

Partnering for a **safer** Internet: the role of APIs

To protect users and other organisations from bad actors, we build and share Application Programming Interfaces (APIs) that detect and block online threats, and help make the internet safer for everyone.



BUILD

We create APIs for safety to protect in 3 key areas.

Some examples of these APIs are:

Child Safety

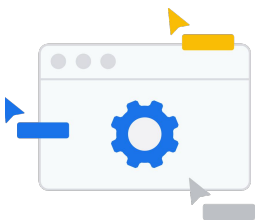
- Content Safety API
- CSAI Match API
- Hash Matching API

Security

- Safe Browsing API
- Project Shield API
- VirusTotal API

Information Quality

- Perspective API
- Vision API
- FactCheck API
- Civics Information API
- Text Moderation



SCALE

To scale our impact, we share our APIs for safety with partners:

46 million

In 2022, the Cloud Armor API blocked one of the largest layer 7 DDoS attacks in history at 46 million requests per second.

2 billion

The Perspective API is called by 1000+ partners almost 2 billion times a day.

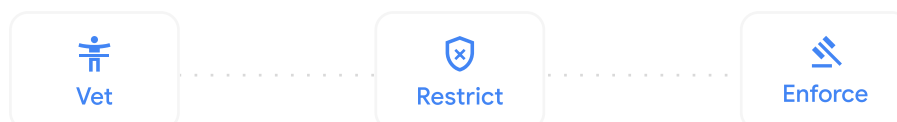
4 billion

The Child Safety Toolkit (the Content Safety API + CSAI Match API) processed over 4 billion images and videos in the last 30 days.



LIMIT

To ensure that only trusted partners have access to our APIs and that they are used in a compliant and safe manner, we take 3 critical steps:



We limit access of our safety APIs to carefully vetted partners, restrict unauthorised usage of these tools, and strictly enforce our terms of service policies. Empowering our trusted partners with safety APIs allows us to work together, and create a safer internet for everyone.

Google's APIs for safety



Child Safety APIs

Child Safety Toolkit

Content Safety

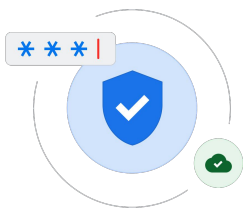
Google's proprietary model helps partners classify and prioritise billions of child sexual abuse material images. It uses ML classifiers to help identify never-before-seen CSAM imagery for fast review, removal and reporting.

CSAI Match

YouTube's proprietary technology (and the first technology to use hash matching) to identify child sexual abuse imagery (CSAI) videos and systematically flag these for partners to review, confirm, report and act upon.

Hash Matching

We provide Google's Hash Matching API to NCMEC to help them prioritise and review CyberTipline reports more efficiently, allowing them to hone in on those reports involving children who need immediate help.



Security APIs

Google's WAAP Solution

Cloud Armor

Protects websites, services and APIs against advanced, targeted and automated distributed denial of service (DDoS) attacks, which censor information.

reCAPTCHA

Defends websites from fraudulent activity, spam and other abuses. reCAPTCHA Enterprise uses an adaptive risk analysis engine to keep automated software from engaging in abusive activities.

Apigee

Google Cloud's full lifecycle API management platform enables businesses to design, secure, deploy, monitor and scale APIs.

Safe Browsing

Our Safe Browsing API allows client applications to check URLs against our constantly updated lists of unsafe web resources. This tool protects 5 billion devices a day by warning users against sites that host malware or unwanted software.

VirusTotal

Allows developers to submit files or URLs for analysis and receive a report on whether or not they are infected with malware.



Information quality APIs

Perspective

An open-source API that uses machine learning to identify 'toxic' comments. Used not only to facilitate interactions between humans, but also to help promote high-quality human interactions with large language models and generative AI.

Vision

Allows developers to easily integrate vision detection features within their applications. This includes image labelling, face and landmark detection, optical character recognition (OCR) and tagging of explicit content.

FactCheck

Aims to facilitate the work of fact checkers, journalists and researchers in discovering what has and hasn't been debunked globally. With this tool, users can search claims from among 300 thousand fact checks from reputable publishers.

Civic Information

The Civic Information API allows developers to build applications that let citizens and voters in the US know about their political representation, ballot information and voting locations. During supported elections, voters can look up their polling place, early vote and dropbox locations, and candidate information.

Text Moderation

Google's Text Moderation tool, available through the Cloud Natural Language API, helps organisations scan for sensitive or harmful content. This tool is able to identify a wide range of harmful content, including hate speech, bullying and sexual harassment.

Every day
you're safer
with Google

We invest in building and sharing APIs that detect and block online threats to ensure that **every day you're safer with Google**. Visit our [Safety Centre](#) to learn more about how we keep more people safer online than anyone else.