

# Pakikipagtulungan para sa **Mas Ligtas** na Internet: Ang Tungkulin ng Mga API

Para protektahan ang mga user at ibang organisasyon mula sa mga bad actor, bumubuo at nagbabahagi kami ng Mga Application Programming Interface (API) na nagde-detect at nagba-block ng mga online na banta, at tumutulong na gawing mas ligtas ang internet para sa lahat.



## PAGBUO

Gumagawa kami ng mga safety API para magbigay ng proteksyon sa tatlong pangunahing bahagi. Ang ilang halimbawa ng mga API na ito ay ang mga sumusunod:

### Kaligtasan ng Bata

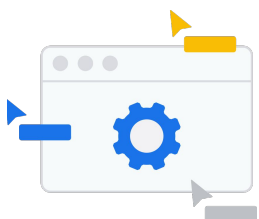
- Content Safety API
- CSAI Match API
- Hash Matching API

### Seguridad

- Safe Browsing API
- Project Shield API
- VirusTotal API

### Kalidad ng Impormasyon

- Perspective API
- Vision API
- FactCheck API
- Civics Information API
- Pag-moderate ng Text



## PAGPAPALAWAK

Para palawakin ang aming epekto, ibinabahagi namin ang aming mga safety API sa mga partner:

### 46M

Noong 2022, na-block ng Cloud Armor API ang isa sa mga pinakamalaking Layer 7 na DDoS attack sa kasaysayan, na may 46M request bawat segundo.

### 2B

Ang Perspective API ay kino-call ng 1000+ partner nang halos 2B beses sa isang araw.

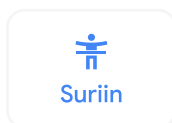
### 4B

Ang Toolkit para sa Kaligtasan ng Bata (Content Safety API + CSAI Match API) ay nagproseso ng 4B+ larawan at video sa loob ng huling 30 araw.

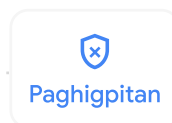


## LIMITASYON

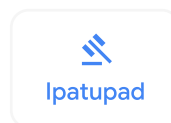
Para masiguradong mga pinagkakatiwalaang partner lang ang may access sa aming mga API at ginagamit ang mga ito sa paraang nakakasunod at ligtas, nagsasagawa kami ng tatlong kritikal na hakbang:



Suriin



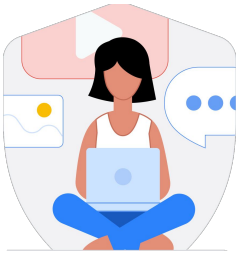
Paghigpitan



Ipatupad

Nililimitahan namin ang pag-access ng aming mga safety API sa mga maingat na sinuring partner, pinaghihigpitan ang hindi awtorisadong paggamit ng mga tool na ito, at mahigpit na ipinapatupad ang aming mga patakaran sa mga tuntunin ng serbisyo. Sa pamamagitan ng pagbibigay-lakas sa aming mga pinagkakatiwalaang partner gamit ang mga safety API, nagagawa naming magtulongan at bumuo ng mas ligtas na internet para sa lahat.

# Mga Safety API ng Google



## Mga API para sa Kaligtasan ng Bata

### Toolkit para sa Kaligtasan ng Bata

#### Kaligtasan ng Content

Tumutulong ang mga pinagmamay-ariang modelo ng Google sa mga partner na uriin at isapriyordad ang bilyon-bilyong larawan ng Materyal na Nagpapakita ng Sekswal na Pang-aabuso sa Bata. Gumagamit ito ng mga ML classifier para makatulong sa pagtukoy ng mga hindi pa nakita dating koleksyon ng imahe ng CSAM para sa mabilis na pagsusuri, pag-aalis, at pag-uulat.

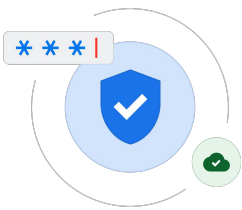
#### CSAI Match

Ang pinagmamay-ariang teknolohiya ng YouTube (at ang unang teknolohiyang gumamit ng hash-matching) para sa pagtukoy ng mga video na nagpapakita ng Koleksyon ng Imahe ng Pang-aabusong Sekswal sa Kabataan (Child Sexual Abuse Imagery o CSAI) at systematic na pag-flag na susuriin, kukumpirmahin, iuulat, at aaksyunan ng mga partner.

#### Hash Matching

Ibibigay namin ang Hash Matching API ng Google sa NCMEC para tulungan silang isapriyordad at suriin ang mga ulat sa CyberTipline nang mas mahusay, na nagbibigay-daan sa kanilang tumuon sa mga ulat na kinasasangkutan ng mga batang kailangan ng agarang tulong.

## Mga Security API



### WAAP Solution ng Google

#### Cloud Armor

Pinoprotektahan ang mga website, serbisyo, at API laban sa mga advanced na naka-target at naka-automate na Distributed Denial of Service (DDoS) attack, na nagse-censor ng impormasyon.

#### reCAPTCHA

Dinedepensahan ang mga website laban sa mapanlokong aktibidad, spam, at iba pang pang-aabuso. Gumagamit ang reCAPTCHA Enterprise ng umaangkop na risk analysis engine para panatilihin hindi nakikilahok sa mga mapang-abusong aktibidad ang mga naka-automate na software.

#### Apigee

Nagbibigay-daan ang platform ng Google Cloud sa pamamahala ng buong lifecycle ng API sa mga negosyo na magdisenyo, mag-secure, sumubaybay, at magpalawak ng mga API.

#### Ligtas na Pag-browse

Sa pamamagitan ng aming Safe Browsing API, nagagawa ng mga client application na alamin kung kasama ang mga URL sa aming palaging ina-update na mga listahan ng mga hindi ligtas sa resource sa web. Nakakapagbigay ang tool na ito ng proteksyon sa

5 bilyong device kada araw sa pamamagitan ng pagbababala sa mga user laban sa mga site na nagho-host ng malware o hindi gustong software.

#### VirusTotal

Nagbibigay-daan sa mga developer na magsumite ng mga file o URL para sa pagsusuri at makatanggap ng ulat sa kung na-infect ba ng malware ang mga ito o hindi.



## Mga API para sa Kalidad ng Impormasyon

### Perspective

Open-source na API na gumagamit ng machine learning para magtukoy ng mga "toxic" na komento. Ginagamit hindi lang para pangasiwaan ang mga interaction sa pagitan ng mga tao, pero para din makatulong sa pagtataguyod ng mga de-kalidad na interaction ng mga tao gamit ang mga large language model at generative AI.

### Vision

Nagbibigay-daan sa mga developer na walang hirap na mag-integrate ng mga feature ng vision detection sa kanilang mga application. Kasama rito ang pag-label ng larawan, pag-detect ng mukha at landmark, optical character recognition (OCR), at pag-tag ng explicit na content.

### FactCheck

Nilalayong pangasiwaan ang trabaho ng mga fact checker, mamamahayag, at mananaliksik sa pagtuklas ng kung ano ang na-debunk at hindi na-debunk sa buong mundo. Gamit ang tool na ito, mahahanap ng mga user ang mga claim mula sa 300K fact check mula sa mga mapagkakatiwalaang publisher.

### Civic Information

Nagbibigay-daan sa mga developer ang Civic Information API na bumuo ng mga application na nagbibigay ng impormasyon sa mga mamamayan at botante sa U.S. tungkol sa kanilang politikal na representasyon, impormasyon ng balota, at mga lokasyon ng pagboto. Sa panahon ng mga sinusuportahang halalan, mahahanap ng mga botante ang kanilang presinto, mga lokasyon ng maagang pagboto at dropbox, at impormasyon ng kandidato.

### Pag-moderate ng Text

Ang tool sa Pag-moderate ng Text ng Google, na available sa pamamagitan ng Cloud Natural Language API, ay nakakatulong sa mga organisasyon na mag-scan para sa sensitibo o mapaminsalang content. Natutukoy ng tool na ito ang isang malawak na hanay ng mapaminsalang content, kasama ang mapoot na salita, pambu-bully, at sekswal na panliligalig.

**Araw-araw  
mas ligtas ka  
sa Google**

Namumuhunan kami sa pagbuo at pagbabahagi ng mga API na nagde-detect at nagba-block ng mga online na pagbabanta para matiyak na **araw-araw ay mas ligtas ka sa Google**. Bisitahin ang aming [Safety Center](#) para matuto nang higit pa kung paano namin pinapanatiling mas ligtas online ang mas maraming tao.