

Google AI와 함께 지켜 나가는 안전한 온라인 세상

오늘도 Google과 함께라면 안전합니다

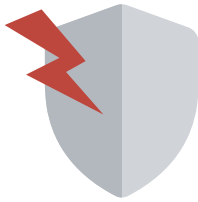


기본으로 내장된 보안



Google은 혁신적이면서도 책임감 있는 접근 방식으로
AI 기술을 활용하여 사람들의 삶을 개선하고자 노력하고 있습니다.
Google은 기본적으로 제품을 **안전하게** 보호하는 **업계 최고의**
AI 보안 인프라를 활용하며, 엄격한 테스트 및 모니터링을 통해 이용자를 보호합니다.

잠재적 위험



1 AI 레드팀

- 보안 전문가들로 구성된 **Google의 레드팀**은 실제 공격이 발생하기 전에 자체 AI 시스템을 해킹하여 취약점을 사전에 탐지하고 해결합니다.



Google Bug Hunters

2 취약점 보상 프로그램

- Google은 **외부의 보안 연구자들이** Google 운영 체제 및 시스템의 취약점을 파악하도록 장려하고 그에 대한 **보상**을 제공합니다.



+1억 건의
피싱 메일 차단



+50억 대의 기기를
유해 앱으로부터 보호



+1천억 개의
설치된 앱 유해성 검사

개인정보 보호를 위한 설계

Google은 책임감 있는 데이터 관행을 엄격하게 준수하고,
제품 작동 방식에 대한 투명성을 제공하며,
이용자가 직접 정보를 제어할 수 있도록 합니다.



① 투명성

- 이용자는 Gemini 앱 개인정보 보호 허브를 통해 Google의 개인정보 보호 관련 주요 내용을 한 곳에서 확인할 수 있습니다.
- Google은 연례 AI 원칙* 준수 보고서를 발행하여 AI 모델의 연구 및 개발 시 적용하는 윤리성 검토 내역을 공개하고, 책임감 있는 AI를 위한 세부 정보를 공유합니다.



② 데이터 최소화 및 간편한 설정

- Google은 머신러닝 기술을 개선하는 데 사용되기를 원하지 않는 정보를 Gemini와의 대화 시 입력하지 않도록 이용자에게 안내합니다.
- 이용자는 간편한 설정을 통해 Gemini 앱 활동 기록을 비활성화하고, Gemini 앱 내 대화 내용을 검토하거나 삭제할 수 있습니다.



*Google의 AI 원칙 확인하기

안전한 콘텐츠를 위한 노력

Google은 AI가 발전함에 따라 강력한 콘텐츠 정책을 뒷받침하는
고도화된 안전 장치를 지속적으로 구축해 나가고 있습니다.

Google은 콘텐츠의 잠재적 유해성 및 편향성을 방지하기 위해
엄격한 테스트와 모니터링을 실시합니다.



◆ Gemini란 무엇인가요?

Gemini는 구글 AI에서 개발한 대규모
언어 모델 제품군입니다.



1 Google의 생성형 AI, Gemini의 더블 체크 기능

- 이용자는 더블 체크 기능을 통해 Google 검색을 사용하여 Gemini의 답변과 유사하거나 다를 수 있는 콘텐츠를 찾아내어 Gemini의 답변에서 부정확할 가능성이 있는 내용을 식별할 수 있습니다.



Watermarked

Non-watermarked

2 신스ID (SynthID) 워터마킹 도구

- AI 생성 이미지, 오디오, 텍스트 또는 비디오에 사람들이 인식할 수 없는 디지털 워터마크를 직접 삽입하여 AI 생성 콘텐츠를 식별할 수 있도록 합니다.
- 신뢰할 수 있는 AI 식별 도구를 개발하여 이용자에게 AI 생성 콘텐츠에 대한 정보를 제공하기 위한 노력입니다.



안전한 온라인 환경을 위한 Google의 노력에 대해 더 많은 정보는
g.co/SafetyCenter/kr에서 확인해 보세요.