PRODUCT INTRODUCTION

WhiteHat

맥락으로 판단하는 AI 보안 — WhiteHat Mail · WhiteHat Code
클라우드부터 온프레미스 (sLLM 학습) 까지

WhiteHat · whitehat.ai.kr

ONE PAGE

한 장 요약 — 왜 메일 보안 , 왜 코드 보안인가

📧 메일 보안이 중요한 이유

BEC(임원 · 거래처 사칭) 피해 연 $2.77B

재무팀 메일 한 통에 속아 회사 계좌에서 수천만 ~ 수억 원이 이체되고 , 알아챘을 땐 이미 늦다 .

피싱 메일의 82.6% 가 AI 생성

오타 · 어색한 문장으로 가짜를 구분하던 시절은 끝났다 — 담당자가 봐도 진짜 같다 .

악성 링크 · 첨부 없는 공격 다수

백신 · 스팸필터가 볼 게 없다 — " 정상적인 요청처럼 보이는 메일 " 이 가장 위험하다 .

→ WhiteMail: 7 에이전트가 의도 · 관계 (맥락) 로 조사하고 근거를 공개

💻 코드 보안이 중요한 이유

AI 생성 코드의 45% 가 보안 취약

바이브 코딩으로 빠르게 만든 서비스가 , 배포 첫날부터 뚫릴 준비가 되어 있다 .

환각 패키지 (존재하지 않는 라이브러리) 19.7%

AI 가 지어낸 패키지명을 그대로 설치하면 , 공격자가 선점해둔 악성 코드가 그대로 실행된다 .

유출 시크릿 1 건 = 인프라 전체 노출

AI 가 커밋에 남긴 API 키 한 줄로 DB· 서버 전체가 공격자 손에 넘어간다 .

→ WhiteHat Code: 정적 +SCA+ 환각 +LLM 딥스캔으로 배포 전 차단

하나의 철학으로 풉니다 — 맥락으로 판단하고 · 근거를 설명하고 · 데이터는 조직 안에 (클라우드 → 온프레미스)

WHY SECURITY

보안이 왜 중요한가

① 돈이 직접 샌다

BEC 송금 사기 한 건이 수천만 ~ 수억. 유출된 코드 시크릿 하나로 전체 인프라가 뚫립니다.

② 신뢰가 무너진다

고객 · 거래처 정보 유출은 브랜드 신뢰와 계약을 잃게 만듭니다. 회복은 훨씬 비쌉니다.

③ 규제 · 법적 책임

개인정보보호법 · 전자금융거래법 등 — 사고 시 과징금과 신고 의무가 따릅니다.

④ AI 가 판을 바꿨다

AI 로 공격은 더 싸고 정교해지고, AI 로 짠 코드엔 취약점이 더 많습니다. 방어도 AI 로.

사고가 난 뒤 수습하는 비용보다, 배포 전에 잡는 비용이 훨씬 쌉니다 — 그래서 WhiteHat 입니다.

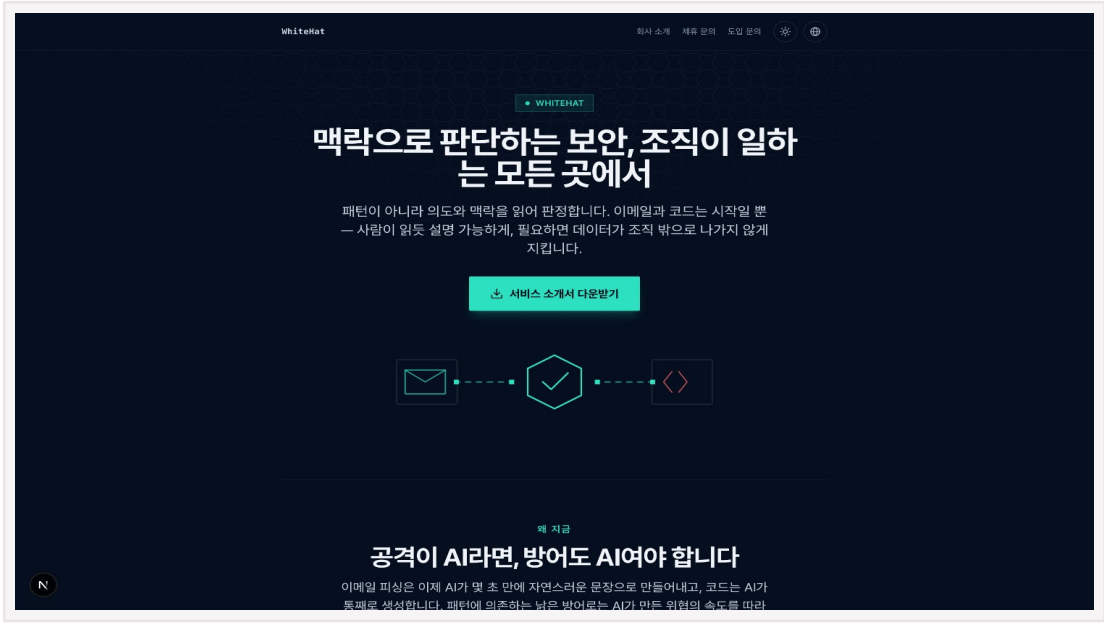
PART 1

WhiteHat Mail

AI 이메일 보안 — 7 개 전문 에이전트가 사람 분석가처럼 메일을 조사합니다

한 줄 소개

- **패턴이 아니라 맥락 · 의도로 판단하는 AI 이메일 보안**
- BEC(임원 사칭) · 피싱 · 큐싱 (QR) · 계정 탈취를 조사
- 판정 근거를 에이전트 · 신호 단위까지 전부 공개 (설명가능성)
- 무료로 시작 · 클라우드 또는 완전 온프레미스 배포



AI 가 만든 공격은 패턴이 없다 — 맥락으로 본다

whitehat.ai.kr 랜딩

THE PROBLEM

AI 가 이메일 공격을 바꿨다

BEC

임원 · 거래처 사칭
송금 · 계좌 변경 유도

QR

큐싱 — 이미지 QR 로
로그인 / 결제 유도

AI

AI 가 문장 · 도메인을
정교하게 위조

- 표면 시그니처 (발신지 · 키워드) 만 보는 필터는 AI 가 만든 자연스러운 문장을 못 잡는다 .
- 진짜 신호는 " 맥락 " — 처음 보는 발신자가 긴급 송금을 요구하는가 , 표시이름과 도메인이 어긋나는가 .
- 사람 분석가가 하던 " 조사 " 를 AI 에이전트가 대신한다 .

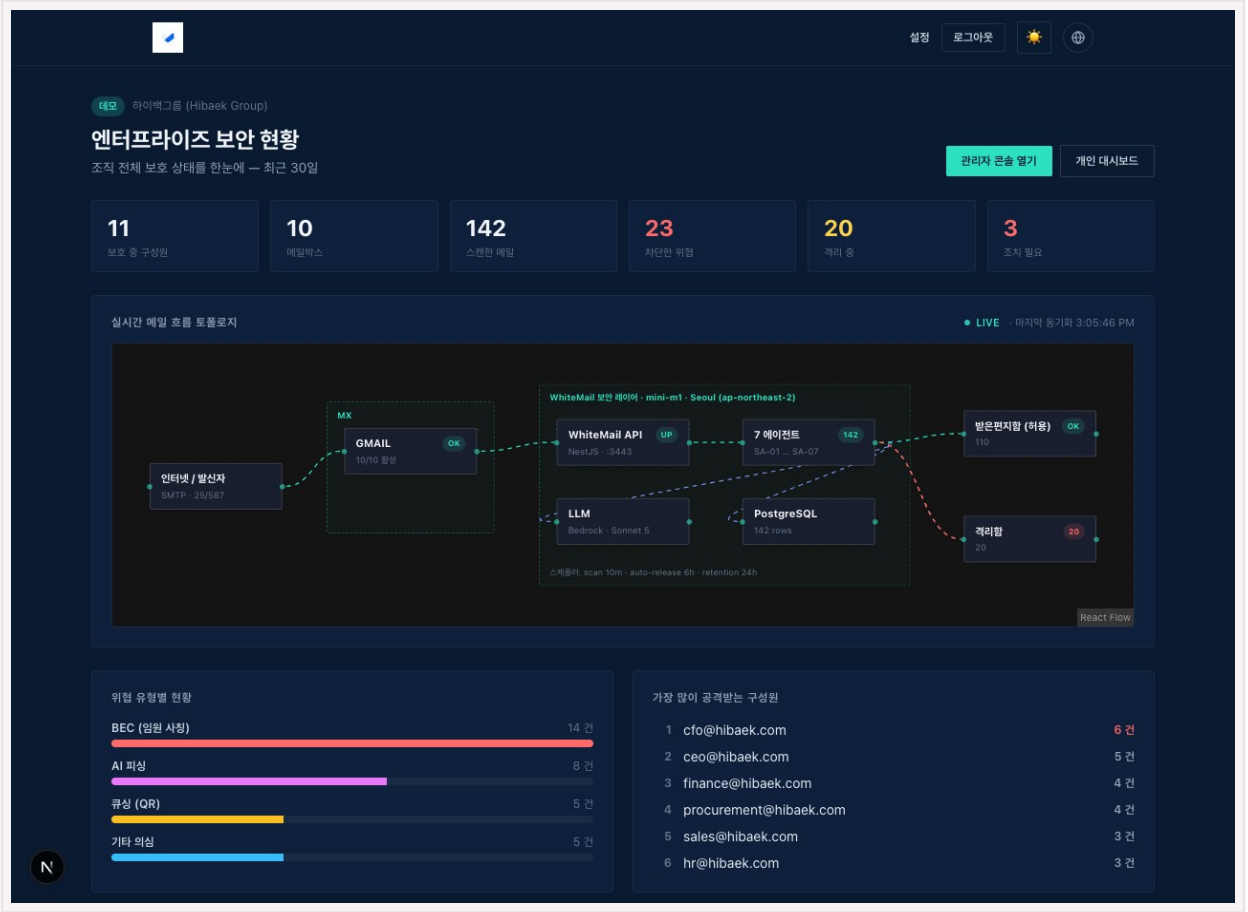
APPROACH

맥락으로 판단하는 7- 에이전트 조사

7- 에이전트 병렬 조사	메일 1 건을 7 개 전문 에이전트가 동시에 " 조사 "
전문 분담	신원 · 인프라 · 링크 · 의도 · 행동 · 첨부 · 평판 — 한 축만 깊게
RAG 근거 보강	위협 인텔리전스 · CVE · 공격사례를 벡터 검색으로 부착
상관보정 집계	신호를 집계 · 보정해 판정하고 , 근거를 그대로 노출

AT A GLANCE

조직 전체 보안 현황 — 한눈에



엔터프라이즈 대시보드 · 실시간 메일 흐름 토폴로지 · 위협 유형 / 표적 통계

AGENTS

7 개 전문 에이전트

SA-01 신원

사칭 · 표시이름 / 브랜드 불일치 · Reply-To/Return-Path

SA-02 인프라

SPF/DKIM/DMARC · 유사 도메인 · 의심 TLD · ARC/BIMI

SA-03 링크 · 비주얼

표시 URL 불일치 · 단축 / 리다이렉트 · 퓨니코드 · QR 큐싱

SA-04 의도 (LLM)

긴급 송금 · 계좌 변경 · 자격증명 탈취 · 압박

SA-05 행동 · 관계

첫 접촉 · 반복 위협 · VEC · 아는 연락처 사칭

SA-06 첨부

실행파일 · 이중확장자 · 매크로 · HTML · 액티브 PDF

SA-07 평판

악성 도메인 /URL 피드 · 신생 도메인

+ 아웃바운드 DLP

발신 유출 — RRN· 카드 · 계좌 · 시크릿 · 대량 수신

SA-01 — 신원 분석

보낸 사람이 진짜 그 사람인지 확인합니다 .

탐지 순서

발신 헤더 파싱
From· 표시이름 · 주소

→

표시이름 ↔ 도메인
사칭 · 브랜드 대조

→

Reply-To/Return-Path
회신 경로 검사

→

신원 신호
DISPLAYNAME/REPLYTO

이렇게 잡습니다 예 : " 김정훈 대표이사 " <ceo.hibaek@gmail.com> — 표시이름은 임원인데 실제 도메인은 회사와 무관 , Reply-To 까지 다름 .

왜 중요한가 가장 흔한 사기 (임원 사칭) 가 여기서 걸립니다 .

SA-02 — 인프라 · 인증 분석

메일이 진짜 그 도메인에서 왔는지 확인합니다 .

탐지 순서

인증 결과 조회
SPF·DKIM·DMARC

→

정렬 판정
alignment

→

도메인 검사
유사 · 신생 · 의심 TLD

→

인프라 신호
SPF_FAIL 등

이렇게 잡습니다 예 : hibaeg.com(유사 도메인) + SPF 인증 실패 + DMARC p=none 우회 .

왜 중요한가 발신 도메인 위조 · 유사 도메인 스푸핑을 잡습니다 .

SA-03 — 링크 · 비주얼 분석

눌렀을 때 어디로 가는지 , 숨긴 링크가 있는지 봅니다 .

탐지 순서

링크 · 이미지 추출
URL· 첨부 이미지

→

표시 ↔ 실제 대조
+ QR 디코드

→

위장 검사
퓨니코드 · 단축 · 리다이렉트

→

링크 신호
PUNYCODE·QR

이렇게 잡습니다 예 : " 결제 확인 " 링크가 실제로는 다른 도메인 · 이미지 QR 을 디코드하니 퓨니코드 피싱 URL.

왜 중요한가 큐싱 (QR)· 퓨니코드처럼 눈으로 못 보는 함정을 대신 확인합니다 .

SA-04 — 의도 분석 (LLM)

이 메일이 결국 무엇을 시키려는지 읽습니다 .

탐지 순서

본문 정규화
제목 · 본문 정리

→

sLLM 의도 분석
무엇을 요구하나

→

위협 인텔 RAG
유사 수법 대조

→

의도 신호
송금 · 탈취 · 압박

이렇게 잡습니다 예 : " 오늘까지 신규 계좌로 송금하고 , 외부 확인은 하지 마세요 " — 전형적인 BEC 압박 정황 .

왜 중요한가 문장이 자연스러워도 " 의도 " 로 판단 — 여기에 학습한 sLLM 이 씁니다 (온프렘 가능).

SA-05 — 행동 · 관계 분석

평소 관계 · 이력과 비교해 이상한지 봅니다 .

탐지 순서

발신자 이력 조회
과거 교신 · 플래그

→

관계 판정
첫 접촉 · 반복 위협

→

VEC 검사
거래처 계좌 변경

→

행동 신호
FIRST_CONTACT 등

이렇게 잡습니다 예 : 오래된 거래처가 갑자기 계좌를 바꾸거나 , 처음 연락한 발신자가 긴급 송금을 요구 .

왜 중요한가 메일 한 통이 아니라 " 관계의 맥락 " 으로 판단합니다 .

SA-06 — 첨부 분석

첨부파일이 위험한 유형인지 검사합니다 .

탐지 순서

첨부 추출	→	유형 검사	→	심층 검사	→	첨부 신호
파일 · 타입		확장자 ·MIME· 매크로		압축 · 컨테이너 (ISO/OneNote)		EXECUTABLE 등

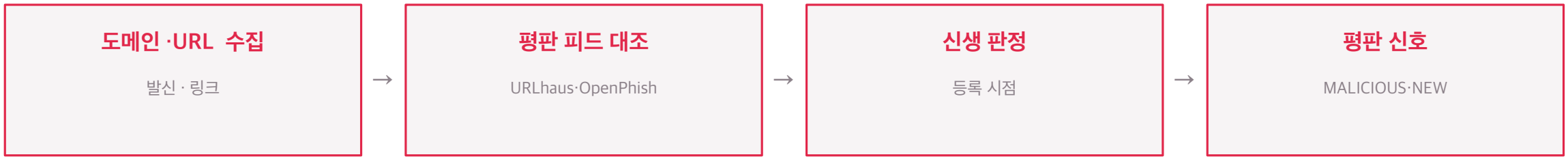
이렇게 잡습니다 예 : 발주 .docm(매크로) + 자료 .pdf.js(이중확장자) 를 함께 첨부 .

왜 중요한가 악성 실행 · 매크로 · 이중확장자 등 첨부로 숨긴 공격을 잡습니다 .

SA-07 — 평판 분석

이미 알려진 나쁜 도메인 /URL 인지 대조합니다 .

탐지 순서

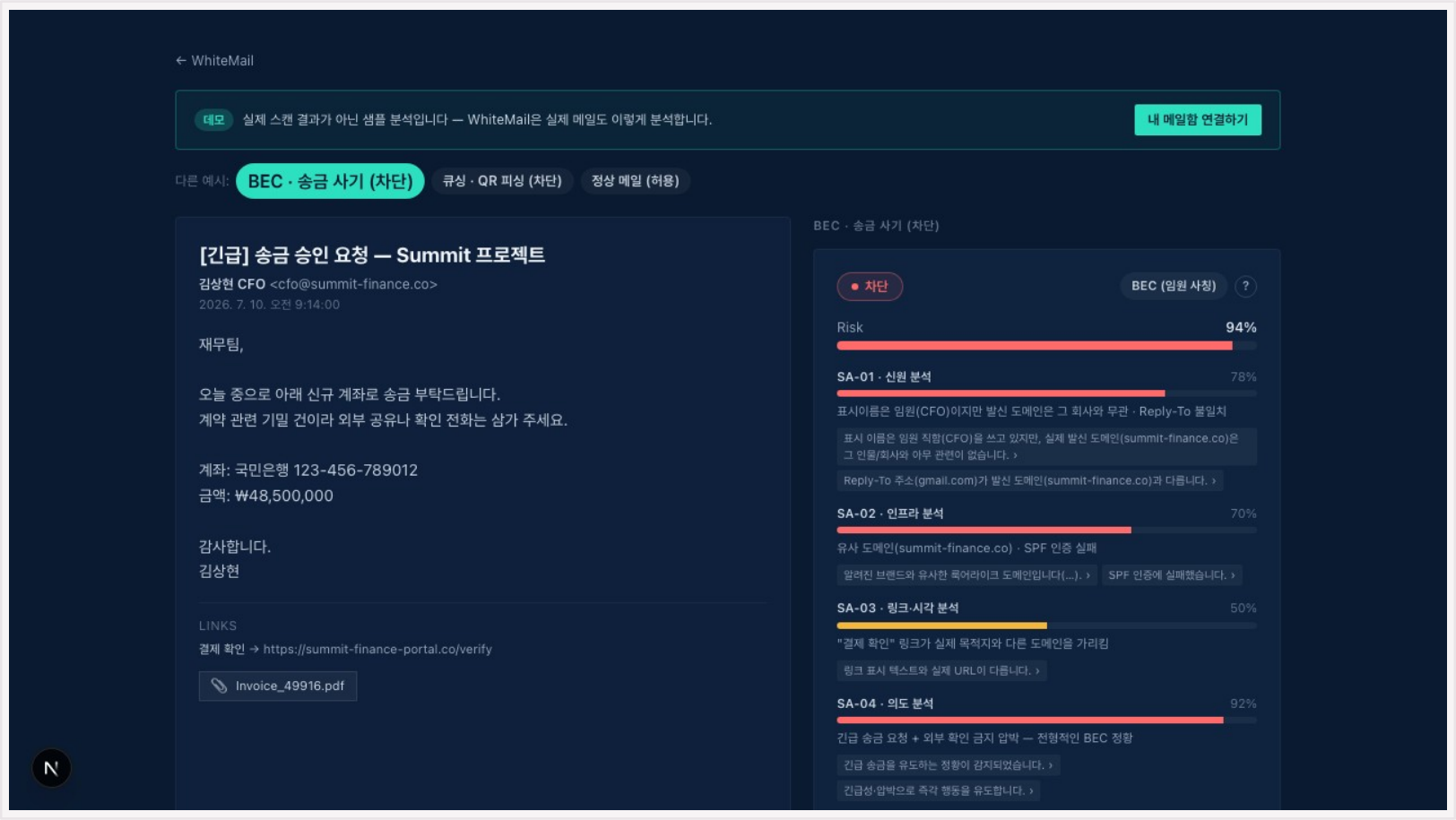


이렇게 잡습니다 예 : 등록한 지 몇 시간 안 된 도메인 · 피드에 올라온 악성 URL 을 단축으로 우회 .

왜 중요한가 전 세계 위협 인텔로 " 이미 나쁜 곳 " 을 즉시 걸러냅니다 .

DEMO

데모 — BEC 송금 사기를 이렇게 조사한다

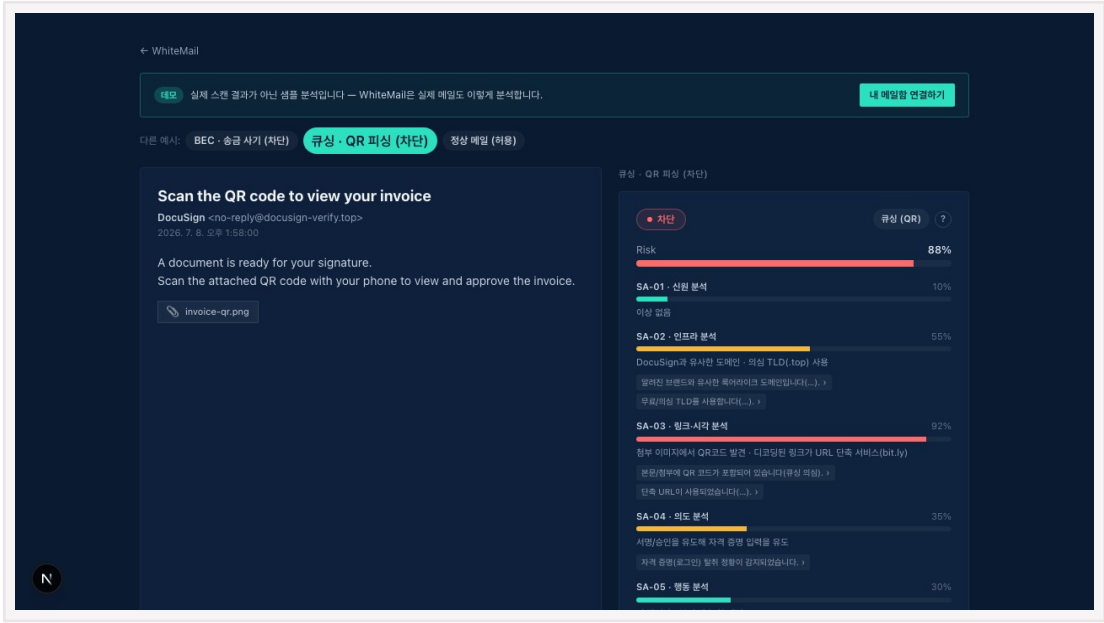


Risk 94% · SA-01 신원 (78%)·SA-04 의도 (92%) 등 각 신호와 근거를 그대로 공개

DEMO

데모 — 큐싱 (QR) 피싱

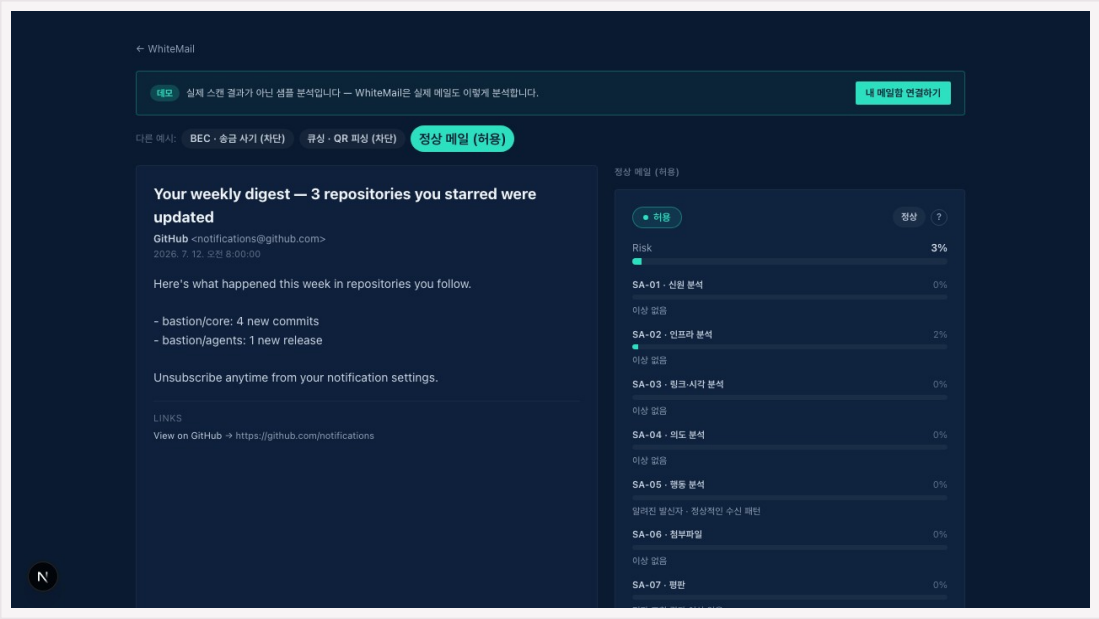
- SA-03 가 이미지 첨부의 QR 을 디코드해 페이로드 URL 을 추출
- 표시된 안내와 실제 목적지 도메인 불일치를 함께 판정
- 본문에 링크가 없어도 이미지 속 QR 로 숨긴 피싱을 잡음



큐싱 · QR 피싱 데모

DEMO

데모 — 정상 메일은 통과



정상 메일 (허용) 데모

- 정상 메일은 낮은 위험도로 통과 — 오탐을 억제
- 각 에이전트의 " 이상 없음 " 근거도 동일하게 공개
- 대조군 (정상 샘플) 으로 오탐 경계선을 상시 점검

EXPLAINABILITY

왜 이렇게 판정했는가 — 전부 공개

판정 구성	에이전트별 점수 + 신호 코드 + 사람이 읽는 근거 문장
드릴다운	신호를 클릭하면 " 왜 그 도메인이 다른지 " 실제 값까지 표시
블랙박스 없음	점수 하나가 아니라 조사 과정을 그대로 노출
활용	보안팀 검증 · 튜닝 · 사용자 교육 자료로 바로 사용

VS. LEGACY

기존 메일 게이트웨이와 무엇이 다른가

비교 축	기존 게이트웨이 (SEG)	WhiteMail
판단 기준	시그니처 · 발신지 평판 · 키워드	발신자의 의도와 관계 — 맥락으로 판단
잡는 위협	알려진 패턴의 스팸 · 악성코드	BEC · AI 피싱 · 큐싱 · VEC — 패턴 없는 표적 공격
판정 설명	점수 하나 · 차단 로그	에이전트별 근거 문장을 신호 단위까지 전부 공개
대응	스팸함 이동에서 종료	자동 격리 · 캠페인 스윕 · 사람 (보안팀) 승인 개입
도입	MX 변경 필요한 게이트웨이	API 5 분 연결 · 온프레미스 · 하이브리드

패턴이 없는 공격의 시대 — 판단 기준 자체가 달라야 잡힌다

근거로 판단하는 AI — RAG

① 임베딩
hash·local·Titan·OpenAI

→

② 벡터 검색
pgvector 코사인

→

③ 근거 주입
유사 공격 ·CVE

→

④ 판정
sLLM 추론

위협 인텔 RAG

SA-04 가 유사 공격 패턴을 검색해 의도 판단을 보강 .

공격 사례 RAG

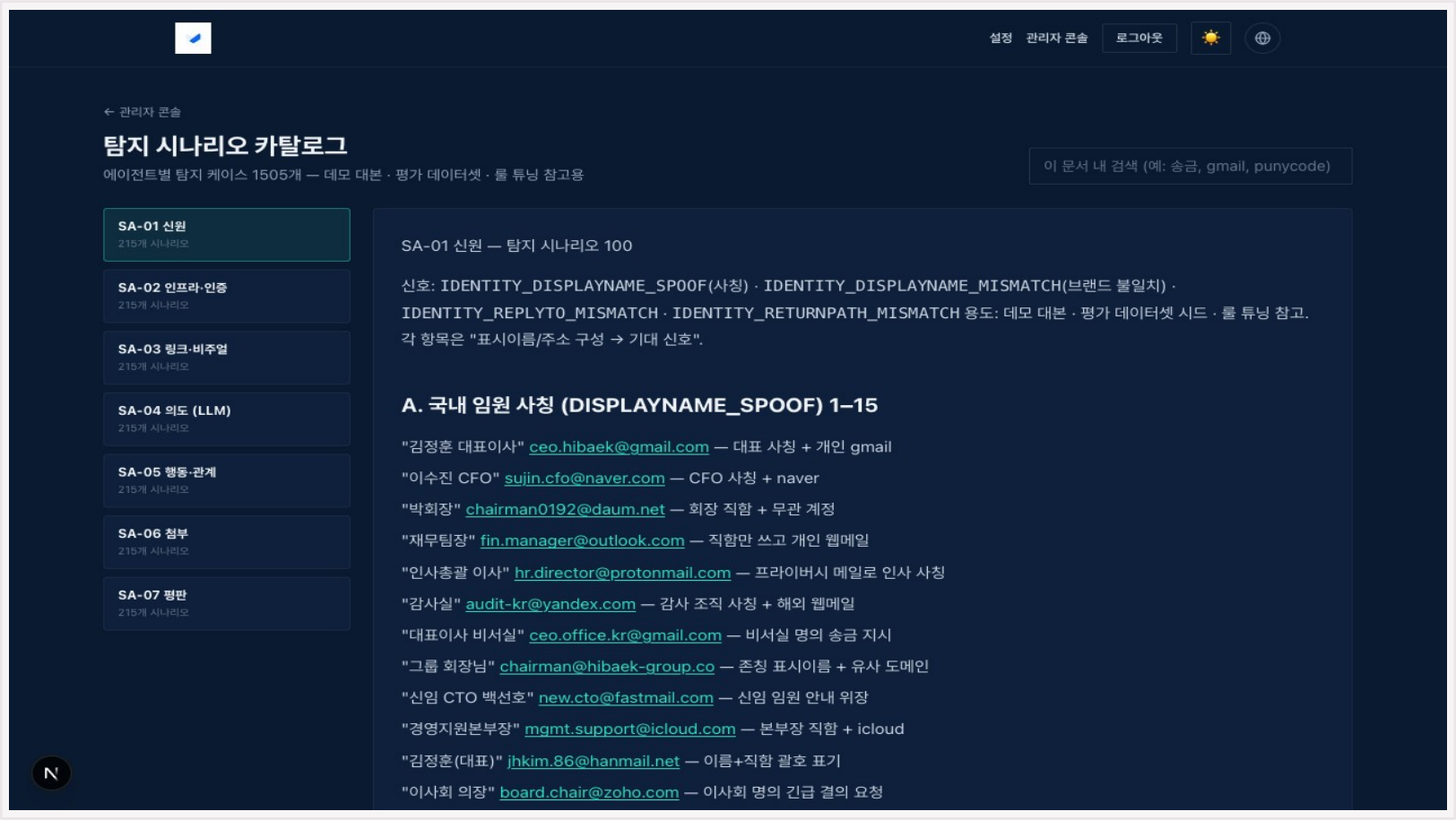
1,505 종 공격 시나리오에서 유사 사례를 부착 (relatedAttacks).

CVE RAG

메일이 CVE 를 참조하면 관련 CVE 를 조회 (relatedCves).

CORPUS

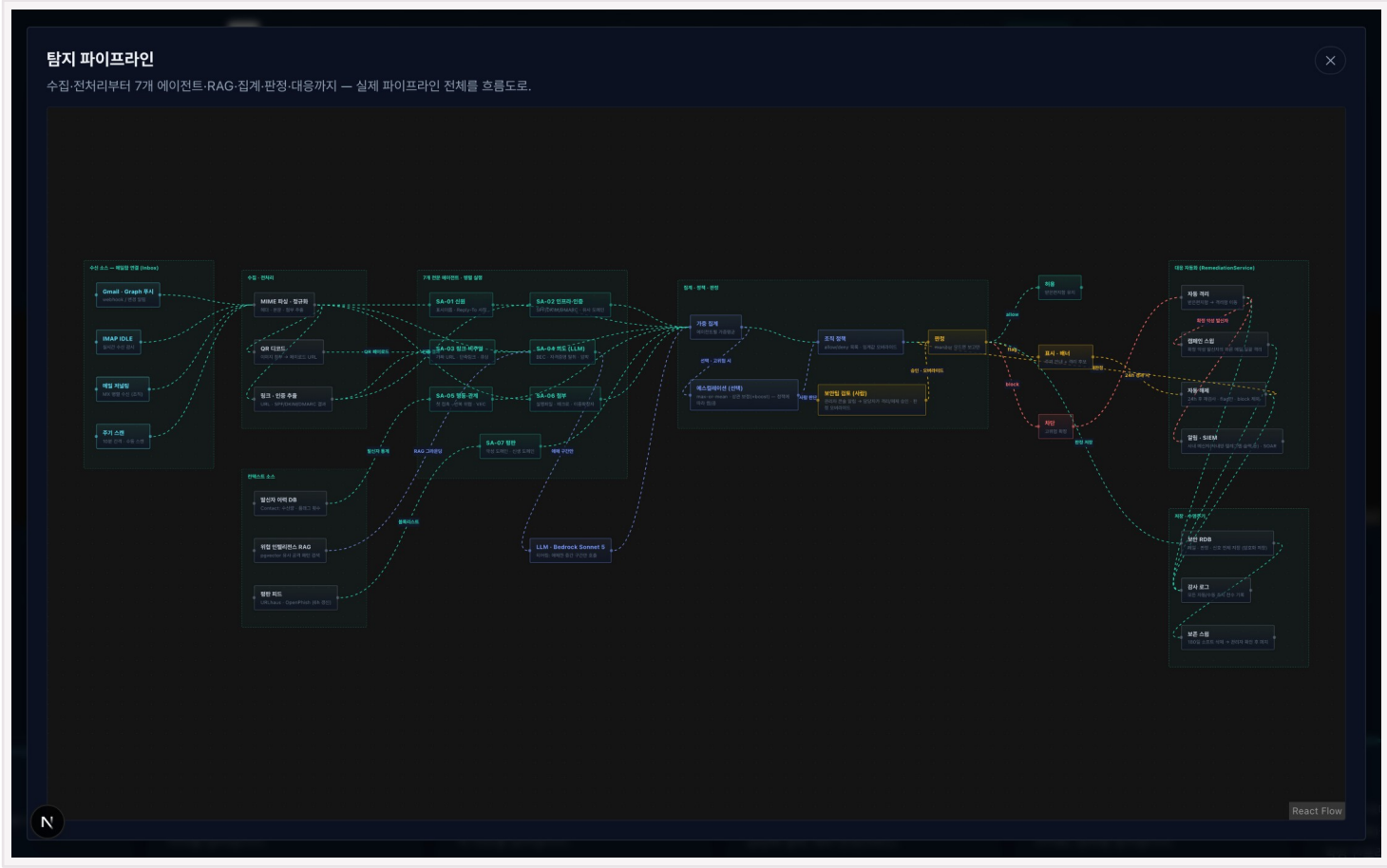
1,505 종 공격 시나리오 카탈로그



SA-01~07 × 215 — 데모 대본 · 평가 데이터셋 · 룰 튜닝 · 공격사례 RAG 코퍼스

PIPELINE

탐지 파이프라인 — 수집부터 대응까지



수신 소스 (Inbox) → 전처리 → 7 에이전트 + RAG → 집계 → 에스컬레이션 (선택) · 보안팀 검토 (사람) → 판정 → 대응 → 저장

에이전트 하나하나가 설명된다 — SA-01 신원

입력 → 검사 → 신호 종합 → 점수 + 근거 흐름도 + 어려운 용어 풀이 (용어사전 ·Wikipedia 연결)

CUSTOMIZE

조직 · 부서별 파이프라인 커스터마이즈

JSON 정의	탐지 파이프라인을 코드가 아니라 JSON 데이터로 정의
조직별 렌더	조직 설정에 저장하면 그 조직 콘솔에 그대로 렌더
부서별 구성	재무본부 강화 규칙 등 부서마다 다른 파이프라인
룰 토글	탐지 룰을 조직 단위로 켜고 끄면 재집계에 즉시 반영

RESPONSE

대응 자동화

- **정책 기반 자동 격리 / 해제**
- 캠페인 스윕 — 동일 위협을 일괄 대응
- 주기 스캔 (10 분) · 실시간 IMAP IDLE · 메일 저널링
- SIEM · Webhook 로 외부 보안 스택 연동
- Monitor-only(저널링) 모드 — 조직 정책에 따라 자동 조치 보류

수신함 · 스캔 현황

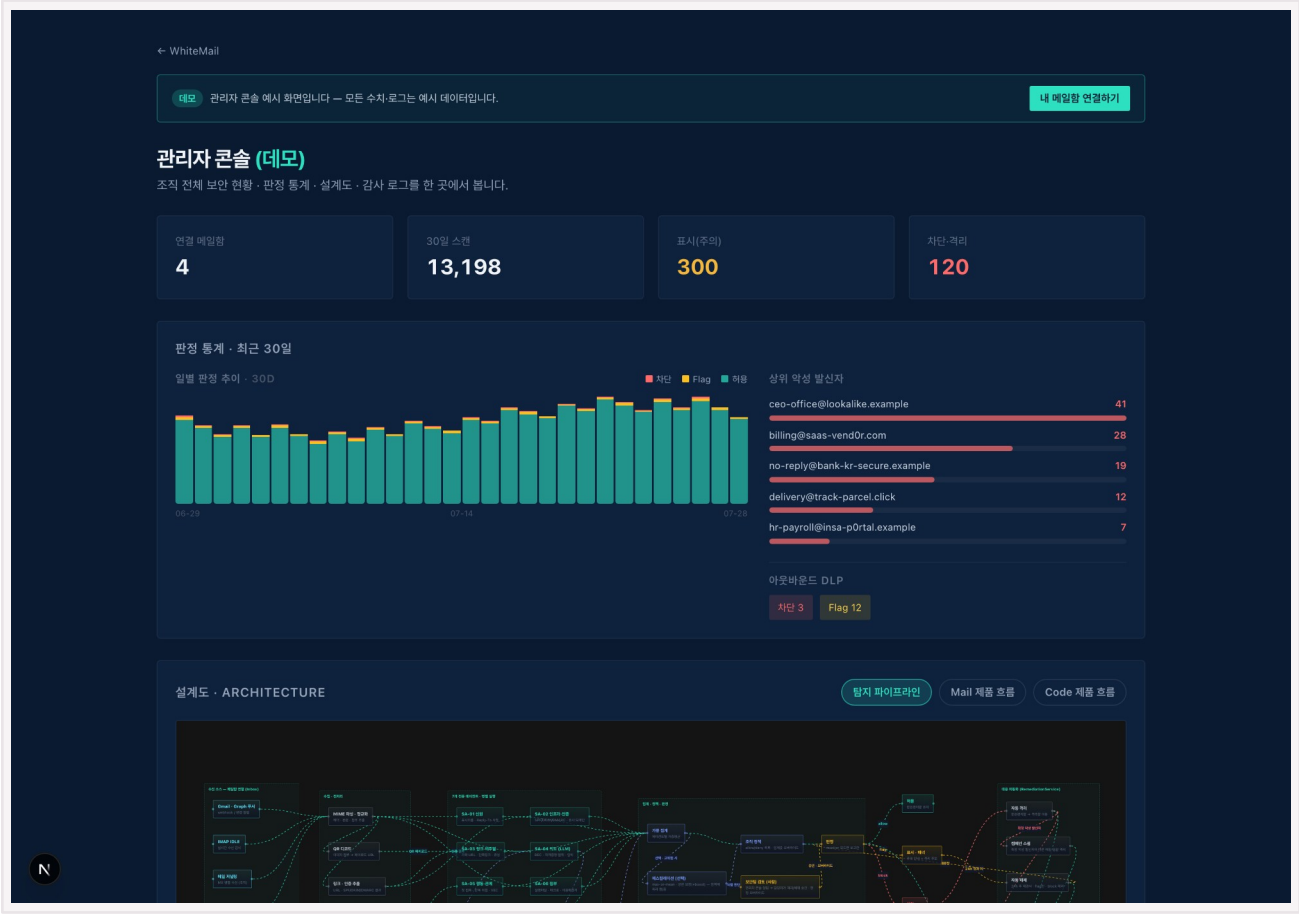
OUTBOUND DLP

나가는 메일의 데이터 유출도 본다

발신 감사	아웃바운드 DLP — 나가는 메일의 민감정보 유출 분석
탐지 항목	주민번호 (날짜 검증) · 카드 (Luhn) · 계좌 · 시크릿 키 · 기밀 키워드
행위 탐지	개인 웹메일 대량 발송 · 대량 수신자 · 대용량 첨부
프라이버시	매칭된 값은 마스킹 — 로그에 원문 미기록

ADMIN

관리자 콘솔 — 한 곳에서 운영



조직 · 설계도 (파이프라인 / 배포 / 시퀀스) · 탐지 룰 · 에이전트 케이스 · 통계

이메일 통계 — 무엇이 오고 , 누가 노리나

30 일 추세	스캔량 · 차단 · 격리를 유형별로
표적 분석	Top 발신자 · 가장 많이 공격받는 구성원
위협 분포	BEC · AI 피싱 · 큐싱 · 기타
아웃바운드 요약	나가는 유출 시그널 집계

LIVE TOPOLOGY

실시간 메일 흐름 — 서버 ·MX 를 눈으로

흐름	인터넷 → MX → WhiteMail API → 7 에이전트 → 받은편지함 / 격리함
실시간 렌더	React Flow — 각 노드 상태 (OK/UP) · 처리량 표시
인프라 가시화	LLM(Bedrock/ 로컬) · 보안 RDB 위치를 한눈에
위치 · 스케줄	리전 (예 : Seoul ap-northeast-2) · 스케줄 표기

ON-PREMISE

sLLM 학습 → 온프레미스에도 그대로

① 도메인 학습

한국어 공격 · 의도 데이터로 sLLM 파인튜닝

→

② 온프렘 서빙

고객 폐쇄망에 vLLM 으로 배포

→

③ 로컬 판정

본문이 밖으로 안 나감 (sovereign)

LLM_PROVIDER=local · EMBED_PROVIDER=hash|local — 의도 모델 · 임베더가 모두 로컬

/api/config 의 sovereign=true 로 " 본문이 외부로 나가지 않음 " 을 검증

진짜 에어갭은 내부 메일서버 IMAP 로 연결 — 인터넷 없이 완결

금융 · 공공 · 망분리 환경 : 본문을 클라우드에 보내지 않고도 맥락 기반 보안이 가능

DATA SOVEREIGNTY

데이터 주권 — 지원이 아니라 절차로

완전 온프렘 / 에어갭

로컬 sLLM + 로컬 임베더 + pgvector — 본문이 배포 경계를 안 벗어남 .

sovereign 검증

/api/config 로 외부 전송 0 을 배포 시 체크리스트로 확인 .

보존 · 파기

사용자별 보존기간 · 만료 소프트삭제 · 로그 리댁션 (본문 · 시크릿 미기록).

보안 운영 절차

반입 무결성 검증 · 최소권한 · 접근통제 · security.txt.

참고 : docs/deployment/on-prem-security-procedures.md — 도입 심사 · 제안서 자산

DEPLOYMENT

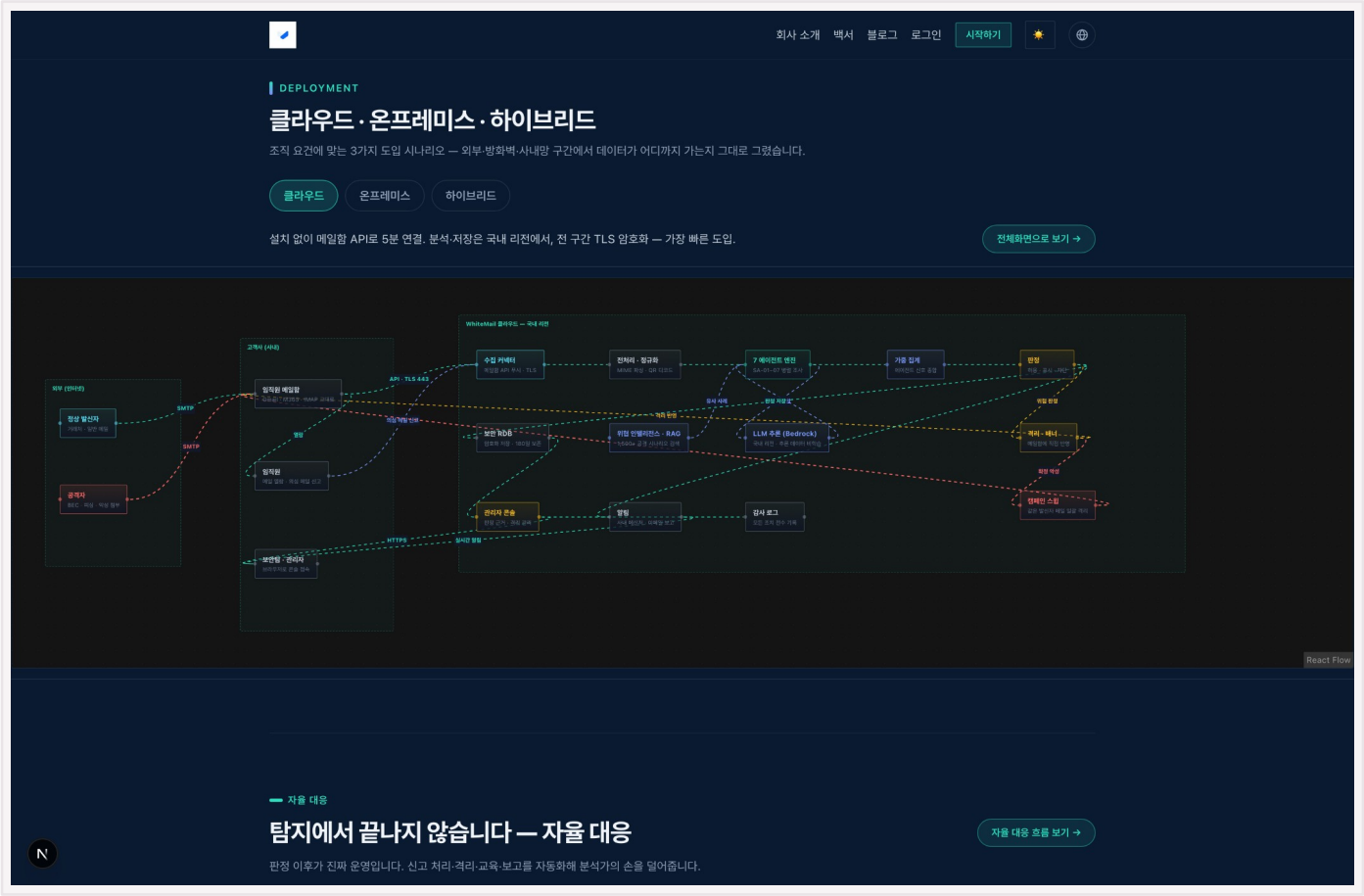
도입 방식 — 세 가지

클라우드 (SaaS)	무료로 즉시 시작 — OAuth 로 메일함 연결, 국내 리전
온프레미스	고객 폐쇄망에 sLLM· 임베더까지 로컬 배포 — 메일 비반출
하이브리드	메일은 회사 안에, 분석은 클라우드에 — 보안 최적화

OEM · 화이트라벨 · 리셀러 · 온프렘 구축 파트너십 환영

TOPOLOGY ①

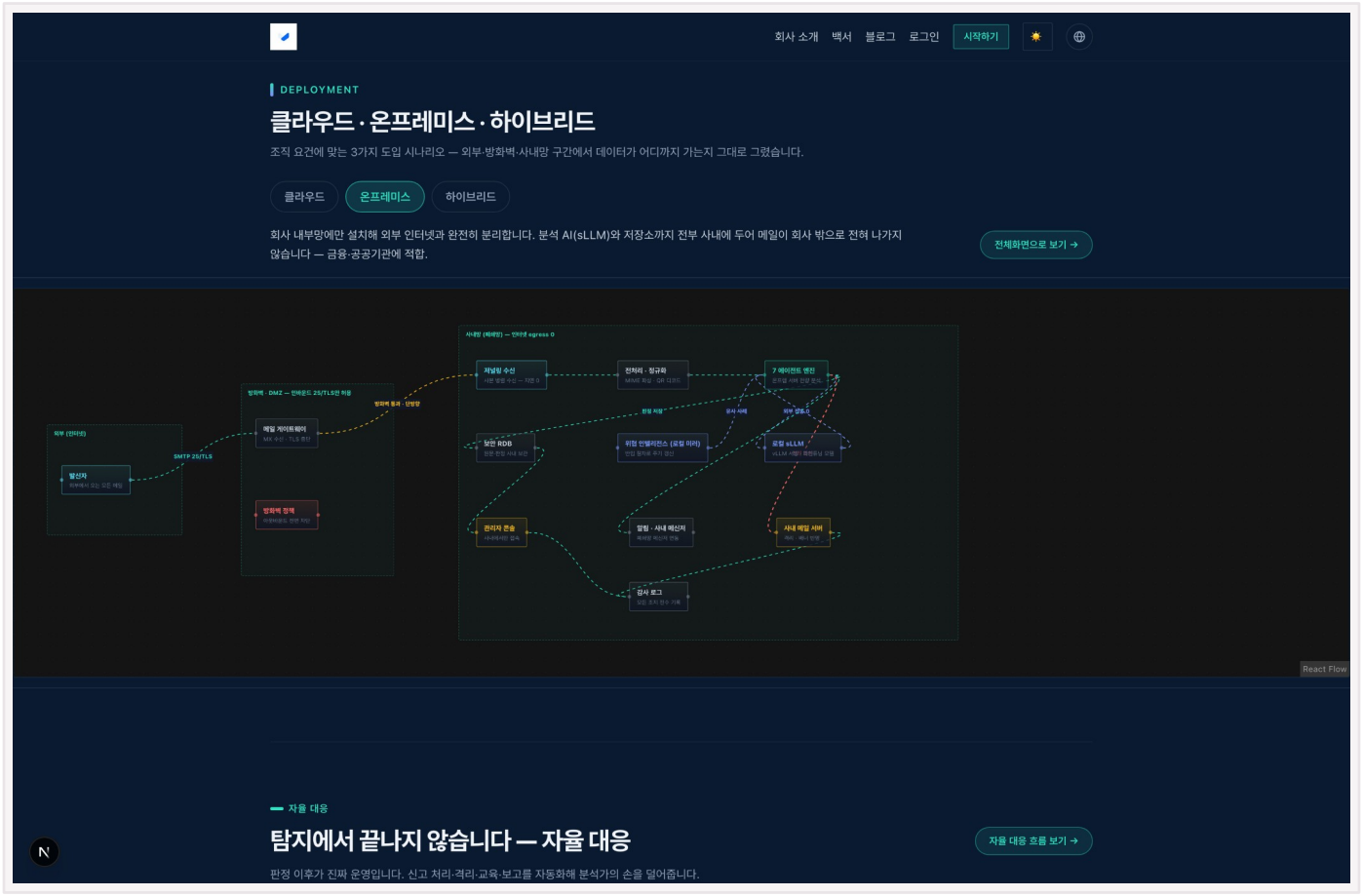
클라우드 — 설치 0, 메일함 API 로 5 분 연결



외부 → 고객사 메일함 → (API·TLS) 국내 리전 : 전처리 → 7 에이전트 → 집계 · 판정 → 격리 · 스윕 → 감사

TOPOLOGY ②

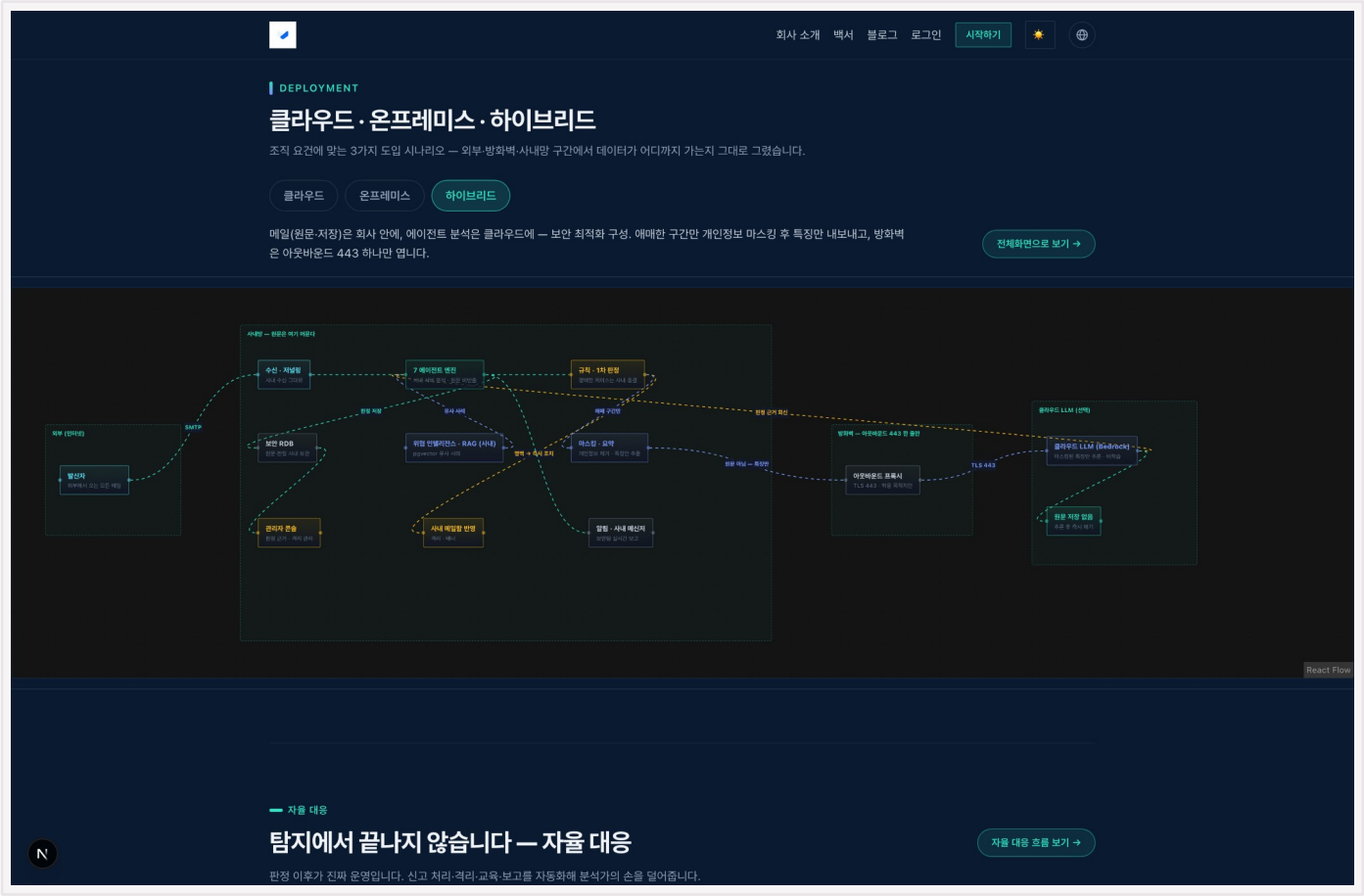
온프레미스 — 폐쇄망 완결 , 인터넷 egress 0



외부 → 방화벽 ·DMZ(인바운드 25/TLS 만) → 사내망 : 저널링 → 엔진 → 로컬 sLLM · 보안 RDB — 원문 비반출

TOPOLOGY ③

하이브리드 — 메일은 회사에 , 분석은 클라우드에



원문 · 저장은 사내 · 애매 구간만 마스킹 후 특징만 443 으로 → Bedrock 추론 (원문 저장 없음 , 즉시 폐기)

적용 비용 — 도입 방식별

도입 방식	인프라	종량 과금	도입비 · 구축 비용
클라우드 (SaaS)	WhiteHat 클라우드 — 국내 리전	일반 처리 건당 1원 · 딥스캔 건당 10 원	없음 — 설치 0 · 5 분 연결 · 무료 체험
회사 인프라 제공형	고객사 제공 서버 /VM 에 설치	일반 처리 건당 1원 · 딥스캔 건당 10 원	도입비 별도 협의
온프레미스 구축형	폐쇄망 풀 구축 (sLLM 포함)	연 라이선스 — 별도 협의	인프라 구축 비용 별도 협의
하이브리드	사내 + 클라우드 LLM	일반 처리 건당 1원 · 딥스캔 건당 10 원	도입비 · 부분 구축 비용 별도 협의

기준 단가 : 일반 처리 1원 · LLM 딥스캔 10 원 (위험 신호 메일에만 자동 적용 , 통상 5~10% · 조직 설정으로 상한 조절) — 100 명 조직 월 40~60 만 원 수준 . 도입비 · 인프라 구축 비용은 별도이며 규모 ·SLA 에 따라 최종 협의 .

SUMMARY

WhiteHat Mail 요약

탐지	7- 에이전트 맥락 조사 + RAG 근거 보강 + 설명가능한 판정
대응 · 운영	자동 격리 · 스윕 ·SIEM · 아웃바운드 DLP · 부서별 커스터마이즈
데이터 주권	sLLM 학습으로 완전 온프레미스 — 본문이 밖으로 안 나감
도입	무료 시작 → 조직 / 엔터프라이즈 → 온프렘 구축

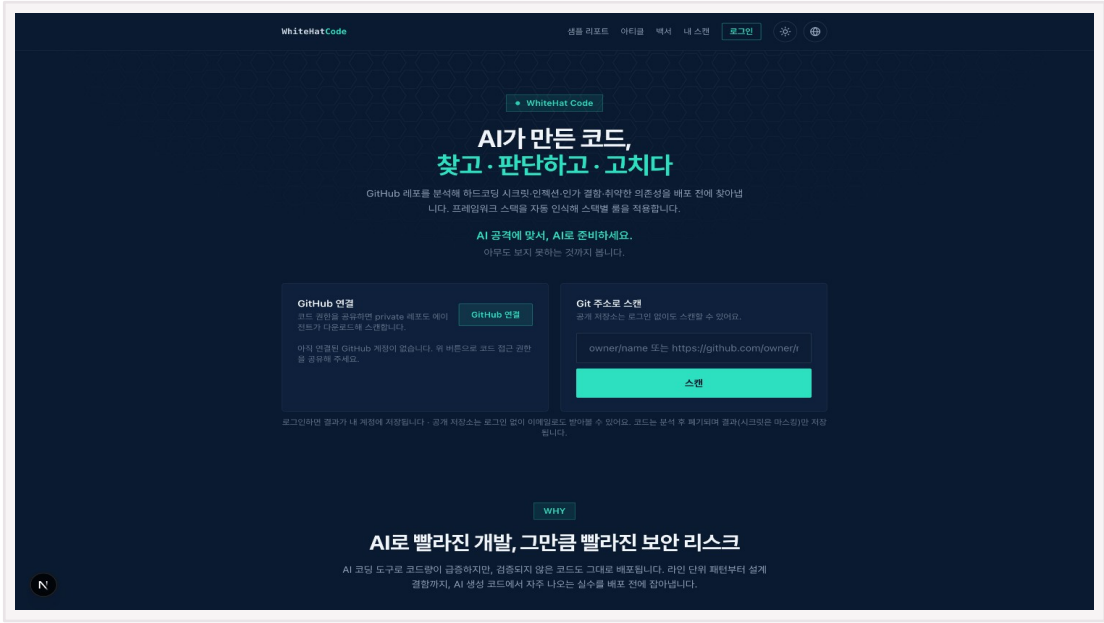
PART 2

WhiteHat Code

AI 코드 보안 스캔 — 바이브 코딩 (AI 생성 코드) 까지 배포 전에 잡는다

한 줄 소개

- **AI 가 만든 코드 — 찾고 · 판단하고 · 고치다**
- GitHub 연결 즉시 자동 스캔 · Git 주소만으로 로그인 없이 스캔
- 정적 규칙 + OSS 취약점 (SCA) + 환각 패키지 + LLM 딥스캔
- 소스는 분석 후 파기 · 결과 (시크릿 마스킹) 만 저장



AI 공격에 맞서 , AI 로 준비하세요

WhiteHat Code 랜딩

THE PROBLEM

바이브 코딩은 빠르지만 안전하지 않다

45%

AI 생성 코드의
OWASP Top 10 취약점 (Veracode)

19.7%

LLM 이 추천한
존재하지 않는 패키지 (USENIX 2025)

↑

AI 코딩 도구 기인 CVE
급증 (Georgia Tech 2026)

- 검증 없이 배포되는 AI 코드 — 라인 단위 취약점부터 설계 결함까지 .
- 환각 패키지를 공격자가 선점하는 slopsquatting 은 인정된 신종 공급망 위협 .
- 우리는 규칙 · 의존성 · 환각 ·LLM 딥스캔을 한 번에 돌린다 .

두 개의 전략 , 하나의 판정

바이브 코딩 보안 + 기존 코드 보안 → 공격 / 방어 관점으로 수렴 → 취약점→코드 보완

FLOW

스캔 흐름

입력 · 클론
GitHub·PR· 임시 클론

→

정적 ·SCA
SAST·OSV· 환각 패키지

→

LLM 딥스캔
map-reduce + RAG

→

판정
심각도 산정

→

PR 게이트
승인정도 · 리포트 ·SARIF

소스코드는 스캔 후 파기 · 외부 워커 모드에서는 API 호스트에 닿지도 않음

정적 SAST

시크릿 하드코딩 · 인젝션 · 인증 · 암호 · 전송 · 설정 규칙 , 스택 자동 인식 .

SCA / OSV

의존성 취약점 — 로컬 OSV/CVE 미러로 버전 매칭 (온프렘 가능).

구조 추출

엔드포인트 · 스택 · 시퀀스 / 유스케이스 다이어그램 자동 생성 .

NEW

환각 패키지 탐지 (slopsquatting)

- **선언됐으나 레지스트리에 실존하지 않는 의존성을 플래그**
- npm·PyPI·Go 프록시에 실존 확인 → 없으면 slopsquatting 위험
- 공격자가 그 이름을 선점해 악성 코드를 심을 수 있음
- 에어갭이면 자동 스킵 (오탐 방지) · 경쟁사가 말로만 하는 축을 실제 탐지로

"react"= 정상 , 지어낸 이름은 high 로 플래그 (실증 완료)

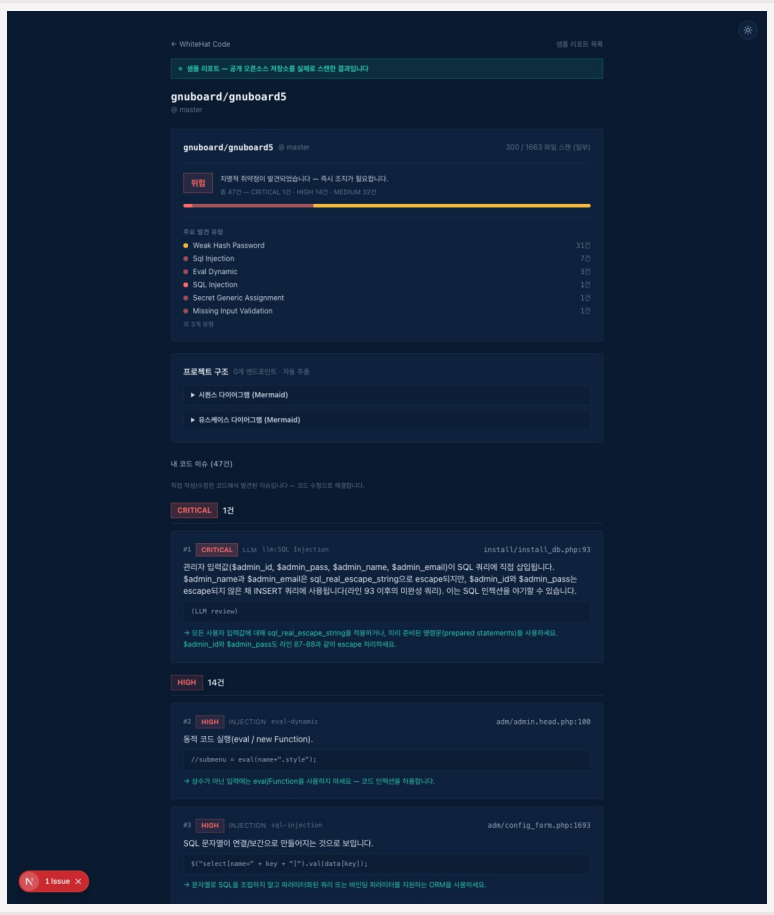
전략도의 " 환각 패키지 [NEW]" 노드

LLM 딥스캔 — 규칙이 못 보는 로직 · 인가 버그

map-reduce	보안 관련 파일만 상한 내 청크로 분석 → 병합
인가 감사	추출된 엔드포인트를 근거로 authz · IDOR · 입력검증 점검
RAG 근거	CVE· 공격사례를 주입 — " 실제로 일치할 때만 인용 "(환각 방지)
티어링	경량 ~Opus 모델 선택으로 정밀도 · 비용 조절

REAL REPORT

실제 스캔 리포트 — gnuboard5

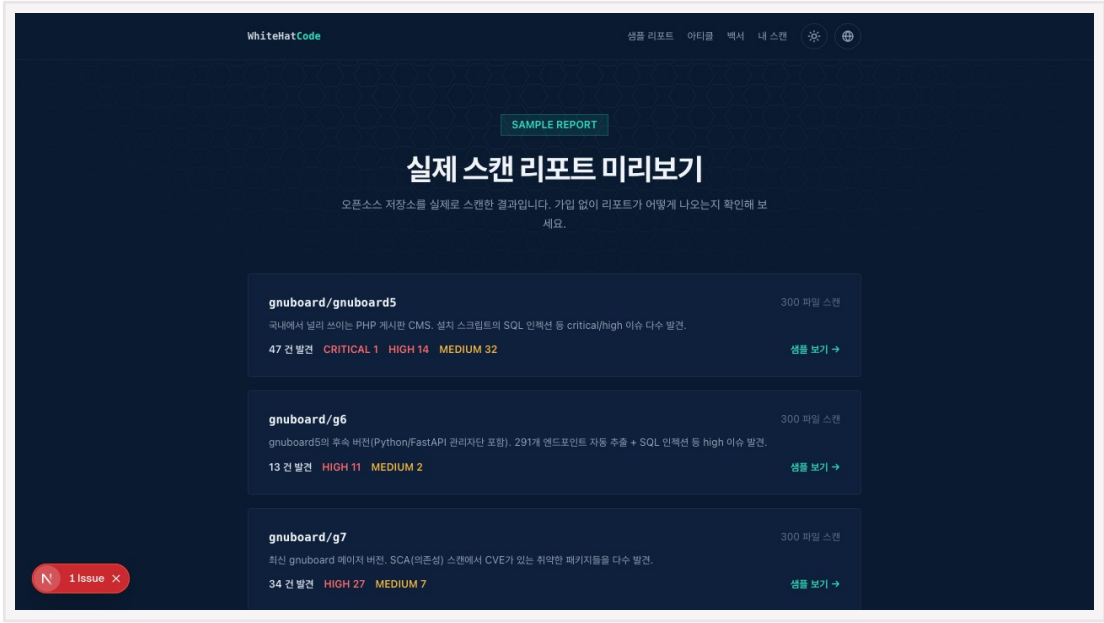


CRITICAL SQL 인젝션 · LLM 이 원인 · 라인 · 수정안까지 한국어로 설명

SAMPLE

가입 없이 보는 샘플 리포트

- gnuboard5 · g6 · g7 — 실제 오픈소스 스캔 결과 공개
- CRITICAL/HIGH/MEDIUM 심각도 분포와 발견 유형
- SCA(의존성) 에서 CVE 있는 취약 패키지도 발견
- 가입 없이 리포트가 어떻게 나오는지 미리보기



샘플 리포트 목록

동적 러너 — 실제 실행 관점

RUNNER 티어	샌드박스에서 코드를 실행해 익스플로잇 관점 확인
런타임 재현	정적 분석이 놓치는 런타임 취약점을 재현
안전 설계	격리된 실행 환경 (untrusted code) 에서만 실행

국내 경쟁사는 정적 중심 — 동적 실행은 잠재 차별점

PR GATE

승인 정도 — 판정이 PR 을 얼마나 막을지

Check Run 게이트	PR 마다 자동 스캔 결과를 게이트로 — CI 보안 게이트
advisory	차단 없음 , 리포트만 — neutral
approval	차단 이슈면 사람 승인 필요 — action_required
auto (기본)	critical/high 면 자동 차단 — failure

SARIF 내보내기로 GitHub 코드스캔 연동

CODE TRANSMISSION

코드 전달도 정책으로 — 외부 LLM= 노출 아님

local-only

외부 전송 0 — 로컬 sLLM 일 때만 딥스캔 . 금융 · 공공 · 망분리 .

redacted

선별 청크 전송하되 전송 전 시크릿 · 자격증명 마스킹 .

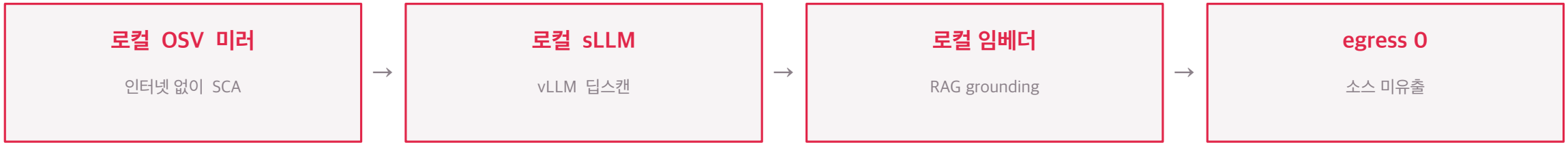
snippets (기본)

보안 관련 파일만 상한 내 청크로 분석 목적 전송 .

공통 : 소스 영구 저장 없음 (스캔 후 파기) · 엔터프라이즈 LLM 은 추론 데이터 비학습 (제공사 정책)

ON-PREMISE

sLLM 딥스캔 — 폐쇄망 안에서 완결



OSV_SOURCE=local · LLM_PROVIDER=local · EMBED_PROVIDER=hash|local

경쟁사는 대개 정적 스캔만 로컬 , AI 딥스캔은 클라우드 — 우리는 딥스캔 ·RAG 까지 로컬 완결

sLLM 은 도메인 데이터로 파인튜닝 (H200) 해 정밀도 확보 — 온프렘 배포

폐쇄망 완결성이 실질 해자 — 데이터 주권은 입장권 , 완결성이 차별점

STRUCTURE

구조 추출 — 스캔이 곧 문서가 된다

결정적 추출 (0 토큰)	REST 엔드포인트 자동 매핑
다이어그램 자동 생성	시퀀스 · 유스케이스 · API (Mermaid)
스택 인식	예 : Supabase 감지 → 스택별 규칙 게이트 + LLM 컨텍스트
이중 활용	내부 문서 + LLM 딥스캔의 근거 (grounding)

INTEGRATION

GitHub 연동 · CI 게이트 · SARIF

GitHub App	설치 즉시 PR 마다 자동 스캔 + Check Run 게이트
Private 레포	설치 토큰으로 안전하게 클론 — 스캔 후 파기
SARIF	GitHub 코드스캔 등 표준 도구와 통합
외부 워커 모드	소스가 API 호스트에 닿지 않고 워커에서만 처리

코드 생성 시점 보안 — MCP 로 IDE 에 붙인다 (별도 구축 가능)

개념	메일을 " 발송할 때 " 감사하듯 , 코드는 " 생성할 때 " 검사
MCP 서버	기존 엔진 (규칙 ·SCA· 환각 · 딥스캔) 을 Claude·Cursor·Copilot 에 노출
즉시 경고	취약 패턴 · 환각 패키지를 생성 즉시 경고 — 커밋 전 차단
제공 형태	별도 구축 (맞춤 개발) — 온프렘 MCP 까지 가능

COVERAGE

지원 범위

정적 규칙	시크릿 · 인젝션 · 인증 · 암호 · 전송 · 설정 카테고리
SCA 생태계	npm · PyPI · Go — OSV/CVE, 로컬 미러 가능
환각 패키지	npm · PyPI · Go 레지스트리에서 실존 확인
LLM 딥스캔	스택 무관 · 리포트 언어 한국어 / 영어 / 일본어

온프렘 도입 — 인터넷 없이 스캔하기

① 미러 반입	OSV/CVE 를 온라인 스테이징에서 받아 해시 검증 후 폐쇄망 적재
② 로컬 서빙	sLLM· 임베더 온프렘 서빙 (vLLM) — 딥스캔 ·RAG 를 로컬에서
③ 기동	OSV_SOURCE=local · LLM_PROVIDER=local → 인터넷 연결 0
④ 검증	배포 체크리스트 — sovereign · 미러 건수 · 소스 비저장 확인

PRICING

가격 사다리 — 개발자부터 엔터프라이즈까지

FREE	자동 스캔 — 무료
EXPERT	자동 스캔 + 전문가 분석 — 15 만원 / $100
RUNNER	+ 로컬 러너 실행 — 150 만원 / $1,000
FULL	+ 문제 해결 · 컨설팅 · 배포 — 별도 협의
증분 스캔 (CI 상시)	PR· 커밋마다 변경분만 상시 스캔 — 별도 협의

무료 셀프서브로 진입 — 엔터프라이즈 영업 중심 경쟁사와 차별

경쟁 우위

축	WhiteHat	통상 경쟁 제품
폐쇄망 완결	SCA+LLM 딥스캔 +RAG 까지 온프렘 (sLLM)	정적 스캔만 로컬 , AI 는 클라우드가 다수
환각 패키지	slopsquatting 실제 탐지 신호	언급 / 마케팅 위주
설명가능성	원인 · 라인 · 수정안까지 한국어로	취약점 나열 위주
가격 진입	무료 셀프서브 → 티어 사다리	엔터프라이즈 영업 중심

국내 Sparrow·Sentrix, 해외 Snyk·Endor 대비 — 완결성 · 한국어 · 설명가능성 · 가격의 결합

SUMMARY

WhiteHat Code 요약

커버리지	정적 + SCA + 환각 패키지 + LLM 딥스캔 — 바이브 코딩까지
설명가능성	원인 · 라인 · 수정안까지 한국어 리포트 · SARIF · PR 게이트
데이터 주권	sLLM 딥스캔 · 로컬 OSV 미러로 폐쇄망 완결 (온프렘)
가격 사다리	무료 셀프서브 → EXPERT → RUNNER → FULL

인프라 구성 — Cloud · 온프레미스 모두

CLOUD · SaaS

사용자 · 메일함 / 저장소	→	whitehat.ai.kr	→	Bedrock	→	매니지드 DB · 미러
OAuth · GitHub		web · api		Sonnet + Titan · Seoul		Postgres · OSV

ON-PREMISE · 폐쇄망 (egress 0)

내부 메일 · 저장소	→	WhiteHat 온프렘	→	로컬 sLLM · 임베더	→	로컬 OSV 미러 · 내부 DB
IMAP · 내부 Git		동일 코드베이스		vLLM · e5/hash		Postgres · pgvector

같은 코드베이스 — LLM_PROVIDER · EMBED_PROVIDER · OSV_SOURCE 만 전환 . Cloud= 편의 , 온프렘 = 본문 · 소스 미유출 (sovereign).

시장 배경 — 검증된 , 자본이 몰리는 수요

$28M

이메일 보안 Ocean
시리즈 A (Lightspeed)

$93M

AI 코드 보안 Endor Labs
시리즈 B

45% · 19.7%

AI 코드 취약 · 환각 패키지
(Veracode · USENIX)

- 이메일 · 코드 양쪽에서 "AI 로 정교해진 위협 " 이라는 같은 문제에 자본과 수요가 몰린다 .
- 우리는 두 시장을 하나의 팀 · 기술 (맥락 판단 + RAG + 데이터 주권) 로 함께 겨냥한다 .
- 아직 초기 성장 구간 — 한국어 · 온프렘 · 설명가능성으로 국내외 틈을 파고든다 .

LANDSCAPE

경쟁 지형 — 우리는 어디에 있나

플레이어	영역	강점 (공개 자료)	우리와의 차이
Ocean (ocean.security)	이메일	에이전틱 이메일 보안 · 시리즈 A $28M · 월 10 억 통 +	클라우드 · 영어 중심 — 온프렘 · 한국어 없음
Sparrow (스패로우)	코드	국내 SAST/SCA 업력 · MCP 연동 · 대기업 · 공공	AI 딥스캔 · 환각 탐지의 실제 신호는 우리가 앞섬
Sentrix CodeScan	코드	국내 AI SAST · 13 개 언어 OWASP + 수정 코드	폐쇄망 AI 완결 · 이메일 통합 없음
Snyk · Endor Labs	코드	해외 대형 · Endor 시리즈 B $93M	온프렘 AI 완결 · 한국어 리포트 약함

WhiteHat 의 자리 이메일 + 코드를 한 팀 · 한 기술로 · 폐쇄망 완결 (sLLM) · 한국어 · 설명가능성 · 무료 셀프서브 가격 사다리 — 경쟁사가 하나씩 가진 강점을 결합 .

GET STARTED

WhiteHat

무료로 시작 · 온프레미스 구축 문의 · 제휴 문의
whitehat.ai.kr

연락처 · WhiteHat 백선호 상무 · whitehat.ai.kr

CONTACT

도입 · 제휴 문의

도입 상담

무료 체험 → PoC → 조직 / 엔터프라이즈 — 규모에 맞는 구성을 제안드립니다.

제휴 파트너

OEM · 화이트라벨 · 리셀러 · 온프렘 구축 · 채널 영업 — 상시 열려 있습니다.

문의 white@whitehat.ai.kr · WhiteHat 백선호 상무

whitehat.ai.kr · mail.whitehat.ai.kr · scan.whitehat.ai.kr