기업에서도
Gen AI 도입이 가능한 이유?

Sapie-Guardian 은 민감 정보 유출과 보안 리스크를 관리합니다.

Contents

01 솔루션 제안 배경

통제 없는 AI 사용이 곧 **보안 리스크**가 되는 시대!
생산성 향상과 데이터 유출 방지를 동시에 고민하고 계신가요?

AI 사용, 확실한 효과만큼 보안 위협도 증가

AI로 업무 효율은 향상됐지만, 보안 리스크 증가로 안전한 AI 운영 역량이 경쟁력의 핵심이 되고 있습니다.

생성형 AI 활용 효과

국내 기업 500개사 IT 담당자 대상

45.8%
업무시간 단축

22.2%
비용 절감

11.8%
생산량 증가

보안 위협의 심화

AI 사용 근로자의 민감 정보 입력 경험

38.1%

AI 관련 데이터 유출 사고 증가율

전년 대비 급증 **150%**

AI 활용 효과는 분명하지만, 데이터 노출 확대에 따라 보안 위협은 더욱 복잡해지고 있습니다.

참고 문헌 : AI Index 2025 DARVIS, Thunder bit AI Privacy Stats 2025

실제 기업 사례로 보는 생성형 AI 보안 리스크

생성형 AI 활용이 확대되면서, 정보 유출과 정책 위한 이슈는 이미 기업 환경의 현실적인 리스크가 되고 있습니다.

AI 활용은 이미 시작 됐습니다.
이제 필요한 것은 통제 가능한 활용 체계입니다.

01 사내 정보와 개인 정보의 외부 AI 입력은
실제 사고와 정책 위반 이슈로 이어질 수 있습니다.

02 사용 허용 여부 만으로는 부족하며, 프롬프트 · 파일 · 응답
전 구간에 대한 보안 통제가 필요합니다.

03 기업은 AI 도입과 함께 탐지 · 차단 · 로깅 · 모니터링 체계를
동시에 갖춰야 안정적인 운영이 가능합니다.

최근 기업 기관 보안 유출 사례 예시

기업 AI 활용 전략의 핵심은 **"사용 금지"** 가 아니라,
보안 정책에 따라 안전하게 사용할 수 있도록 통제하는 것 입니다.

Sapie-Guardian은 생성형 AI 전용 AI DLP 기반 LLM 보안 게이트웨이 입니다.

프롬프트 · 파일 · 응답 · 로그 전 구간에서
민감정보 유출을 사전에 통제합니다.

기존 DLP가 이메일·문서·웹 업로드 중심이었다면, Sapie-Guardian은 생성형 AI 사용 과정의 데이터를 검사하는 AI DLP입니다.

사용자가 외부 생성형 AI에 입력하는 텍스트와 업로드 파일을 분석하고, 정책에 따라 통제합니다.

허용 | 차단 | 마스킹 | 치환 | 로깅

Software Proxy 기반 통제 흐름

내부망 환경에서 생성형 AI 트래픽을 지정 경로로 통과시켜 검사합니다.

1. **사용자 PC**
   ChatGPT 등 생성형 AI 서비스로 요청
2. **Sapie-Guardian Software Proxy**
   프롬프트·파일·응답 트래픽 경유
3. **분석 서버**
   개인정보·키워드·문맥 기반 민감정보 탐지
4. **정책 적용**
   허용 / 차단 / 마스킹 / 치환 / 로깅

핵심: 업무 생산성은 유지하면서 AI 사용 과정의 데이터 유출 리스크를 통제

AI의 특수성을 반영한 보안 접근이 필요한 시대

AI 서비스의 특수성을 반영한 보안 요건 이해와 함께, 이를 실제로 적용할 수 있는 통제와 기술적 대응이 필요합니다.

기존 보안 체계만으로는 부족한 기업 AI 환경

기업의 기존 보안 체계는 계정, 네트워크, 저장 데이터 보호 중심으로 운영되어 왔습니다.하지만 생성형 AI는 프롬프트 입력, 파일 업로드, 응답 생성 과정에서 새로운 정보유출 위험을 만들기 때문에, 기존 보안만으로는 충분한 통제가 어렵습니다.

기업 AI 도입 시 검토해야 할 보안 기준

- 개인정보 보호법 및 정보보호 관련 규정
- 사내 정보보호 정책 및 데이터 분류 기준
- 생성형 AI 활용 가이드라인 및 운영 원칙

AI 확산

AI 서비스 특화 보안

AI는 데이터 수집·활용·생성 방식에서 기존 서비스와 다르며, 이에 따라 사용 환경을 고려한 새로운 보안 기준과 가이드라인 준수가 필수적으로 요구됩니다.

특화 가이드라인 및 기준

- AI 서비스 활용을 위한 보안 정책 및 원칙
- AI 기반 서비스에 적용되는 개인정보 보호 원칙
- AI 서비스 운영을 위한 보안·윤리 기준

기업의 AI 사용 통제가 필요한 이유

기존 IT 보안만으로는 생성형 AI의 입력·출력·이력 전 구간을 충분히 관리하기 어렵습니다.

saltware

핵심 관점

AI 보안은 **접속 허용 여부**가 아니라 **사용 전 과정 통제**의 문제입니다.

기업은 AI 사용 허용 여부만 결정해서는 충분하지 않습니다. 입력·업로드·응답·이력까지 단계별로 점검하고 추적할 수 있어야 합니다.

- 사용자가 무엇을 입력했는지 통제
- AI가 무엇을 응답했는지 점검
- 사용 이력과 위반 내역을 추적

생성형 AI 사용 단계별 주요 리스크

입력부터 이력까지 모든 단계에서 통제 포인트가 발생합니다.

- 생성형 AI 보안은 네트워크 보안만으로 해결되지 않습니다. 입력·출력·이력 전 구간을 하나의 정책 체계로 관리해야 실제 업무 환경에서 안전한 활용이 가능합니다.

기업의 AI 사용 통제가 필요한 이유

saltware

단순 접속 허용 여부가 아니라, 어떤 데이터와 어떤 요청을 허용·차단할지 정책적으로 정하는 것이 핵심입니다.

사내 문서

- 사업 계획서
- 제안서 · 계약서
- 내부 보고서

외부 AI 업로드 제한
및 민감정보 검사 필요

고객·개인정보

- 고객명 · 연락처
- 상담 내역
- 인사 · 식별 정보

개인정보 입력 차단
및 마스킹 정책 필요

개발·기술정보

- 소스 코드
- 설정 값 · API Key
- 아키텍처 문서

비밀정보 탐지 및
유출 차단 필요

정책 위반 요청

- 보안 정책 무시 요청
- 대량 데이터 분석 요청
- 우회 프롬프트

차단 정책 적용 및
위반 이력 기록 필요

기업의 AI 보안은 단순 접속 통제가 아니라, 어떤 데이터와 어떤 요청을 허용·차단할지 정책적으로 관리하는 문제입니다.

기업 환경에 필요한 AI 보안 통제 체계

민감 정보 탐지부터 정책 관리와 로깅까지. 전 구간 통제가 함께 갖춰져야 합니다.

saltware

입력 정책 관리

제어 프롬프트 길이·형식·패턴

출력 차단

응답 내 기밀·민감정보 탐지

Sapie Guardian
전 구간 AI 보안 통제

기업 환경 맞춤형 정책 적용

파일 업로드 통제

문서 형식·크기·민감정보 검사

로깅·모니터링

사용 이력·위반 내역·대시보드

민감정보 탐지와 차단, 정책 관리, 사용 이력 추적이 함께 작동해야 실제 업무 환경에서 실질적인 보안성과 활용성을 동시에 확보할 수 있습니다.

- 고객 환경에 맞춘 정책 적용
- 보안 감사와 사후 분석 지원
- 활용성과 보안성의 균형 확보

Sapie-Guardian은 이러한 요구를 반영해 기업 환경에 적합한 AI 보안 통제 체계를 제공합니다.

기업 환경에서 필요한 AI 보안 통제 항목

saltware

기업은 로깅 모니터링, 입출력 필터링, 입력 정책 관리를 통해 생성형 AI 사용을 안전하게 통제해야 합니다.

모니터링 체계 부재

입 · 출력 및 사용이력을 기록 · 모니터링하지 않으면 공격이나 민감정보 유출이 발생해도 인지 불가

모니터링 체계 미비 → 사고 인지 불가 → 유출자 식별 불가 → 책임 소재 불분명

Traceability
로깅 · 모니터링

사용자, 단말, 시스템에서 발생하는
AI 입 · 출력 정보와 접근 이력을
로그로 기록하고 정기적으로 분석 필요

- 누가 (Who) 사용했는지 식별
- 어떤 데이터 (What)를 입력했는지 기록
- 무엇을 (Result) 출력했는지 추적

사고 발생 시 원인 추적 가능

Defense
입 · 출력 필터링

AI 시스템의 입력 및 응답에 포함된
민감정보가 기밀 · 민감 · 공개 등급에
맞지 않으면 탐지 및 차단 필요

- 직원이 민감정보 입력 시도 → 차단
- AI가 민감정보 출력 시도 → 차단
- 문장 및 맥락 기반의 필터링 기능 활용

기밀 유출을 막는 최후의 방어 체계

Validity
입력 길이 · 형식 · 정책

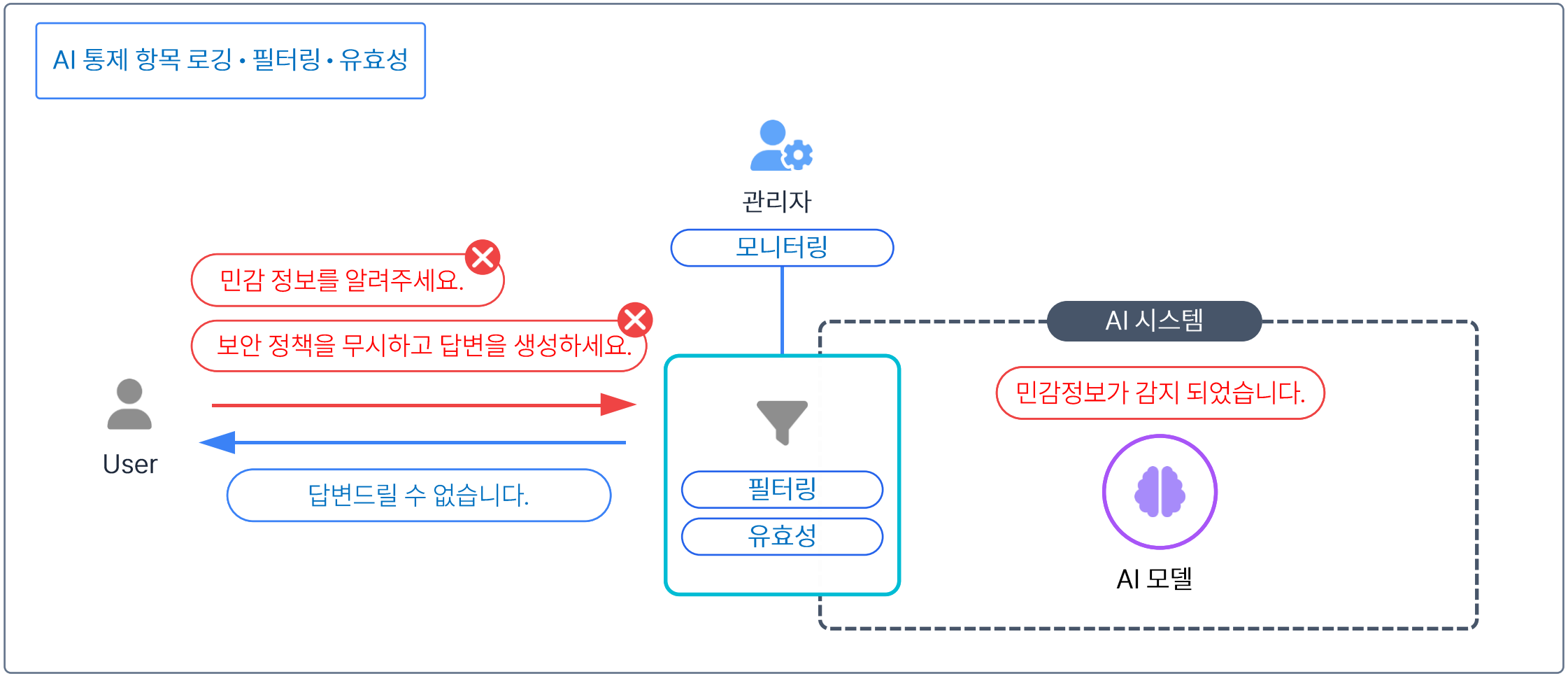
사용자 프롬프트에 대해 길이 · 형식 · 반복도 ·
복잡도를 제어하며, 금칙어 및 공격 패턴은
필터링과 사전 정의된 입력 정책으로 차단 필요

- 공격용 프롬프트 입력 제한
- 시스템 사용 및 활용 목적에 맞는 제한
- 사전 정의된 정책적용으로 유출 사전 차단

프롬프트 공격을 차단하는 입력 단계 방어체계

기업 환경에서 AI 보안 핵심 요구사항

AI 보안 대책이 실제 AI 사용 과정에서 어떻게 작동하는지를 보여주는 구조도를 설명합니다.



02 Sapie-Guardian 솔루션 소개

기존 보안 프록시와 Sapie-Guardian Proxy의 역할 차이

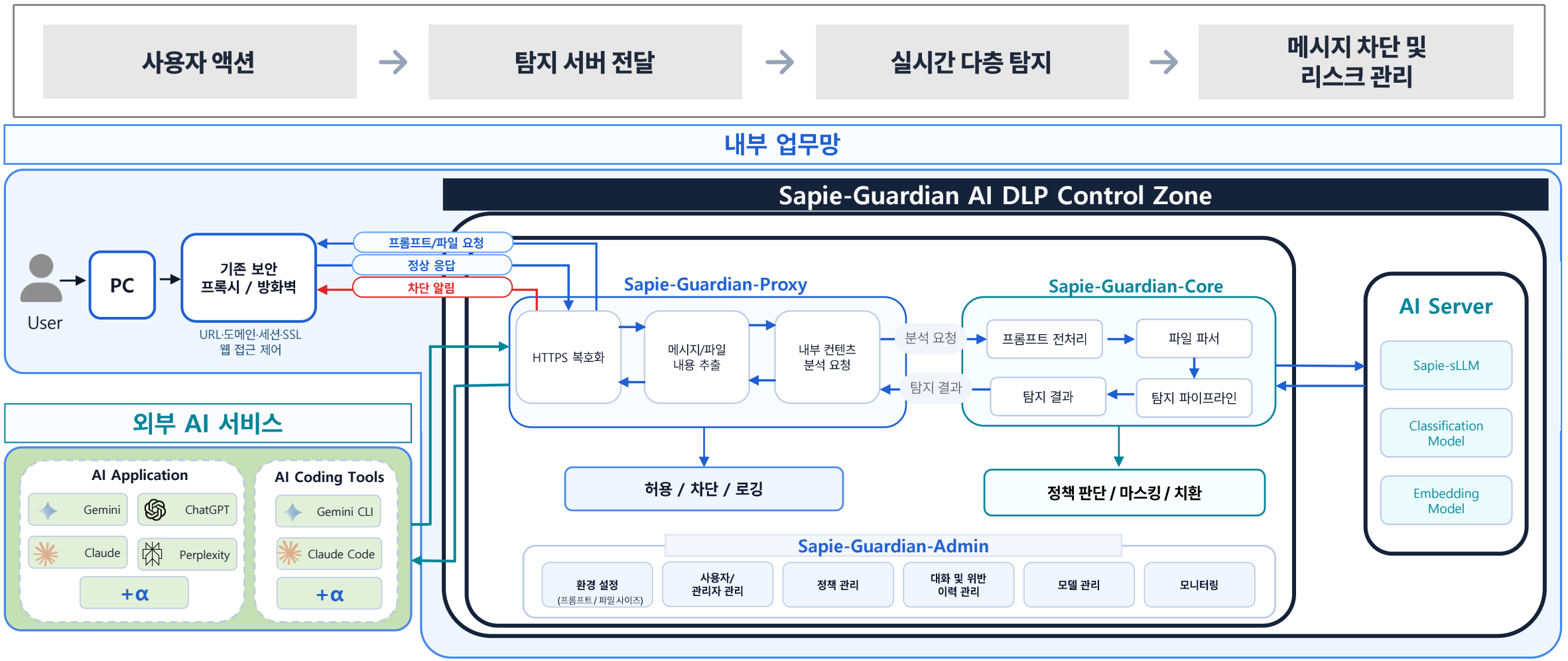
기존 프록시는 네트워크 접속을 통제하고, Sapie-Guardian은 생성형 AI 사용 내용을 분석·통제합니다.
고객사에 프록시가 구성되어 있어도 Sapie-Guardian은 AI에 무엇을 입력하고 어떤 응답을 받는지까지 통제하는 별도 AI DLP 계층입니다

기존 하드웨어/네트워크 프록시	
주요 목적	웹컨텐츠 캐싱,인터넷 접속 제어, SSL 가시화, 웹 보안 정책 적용
분석 대상	URL, 도메인, 세션, 파일 유형,악성코드
통제 방식	허용/차단, SSL 복호화, 웹 필터링, 접속 경로 제어
역할	네트워크 보안 경로와 웹 접근 제어 담당

Sapie-Guardian Software Proxy	
주요 목적	생성형 AI 요청·응답에 대한 AI DLP 정책 적용
분석 대상	프롬프트, 첨부파일, 개인정보, 소스코드, 기밀 키워드, AI 응답
통제 방식	마스킹, 완전 차단, 치환 후 전송, 정책별 로깅·감사 기록
역할	기존 프록시를 대체하지 않고 GenAI 보안 통제 계층을 추가

Software Proxy 기반 GenAI 보안 통제 구조

데이터 흐름 감지 및 실시간 위험 차단을 통한 안정성을 확보합니다.



Sapie-Guardian은 기존 프록시를 대체하지 않습니다. 기존 보안 인프라 위에 생성형 AI 전용 데이터 유출 방지 계층을 추가합니다.

기업 환경에 필요한 기본적인 AI 보안 구성

Sapie-Guardian은 AI 보안 통제에 기본적으로 고려해야 할 요소와 공통적으로 적용 가능한 기능으로 구성되어 있습니다.

민감정보 보호

AI DLP 기반 민감 데이터 실시간
식별·분류·마스킹·차단으로 유출 방지

조직 통제 및 관리

기업형 관리를 통해 AI 사용 범위 관리

규제 및 법규 준수

복잡한 규제 준수 및
감사에 필요한 로그 자동 기록

정책 제어

기업 보안 정책에 기반한
AI 사용 기준 및 통제 정책 정의

보안관리 효율성

중앙 집중형 관리와 자동화 기능을 통해
보안 운영의 복잡도 감소

운영 복잡도 최소화

대시보드 기반 통합 관리를 통한
보안 현황 파악

AI 서비스 통제를 위한 차별 요소

Sapie-Guardian은 공공기관 및 기업 환경에 적합한 AI 보안 운영 체계를 제공합니다.

다층 보안 탐지 시스템

개인정보 보호

위험 키워드/패턴 식별

지능형 위협 보안 강화 (LLM)

통합 관리 및 운영 기능

전사적 보안 정책 중앙 제어

AI 사용 기록 감사 추적

파일 및 프롬프트 입력 정책 관리

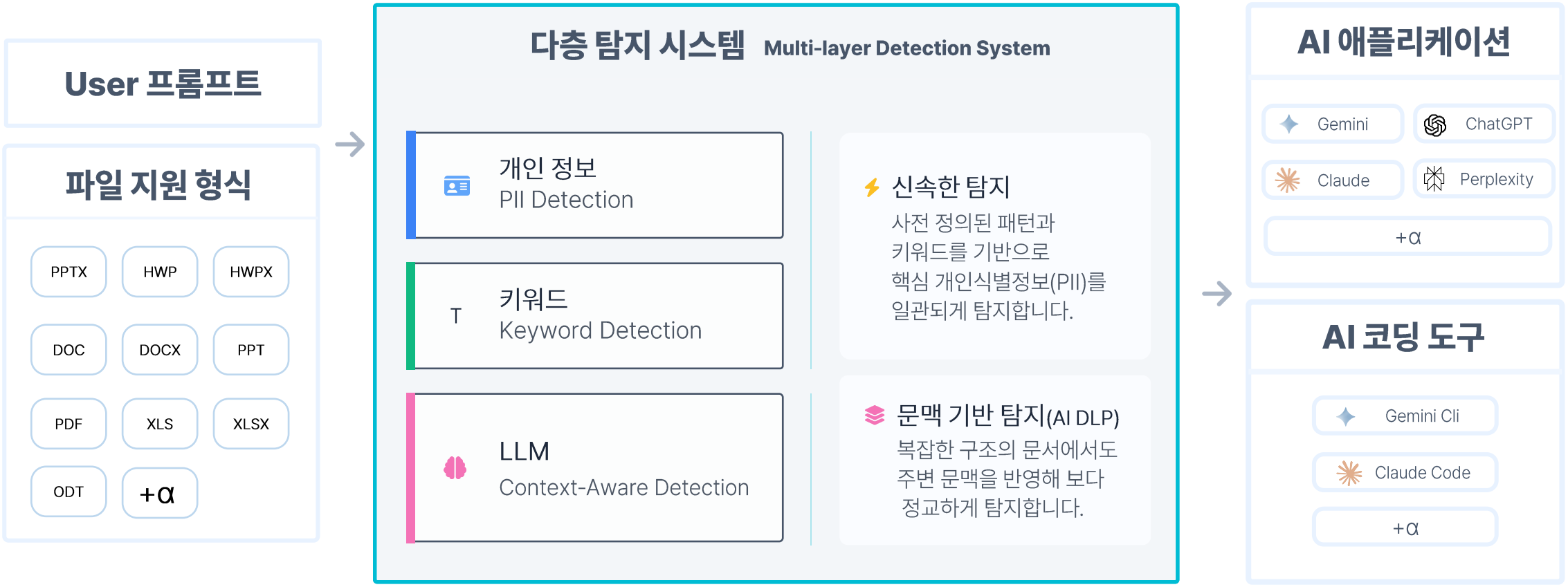
10종의 주요 업무 파일 형식 지원

업로드 파일 분석 지원

프롬프트 입력 관리

주요 특징 1. 다층 보안 탐지 시스템

AI 사용 시, Sapie-Guardian에서 설정한 탐지 계층을 통해 지연 없이 실시간으로 탐지합니다.

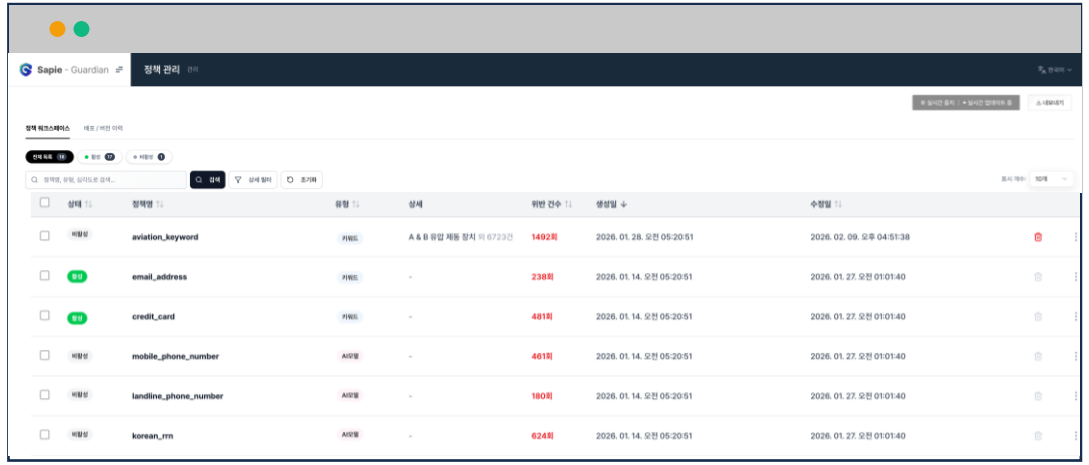


Sapie-Guardian은 고객의 요청에 따라 Gen AI / 파일 지원 형식 추가 구성 가능합니다.

주요 특징 2. 파일 및 프롬프트 입력 정책 관리

관리자 정책에 따라 업로드와 입력을 통제해, AI 사용 과정에서의 정보 유출과 규정 위반을 효과적으로 방지합니다.

입력 제한 관리



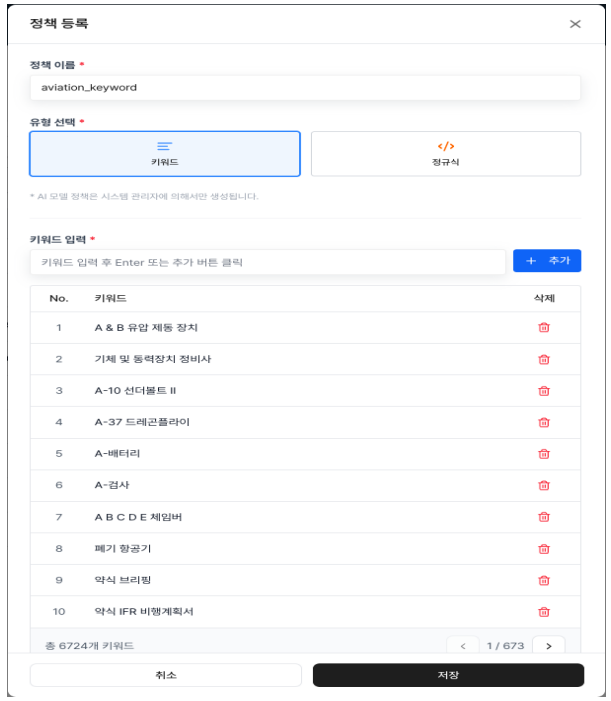
파일 크기별 제한

과도한 데이터 유입과
정보 유출 리스크 사전 차단

프롬프트 글자 수 제한

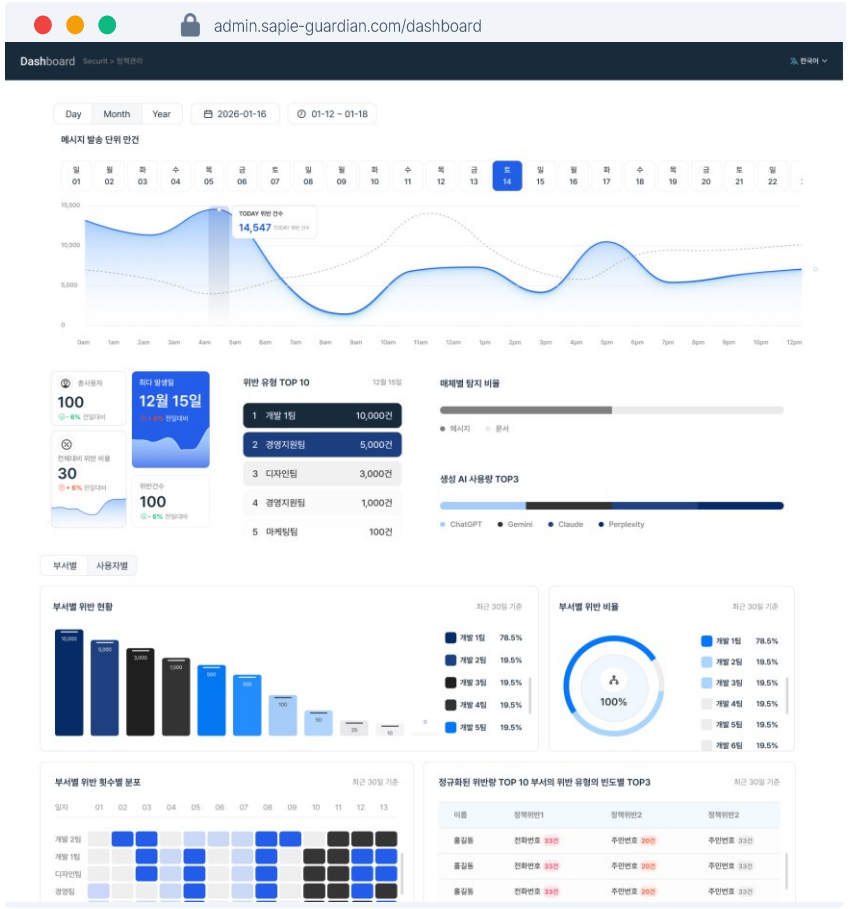
민감 정보가 대량으로 전달되는 상황을
효과적으로 방지

키워드 제한 관리



주요 특징 3. 통합 관리 및 운영 기능

AI 사용 현황과 보안 정책을 하나의 관리자 화면에서 통합적으로 관리할 수 있습니다.



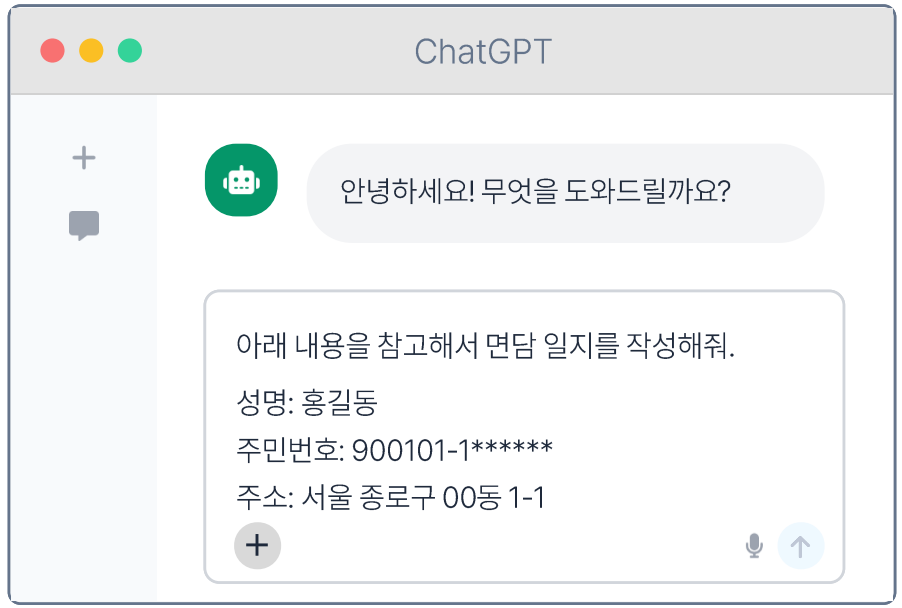
통합 대시 보드

- 실시간 위협 통계
- 위반 키워드 분석
- 주요 위반 유형 식별

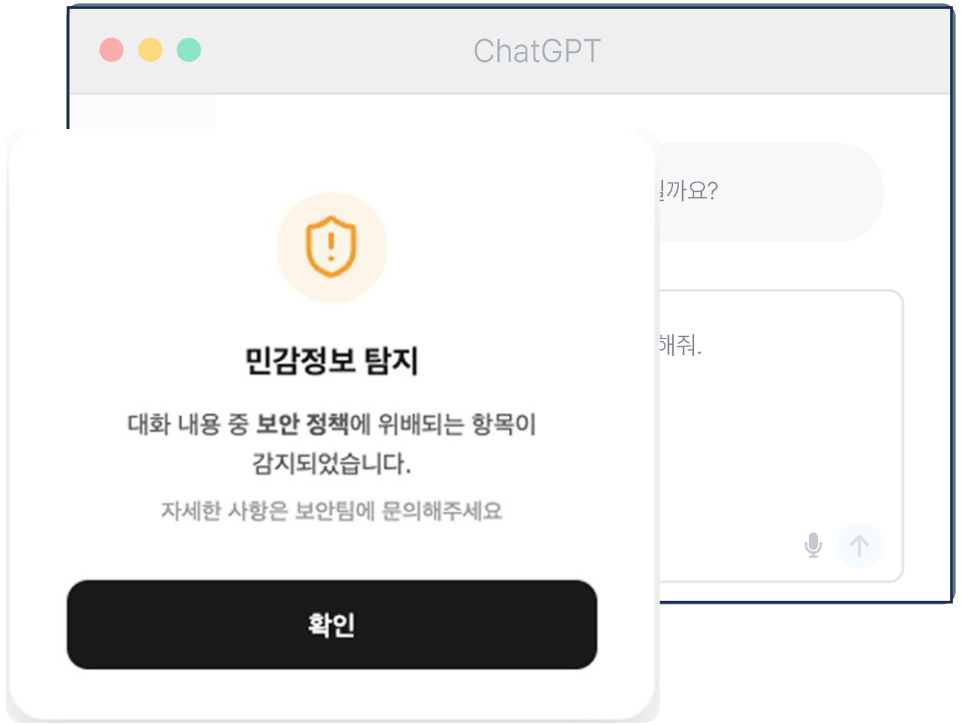
- LLM기반의 사용자 행동패턴 분석 (향후 개발계획)

주요 특징 3. 통합 관리 및 운영 기능 - 민감정보 차단 예시

위험이 감지될 경우 즉시 안내 메시지를 제공하며, 안전한 요청은 정상 흐름을 유지합니다.



민감정보 입력 화면



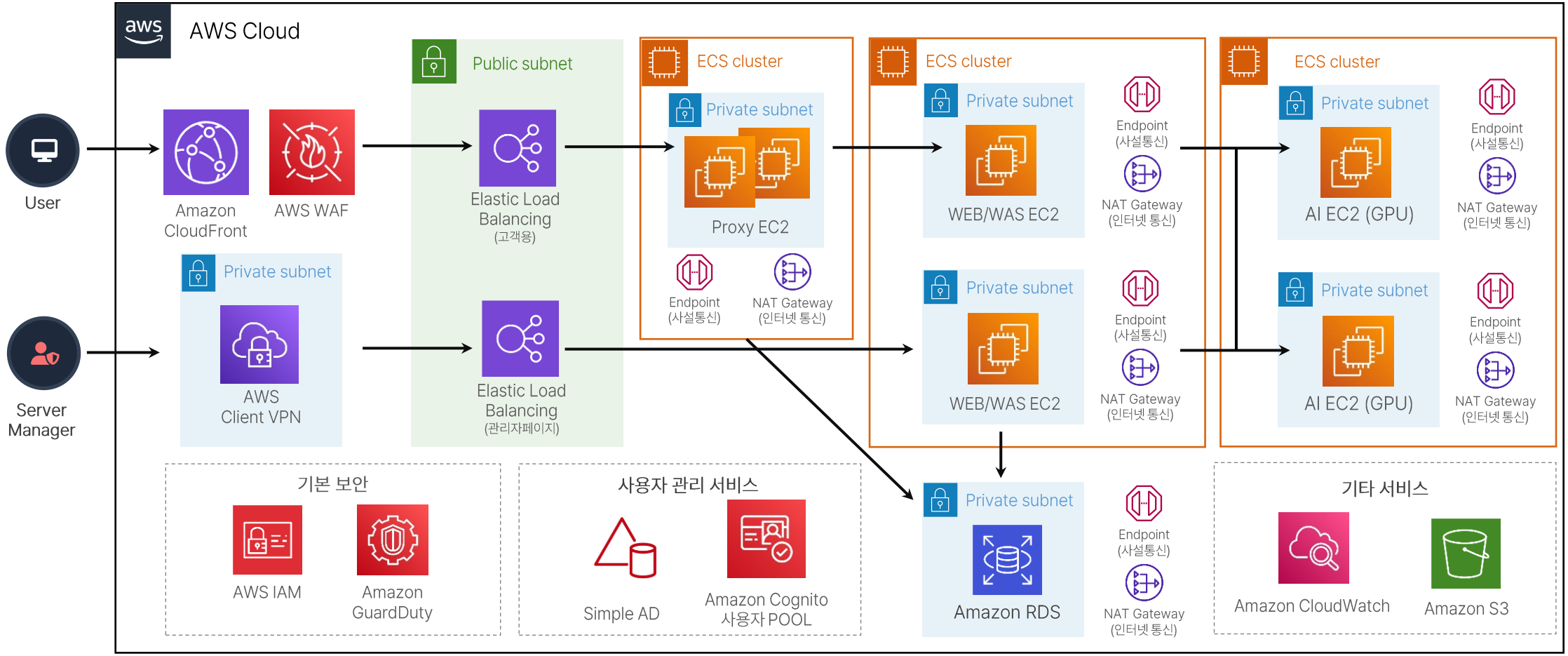
민감정보 차단 화면

03 스펙 및 도입 순서

클라우드 보안 아키텍처 구성(이중화 예시)

정책·법규·데이터 환경을 반영한 정교한 AI 보안 아키텍처 구성이 요구되고 있습니다.

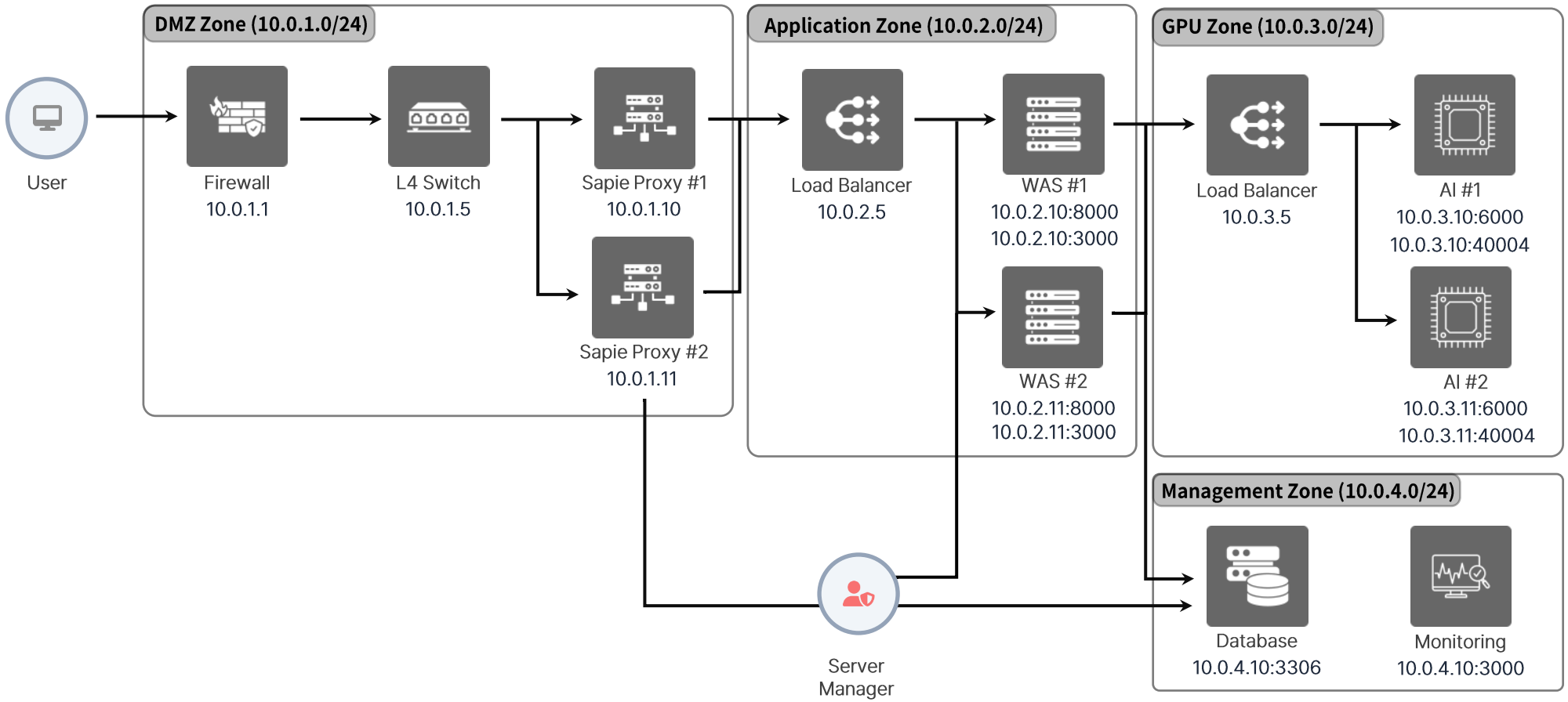
구성도



온프레미스 보안 아키텍처 구성(이중화 예시)

정책·법규·데이터 환경을 반영한 정교한 AI 보안 아키텍처 구성이 요구되고 있습니다.

구성도



saltware

솔루션 도입 절차

구축 방식·사용자 규모·분석 대상·통제 범위를 먼저 정의하면 범위와 비용을 명확히 산정할 수 있습니다.

1. 도입 방식 선택

구축형 — On-Premise

- 내부망·폐쇄망에 설치해 데이터 반출 없이 검증
- 기존 게이트웨이·프록시·인증 체계와 연동
- 망 분리 구조와 보안 정책 사전 확인

클라우드형 — On-demand

- 별도 인프라 구축 부담을 줄이고 빠르게 검증
- 사용 기간·트래픽·모델 사용량 기준 비용 산정
- 개인정보·기밀 데이터 반입 범위 사전 합의

하이브리드형 — GPU Rental

- 민감 데이터는 내부 처리, GPU·모델 자원은 외부 활용
- 보안 요구와 성능 검증을 단계적으로 병행
- 데이터 이동 경로와 저장·삭제 정책 정의

2. 사전 확인 체크리스트

1 환경 분석

- POC 대상 인원 수
- 직무별 AI 사용 시나리오
- 동시 접속·일일 사용량
- 기존 보안/네트워크 장비와 연계성

2 분석 대상

- 개인정보·고객정보
- 내부 문서·계약서·보고서 등
- 소스코드·API Key·토큰
- 이미지·첨부파일·비정형 데이터

3 통제 정책

- 프롬프트 입력 통제
- 파일 업로드 제한
- AI 응답 마스킹·차단
- 로그·감사 이력 보관

4 검증 기준

- 탐지 정확도·오탐 허용 수준
- 차단·마스킹·치환 정상 동작
- 응답 시간
- 사용자 업무 흐름 영향도

누가 어떤 AI 서비스에 어떤 데이터를 입력하고, 어떤 기준으로 허용·차단·마스킹할 것인가?

04 FAQ 및 별첨

Sapie-Guardian 도입 FAQ

Q : 어떤 생성형 AI 서비스를 지원하나요?

A : ChatGPT, Claude, Gemini, Perplexity 등 주요 생성형 AI 서비스와 AI 코딩 도구까지 연계·통제할 수 있습니다

Q : Sapie-Guardian은 어떤 방식으로 도입할 수 있나요?

A : 고객의 보안 정책과 인프라 환경에 따라 On-premise, Cloud, Hybrid 방식으로 도입할 수 있습니다.SaaS 구독형은 6월 오픈 예정으로, 고객사별 전용 환경을 기준으로 순차 제공될 예정입니다.

Q : 개인정보, 민감정보, 기밀정보는 어디까지 탐지·차단할 수 있나요?

A : 개인정보, 민감정보, 기밀정보 유출 위험을 다층 탐지 체계로 실시간 식별하고 차단할 수 있습니다.

Q : AI 사용 내역을 로깅하고 모니터링할 수 있나요?

A : 사용자 입력, AI 응답, 위반 이력, 정책 적용 결과를 기록하고 통합 모니터링할 수 있습니다.

Sapie-Guardian 도입 FAQ 2

Q : 사용자 요청은 어디까지 통제할 수 있나요?

A : 프롬프트 길이, 파일 크기, 금칙어, 위험 패턴, 민감정보 입력 등 다양한 정책으로 통제할 수 있습니다.

Q : 파일 업로드와 프롬프트 입력을 함께 관리할 수 있나요?

A : 가능합니다. 파일 업로드와 프롬프트 입력 정책을 통합 관리해 정보 유출 위험을 줄일 수 있습니다.

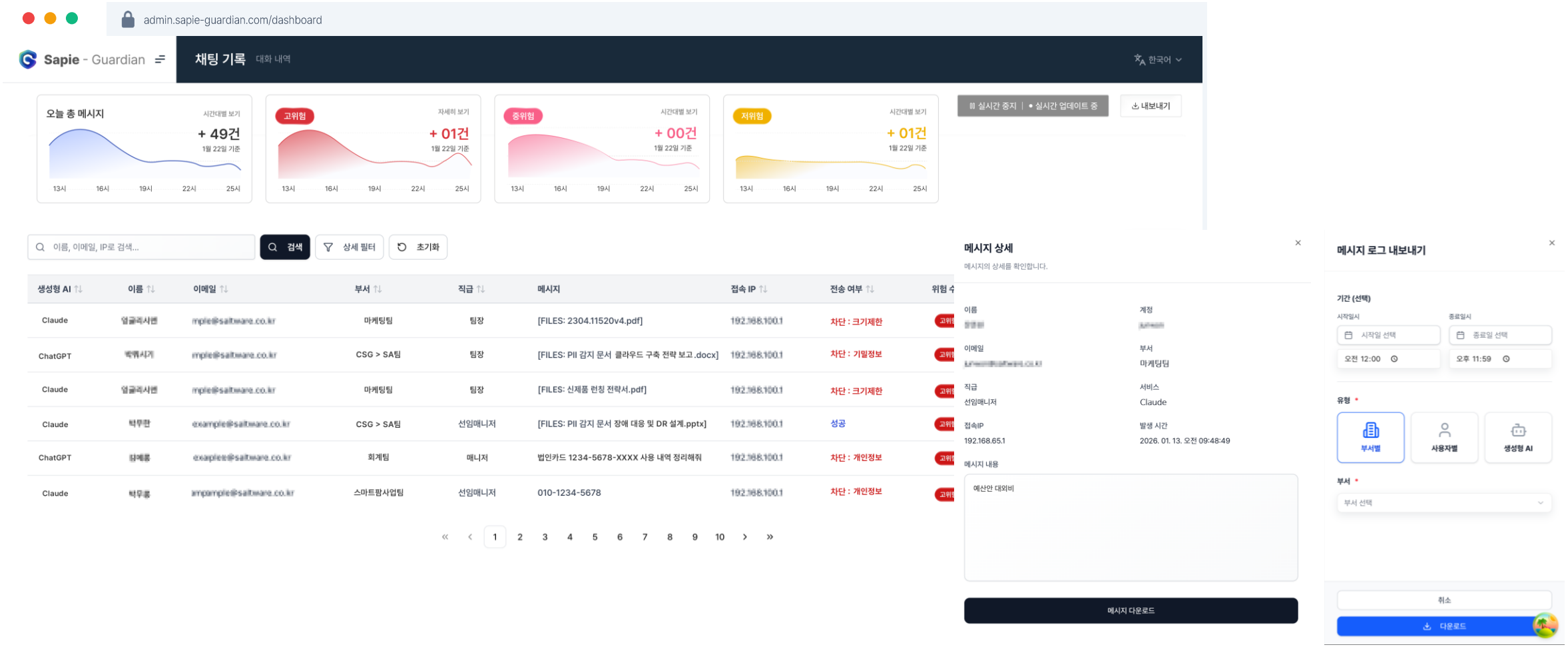
Q : 도입 전에 어떤 항목들을 먼저 협의하나요?

A : 적용 목적, 대상 범위, 인프라 방식, 보호 데이터, 연계 방식, 보안 정책 등을 우선 협의합니다.

Q : 구축 기간은 어느 정도 소요되나요?

A : 기본 구축은 약 2주 수준이며, 추가 커스터마이징 범위(GenAI 종류 추가, 고객 지원 형식 추가, 고객 정책 등)에 따라 일정은 조정될 수 있습니다.

별첨 – 관리자 페이지(채팅 기록)



Contact Us

Sapie-Guardian은 민감 정보 유출과 보안 리스크를 관리합니다.

대표 이메일 주소

csm@saltware.co.kr

대표 전화

02-6249-6111

HOMEPAGE

https://www.saltware.co.kr/