1,700여개 고객사
종합 정보보호 전문기업의 28년 노하우

# Safe AI

"AI 사용을 쉽고 안전하게"

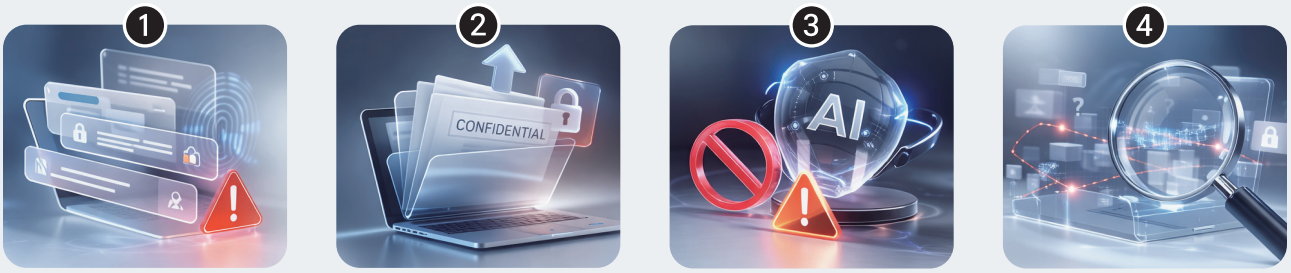
생성형 AI 게이트웨이 솔루션


SOOSAN INT

## AI 보안 대책 마련의 필요성

### 생성형 AI의 다양한 보안 위험 유형

**"내부 부주의, 서비스 취약점, 데이터 관리 부주의 등 보안 위협은 다층적입니다."**




**민감 정보 프롬프트 입력**
개인/계정/계약/재무 정보 등
민감 정보 AI 프롬프트 입력

**기밀 파일 업로드**
문서/로그/소스코드 등
중요 정보 LLM 업로드

**정책 위반 도구 사용**
공식 허용·인증되지 않은
제공자 불분명 AI 도구 사용

**감사/추적 어려움**
누가/언제/무엇을/어디로
전송했는지 증명·추적 한계

## 기존 보안 체계의 한계

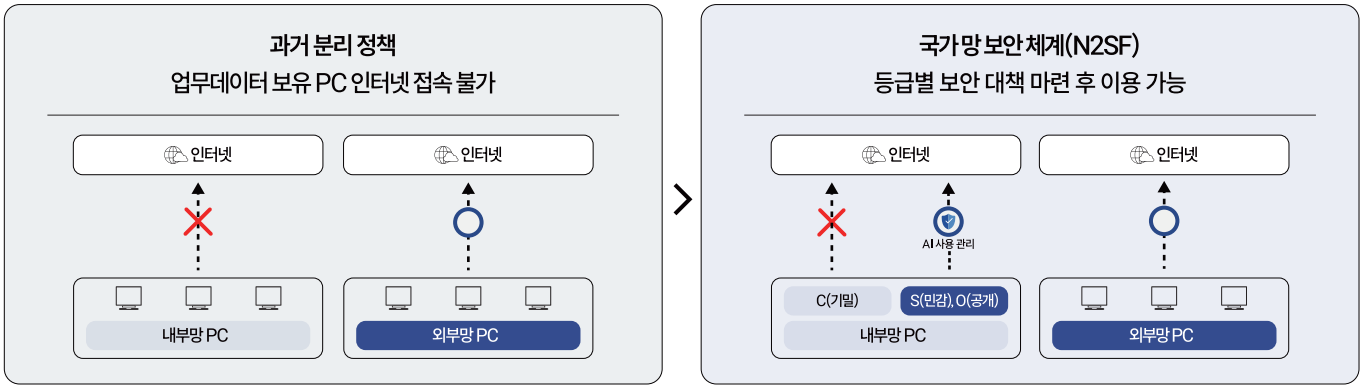
**생성형 AI 원천 봉쇄 불가**
업무 생산성 및 경쟁력 저하,
사용자 민원 증가

**개인/기밀정보 통제 불가**
사용자가 프롬프트에
입력하는 데이터 통제 한계

**암호화 트래픽 확인 불가**
암호화 트래픽으로 인해
완전한 탐지·차단 어려움

**모니터링 및 관리 감독 불가**
누가, 어떤 서비스를
어떻게 이용했는지 알 수 없음

**N2SF 정책 변화** 최신 정책 변화 가이드라인(N2SF)에 따라 생성형 AI 서비스 이용 시 '보안 대책 마련' 필수

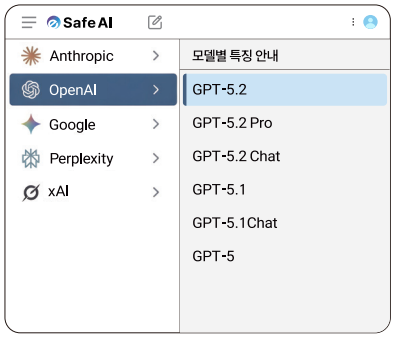



- 내부망 데이터를 외부 SaaS에 전송 불가

- 취급정보 기밀등급 분류 및 승인된 경우 외부 SaaS 이용 가능
  ※ 단, 일정 수준 이상의 보안조치 적용 반드시 필요

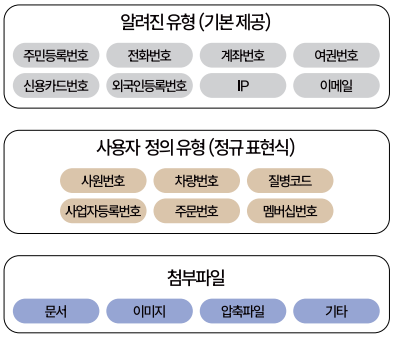
## 멀티 LLM 플랫폼 + 프롬프트/파일 보안 + AI 연동 관리 = eSafe AI

### 사용자 생산성 Up!




- 한 화면 안에서 다양한 LLM 모델 이용
- eSafe AI가 제공하는 보안 웹 인터페이스 사용

### 보안 관리자 보안성 Up!




- 개인정보, 키워드, 첨부파일, 이미지의 내용도 탐지/차단
- 모든 데이터와 보안 이벤트 로깅

### AI 담당자 효율성 Up!


- 생성형 AI 활용 현황 실시간 모니터링
- 정책 템플릿 기능
- AI별 사용량 토큰 관리

### AI 보안 기능

- ☑ 프롬프트 속 민감정보/정책위반 요소 탐지
- ☑ 부서/업무별 정책(허용 LLM, 보안 규칙) 차등 적용
- ☑ LLM 이용 상세 내역 로깅 및 감사/추적

### 구성 특징

- ☑ On-premise 환경에 게이트웨이 구축 & SaaS LLM과 통신
- ☑ 외부 네트워크에서 LLM 이용 내역에 대한 접근 차단

## 제품 개요

‘eSafe AI’는 안전한 생성형 AI 서비스만 사용할 수 있도록 관리하고, N2SF 가이드라인을 준수하여 생성형 AI 사용 과정에서 발생할 수 있는 민감정보 유출을 실시간으로 저장·차단합니다. 또한 업무에 활용되는 생성형 AI의 사용 비용을 효과적으로 제어할 수 있는 AI 업무혁신 전문 솔루션입니다.

### 안전한 모델만 골라쓰기

여러 종류의 특화된 LLM을 하나의 화면에서 자유롭게 사용

### 민감정보 필터링

프롬프트, 첨부파일(이미지 등)에 포함된 민감정보를 실시간으로 저장하고 탐지/차단

### 실시간 대시보드

누가, 어떤 LLM을, 얼마나 안전하게 사용하는 중인지 일목요연하게 정리

### 사용자별 정책 설정

사용자, 부서별로 차별화된 접근 권한 및 보안 거버넌스 정책 수립 가능

### 통합 제어 체계

웹브라우징 방식, API 방식 모두 지원하여 SDK, 플러그인 방식도 통제 가능

### AI 사용 예산 관리

사용자/부서/조직에서 발생한 총 비용에 대한 실시간 모니터링 및 임계치 지정

# 주요 기능



### 1 멀티 LLM 사용 편의성

- 웹브라우저만 있으면 접속 준비 완료
- 이용 중 원클릭 모델 전환 가능
- 복수 LLM 동시 사용 및 응답 비교
- 수백 종의 LLM 연동 지원
- 신규 LLM에 대한 손쉬운 연동 지원

### 2 데이터 보안

- 민감 정보 자동 탐지 및 마스킹
- 정책 위반 데이터 차단
- LLM으로 프롬프트 전송 전 검증
- 사용자별 직접 접근 차단
- 프롬프트 내용 및 첨부파일까지 모두 로깅

### 3 비용 및 사용량 관리

- 실시간 비용 리포트
- 총 사용량 임계치 설정
- 예산 설정 및 알림
- 사용자/팀 단위 사용량 제한

### 4 모니터링 및 로그 관측

- 실시간 사용 현황 대시보드
- 외부 LLM API Key 중앙 관리
- 모든 요청/응답 로그 기록
- 감사(Audit) 추적
- 분석 및 리포트 생성

### 5 AI 에이전트 사용 통제

- Claude Code 등 AI 에이전트와 연동 가능
- AI 에이전트가 LLM에 보내는 모든 데이터 기록
- 도구 호출 시 전달되는 데이터 기록
- 민감 정보 포함 여부 검사

### 6 사용자 및 권한 관리

- 조직/팀/사용자 단위 계정 관리
- 계층적 권한 구조
- 사용자/팀 단위 모델 사용 권한 제어
- 정책 기반 접근 제어

# 프롬프트 데이터 보안 및 유출 방지

## 사용자 프롬프트 차단 정책

**INPUT 사용자 원본 입력** — USER

사용자가 입력한 프롬프트

"다음 [대외비] 신제품 로드맵 자료를 영어로 요약해 줘."

첨부파일
2026_Roadmap_확정안.docx

⚠ 정책 위반 키워드, 첨부파일 검출

**PROCESS eSafe AI 정책 엔진** — GATEWAY

- 정 책 명 대외비 키워드 차단
- 위반 대상 본문·첨부파일
- 위반 사유 키워드 "대외비" + 기밀 분류 문서
- 적 용 자 전사·R&D본부 강화 정책

LLM 미전송·위반 사유 사용 이력에 기록

**OUTPUT 사용자 화면 표시** — NOTICE

⚠ 사용자의 요청이 eSafe AI 보안 정책에 의해 차단되었습니다.

[정 책 명] 대외비 키워드 차단
[위반 대상] 본문·첨부파일
[위반 사유] 패턴 일치

정책 위반 전송 차단 및 사용자 알림

## 민감 정보 마스킹 정책

**INPUT 사용자 원본 입력** — USER

사용자가 입력한 프롬프트

"010-1234-5678 과 010-1999-4321 중 하나를 선택해 줘."

⚠ 휴대전화번호 2건 탐지

**PROCESS eSafe AI 마스킹 엔진** — GATEWAY

{MASK1} 010-1234-5678

{MASK2} 010-1999-4321

LLM 전송 본문
"{MASK1}과 {MASK2} 중 하나를 선택해 줘."

매핑 테이블·자사 어플라이언스 내부 보관

**OUTPUT 사용자 화면 표시** — RESTORED

실제 LLM 응답(마스킹)
"{MASK2}를 선택했습니다."

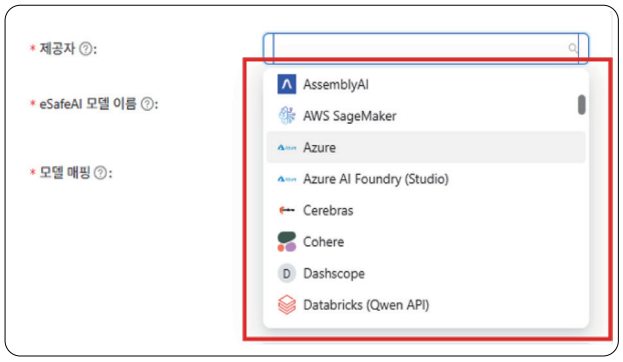
사용자에게 보여지는 LLM 응답(원본 복원)
"010-1999-4321 을 선택했습니다."

사용자 화면에서는 원본으로 복원하여 표시

- 주민등록번호, 신용카드 번호, 전화번호 등 민감한 개인정보를 자동으로 식별하고 마스킹 처리하여 외부 유출을 원천 차단
- 기업의 기밀 문서 패턴, 소스 코드, 영업 비밀 등 중요 정보에 대한 커스텀 패턴을 정의하여 정교한 필터링 체계 구축

## 다양한 LLM 연동 지원

### 다양한 LLM 제공사 모델과의 손쉬운 연동 기능 제공




다양한 LLM 제공사 목록(예시)


기 보유한 API Key 활용(예시)

- 대다수의 LLM 모델 제공 및 신규 LLM도 빠른 시일 내 등록 가능, 업무 환경에 최적화된 다양한 AI 모델을 유연하게 연동·사용
- 기존 보유하고 있는 LLM 제공사의 API Key를 등록하여 eSafe AI 내에서 해당 모델을 간편하게 등록·관리

## AI 가시성 확보를 통한 내역 관리성 Up!

### LLM 사용 모니터링 및 로그 데이터 제공

서비스 로그
총 317개의 로그가 조회되었습니다.
21 - 40 / 317

| | No | 사용자 이름 | 발생일시 | 결과 | 정책 | 발신자 | 본문 | 대화 |
|---|---|---|---|---|---|---|---|---|
| > | 25 | test1 | 2026.01.21 16:45 | 감시 | 감시테스트 | User | ㅎㅇ | |
| > | 26 | test1 | 2026.01.21 16:45 | | | AI | 오늘도 고생 많으셨어요! 퇴근하고 편히 쉬세요. 좋은 저녁 보내시길 바랍니다! | |
| > | 27 | test1 | 2026.01.21 16:45 | 감시 | 감시테스트 | User | 퇴근하자 | |
| > | 28 | test1 | 2026.01.21 16:45 | 차단 키워드 본문 | 정책테스트 | User | 마스킹 원본 | |
| > | 29 | test1 | 2026.01.21 16:45 | | | AI | 오늘 하루 고생 많으셨어요! 편안한 저녁 되시고, 퇴근 후에는 꼭 휴식도 잘 취하세요. | |
| > | 30 | test1 | 2026.01.21 16:45 | 감시 | 감시테스트 | User | 퇴근하자 | |

## 실시간 통합 관제로 관리자 편의성 Up!

**보안 이벤트**
- 차단/경고 이벤트 모니터링
- 위반 유형별 통계

**사용 이력**
- 사용자별/부서별 이력 조회
- 이상 사용 탐지 알림

**사용량/추이**
- 전체 AI 사용량 실시간 집계
- 일별/주별/월별 추이

**차단 원인 상세**
- 위반 정책명·위반 상세
- 탐지 개인정보 유형 표시

**상세 로그**
- 사용 시간·사용자·모델
- 입출력 요약·첨부파일 정보

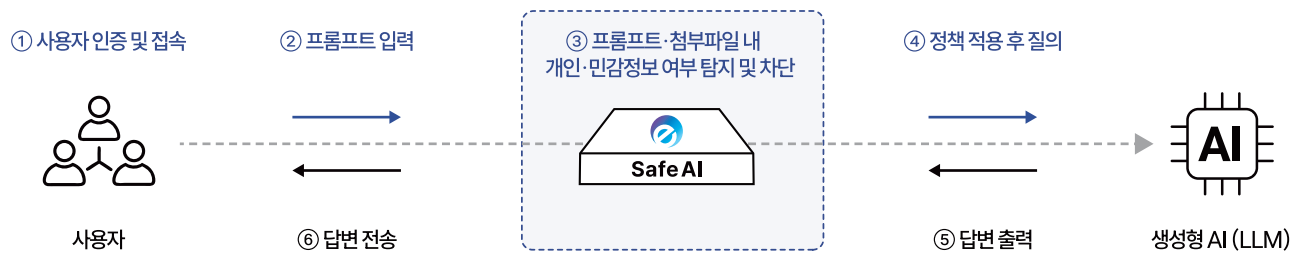
**비용 현황**
- AI별 토큰 사용량 추적
- 예산 대비 사용률

정책 탐지/차단 및 실시간 AI 사용 현황을 한눈에 파악. 보안 사고 시 사후 추적 완벽 지원!

Safe AI

1,700여개 고객사에게 인정받은 28년 노하우로 여러분의 네트워크 보안을 책임집니다.

## 시스템 구성도




- AI-DLP 및 가이드라인을 기반으로 민감 정보를 차단하고, 프롬프트 위협을 실시간 모니터링하는 직관적인 대시보드 제공
- 공공·금융기관의 보안성 검토와 N2SF 요건에 최적화된 화면 제공

## 도입 효과

### eSafe AI 도입 4대 기대효과

**① 보안 강화**
- 민감정보 유출 원천 차단
- AI 사용 제어 및 Shadow AI 통제

**② 컴플라이언스**
- 개인정보보호법, ISMS-P 감사 대응 체계 확보

**③ 운영 효율**
- Agentless 즉시 도입
- IT팀 관리 부담 최소화

**④ 비용 최적화**
- 사용자/부서별 임계치 조정
- 사용량 가시화로 AI 사용 예산 절감

SOOSAN INT

(주)수산아이앤티 | 서울시 강남구 밤고개로 5길 13 수산빌딩 | https://www.soosanint.com | 02-750-0843 | ask@soosan.co.kr