

CARU Releases a Risk Matrix for Using Generative AI with Kids

Gonzalo E. Mon

November 1, 2025

The Children's Advertising Review Unit (or "CARU") recently released [Generative AI & Kids: A Risk Matrix for Brands & Policymakers](#), a framework to help companies that market to children mitigate potential risks related to their use of AI. Following are eight categories of risk CARU identified, as well as a high-level overview of some of the steps companies should consider to mitigate those risks.

1. **Misleading and Deceptive Advertising:** Companies should ensure that the use of AI doesn't mislead children about how products perform or blur the line between what's real and imaginary. They should build effective governance policies and review contracts with vendors to ensure their policies are aligned.
2. **Influencers and Endorsers:** Companies that use virtual influencers, avatars, or chatbots should review AI outputs to ensure any claims are supported. They should implement robust influencer review processes and use clear disclosures so that children who interact with chatbots know they aren't interacting with a real person. Companies should also build guardrails so children can't share personal data with virtual influencers, avatars, or chatbots.
3. **Privacy and Data Protection:** Companies should implement privacy-by-design, limit data collection, and obtain verifiable parental consent before collecting personal information from children. In addition to complying with laws such as COPPA, CARU recommends that companies not permit children's data to be collected, used, or disclosed by an AI model, including for training purposes.
4. **Safe and Responsible Use of AI:** Companies should know the source of their data, ensure humans are in the loop, and conduct regular bias impact assessments. They should also ensure that ads or events do not entice children to interact with strangers or target them for products that could isolate them or increase bullying.
5. **Mental Health and Development:** Companies should design with a focus on mental well-being, promote healthy screen use, and avoid addictive UX design and patterns. They should monitor the emotional impact of features on children and avoid designing chatbots to mimic human interaction.
6. **Manipulation and Commercialization:** Companies should clearly label ads and restrict behavioral targeting for children. They should implement ethical design practices and disable nudging and notification techniques for children.
7. **Exposure to Harmful Content:** Companies should use age-tiered filters and consider device-level age verification tools with shared family devices to signal to apps when a child is under 13. They should audit their moderation systems and allow users to flag or report problematic content.
8. **Lack of Transparency and Explainability:** When companies use algorithms, they should explain how AI models make decisions and how they are trained in a manner that is easy to

understand. Companies should provide regular reminders that a user is interacting with AI-based technology and ensure that they have clear consent mechanisms.

For companies using AI to market to children, the Risk Matrix should serve as a helpful roadmap of risks they should consider and steps they should take to mitigate those risks. It's likely also a roadmap of issues that CARU will focus on when it brings cases related to AI. At the end of the report, CARU invites readers to learn more about two cases they've already brought in this space. Odds are that we'll soon see more.