BBB National Programs
Children's Advertising Review Unit®

Center for Industry Self-Regulation

Generative AI & Kids:

A Risk Matrix for Brands & Policymakers

The Children's Advertising Review Unit (CARU) AI Working Group will continue its efforts under BBB National Programs' 501c(3), the Center for Industry Self-Regulation. Learn more at industryselfregulation.org/incubator.

Contents

Executive Summary

Artificial intelligence (AI) has been around for decades. Put simply, AI refers to a broad branch of computer science concerned with creating machines that can learn, make decisions, and perform tasks to a human-like level. AI functions by taking in data and using an iterative processing system and different algorithms to learn from patterns found in the data and then react to it in a specific way. Technologies such as the Roomba, IBM's Deep Blue and Watson, digital assistants such as Siri and Alexa, search engines, and social media platforms are all examples of AI that many people use daily.

More recently, the modern era of AI has quickly evolved to include generative AI, a type of AI named for its ability to generate innovative and realistic images, text, videos, and other media in response to a user's prompt or request. Generative AI creates content that, in the past, could only come from humans. Unlike past systems that were coded to respond to a set inquiry, generative AI continues to learn from and is trained on vast amounts of inputs (documents, photos, and more) that already exist online.

Generative AI evolves as it continues to train on more data. In the children's space, some examples are smart toys, smart speakers, toys that can provide personal responses to video games in the metaverse, and interfaces or tools that can provide personalized recommendations for movies, music, games, and toys based on a child's interests and preferences.

While generative AI has many benefits, *it also poses a myriad of risks,* especially for children, who are vulnerable to advertising messages and practices due to their limited knowledge, experience, sophistication, and maturity.

With this concern in mind, BBB National Programs' Children's Advertising Review Unit (CARU) issued a compliance warning in May 2024 to address the applicability of our Advertising and Privacy Guidelines to the use of generative AI in the children's space. In the compliance warning, CARU emphasized and reminded brands and industry that its guidelines apply to all online advertising, in any medium, directed to children under the age of 13 and any online collection of personal information from children in any online service that uses generative AI.

More recently, CARU convened a working group of industry experts to consider the potential risks that generative AI poses to children and tweens, such as misleading or deceptive advertising that uses AI, AI tools that collect personal, sensitive data from children, AI systems that produce biased and discriminatory output, and the potential that addictive AI tools may cause mental health issues or other harms to children.

The overall goal of our working group discussions is to develop guardrails and guidance for the use of generative AI in the children's space. This diverse Working Group of CARU Supporters, made up of well-respected global industry professionals representing toy, gaming, network, food, streaming, adtech, and mobile brands popular with children, joined with the CARU team to discuss online advertising, privacy, and safety issues brands and companies face when designing and developing their online services directed to children using generative AI.

As Dustin Ford, a member of the Working Group representing Hasbro, said, "Given the rapidly evolving landscape of AI, CARU's AI Working Group is timely, covering relevant and important topics in a valuable 'round table' format that encourages group discussion and thoughtful analysis."

Another member of the Working Group, Christopher Adams, Vice President at Nickelodeon, described his experience, stating, "CARU's AI Working Group, discussing the use of AI in the child-directed marketplace, is conducted in a helpful manner and hearing from external speakers' first-hand knowledge has been a great learning experience for me."

Using feedback and learnings from those discussions, CARU has organized the risks AI can pose to children into the following categories:

- Misleading and deceptive advertising
- Deceptive influencer and endorser practices
- Privacy invasions and data protection risks
- Bias and discrimination in the use of AI
- Harms to mental health and development
- Manipulation and over-commercialization
- Exposure to harmful content
- Lack of transparency

The **Generative AI & Kids: A Risk Matrix for Brands & Policymakers** is designed to help companies developing or deploying AI systems to identify and mitigate risks specific to children. It is also designed to help consumers understand the risks AI can pose to children and how to respond with informed, protective actions. The matrix maps key risk areas, real-life examples, potential harms, corporate responsibilities, and practical steps companies can take based on trusted guidance and leading global frameworks.

This matrix is intended to be a set of realistic, actionable considerations and best practices. It does not intend to resolve or respond to all the complex scenarios of using generative AI in the children's space, as the technology is evolving and each situation is unique and can be complex.

Misleading & Deceptive Advertising

IF YOU ARE...

Using AI to develop ads directed to children

YOU POTENTIALLY RISK...

Violation of the CARU Advertising Guidelines and Section 5 of the FTC Act

Helping a child establish a parasocial relationship with fake or imaginary characters, chatbots, or social companions

Developing unsafe or inappropriate advertising content, disinformation/misinformation

A child's developing an overreliance on technology for advice and companionship

Erosion of brand safety, trust, and reputation

IP threat or loss

Legal liability

ASK YOURSELF...

- Does the ad contain enhanced product imagery or performance claims (express or implied)?
- Does the ad blur the distinction between real and imaginary or fantasy experiences?
- Does the ad use AI-generated deep fakes or simulated elements, including the simulation of realistic people, places, or things?
- Does the ad use AI-powered voice cloning techniques and chatbots?

GUIDANCE/MITIGATIONS

- Ensure the AI enhancement does not mislead as to product characteristics, function, or performance
- Ensure that the ad does not mislead or blur the distinction between what is real and what is imaginary
- Build effective governance and compliance advertising policies
- Review contracts with third-party vendors, especially data and ad placement agreements, to ensure they have compliant AI advertising and privacy practices

RELEVANT FRAMEWORKS

CARU Advertising Guidelines

FTC Policy Statement on Deception

FTC Policy Statement Regarding Advertising Substantiation

Section 5 of the FTC Act

Influencers & Endorsers

IF YOU ARE...

Using child-directed social media, virtual influencers, digital avatars, or chatbots

YOU POTENTIALLY RISK...

- Violation of the CARU Advertising Guidelines and Section 5 of the FTC Act
- A child oversharing personal and private information with virtual influencers, avatars, and chatbots
- Helping a child establish a parasocial relationship with fake or imaginary characters, chatbots, or social companions
- Erosion of brand safety, trust, and reputation
- IP threat or loss
- Legal liability

ASK YOURSELF...

- Are the AI outputs misleading, deceptive, unsafe, or inappropriate for children?
- Can the child determine whether the AI influencer is a real person?
- Are you ensuring that your influencers are not making false or deceptive claims?
- Can a child determine that the chatbot is not a real person?

GUIDANCE/MITIGATIONS

- Thoroughly review any AI output and ensure there is substantiation for any AI-generated claims, including AI created visuals (which are claims), chatbots, avatars, etc.
- Develop a robust influencer review process that includes AI-generated content, including chatbots and AI-created avatars
- Use clear and meaningful disclosures to tell children they are communicating with an AI-generated tool and not a real person
- Develop moderation tools to ensure that children are not sharing personal information with the virtual influencer, avatar, or chatbot and are using them in a safe and responsible manner
- Ensure your advertising agreements do not promote questionable influencers or unknown content
- Beware of ad placements on live streams that involve child influencers
- Review your procurement and advertising spaces when it comes to advertising with questionable influencers or unknown content
- Test technology often to mitigate risks of bias and misinformation

RELEVANT FRAMEWORKS

CARU Advertising Guidelines

FTC's Children's Online Privacy Protection Act (COPPA) Rule

FTC Endorsement Guides and FAQs

FTC Policy Statement on Deception

Privacy & Data Protection

IF YOU ARE...

Using or creating AI-powered apps, AI toys, smart devices, voice assistants, learning tools, or educational apps

YOU POTENTIALLY RISK...

- Violation of the CARU Privacy Guidelines and the FTC's COPPA Rule
- Erosion of brand safety, trust, and reputation
- IP threat or loss
- Legal liability
- Data misuse, including how a third party may use AI prompts from children as training data
- Identity theft
- Profiling and deception
- Potential for child surveillance

ASK YOURSELF...

- Am I collecting, using, or disclosing children's personal information?
- Are my AI prompts and outputs collecting, using, or disclosing children's data or personal information?
- Am I aware of how third parties are using AI inputs and outputs from children, including whether they are used for AI-training purposes?

GUIDANCE/MITIGATIONS

- Embed privacy-by-design
- Limit data collection
- Obtain verifiable parental consent and provide direct notice to parents prior to any collection of personal information from children
- Be aware of the requirements of the COPPA Rule and align yourself with a trusted COPPA Safe Harbor like CARU
- Ensure secure storage and processing and end-to-end encryption
- Enable high privacy settings by default
- Best Practice: Do not permit children's data to be collected, used, or disclosed by an AI model, including for training purposes
- Communicate policies through mandatory training

RELEVANT FRAMEWORKS

CARU's Privacy Guidelines

FTC's Children's Online Privacy Protection Act (COPPA) Rule and the FAQs

Safe & Responsible Use of AI

IF YOU ARE...

Creating or using AI products directed to children, including chatbots, social companion apps, avatars, and toys

YOU POTENTIALLY RISK...

- Violation of the CARU Advertising and Privacy Guidelines and COPPA
- Bias and discrimination: Unfair or disparate treatment based on race, gender, or ability
- Brand safety and reputation: some AI-generated output uses popular children's characters in biased, unsafe, and inappropriate manners
- Online privacy and safety issues
- Profiling and deception
- Grooming
- Fraud, misinformation, or disinformation
- Cyberbullying
- Parasocial relationships
- Child sexual abuse material
- Overreliance on AI

ASK YOURSELF...

- Are we using facial recognition technology?
- Are we using content filtering and content moderation tools?
- Is my online service allowing children to create their own avatars (uploading photo)?
- Are we using the microphone and obtaining verifiable parental consent prior to any collection of children's personal information?
- Are we using the camera and obtaining verifiable parental consent prior to any collection of children's personal information?

GUIDANCE/MITIGATIONS

- Ensure human-in-the-loop oversight
- Know the genesis of and diversify the training data
- Conduct regular bias impact assessments
- Be mindful of the vendors you use and have a comprehensive vetting process in place
- Ensure ads or events do not send or entice children to interact with strangers or target them for products that could isolate or increase bullying
- Consider spearheading an AI Greenlight Committee for new AI products

RELEVANT FRAMEWORKS

CARU Advertising Guidelines

FTC's Children's Online Privacy Protection Act (COPPA) Rule

FTC Policy Statement on Deception

Section 5 of the FTC Act

Notable FTC cases EPIC Games and NGL

UNICEF AI for Children

OECD AI Principles

Mental Health & Development

IF YOU ARE...

Developing or using chatbots, social companions, virtual influencers, recommendation engines in social or gaming platforms, social media algorithms, or endless video feeds

YOU POTENTIALLY RISK...

- Violation of the CARU Advertising and Privacy Guidelines
- Erosion of brand reputation and trust
- Legal liability
- IP theft or loss
- Creating an environment for social isolation, body image issues, low self-esteem, anxiety, attention disruption, addictive behaviors, anxiety, reduced attention span, emotional manipulation
- Overreliance on AI undermining critical cognitive development and thinking skills

ASK YOURSELF...

- Am I creating toys, apps, or products and services that take the place of human interaction?

GUIDANCE/MITIGATIONS

- Avoid addictive UX design and patterns
- Use effective human and AI moderation tools
- Implement digital wellness features
- Monitor emotional impacts
- Promote healthy screen use
- Design and test regularly for well-being
- Avoid designing chatbots to mimic human interaction

RELEVANT FRAMEWORKS

CARU Advertising Guidelines

Section 5 of the FTC Act

APA Guidelines

UNICEF AI for Children

Common Sense Media

Manipulation & Commercialization

IF YOU ARE...

Using AI for personalized ads, gamified purchases, AI influencers, targeted ads in games, YouTube, apps, or selling or targeting ads to children

YOU POTENTIALLY RISK...

- Violation of the CARU Advertising and Privacy Guidelines, COPPA, and FTC Endorsement Guides
- Unethical targeting
- Loss of autonomy
- Excessive spending
- Overconsumption
- Manipulation or pressure for kids to buy things or watch ads
- Hidden marketing

ASK YOURSELF...

- Am I collecting, using, or sharing children's personal information to create personalized marketing or ads in games or apps?

GUIDANCE/MITIGATIONS

- Restrict behavioral targeting for children
- Provide transparent and clear disclosures of advertising
- Disable nudging and notification techniques
- Implement ethical design practices

RELEVANT FRAMEWORKS

CARU Advertising Guidelines
FTC's Children's Online Privacy Protection Act (COPPA) Rule
FTC Policy Statement on Deception
FTC Policy Statement Regarding Advertising Substantiation
Section 5 of the FTC Act
UK ICO Children's Code
FTC Endorsement Guides and FAQs

Exposure to Harmful Content

IF YOU ARE...

Using AI-generated content or videos, user-generated platforms, AI chatbots or apps, or game forums

YOU POTENTIALLY RISK...

- Violation of the CARU Advertising and Privacy Guidelines
- Access to age-inappropriate, false, or unsafe content, such as sexual imagery, alcohol or tobacco, violence, misinformation and disinformation, deepfakes
- Child sexual abuse material

ASK YOURSELF...

- Are we using responsible and trusted human and AI content moderation tools?
- Do we monitor AI content moderation failures in videos, AI apps and game forums, synthetic media, chatbots, and AI voice-cloning?

GUIDANCE/MITIGATIONS

- Use age-tiered filters
- Consider device-level age verification tools with shared family devices to signal to apps when a child is U13
- Strengthen and frequently audit your AI and human moderation systems
- Enable reporting and flagging systems
- Ensure you are using verified content sources
- Create AI transparency and explainability tools

RELEVANT FRAMEWORKS

CARU Advertising Guidelines

CARU's Privacy Guidelines

IEEE Trustworthy AI

EU Digital Services Act

Family Online Safety Institute

Lack of Transparency & Explainability

IF YOU ARE...

Using algorithms

YOU POTENTIALLY RISK...

- Violation of the CARU Advertising Guidelines
- Erosion of brand safety, trust, and reputation
- Misleading consumers about how decisions are made or if they are accurate

ASK YOURSELF...

- Are the algorithms we use in social media, apps, chatbots, or social companions opaque?

GUIDANCE/MITIGATIONS

- Implement explainability tools
- Provide regular reminders that a user is interacting with a non-human, AI-based technology
- Offer child and consumer-friendly privacy policies, direct notice, and other disclosures for families
- Ensure clear opt-in and opt-out consent mechanisms are in place as well as a right to delete and appeal

RELEVANT FRAMEWORKS

CARU Advertising Guidelines

FTC Policy Statement on Deception

ISO/IEC 23894

OECD AI Principles

Smart Choices: A Parents' Guide to Navigating the Risks of Generative AI

In a world increasingly shaped by artificial intelligence (AI), today's parents and caregivers face new challenges—and opportunities—when it comes to guiding their children through a digital landscape that is evolving faster than ever. This tip sheet provides a place to start a conversation with your child.

The Known Risks	What You Can Do
AI tools may sound smart—but they can also generate inaccurate, misleading, or deceptive advertising content and encourage unwanted purchases.	• Use available filters and tools. Be aware that choices are often highlighted to nudge users to select the least restrictive settings. Choose what is best for your household. • Teach children how to recognize ads and persuasive intent (i.e. what not to click on that may result in an unwanted purchase or redirect your child elsewhere). • Consider ad-free or paid versions of streaming services, mobile games, and other apps or platforms. • Turn off in-app purchases and ensure your credit card is not set as a default payment. • Align your child's viewing and playing habits with trusted brands and companies.
Many AI tools (such as AI-powered apps, AI toys, smart devices, voice assistants, learning tools, and/or educational apps) collect user data upon login. Kids often enter personal, sensitive data without knowing it.	• Turn off mics and cameras when not in use or necessary. • Consider requiring approval for your child's app downloads. • Discuss good digital citizenship and what your child shares online. • Use child accounts with high privacy settings by default. • Read online notices and privacy policies. In particular, you want to know WHAT data is being collected, WHY it is being collected, and WHO it is being shared with. • Ask questions of companies before you allow your child to use the online service.
AI systems are trained on human data—which includes bias. That means tools may unintentionally favor certain languages, cultural norms, or demographics.	• Talk to your child about AI and bias. • Learn how to fact-check the output of the AI tool and share your findings with your child, teaching them to question the veracity of AI. • Learn about how AI tools are used in media and advertising directed to children. • Look for trusted, reviewed, and transparent platforms and tools, particularly those that permit reporting inaccuracies.

Smart Choices: A Parents' Guide to Navigating the Risks of Generative AI

The Known Risks	What You Can Do
Sites, tools, platforms, and apps that use AI can be addictive and cause various mental health issues, especially when tools offer instant results, images, and games or take the place of human interaction.	• Set screen time limits for AI tools. • Co-view with your child and talk about potential harms. • Create safe spaces for children to ask questions at home and at school. • Talk about online experiences regularly.
While many platforms offer filters and content moderation tools, AI can still generate mature, unsafe, or inappropriate results.	• Use available filters and tools. • Choose trusted online sites, services, and platforms. • Encourage conversations about things your children see or experience online. • Familiarize yourself with reporting tools and teach your child how to use them in the event of potentially dangerous or inappropriate activity.

Learn more about the Children's Advertising Review Unit and two of its cases related to AI in the children's space: KidGeni (August 2024) and Buddy AI (February 2025).

BBB National Programs
Children's Advertising Review Unit®

BBB National Programs' Children's Advertising Review Unit (CARU) helps companies comply with laws and guidelines that protect children from deceptive or inappropriate advertising and ensure that, in an online environment, children's data is collected and handled responsibly. CARU seeks change through the voluntary cooperation of companies and where relevant, enforcement action. CARU is the nation's first Safe Harbor Program under the Children's Online Privacy Protection Act (COPPA), ensuring that participating companies' products and services are compliant with the complex federal law on children's privacy.

Center for Industry Self-Regulation

The Children's Advertising Review Unit (CARU) AI Working Group will continue its efforts under BBB National Programs' Center for Industry Self-Regulation (CISR). CISR is dedicated to education and research that supports the development and implementation of fair, future-proof best practices, and to the education of policymakers and the public regarding industry self-regulation as viable alternative to government regulation.