

AGs Protect Children from AI (and Chainsaws)

Paul L. Singer, Abigail Stempson, Beth Bolen Chun

April 17, 2024

Last week in South Carolina, AGs, staff, and members of the community gathered for the AI and Preventing Child Exploitation Seminar, presented jointly by the Attorney General Alliance (AGA) and the National Association of Attorneys General (NAAG). Sessions focused on robocalls, online platforms, youth digital wellness and mental health, and the potential benefits of AI.

Attorney General Perspective

The first panel “AI and Exploitation of Children” featured South Carolina Attorney General Alan Wilson and members of his staff, Whitney Michael, Senior Advisor, Joseph Spate, Assistant Deputy Solicitor General, and Kyle Senn, Senior Assistant Attorney General. This panel provided an excellent summary of the perspective of State AGs on combatting child exploitation and how AI can both harm and benefit society.

AG Wilson explained that social media and AI are replacing tobacco and opioids as the new bipartisan issues, with AGs, including Oregon Attorney General Ellen Rosenblum, working to keep the topics at the forefront. He explained that providing personal information over the internet is now expected and natural, and while our ability to protect ourselves has increased, so has the ability to hack. Unfortunately, predators also take advantage of the fact that children are comfortable providing information over the internet, and they use a variety of online platforms to exploit people on the Internet.

AG Wilson said it best when he compared AI to a chainsaw – a valuable tool in the hands of a lumberjack, but a deadly weapon in the hands of Jason Voorhees. He touted the joint [54 state and territory letter](#) to Congress spearheaded in part by his office. The letter, cosponsored by Oregon, North Carolina, and Mississippi, asked Congress to help evolve the legal landscape in light of changing technology, which he described as both amazing but capable of incomprehensible feats. AGs are working together to fill in the gaps in current laws to prevent and enforce against a variety of ways AI can be used to create child sexual abuse material (CSAM). In the wake of this letter, AG Wilson explained that Congress is now setting up an ad hoc committee to study AI, and other bipartisan bills are dropping.

Industry Thoughts

We heard on other panels from industry representatives how they are working to address child exploitation. One gaming platform described a range of tools including AI moderation in combination with human moderators to help combat child exploitation. It uses automated chat filtering for personal information and machine-learning to remove inappropriate language in violation of community standards. The platform scans each image upload using AI to ensure it is appropriate and compares it to hashed National Center for Missing and Exploited Children (NCMEC) databases. The

platform does not allow images of real life people and provides account monitoring by parents for users under 18. Finally, the platform reports to the FBI and NCMEC using automated tools and escalates review of “trusted flagger” reports. They use a law enforcement response tool to speed up subpoena response times.

One AI platform described its safety-first principles as it seeks to benefit humanity. Policies outline the appropriate use, and the company constantly evaluates risk from pre-training to launch to ongoing monitoring. Pre-training excludes adult content, dark web, payroll and other content from data aggregators. Post training, automated and human evaluators work to tune the AI so it behaves in accordance with policies, such as refusing to answer when appropriate to avoid providing harmful material or personal information.

Conclusion

AGs, social and online platforms, and AI programs themselves are working to combat the dark side of AI including child exploitation. However, if third-party platforms or AI companies themselves fail to implement appropriate safeguards for children, it is likely they will encounter an AG inquiry in the civil or criminal realm.