Why Safer?

• Detection service installed on client infrastructure with user privacy in mind
• Comprehensive coverage to identify both known and unknown CSAM at scale
• Increase Trust & Safety team’s efficiency.
• Break down silos and leverage community knowledge
• Scale platform detection with user privacy in mind

Product overview

Safer is a complete solution that automates the detection and reporting of child sexual abuse material (CSAM) from content hosting platforms at scale, allowing companies to protect their communities while supporting victims of child sexual abuse by stopping the circulation of their abuse content.

Product features

Hash matching
Remove the threat of known CSAM on your platform with the power of Safer hash matching, fueled by our growing collaborative dataset of over 10 million hashes

CSAM Classifier
With Thorn’s CSAM Classifier integrated into Safer, customers can now detect potentially new, unknown CSAM, as well as improve mitigation of false positives and streamline perceptual matches

Video Detection
Safer detects abuse imagery, as well as video content, keeping your platform safety ahead of the curve as the number of reported CSAM video files reported to NCMEC continue to grow

Reporting API
Reporting API: Leverage intelligence of CSAM flagged by Safer’s Detection services to streamline reporting and removal based on company policy and internal processes

Additional Resources

• Learn More
• Safer for AWS
• Safer: Building the internet we deserve
**How it works**

Identify known and unknown CSAM with perceptual hashing and machine learning algorithms. Queue flagged content for review with content moderation tools built with employee wellness in mind. Review and report verified CSAM and securely store content in accordance with regulatory obligations. Contribute hashes of new content to the Safer community.

**Differentiators**

- Safer leverages a unique machine learning classification model that targets unknown/unreported files that are suspected to be CSAM.
- SaferList, a service for Safer customers that facilitates information-sharing across the industry and strengthens the entire ecosystem with every new file found.
- All the child abuse content Safer can find can be reviewed and reported within a single interface, prioritizing getting critical information into the hands of front-line responders so they can identify children faster and get to them sooner.
What our customers are saying

“Being able to scan Flickr in real-time for child abuse content provides our human moderation team critical intelligence.”

- Holly Lawrence, Trust & Safety Manager, Flickr

Data Points

- **18M+**
  Hashes of known CSAM in database

- **300K+**
  CSAM files identified on customer platforms since launching

- **>99%**
  Precision to detect unreported CSAM based on data science bench tests

Solution available in [AWS Marketplace](https://aws.amazon.com/marketplace)