

GOOGLE IN EUROPE

# Using AI to help organizations detect and report child sexual abuse material online

Nikola Todorovic  
Engineering Lead

Abhi Chaudhuri  
Product Manager

Published Sep 3, 2018

Using the internet as a means to spread content that sexually exploits children is one of the worst abuses imaginable. That's why since the early 2000s we've been investing in technology, teams, and working closely with expert organizations, like the [Internet Watch Foundation](#), to fight the spread of child sexual abuse material (CSAM) online. There are also many other organizations of all sizes that are deeply committed to this fight—from civil society groups and specialist NGOs to other technology companies—and we all work to ensure we share the latest technological advancements.

Today we're introducing the next step in this fight: cutting-edge artificial intelligence (AI) that significantly advances our existing technologies to dramatically improve how service providers, NGOs, and other technology companies review this content at scale. By using deep neural networks for image processing, we can now assist reviewers sorting through many images by prioritizing the most likely CSAM content for review. While historical approaches to finding this content have relied exclusively on matching against hashes of known CSAM, the classifier keeps up with offenders by also targeting content that has not been previously confirmed as CSAM. Quick identification of new images means that children who are being sexually abused today are much more likely to be identified and protected from further abuse.

We're making this available for free to NGOs and industry partners via our Content Safety API, a toolkit to increase the capacity to review content in a way that requires fewer people to be exposed to it. Susie Hargreaves OBE, CEO, Internet Watch Foundation, said: "We, and in particular our expert analysts, are excited about the development of an artificial intelligence tool which could help our human experts review material to an even greater scale and keep up with offenders, by targeting imagery that hasn't previously been marked as illegal material. By sharing this new technology, the identification of images could be speeded up, which in turn could make the internet a safer place for both survivors and users."

We agree. This initiative will allow greatly improved speed in review processes of potential CSAM. We've seen firsthand that this system can help a reviewer find and take action on 700% more CSAM content over the same time period.

We've been [investing for years](#) to tackle this challenge, developing technology to detect CSAM in ways that are precise and effective. We've also been [working across the industry](#) and with NGOs to combat CSAM through our work in the [Technology Coalition](#) since 2006, with the [WePROTECT Global Alliance](#), and through industry-wide initiatives aimed at sharing best practices and known hashes of CSAM.

Identifying and fighting the spread of CSAM is an ongoing challenge, and governments, law enforcement, NGOs and industry all have a critically important role in protecting children from this horrific crime. While technology alone is not a panacea for this societal challenge, this work marks a big step forward in helping more organizations do this challenging work at scale. We will continue to invest in technology and organizations to help fight the perpetrators of CSAM and to keep our platforms and our users safe from this type of abhorrent content. We look forward to working alongside even more partners in the industry to help them do the same.

If you're interested in using the Content Safety API service at your organization, please reach out to us via [this form](#). ■

POSTED IN: [GOOGLE IN EUROPE — SAFETY AND SECURITY](#)

- HIDE RELATED ARTICLES



RELATED ARTICLES



Celebrating 15 years of Google Ireland



Meet Alex, the Spanish jobseeker whose app brought work to thousands

Making it easier to control your data, directly in Google products

Opening up opportunities in the



