더 안전한 인터넷을 위한 파트너십: API의 역할

Google은 사용자와 조직을 악의적인 행위자들로부터 보호하기 위해 온라인 위협을 감지하고 차단하는 애플리케이션 프로그래밍 인터페이스(API)를 개발하고 공유할 뿐만 아니라 모두를 위해 더 안전한 인터넷을 만들 수 있도록 지원합니다.

개발

Google은 3가지 핵심 영역의 보호를 위해 안전 API를 개발합니다. 다음은 이러한 API의 몇 가지 예시입니다.

아동 안전

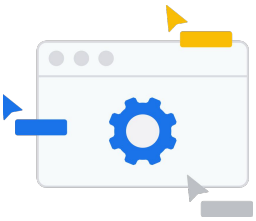
- Content Safety API
- CSAI Match API
- Hash Matching API

보안

- Safe Browsing API
- Project Shield API
- VirusTotal API

정보 품질

- Perspective API
- Vision API
- FactCheck API
- Civics Information API
- 텍스트 검토



확장

Google은 영향력을 높이기 위해 파트너들에게 안전 API를 공유합니다.

460만

2022년에는 Cloud Armor API가 사상 최대 규모(초당 460만 요청)의 레이어 7 DDoS 공격을 차단했습니다.

20억

Perspective API는 1,000여 개의 파트너들에 의해 하루에 20억 회 가까이 호출됩니다.

40억

아동 안전 도구 모음(Content Safety API + CSAI Match API)은 지난 30일간 40억 개가 넘는 이미지와 동영상을 처리했습니다.

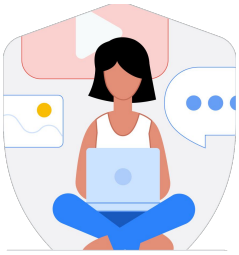
제한

신뢰할 수 있는 파트너만 Google API에 액세스하고 규정을 준수하는 안전한 방식으로 사용할 수 있도록, Google은 3가지의 중요한 단계를 시행합니다.

심사 → 통제 → 시행

Google은 엄격한 심사를 거친 파트너들에게만 안전 API 액세스를 제공하고, 이러한 도구의 무단 사용을 통제하며, 서비스 약관 정책을 엄격하게 시행합니다. Google은 신뢰할 수 있는 파트너들이 Google의 안전 API를 사용할 수 있도록 지원함으로써 파트너들과 협력하여 모두를 위해 더 안전한 인터넷을 만들고 있습니다.

Google의 안전 API



아동 안전 API

아동 안전 도구 모음

Content Safety

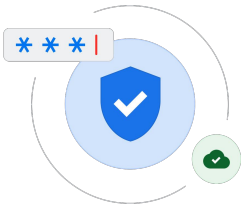
Google의 독자적인 모델은 파트너들이 수십억 개의 아동 성적 학대 콘텐츠 이미지를 분류하고 우선순위를 지정할 수 있도록 지원합니다. 이러한 모델은 ML 분류기를 사용하여 신속한 검토, 삭제, 신고 목적으로 이전에 본 적 없는 CSAM 이미지를 식별합니다.

CSAI Match

아동 성적 학대 이미지(CSAI) 동영상을 식별하고 파트너들이 검토, 확인, 신고 및 조치할 수 있도록 체계적으로 이슈 제기를 다루는 YouTube의 독자적인 기술(이자 해시 매칭을 사용하는 최초의 기술)입니다.

Hash Matching

Google은 NCMEC가 CyberTipline 신고를 효율적으로 검토하고 우선순위를 지정하여 즉각적인 도움이 필요한 아동과 관련된 신고를 빠르게 파악할 수 있도록 Google의 Hash Matching API를 제공합니다.



보안 API

Google의 WAAP 솔루션

Cloud Armor

타겟팅되고 자동화된 DDoS 공격의 정보 검열 및 삭제로부터 웹사이트, 서비스, API를 보호합니다.

reCAPTCHA

허위 행위, 스팸 및 그 밖의 악용 사례로부터 웹사이트를 보호합니다. reCAPTCHA Enterprise는 적응형 위험 분석 엔진을 사용하여 자동화된 소프트웨어가 악용 활동에 사용되지 않도록 보호합니다.

Apigee

Google Cloud의 전체 수명 주기 API 관리 플랫폼은 비즈니스가 API를 설계, 보호, 배포, 모니터링, 확장하도록 지원합니다.

Safe Browsing

Google의 Safe Browsing API는 클라이언트 애플리케이션이 지속적으로 업데이트되는 안전하지 않은 웹 리소스 목록을 사용하여 URL을 확인할 수 있도록 지원합니다. 이 도구는 이용자에게 멀웨어나 원치 않는 소프트웨어를 호스팅하는 사이트를 경고함으로써 매일 50억 대의 기기를 보호합니다.

VirusTotal

개발자가 분석을 위해 파일 또는 URL을 제출하여 멀웨어 감염 여부에 관한 신고를 받아볼 수 있습니다.

정보 품질 API

Perspective

머신러닝을 사용하여 유해한 댓글을 식별하는 오픈소스 API입니다. 사람과 사람 사이의 상호작용을 지원할 뿐만 아니라 대규모 언어 모델과 생성형 AI를 통해 사람들이 양질의 상호작용을 할 수 있도록 돕습니다.

Vision

개발자가 애플리케이션에 비전 감지 기능을 손쉽게 통합할 수 있습니다. 이미지 라벨 지정, 얼굴 및 랜드마크 인식, 광학 문자 인식(OCR), 선정적인 콘텐츠 태그와 같은 기능이 포함됩니다.

FactCheck

팩트체크 전문가, 언론인, 연구자들이 전 세계적으로 사실로 확인된 콘텐츠와 사실이 아닌 것으로 확인된 콘텐츠를 알아볼 수 있도록 지원합니다. 이용자는 이 도구를 사용하여 유명 퍼블리셔들의 30만 건의 팩트체크된 데이터에서 정보를 검색할 수 있습니다.

Civic Information

개발자는 Civic Information API를 사용하여 미국의 시민들과 유권자들이 자신이 지지하는 정당, 투표용지 정보, 투표소에 대해 알아볼 수 있는 애플리케이션을 개발할 수 있습니다. 유권자들은 선거기간 중에 투표소, 사전 투표 장소, 후보자 정보를 찾아볼 수 있습니다.

텍스트 검토

Cloud Natural Language API를 통해 제공되는 Google의 텍스트 검토 도구는 조직에서 민감한 콘텐츠와 유해한 콘텐츠를 검증하도록 지원합니다. 이 도구는 증오심 표현, 괴롭힘, 성희롱을 비롯한 다양한 유해한 콘텐츠를 식별할 수 있습니다.

Google과 함께라면 언제나 안전합니다

Google은 온라인 위협을 감지하고 차단하는 API를 구축하고 공유하는 데 투자하여 **Google과 함께라면 언제나 안전하게 인터넷을 사용할 수 있도록 지원합니다.** Google 안전 센터를 방문하여 Google이 온라인에서 더 많은 사용자를 안전하게 보호하는 방법에 대해 자세히 알아보세요.