beazley

Article

Safeguarding children in the world of AI

Nisar Siddiqui • September 10, 2024

As the capabilities and ubiquity of Generative AI have exploded over the past year, the use of the technology to generate child sexual abuse material (CSAM) has been an issue of growing concern. Since **we first explored this accelerating risk** in 2023, an increase in cases of improper image generation and utlization has raised public awareness and driven greater legislative and regulatory scrutiny, altering the risk profile.

Publicly available AI image generators have changed the ease – and nature – of this activity.

Initially, this was a trend driven by adults seeking to create and share pornographic content, perhaps under the misapprehension that it would not constitute criminal activity. Today, teens' use of deepfake technologies has become more prevalent as social media and mobile apps have made it possible to access and alter images right from their phones. Where once a pedophile might have taken a pornographic image and superimposed a child onto it, access to content generators in tools where kids already post their own likenesses has made it far too easy for them to use photos of their peers to create pornographic and other inappropriate images.

As a result of this shift, what was a speculative risk less than a year ago is now increasingly widespread. There has been significant press coverage focusing on incidents involving schools in particular, with stakeholders questioning how schools are prepared to discourage, monitor, and respond to these occurrences.

The online child harm numbers are staggering.

2023 was the year of most reported CSAM incidents on record, according to the UK-based Internet Watch Foundation. A global deep dive into online abuse from the Childlight Global Child Safety

Institute at the University of Edinburgh has revealed even more troubling statistics. Their **Into the Light Index** was recently released, revealing the unprecedented statistic that 1 in 8 children worldwide have been the victim of non-consensual image offenses (12.6%) and online solicitation (12.5%).

These trends are enabled by the evolving technology. As AI's capabilities improve, it's increasingly possible to create hyper-realistic CSAM images, regardless of efforts to implement guardrails. Many mainstream AI companies claim to have added safeguards making it more difficult to create this material using the newest versions of these tools. But with nothing to prevent people from downloading earlier or unrestricted versions of the tools without these protections in place, this is proving insufficient.

Criminal and civil law is quickly developing around these issues.

Two recent cases which took place in **North Carolina** and **Pennsylvania** involved investigation and charges in relation to the creation of AI deepfakes, where the faces of real children were superimposed on images. Then, in May of 2024, **the first federal charge involving images entirely generated using AI took place in Wisconsin**. This case highlighted a little-tested legal avenue which argues that AI-invented images should be treated in the same way as real-world incidents of recorded child sex abuse, setting legal precedent. We anticipate legislative and regulatory activity around this issue may increase as a result.

On the civil side, parents of minors victimized within youth-serving organizations, such as schools, are righteously angry and looking to sue, bringing claims against the perpetrators, their parents, and even the organizations for negligent supervision and inadequate or delayed response.

Remaining silent about these issues is no longer feasible for youth-serving organizations.

There is a huge need for education and guardrails as schools address the need to raise awareness of the risks and to help staff notice the signs. Counselors and teachers must be trained to be more alert to subtleties; with speed of the essence, they must be prepared to ask the right questions immediately if they become aware of potentially harmful conduct, including the distribution or sharing of images. In the event of an incident, administrators must be prepared to respond quickly and appropriately.

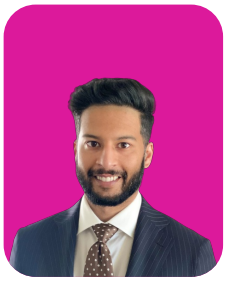
Awareness also must be raised with students and parents. It may be tempting to avoid talking about such difficult subjects with students, but they need to understand the ramifications and consequences of using AI tools to create sexually explicit images.

Administrators should also consider their policies for displaying student likenesses on the school website and social media accounts. Given the ease with which any publicly available image can be altered, robust policies and procedures regarding how, when, and where images are

shared must be developed, maintained, and implemented.

Consider resources like Safeguard to help mitigate this exposure.

No youth-serving organization is immune to the risks of CSAM deepfakes today, and if trends continue, many more institutions will soon find themselves at the center of an incident. Key resources to manage these risks, including advanced risk management and proactive crisis response services, are available with every **Safeguard** policy, and our team is always happy to speak with brokers about how to help their clients best manage this growing concern.



Nisar Siddiqui

Underwriter - Safeguard

BZSL 228.

The information set forth in this communication is intended as general risk management information. Beazley does not render legal services or advice. It should not be construed or relied upon as legal advice and is not intended as a substitute for consultation with counsel. Although reasonable care has been taken in preparing the information set forth in this communication, Beazley accepts no responsibility for any errors it may contain or for any losses allegedly attributable to this information. The descriptions contained in this communication are for preliminary informational purposes only. The product is available in the US on a surplus lines basis only, through either Beazley Excess and Surplus Insurance, Inc. or a licensed surplus lines broker underwritten by Beazley syndicates at Lloyd's. The exact coverage afforded by the product described herein is subject to and governed by the terms and conditions of each policy issued. The publication and delivery of the information contained herein is not intended as a solicitation for the purchase of insurance on any US risk. Beazley USA Services, Inc. is licensed and regulated by insurance regulatory authorities in the respective states of the US and transacts business in the State of California as Beazley Insurance Services (License#: 0G55497). Non-insurance products and services are provided by non-insurance company Beazley affiliates or independent third parties. Separate terms and conditions may apply.

beazley

©Beazley Group | LLOYD's Underwriters