback should continually be collected from users, even post-deployment. In the case of tools that may also be used by youth, companies should make efforts to create accessible and age-appropriate feedback mechanisms as well. This will provide helpful insights into any features that need to be adjusted or removed if in violation of privacy protections.

**Technologists should prioritize ongoing maintenance of genAI tools.**[80] This can include use of privacy impact assessments (PIAs), risk-assessment frameworks, executive accountability processes, and data protection impact assessments (DPIAs). Companies should also resource proactive monitoring for new trends and potential emergent harms that may elude existing enforcement systems. In the rush to build and introduce new features and products to the market, technologists and companies should recognize that prioritizing privacy-by-design from initial research and development stages will build a strong foundation for future products and its users to thrive.

## For educators and parents: Protecting youth privacy in their exploration and use of genAI tools

**For school administration, institute privacy impact assessments (PIAs) and develop robust cybersecurity.** These safeguards are important in determining which technologies to approve for use in schools. Before approving and procuring any technologies for student and educator use, PIAs should be conducted by a privacy expert/officer. School boards interviewed for this study including the Ottawa Catholic District School Board and the York Region District School Board. Each had their own PIAs, vetting approaches, and tools to explain different "levels" of approvals for educator use. While each school board had different grading methods and allowances for educator use of certain genAI tools, both stated the commitment to using PIAs before releasing genAI tools was non-negotiable.[81] This allows for longevity of the products procured, and takes the weight off of the educator's shoulders in assessing these tools. A dedicated cybersecurity team and robust cybersecurity plans should also be developed and kept updated, to proactively protect schools from new cybersecurity risks that may be introduced with any newly approved technologies.[82]

School boards looking to approve and implement new technologies should refer to legislation (for example Bill 194 in Ontario), and privacy-by-design principles. Beyond privacy risks, educators should also be informed of harms and risks associated with the use of genAI tools (e.g. potential for biased outcomes, hallucinations),[83] and be trained to act to reduce the risk of these harms (e.g. start conversations with students about why an output was generated; engage in critical thinking and fact-checking of outcomes).

In addition, school administration should review a school's resource availability and capacity to pilot tools prior to adoption, to consider structural power imbalances and inequity. Ask questions like: Can teachers co-develop the testing process for the tools? Can parent-teacher associations (PTAs) and students contribute to and contest this process? These questions will help school boards meaningfully assess tools that educators and students will use.

**For teachers and students in schools, foundational digital and media literacy training can be extended to include AI literacy.** School boards can provide opportunities for educators to learn about the board's plan to implement and approve different technologies, and to be informed about capabilities, privacy risks, and tips before using the tools. Resources like Stanford's CRAFT AI Literacy offer a great foundation for educators to help students explore and question AI. The Dais' Digital Literacy Toolkit (set to release in March 2025) offers digital literacy lesson plans for educators on topics like deepfakes, synthetic media, misinformation, and disinformation.[84]

**Offer standardized guides and "explainers" about genAI and privacy.** When updated iteratively to reflect the rapid change in the technology, these guides can offer a do-it-yourself resource for educators, students, and parents to use and reference. School boards like the Ottawa Catholic School Board[85] have developed their own guidance, creating both internal and externally accessible resources on their websites. The GenAI and Youth Privacy Tipsheet paired infographic likewise is one of these quick resources for educators. Resources like TeachAI's "AI Guidance for Schools Toolkit" provides an easy-to-use framework for schools looking to adopt AI tools responsibly and transparently.[86] Ensuring that educators are familiar and comfortable with the introduction of these tools enables students to be fully supported and to understand the safety precautions (e.g. not entering personal information or information beyond what is necessary) around using genAI tools. Training options and support to use alternative technologies should also be made available for educators who may not yet feel comfortable using these tools.

Whether through education or personal contexts, youth are increasingly interacting with genAI technologies in their daily lives. Tools offered to students and educators provide potential benefits, from personalized and iterative lesson plans and AI tutors to chatbots offering human-like interactions and companionship to users. For teachers and education administrators as well as for parents, there are steps that can be taken to mitigate the privacy risks that come with the use of these tools.

While at home, **parents should practice the following three steps to manage privacy risks** in the use of genAI tools by their children (and their own use).

1. **Read and understand the privacy notices of genAI tools you or your child might use.** While some privacy notices may be simpler to read than others, ensuring that you and your child understand and are comfortable with the data that will be collected, and how it will be used and stored, is crucial to fully providing consent before using a tool.

2. **Stay up to date with your digital, media, and AI literacy skills.** Knowing at a high level how generative AI works, as well as its capabilities, its risks and harms (e.g. bias, data collection, hallucinations), and how you can safely and effectively use the tools to your benefit will allow you to support your child's exploration and use of the tool. Staying up to date will equip you with the skills to navigate newer emerging features and products. Some of these resources can be found at this footnote.[87]

3. **Talk with your child about genAI tools and features they use on a regular basis.** Regardless of the tools your child uses, whether integrated genAI features on social media platforms like Snapchat's My AI and Meta AI via Facebook or Instagram, chatbots like Character.ai, or large language models (LLMs) like OpenAI's ChatGPT, Google's Gemini, or Microsoft's Copilot, knowing which tools your child is using means you can guide them in thinking through the opportunities and risks associated with each tools, and to advance in your AI literacy together.

**For practical tips, see the accompanying infographic: GenAI and Youth Privacy Tipsheet.**

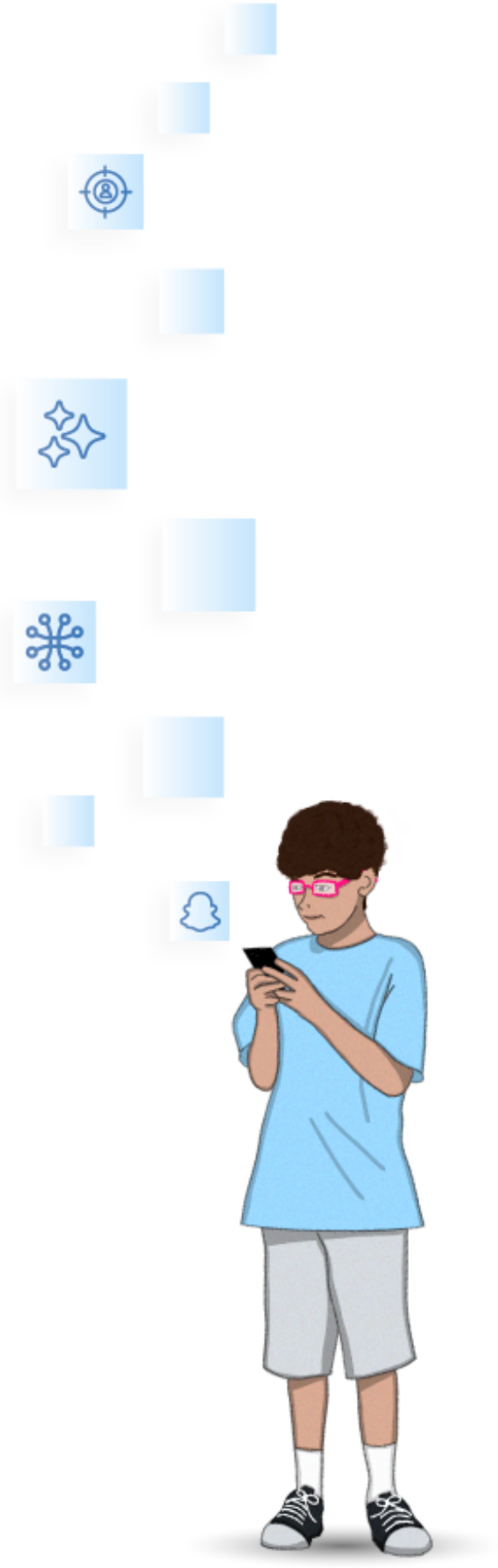
## Summary

- Privacy-by-design principles are a useful foundation for the development of genAI tools. In relation to genAI, companies should anonymize or pseudonymize data, minimize data being collected, conduct regular audits and third-party vetting, and clearly communicate what types of data are collected and for what purpose.

- Policymakers should prioritize and make specific arrangements to protect youth privacy. Privacy protections should be accessible to youth, have accountability mechanisms in place, and be co-developed with youth and relevant subject-matter experts.

- Technologists are core decision-makers in the development stage of any technology. Make use of, and actively review effective privacy-enhancing techniques, conduct ongoing maintenance in the form of privacy impact assessments (PIAs), and provide agile feedback and reporting mechanisms for users.

- Educators and parents play a key role in a child's development and understanding of technology. Be digitally fluent by understanding the capabilities and limitations of the technology to guide youth in their exploration and use of genAI tools.



# 5 Conclusion

Ensuring that the wider Canadian population benefits from emerging technologies like generative AI requires the development of corresponding guardrails to protect those who are at disproportionate risk of harm. In the case of generative AI, children and teenagers using these tools require special consideration. As such, this report assessed the privacy risks that children and teenagers face when using genAI technologies. We propose policy and technical interventions, as well as user best practices for parents, teachers, children, and teenagers.

The privacy-related harms that come along with generative AI tools include, but are not limited to, non-consensual collection and use of user data for alternate purposes, insufficient security measures leading to data breaches, CSAM content risks, and the potential for sensitive information to be disclosed due to differing perceptions of privacy and technology as a minor. While genAI technologies fall in scope with privacy laws in Canada, government ministries and school boards nationwide have also been releasing varying guidelines within the education sector. This indicates a growing need for practical, sector-specific guidelines to be developed and issued, to support users and organizations looking to adopt and use genAI technologies in different settings.

Educators, parents, children, and teenagers should be equipped with the knowledge and resources they need to responsibly and safely adopt and use new tools both at school and at home. As genAI tools continue to advance at lightning speed, it is crucial for technologists and policymakers to collaborate and discuss the best way forward to build out privacy-protecting policy interventions that are comprehensive and adaptable to the changing technology landscape.

# Endnotes


**1** Nomisha Kurian, "'No, Alexa, No!': Designing Child-Safe AI and Protecting Children from the Risks of the 'Empathy Gap' in Large Language Models," *Learning, Media, and Technology* (2024): 1-14, doi:10.1080/17439884.2024.2367052.

**2** Sara M. Grimes, Alissa N. Antle, Valerie Steeves, and Natalie Coulter, "Responsible AI and Children: Insights, Implications, and Best Practices," CIFAR, April 2024, https://cifar.ca/wp-content/uploads/2024/04/CIFAR-Responsible-AI-and-Children-EN_Final.pdf. As defined in this CIFAR report, dark patterns refer to "elements that draw on behavioural science, user data, and predictive algorithms to manipulate users into doing things they don't want to do."

**3** Yena Lee, "Thorn and All Tech Is Human Forge Generative AI Principles with AI Leaders to Enact Strong Child Safety Commitments," Thorn, July 16, 2024, https://www.thorn.org/blog/generative-ai-principles/.

**4** Siladitya Ray, "Samsung Bans ChatGPT Among Employees After Sensitive Code Leak," *Forbes*, May 2, 2023, https://www.forbes.com/sites/siladityaray/2023/05/02/samsung-bans-chatgpt-and-other-chatbots-for-employees-after-sensitive-code-leak/; James Vincent. "Apple Restricts Employees from Using ChatGPT over Fear of Data Leaks," *The Verge*, May 19, 2023, https://www.theverge.com/2023/5/19/23729619/apple-bans-chatgpt-openai-fears-data-leak; Catherine Thorbecke, "Don't Tell Anything to a Chatbot You Want to Keep Private," *CNN Business*, April 6, 2023, https://www.cnn.com/2023/04/06/tech/chatgpt-ai-privacy-concerns/index.html.

**5** "AI Language Models: Technological, Socio-Economic and Policy Considerations," OECD Digital Economy Papers, vol. 352, April 13, 2023, https://doi.org/10.1787/13d38f92-en.

**6** Elana Zeide and Helen Nissenbaum, "Learner Privacy in MOOCs and Virtual Education," *Theory and Research in Education* 16, no. 3 (2018): 280–307, https://doi.org/10.1177/1477878518815340.

**7** Aisha Malik, "Snapchat Taps Google's Gemini to Power its Chatbot's Generative AI Features," TechCrunch, September 24, 2024, https://techcrunch.com/2024/09/24/snapchat-taps-googles-gemini-to-power-its-chatbots-generative-ai-features/.

**8** Bernard Marr, "The Role of Generative AI In Video Game Development," *Forbes*, April 18, 2024, https://www.forbes.com/sites/bernardmarr/2024/04/18/the-role-of-generative-ai-in-video-game-development/; Rebecca Cairns, "'Video Games Are in for Quite a Trip': How Generative AI Could Radically Reshape Gaming," CNN, October 23, 2023, https://www.cnn.com/world/generative-ai-video-games-spc-intl-hnk/index.html.

**9** John Hendel, "Crisis Text Line Ends Data-Sharing Relationship with for-Profit Spinoff," *POLITICO*, January 31, 2022, https://www.politico.com/news/2022/01/31/crisis-text-line-ends-data-sharing-00004001.

**10** Elana Zeide, "The Structural Consequences of Big Data-Driven Education," *Big Data 5*, no. 2 (2017), https://papers.ssrn.com/abstract=2991794; Anjali A. Nambiar, "Securing Student Data in the Age of Generative AI: A Tool for Data Privacy Enhancement in K12 Schools," *MIT Responsible AI for Social Empowerment and Education*, 2024, https://raise.mit.edu/wp-content/uploads/2024/06/Securing-Student-Data-in-the-Age-of-Generative-AI_MIT-RAISE.pdf.

**11** Elana Zeide and Helen Nissenbaum, "Learner Privacy in MOOCs and Virtual Education," *Theory and Research in Education* 16, no. 3 (2018): 280–307, https://doi.org/10.1177/1477878518815340.

**12** Boris Lubarsky, "Re-Identification of 'Anonymized' Data," *Georgetown Law Technology Review*, April 2017, https://georgetownlawtechreview.org/re-identification-of-anonymized-data/GLTR-04-2017/.

**13** Liangyuan Na, Cong Yang, Chi-Cheng Lo, et al., "Feasibility of Reidentifying Individuals in Large National Physical Activity Data Sets From Which Protected Health Information Has Been Removed With Use of Machine Learning," *JAMA Network Open* 1, no. 8, (2018), https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2719130.

**14** Karl Manheim and Lyric Kaplan, "Artificial Intelligence: Risks to Privacy and Democracy," *Yale Journal of Law and Technology* 106, no. 21 (2019), https://yjolt.org/sites/default/files/21_yale_j.l._tech._106_0.pdf.

**15** Laura Sebben, "PowerSchool Data Breach: Toronto Public School Students from 1985 to 2024 Impacted," *CTV News*, January 20, 2025, https://www.ctvnews.ca/toronto/article/powerschool-data-breach-toronto-public-school-students-from-1985-to-2024-impacted/.

**16** "Sharing More About the Recent Incident," character.ai, December 17, 2024, https://blog.character.ai/sharing-more-about-the-recent-incident/.

**17** Mark Keierleber, "Whistleblower: L.A. Schools' Chatbot Misused Student Data as Tech Co. Crumbled," *The 74 Million*, July 1, 2024. https://www.the74million.org/article/whistleblower-l-a-schools-chatbot-misused-student-data-as-tech-co-crumbled/.

**18** Elana Zeide, "The Structural Consequences of Big Data-Driven Education," *Big Data 5* no. 2 (2017), https://papers.ssrn.com/abstract=2991794.

**19** Kathryn C. Montgomery, Jeff Chester and Tijana Milosevic, "Ensuring Young People's Digital Privacy as a Fundamental Right," in *International Handbook of Media Literacy Education* (Routledge, 2017); "Privacy Considerations for Generative AI," University of Illinois Urbana-Champaign, https://cybersecurity.illinois.edu/policies-governance/privacy-considerations-for-generative-ai/.

**20** Shaanan Cohney, Ross Teixeira, Anne Kohlbrenner, Arvind Narayanan, Mihir Kshirsagar, Yan Shvartzshnaider and Madelyn Sanfilippo, "Virtual Classrooms and Real Harms: Remote Learning at U.S. Universities," in *Proceedings of the Seventeenth USENIX Conference on Usable Privacy and Security* (2021): 653–73.

**21** Nicholas D. Santer, Adriana Manago, Allison Starks, and Stephanie M. Reich, "Early Adolescents' Perspectives on Digital Privacy," *Works in Progress*, June 29, 2021, https://wip.mitpress.mit.edu/pub/early-adolescents-perspectives-on-digital-privacy/release/1.

**22** Canada, Communications Security Establishment, "Why You Should Never Give Your Personal Information to AI," Get Cyber Safe, December 6, 2024, https://www.getcybersafe.gc.ca/en/blogs/why-you-should-never-give-your-personal-information-ai; Mozilla Foundation, "How to Protect Your Privacy from ChatGPT and Other AI Chatbots," *Privacy Not Included, July 18, 2024, https://foundation.mozilla.org/en/privacynotincluded/articles/how-to-protect-your-privacy-from-chatgpt-and-

other-ai-chatbots/; Jennifer King and Caroline Meinhardt, *Rethinking Privacy in the AI Era: Policy Provocations for a Data-Centric World*, Stanford University: Human-Centered Artificial Intelligence, February 2024, https://hai.stanford.edu/sites/default/files/2024-02/White-Paper-Rethinking-Privacy-AI-Era.pdf.

**23** Elana Zeide, "The Structural Consequences of Big Data-Driven Education." *Big Data 5*, no. 2 (2017), https://papers.ssrn.com/abstract=2991794.

**24** Michael Gerlich, "AI Tools in Society: Impacts on Cognitive Offloading and the Future of Critical Thinking." Societies 15, no.1 (2025): 6, https://doi.org/10.3390/soc15010006.

**25** Emily M. Bender, Timnit Gebru, Angelina McMillan-Major and Shmargaret Shmitchell, "On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?" in *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (2021): 610–23, https://doi.org/10.1145/3442188.3445922.

**26** Ziwei Ji, Nayeon Lee, Rita Frieske, Tiezheng Yu, Dan Su, Yan Xu, Etsuko Ishii, Ye Jin Bang, Andrea Madotto, and Pascale Fung, "Survey of Hallucination in Natural Language Generation," *ACM Computing Surveys 55*, no. 12 (2023): 248:1-248:38. https://doi.org/10.1145/3571730.

**27** Angus Crawford and Tony Smith, "Illegal Trade in AI Child Sex Abuse Images Exposed," *BBC News*, June 27, 2023, https://www.bbc.com/news/uk-65932372.

**28** Henry Fraser, "Deaths Linked to Chatbots Show We Must Urgently Revisit What Counts as 'High-Risk' AI," *The Conversation*, October 31, 2024, http://theconversation.com/deaths-linked-to-chatbots-show-we-must-urgently-revisit-what-counts-as-high-risk-ai-242289.

**29** "PIPEDA Requirements in Brief," Office of the Privacy Commissioner of Canada, https://www.priv.gc.ca/en/privacy-topics/privacy-laws-in-canada/the-personal-information-protection-and-electronic-documents-act-pipeda/pipeda_brief/; *Personal Information Protection and Electronic Documents Act*, SC 2000, c 5.

**30** "The Privacy Act in Brief," Office of the Privacy Commissioner of Canada, Privacy Act, RSC 1985, c P-21," https://www.priv.gc.ca/en/privacy-topics/privacy-laws-in-canada/the-privacy-act/pa_brief/.

**31** "PIPEDA Requirements in Brief," Office of the Privacy Commissioner of Canada, https://www.priv.gc.ca/en/privacy-topics/privacy-laws-in-canada/the-personal-information-protection-and-electronic-documents-act-pipeda/pipeda_brief/; *Personal Information Protection and Electronic Documents Act*, SC 2000, c 5.

**32** Bill C-27, *An Act to enact the Consumer Privacy Protection Act, the Personal Information and Data Protection Tribunal Act and the Artificial Intelligence and Data Act and to make consequential and related amendments to other Acts*, 1st Sess, 44th Parl, 2022.

**33** "About the OPC," Office of the Privacy Commissioner of Canada, April 8, 2024, https://www.priv.gc.ca/en/about-the-opc/.

**34** "PIPEDA Fair Information Principles," Office of the Privacy Commissioner of Canada, https://www.priv.gc.ca/en/privacy-topics/privacy-laws-in-canada/the-personal-information-protection-and-electronic-documents-act-pipeda/p_principle/.

**35** Bill 194, *Strengthening Cyber Security and Building Trust in the Public Sector Act*, 2024, 1st Sess, 43rd Parl, Ontario, 2024.

**36** Act respecting the protection of personal information in the private sector, P-39.1, Quebec, 2024.

**37** "Considerations for Using AI Tools in K-12 Schools," Ministry of Education and Child Care, https://www2.gov.bc.ca/assets/gov/education/administration/kindergarten-to-grade-12/ai-in-education/considerations-for-using-ai-tools-in-k-12-schools.pdf.

**38** "L'utilisation pédagogique, éthique et légale de l'intelligence artificielle générative – Guide destiné au personnel enseignant – 2024-2025," Ministère de l'Éducation Quebec, https://cdn-contenu.quebec.ca/cdn-contenu/education/Numerique/Guide-utilisation-pedagogique-ethique-legale-IA-personnel-enseignant.pdf.

**39** "Recommended Approaches to Generative Artificial Intelligence," Government of New Brunswick, 2024, https://plhub.nbed.ca/wp-content/uploads/sites/10/2024/02/Recommended-Approaches-to-Generative-AI.pdf.

**40** "WCDSB AI Guidelines," WCDSB, https://innovate.wcdsb.ca/ai/.

**41** "Artificial Intelligence at the OCSB," OCSB, August 23, 2024, https://www.ocsb.ca/why-ocsb/humane-use-of-technology/artificial-intelligence-at-the-ocsb/.

**42** Cheng-chi (Kirin) Chang, 2024, "When AI Remembers Too Much: Reinventing the Right to Be Forgotten for the Generative Age," 19 *Washington Journal of Law*, Technology & Arts 22 (2024), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4868555.

**43** *Generative AI: The Data Protection Implications*, CEDPO AI Working Group, October 16, 2023, https://cedpo.eu/wp-content/uploads/generative-ai-the-data-protection-implications-16-10-2023.pdf.

**44** "Intelligenza artificiale: il Garante blocca ChatGPT. Raccolta illecita di dati personali. Assenza di sistemi per la verifica dell'età dei minori," Garante per la Protezione dei Dati Personali, (March 31, 2023), https://www.garanteprivacy.it:443/home/docweb/-/docweb-display/docweb/9870847; Elvira Pollina and Alvise Armellini, "Italy fines OpenAI over ChatGPT privacy rules breach," *Reuters*, December 20, 2024, https://www.reuters.com/technology/italy-fines-openai-15-million-euros-over-privacy-rules-breach-2024-12-20/.

**45** Mindy Nunez Duffourc, Sara Gerke and Konrad Kollnig, "Privacy of Personal Data in the Generative AI Data Lifecycle," *JIPEL* 13, no. 2 (2024), https://jipel.law.nyu.edu/privacy-of-personal-data-in-the-generative-ai-data-lifecycle/; Tara Deschamps. "Can You Opt Out of Letting Meta's AI Chatbot Use Your Data? It's Not So Simple." *Global News*, June 14, 2024, https://globalnews.ca/news/10566594/meta-ai-opt-out-data-explainer/.

**46** "Guidelines 1/2024 on processing of personal data based on Article 6(1)(f) GDPR," European Data Protection Board, October 8, 2024, https://www.edpb.europa.eu/system/files/2024-10/edpb_guidelines_202401_legitimateinterest_en.pdf.

**47** *Regulation (EU) 2016/679 (General Data Protection Regulation)*, Art. 8.

**48** *Regulation (EU) 2024/1689 (Artificial Intelligence Act)*, Art. 5.

**49** *Regulation (EU) 2022/2065 (Digital Services Act)*; European Commission: Directorate-General for Communications Networks, Content and Technology; The Digital Services Act (DSA) Explained: Measures to Protect Children and Young People Online. Publications Office of the European Union, (2023), https://data.europa.eu/doi/10.2759/576008.

**50** Steve Wood, *Impact of Regulation on Children's Digital Lives*, Digital Futures for Children Centre, LSE and 5Rights Foundation, 2024, https://eprints.lse.ac.uk/123522/1/Impact_of_regulation_on_children_DFC_Research_report_May_2024.pdf.

**51** "Student Data Privacy," California IT in Education, https://www.cite.org/stuprivacy.

**52** "An Overview of the California Student Data Privacy Agreement," CITE, https://www.capousd.org/subsites/Purchasing/documents/Doing-Business/Vendor-Required-Forms-and-Registration/California-Student-Data-Privacy-Agreement.pdf.

**53** *Regulation (EU) 2016/679 (General Data Protection Regulation)*, Art. 5.

**54** *Regulation (EU) 2016/679 (General Data Protection Regulation)*, Recital 38.

**55** "Code Standards," ICO, December 2, 2024, https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/childrens-information/childrens-code-guidance-and-resources/age-appropriate-design-a-code-of-practice-for-online-services/code-standards/.

**56** *Regulation (EU) 2016/679 (General Data Protection Regulation)*, Art. 12.

**57** "Code Standards," ICO, December 2, 2024, https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/childrens-information/childrens-code-guidance-and-resources/age-appropriate-design-a-code-of-practice-for-online-services/code-standards/.

**58** *Regulation (EU) 2022/2065 (Digital Services Act*), Art. 25.

**59** Ibid.

**60** Elvira Pollina and Alvise Armellini, "Italy fines OpenAI over ChatGPT privacy rules breach," *Reuters*, December 20, 2024, https://www.reuters.com/technology/italy-fines-openai-15-million-euros-over-privacy-rules-breach-2024-12-20/.

**61** "Industry's Role in Promoting Kids' Online Health, Safety, and Privacy: Recommended Practices for Industry," National Telecommunications and Information Administration, November 7, 2024, https://www.ntia.gov/report/2024/kids-online-health-and-safety/online-health-and-safety-for-children-and-youth/taskforce-guidance/recommended-practices-for-industry.

**62** "Joint Statement on a Common International Approach to Age Assurance," Office of the Privacy Commissioner of Canada, September 19, 2024, https://www.priv.gc.ca/en/opc-news/speeches-and-statements/2024/js-dc_20240919/.

**63** "Code Standards." ICO, December 2, 2024, https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/childrens-information/childrens-code-guidance-and-resources/age-appropriate-design-a-code-of-practice-for-online-services/code-standards/; "How Does the Right to Be Informed Apply to Children?" ICO, December 4, 2024, https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/childrens-information/children-and-the-uk-gdpr/how-does-the-right-to-be-informed-apply-to-children/.

**64** "Child-Friendly Privacy Notices," DPE Knowledge Bank, April 16, 2021, https://dataprotection.education/freebies/child-friendly-privacy-notices.

**65** Safer Technologies 4 Schools, https://st4s.edu.au/.

**66** *Privacy by Design*, Information and Privacy Commissioner of Ontario, January 2018, https://www.ipc.on.ca/sites/default/files/legacy/2018/01/pbd-1.pdf.

**67** Jaemarie Solyst, Ellia Yang, Shixian Xie, Jessica Hammer, Amy Ogan and Motahhare Eslami, "Children's Overtrust and Shifting Perspectives of Generative AI," *arXiv*, June 29, 2024, https://doi.org/10.48550/arXiv.2404.14511; Giovanni Vespoli, Benedetta Taddei, Enrico Imbimbo, Lisa De Luca and Annalaura Nocentini, "The Concept of Privacy in the Digital World According to Teenagers," Journal of Public Health (2024), doi.org/10.1007/s10389-024-02242-x.

**68** Jasmine Irwin, Alannah Dharamshi and Noah Zon, *Children's Privacy in the Age of Artificial Intelligence*. Canadian Standards Association, 2021, https://www.csagroup.org/article/research/childrens-privacy-in-the-age-of-artificial-intelligence/.

**69** Dominique Kelly and Jacquelyn Burkell, "Identifying and Responding to Privacy Dark Patterns," *FIMS Publications* 385 (2024), https://ir.lib.uwo.ca/context/fimspub/article/1392/viewcontent/Final_Report_Identifying_and_Responding_to_Privacy_Dark_Patterns.pdf.

**70** "Please Note: Using AI Chatbot May Lead to Data Leaks," *Autoriteit Persoonsgegevens*, August 6, 2024, https://www.autoriteitpersoonsgegevens.nl/actueel/let-op-gebruik-ai-chatbot-kan-leiden-tot-datalekken; "OVIC Finds Department Responsible for Breaches of Privacy Through Use of ChatGPT," *Office of the Victorian Information Commissioner,* September 24, 2024, https://ovic.vic.gov.au/mediarelease/ovic-finds-department-responsible-for-breaches-of-privacy-through-use-of-chatgpt/.

**71** "Sharing More About the Recent Incident," *character.ai*, December 17, 2024, https://blog.character.ai/sharing-more-about-the-recent-incident/.

**72** Tony Bradley, "How GenAI Is Becoming A Prime Target For Cyberattacks," *Forbes*, October 10, 2024, https://www.forbes.com/sites/tonybradley/2024/10/10/how-genai-is-becoming-a-prime-target-for-cyberattacks/.

**73** Matt Sutton and Damian Ruck, "Indirect Prompt Injection: Generative AI's Greatest Security Flaw," Centre for Emerging Technology and Security, November 1, 2024, https://cetas.turing.ac.uk/publications/indirect-prompt-injection-generative-ais-greatest-security-flaw; *Artificial Intelligence: Generative AI Training, Development, and Deployment Considerations*. U.S. Government Accountability Office, October 22, 2024, https://www.gao.gov/assets/gao-25-107651.pdf.

**74** Antonio Emanuele Cinà, Kathrin Grosse, Sebastiano Vascon, Ambra Demontis, Battista Biggio, Fabio Roli and Marcello Pelillo, "Backdoor Learning Curves: Explaining Backdoor Poisoning Beyond Influence Functions," *International Journal of Machine Learning and Cybernetics* (2024), doi.org/10.1007/s13042-024-02363-5.

**75** David Cohen, "Data Scientists Targeted by Malicious Hugging Face ML Models with Silent Backdoor," *JFrog* (blog), February 27, 2024, https://jfrog.com/blog/data-scientists-targeted-by-malicious-hugging-face-ml-models-with-silent-backdoor/.

**76** *Generative AI: The Data Protection Implications*, CEDPO AI Working Group, October 16, 2023, https://cedpo.eu/wp-content/uploads/generative-ai-the-data-protection-implications-16-10-2023.pdf; T. Maheshwaran, "Privacy-preserving Computing: Balancing Privacy in the Digital Age," in *Exploring the Frontiers of Artificial Intelligence and Machine Learning Technologies* (2024), doi.org/10.59646/efaimltC10/133; Jordan Wilson and Katharina Koerner, hosts, "Ep: 90: How to tackle AI privacy and governance," Everyday AI (podcast), August 29, 2023, https://www.youreverydayai.com/

how-to-tackle-ai-privacy-and-governance/.

**77** Claudio Novelli, Federico Casolari, Philipp Hacker, Giorgio Spedicato and Luciano Floridi, "Generative AI in EU Law: Liability, Privacy, Intellectual Property, and Cybersecurity," *arXiv*, March 15, 2024, https://doi.org/10.48550/arXiv.2401.07348.

**78** Susan Hao, Piyush Kumar, Sarah Laszlo, Shivani Poddar, Bhaktipriya Radharapu and Renee Shelby, "Safety and Fairness for Content Moderation in Generative Models," *arXiv*, June 9, 2023, https://doi.org/10.48550/arXiv.2306.06135; "Principles for Responsible, Trustworthy and Privacy-Protective Generative AI Technologies," Office of the Privacy Commissioner of Canada, December 7, 2023, https://www.priv.gc.ca/en/privacy-topics/technology/artificial-intelligence/gd_principles_ai/.

**79** Susan Hao, Piyush Kumar, Sarah Laszlo, Shivani Poddar, Bhaktipriya Radharapu, and Renee Shelby, "Safety and Fairness for Content Moderation in Generative Models," *arXiv*, June 9, 2023, https://doi.org/10.48550/arXiv.2306.06135.

**80** *Generative AI: The Data Protection Implications*, CEDPO AI Working Group, October 16, 2023, https://cedpo.eu/wp-content/uploads/generative-ai-the-data-protection-implications-16-10-2023.pdf; *Guidance for Small Custodians on the Use of Artificial Intelligence*, Office of the Information and Privacy Commissioner of Alberta, November 2023, https://oipc.ab.ca/wp-content/uploads/2023/12/Guidance-for-Small-Custodians-on-the-use-of-Artificial-Intelligence-November-2023.pdf; *Generative AI and the EUDPR: First EDPS Orientations for Ensuring Data Protection Compliance When Using Generative AI Systems*, European Data Protection Supervisor, June 2024, https://www.edps.europa.eu/system/files/2024-06/24-06-03_genai_orientations_en.pdf

**81** The OCSB's grading scale for new technologies consists of 1) approved technologies, 2) approved with risk, and 3) not approved. The YRDSB's method uses a colour grading method to grade new technologies, and publishes a list of Green: approved technologies, Purple: use with considerations/educator guidelines provided, Yellow: use with considerations/educator guidelines provided and parent/guardian consent, and Red: not available/approved.

**82** Anjali A. Nambiar, "Securing Student Data in the Age of Generative AI: A Tool for Data Privacy Enhancement in K12 Schools," *MIT Responsible AI for Social Empowerment and Education*, 2024, https://raise.mit.edu/wp-content/uploads/2024/06/Securing-Student-Data-in-the-Age-of-Generative-AI_MIT-RAISE.pdf.

**83** Philippe Lorenz, Karine Perset and Jamie Berryhill, "Initial Policy Considerations for Generative Artificial Intelligence," OECD Artificial Intelligence Papers, No. 1., OECD Publishing, 2023, https://www.oecd-ilibrary.org/science-and-technology/initial-policy-considerations-for-generative-artificial-intelligence_fae2d1e6-en.

**84** "Empowering Students with AI Literacy," CRAFT, https://craft.stanford.edu/; The Dais' Digital Literacy Toolkit is set to release March 2025 on dais.ca.

**85** "AI at the OCSB," Ottawa Catholic School Board, https://www.ocsb.ca/why-ocsb/humane-use-of-technology/artificial-intelligence-at-the-ocsb/.

**86** "AI Guidance for Schools Toolkit," TeachAI, https://www.teachai.org/toolkit.

**87** Delaney Ruston, "5-Step Plan To Help Kids and Teens with AI Companion Chatbots," Screenagers, https://www.screenagersmovie.com/blog/5-step-plan-ai-chatbots; "Teaching Digital Citizenship," Cyber Civics, https://www.cybercivics.com.