

White Paper

AhnLab AI PLUS와 함께하는 에이전틱 AI 보안 혁신

오늘 아침 출근길에 내 업무의 일부를 AI가 알아서 처리해 준다면?

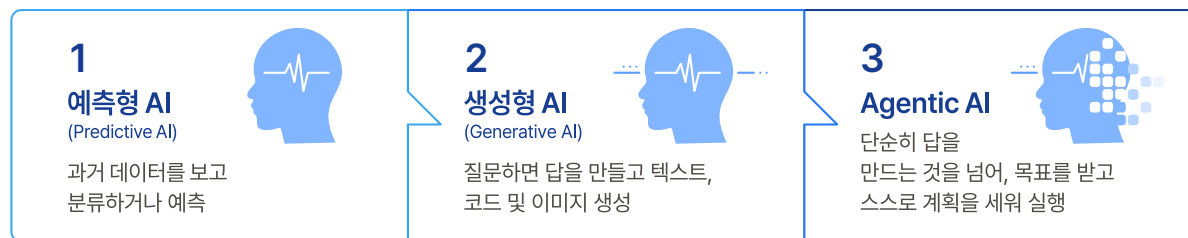
어떤 이는 편하다고 느낄 것이고, 또 다른 이는 AI가 내 직업을 대체할 지 모른다는 불안감을 느낄 수도 있다.

중요한 건, 이런 고민이 이제 상상이 아니라는 점이다. AI는 이제 단순 답변 제시를 넘어, 업무를 함께 처리하는 방향으로 빠르게 진화하고 있다. 그리고 그 중심에는 '에이전틱 AI(Agentic AI)'가 있다.

이번 글에서는 에이전틱 AI의 개념에 대해 살펴보고, 보안 담당자 입장에서 대응 전략과 안랩의 에이전틱 AI 보안 플랫폼 AhnLab AI PLUS의 가치 및 도입효과를 알아본다.

Agentic AI는 무엇인가?

AI가 진화해 온 흐름을 큰 틀에서 보면 다음 세 단계로 나뉜다.



쉽게 말하면, '알려주는 AI'에서 '수행하는 AI'로 넘어가는 것이다. 그리고 에이전틱 AI의 핵심 가치는 다음 네 가지다.

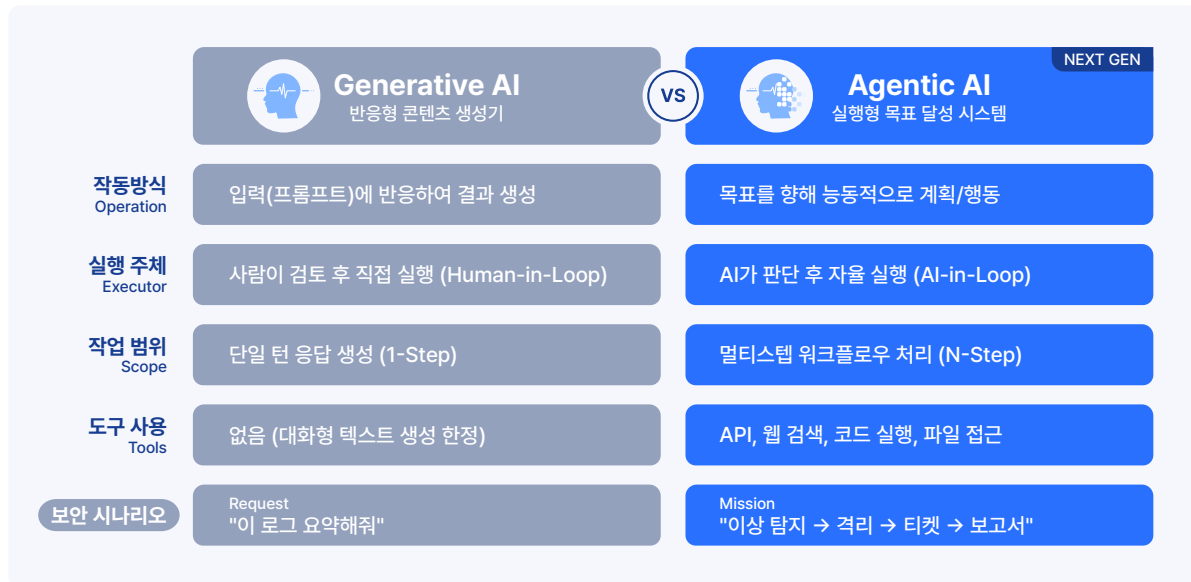
자율성 AI (Autonomy) 사람의 지시를 계속 기다리지 않고 스스로 판단	목표 지향성 AI (Goal-Driven) 목표 달성을 위해 여러 단계의 작업 수행
도구 사용 (Tool Use) 단순 텍스트 생성을 넘어, 외부 시스템과 능동적으로 상호작용	메모리 & 학습 (Memory) 이전 결과를 기억하고 다음 행동에 반영

즉, Agentic AI는 단순 자동화가 아닌 목표 중심의 자율 실행 시스템이다.

Agentic AI는 무엇이 다른가?

기존 생성형 AI와 에이전틱 AI의 가장 큰 차이는 반응형인지 실행형인지 여부다. 생성형 AI는 보통 우리가 질문하면 답을 준다. 여기까지는 사람이 중심이다. 반면, 에이전틱 AI는 사람이 목표를 주면 스스로 순서를 정하고, 필요한 도구를 써서 결과를 만들어낸다.

보안 관점에서 보면, “이 로그 좀 요약해 줘” 수준이었다면, 이제는 “이상 징후를 찾고, 필요한 정보를 모으고, 격리하고, 티켓 만들고, 보고서까지 작성해 줘”라는 흐름으로 가는 것이다. 보안도 과거에 사람 중심의 제어 환경이 었다면, 이제는 AI가 중심이 되어가고 있는 것이다.



[그림 1] 생성형 AI vs 에이전틱 AI

최근 에이전틱 AI에 대한 관심도가 부쩍 높아졌는데, 그 이유는 다음 세 가지의 성숙도가 높아졌기 때문이다.

LLM의 추론 능력 단순 요약을 넘어 복잡한 단계의 문제 처리 가능	에이전트 프레임워크/프로토콜 도구를 연결하고 실행하는 기반 개선	기업 환경이 API 중심으로 전환 보안 솔루션, 티켓 시스템, 클라우드, 인증이 모두 연결 가능한 구조
--	--	--

즉, 기술과 연결 환경이 준비되었고, 실제 수요도 생겨났기 때문이다.

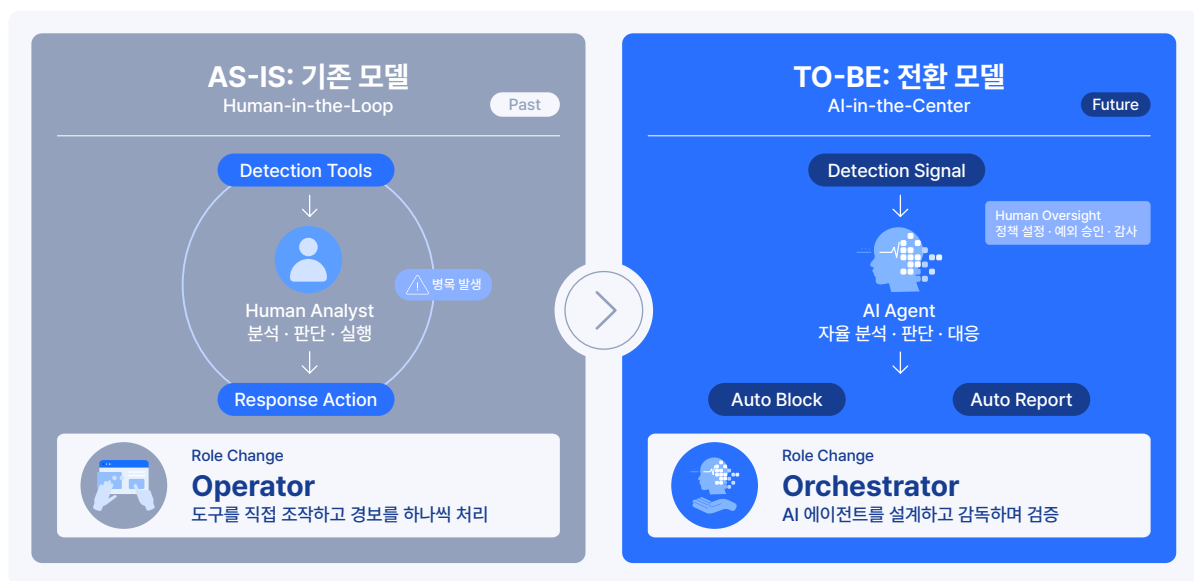
에이전틱 AI에 대한 시장 전망도 굉장히 공격적이다. 시장조사기관 가트너(Gartner)는 전체 브랜드의 60%가 Agentic AI를 통해 1:1 맞춤형 고객 인터랙션을 제공할 것으로 전망했다. 중요한 것은 AI가 ‘보조 기능’을 넘어 업무를 수행하는 ‘조직의 구성 요소’가 된다는 점이다. 이는 조직들의 제품 설명, 운영 방식 등에 대한 고객 기대 수준 자체가 달라질 가능성이 높다는 의미이기도 하다.

AI와 함께하는 보안 담당자의 역할 변화

보안 담당자 입장에서 이러한 변화에 대해 주목해야 할 키워드는 크게 네 가지가 있다.

에이전틱 AI 감독 (Agentic AI Oversight) 보이지 않는 AI 에이전트가 늘어나면서 새로운 공격 표면 발생	AI 전용 인증 및 액세스 관리 (IAM for AI Agents) 사용자 뿐만 아니라 AI 에이전트도 하나의 행위 주체로 관리 필요
AI 주도 SOC (AI-Driven SOC) 반복적인 분석 업무는 AI가 더 많이 맡게 될 전망	선제적 보안 (Preemptive Security) 사고 이후가 아닌 사전 징후 단계 방어

핵심은 운영, 권한, 탐지, 대응 체계가 모두 바뀔 수 있다는 것이다. 그리고, 이는 더 이상 먼 미래의 이야기가 아니다. 그렇다면, 보안 담당자의 역할은 어떻게 바뀔까? 예전에는 사람이 보안 툴을 직접 다루고 알람(alert)을 하나씩 처리하는 운영자(operator) 역할이 메인이었다면, 앞으로는 AI 에이전트를 설계하고, 정책을 정하고, 예외를 승인하고, 결과를 검증하는 오케스트레이터(orchestrator)로서의 역할이 더 중요해질 것이다.



[그림 2] 에이전틱 AI 보안 모델

AI가 사람을 완전히 대체하는 것이 아니라, 사람의 역할이 상향되는 것이다. 반복적인 업무 처리보다 '감독, 판단, 기준 수립 및 검증'의 비중이 커지는 것으로 이해하면 된다. 이제 우리가 실제 접하는 보안 도메인에서 어떤 변화가 예상되는지 알아보자.

#1. 엔드포인트 보안

엔드포인트 보안에서 AI 에이전트 활용 방안은 꽤 명확하다.

자율적인 위협 탐지 및 격리	랜섬웨어 의심 행위가 발견되면 즉시 격리하고, 연관 IoC를 추출해 다른 엔드포인트에 차단 정책 배포
취약점 우선순위 자동 조정	단순 CVE 점수가 아닌 자산 중요도, 공격자 활용 가능성 등을 바탕으로 위험도 평가 및 패치 우선순위 조정
포렌식 데이터 수집 및 보고 자동화	메모리 덤프, 프로세스 트리, 네트워크 연결 이력 등 포렌식 정보를 자동 수집하고 보고서 초안까지 자동 생성

실무적인 관점에서 보면, 엔드포인트 영역에서 AI는 '대응 시간(MTTR) 단축과 위협 분석 리소스 효율화'와 직결된다.

#2. 네트워크 보안

다음으로, 네트워크 보안에서의 활용 포인트는 다음과 같다.

자율적인 세그멘테이션 및 정책 최적화	트래픽 패턴을 지속 분석하고 정책 최적화
동적 제로 트러스트 정책 집행	제로 트러스트 정책을 정적인 룰이 아니라 맥락 기반으로 동적 집행
디도스 & C2 자율 차단	디도스나 C2(Command & Control) 통신처럼 빠르게 막아야 하는 영역 자율 차단 수행
설정 오류 자율 교정	설정 오류를 자동으로 발견하고 교정 (예: 클라우드 환경의 잘못된 보안 그룹, 퍼블릭 노출 스토리지 자동 탐지/교정)

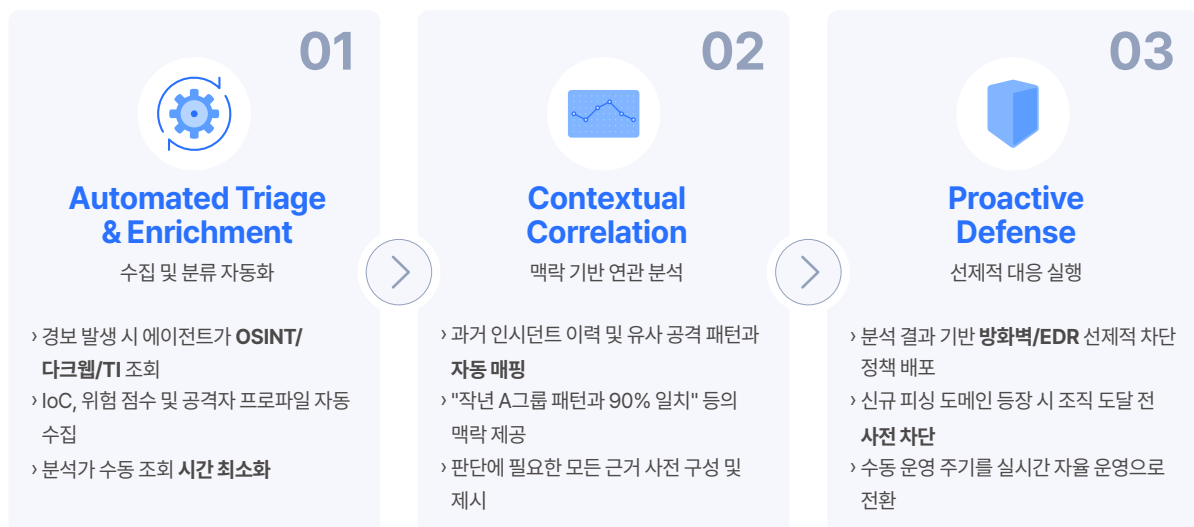
즉, 네트워크 보안에서는 'AI가 정책 운영 자동화와 대응 속도 개선'에서 중추적인 역할을 한다.

#3. 보안 운영 (SOC)

실제 SOC 운영에서는 하나의 AI가 모든 것을 다 처리하는 것이 아니라, 여러 에이전트가 역할을 나눠서 협업하는 구조로 발전하고 있다. 다음은 에이전트 간 협업 구조의 예시이다.

탐지 에이전트	이벤트 탐지 및 이상 징후 식별
위협 인텔리전스 에이전트	IP/hash 평판 조회, 공격자 프로파일링 등으로 맥락 보강
포렌식 에이전트	로그, 메모리 덤프 등 관련 증거 자동 수집
대응 에이전트	내부 영향 최소화를 위한 격리 혹은 차단 수행
보고 에이전트	티켓과 리포트 작성

최근에는 이러한 AI 에이전트 기반 보안 운영을 '에이전틱 SOC(Agentic SOC)'라는 개념으로 정립하고 있다. 이 구조의 장점은 속도와 분업이다. 특히, 사람이 모든 단계를 순서대로 처리할 때 생기는 병목을 줄일 수 있다.



[그림 3] 에이전틱 SOC 워크플로우

아마, 보안 담당자가 실무에서 가장 먼저 체감할 수 있는 영역도 바로 AI 에이전트 기반 보안 운영일 것이다. 과거에는 알림이 발생하면, 담당자가 직접 평판 조회, 과거 이력 및 유사 사례 확인을 진행해 영향도와 대응 방안을 정리해야 했다.

하지만, 에이전틱 AI가 보안 운영에 접목되면 의사결정에 필요한 정보를 분석 및 취합하여 실행 가능한 형태로 제시해준다. 보안 담당자는 정보 조화와 정리에 쓰던 시간을 줄이고 더 중요한 의사결정에 집중할 수 있게 된다.

AI가 불러오는 리스크와 대처 방안

에이전틱 AI는 기업과 보안 담당자들에게 엄청난 혁신의 기회를 제공한다. 문제는 이러한 기회를 공격자들도 동일하게 갖는다는 것이다.

AI 에이전트는 공격자에게도 굉장히 매력적인 도구다. 공격의 속도, 규모 및 정교함의 수준을 비약적으로 올릴 수 있기 때문이다. 공격자들이 AI 에이전트를 활용할 수 있는 방법은 다음과 같다.

AI 기반 자율 공격 캠페인	AI 에이전트를 활용하여 취약점 스캔, 초기 침투, 측면 이동, 데이터 유출까지 공격 체인 자율 수행
스피어 피싱 고도화	공격 대상의 정보를 학습해 맞춤형 메시지를 대량 생성하고 피해자 회신에 실시간 대응하는 '대화형 피싱 에이전트' 동작
취약점 무기화	CVE 정보를 AI 에이전트가 분석하여 익스플로잇 코드 자동 생성 - 패치 배포와 공격 실행 사이 '골든 타임' 단축

AI를 활용해 공격을 고도화하는 것 뿐만 아니라, 조직에서 사용하고 있는 AI 에이전트에 대한 공격을 감행해 피해를 입히는 것 역시 고려해야 한다. 이와 관련된 구체적인 위협은 크게 다섯 가지 정도로 정리할 수 있다.

프롬프트 인젝션 (Prompt Injection)	악의적인 프롬프트를 주입해 본래 LLM의 지침을 우회하여 잘못된 행동 유도
에이전트 하이재킹 (Agent Hijacking)	토큰이나 자격증명을 탈취해 AI 에이전트로 위장하여 시스템 제어권 탈취
툴 포이즈닝 (Tool Poisoning)	AI 에이전트가 사용하는 외부 도구나 플러그인을 오염시켜 실행 조작
메모리 포이즈닝 (Memory Poisoning)	Vector Database 등 장기 메모리를 오염시켜 판단 로직 훼손
과도한 권한 부여 (Excessive Permission)	에이전트에 필요 이상의 권한을 부여하여 한 번의 침해로 피해 확대

종합하면, AI 에이전트를 도입해 잘 활용하는 것도 중요하지만 먼저 'AI가 연결된 권한과 실행 경로를 통제할 수 있는 구조'를 만드는 것이 우선적인 과제다.

에이전틱 AI 시대를 맞아 조직들이 준비해야 할 것은 'AI 보안 & 거버넌스(AI security & governance)'다. 최근, 글로벌 시장에서도 주목 받고 있는 개념이다.

에이전트 인벤토리 관리 (Agent Inventory)	조직에서 운영 중인 모든 AI 에이전트(SaaS 포함) 가시성 확보 및 소유자/관리자 명시
최소 권한 (Least Privilege)	읽기 전용 기본값 설정, 에이전트 접근 네트워크 세그먼트 제한, 태스크 단위 권한 분리
Human-in-the-Loop 기준 정립 (HITL Policy)	자동과 승인 필수 범위 기준 설정 및 에이전트 오작동 시 즉시 차단 가능한 'Kill Switch' 마련
AI 전용 인증 체계 (Machine Identity)	에이전트를 기계 행위자로 보고 별도 자격증명과 감사 로그 체계 운영

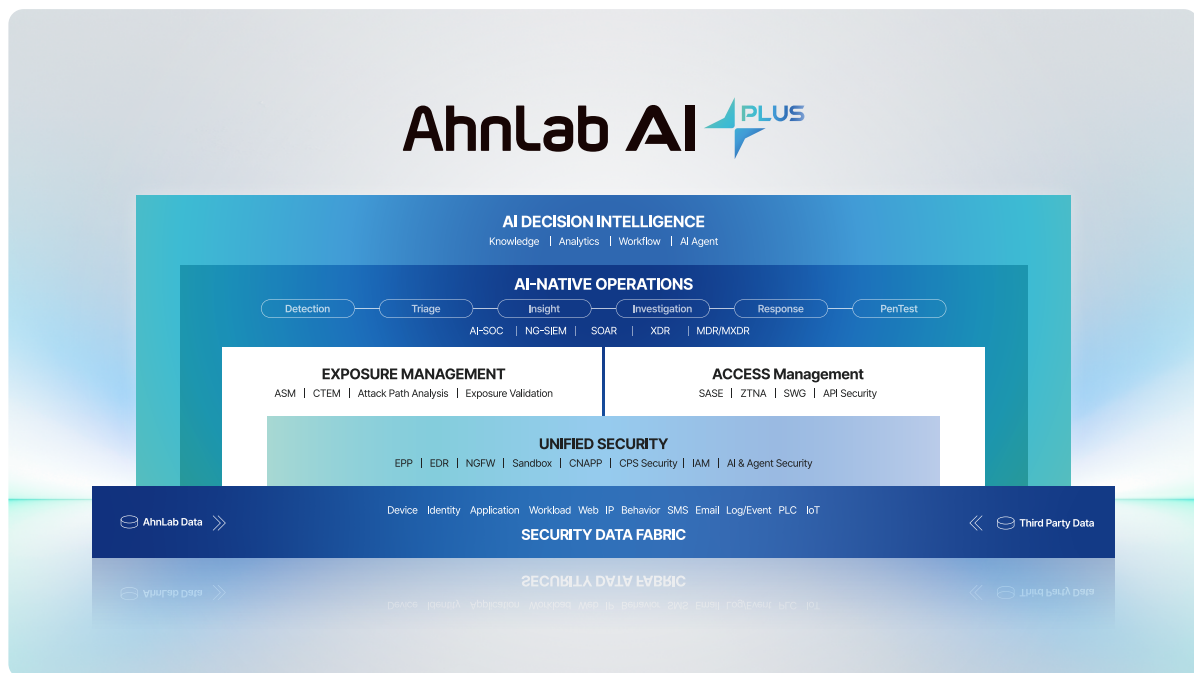
AhnLab AI PLUS와 함께하는 에이전틱 AI 보안 혁신

최근 공격자들의 AI 활용이 점점 더 활발해지면서 방어자의 AI 활용이 필수불가결해졌다. 안랩은 자체 구축 에이전틱 AI 보안 플랫폼 AhnLab AI PLUS를 중심으로 고객들의 AI 보안 대전환에 함께하고 있다.

AhnLab AI PLUS는 안랩의 AI-네이티브 보안 전략과 방향성을 대표하는 통합 브랜드이자, 기술·제품·서비스를 유기적으로 연결하는 플랫폼 및 아키텍처다. 안랩은 AhnLab AI PLUS를 중심으로 기존 제품과 서비스에 AI를 적용해 탐지·분석·판단·대응 역량을 강화하는 'AI-powered Security'에서 제품과 서비스, 보안 운영을 AI 중심으로 설계하는 'AI-native Security'로 진화해 나가고 있다.

AhnLab AI PLUS는 제품·영역별로 분산된 데이터를 하나의 보안 맥락으로 연결하고, 복수의 AI 에이전트가 협업해 보안 운영 과정을 수행하는 에이전틱 AI 기반 아키텍처를 구현한다. AI가 중요한 위협을 이해하고 우선순위를 제시하며 사람이 더 빠르고 정확하게 판단할 수 있도록 지원하는 AI-네이티브 보안 운영 체계를 구축하게 된다.

AhnLab AI PLUS의 구체적인 구조와 각 계층의 역할은 다음과 같다.

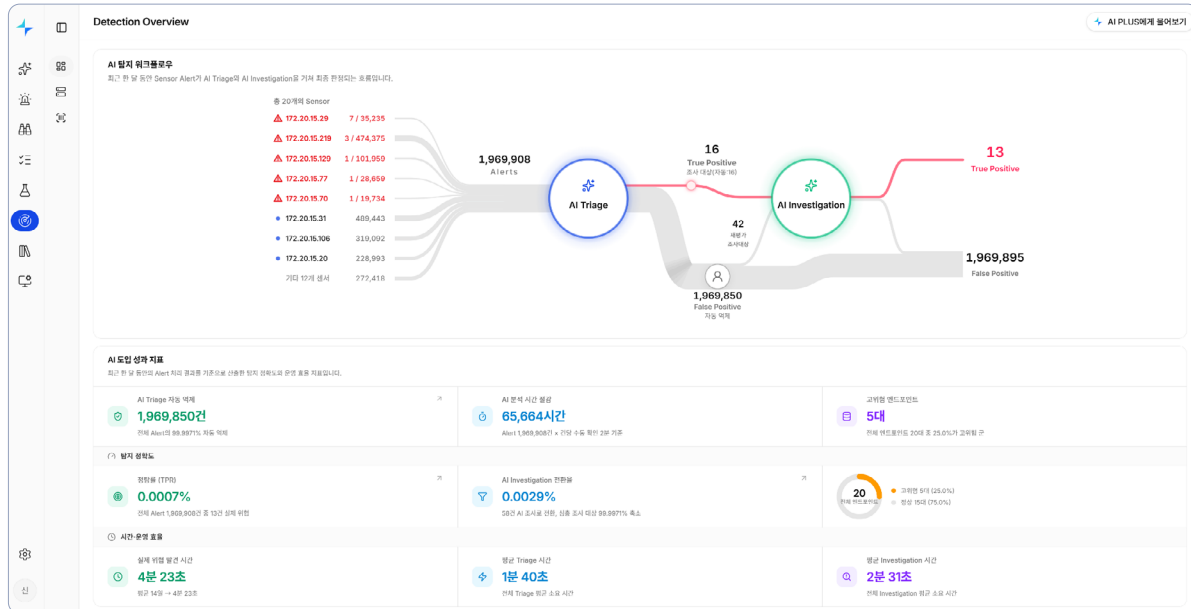


[그림 4] AhnLab AI PLUS 구조도

계층	역할
AI Decision Intelligence	• 단순 위협 분석을 넘어 보안 운영 전반의 우선순위 설정과 의사결정 지원 • 수많은 알람/이벤트 중 무엇이 중요한지 이해하고 판단
AI-Native Operation	• 탐지, 분류, 분석, 대응, 공격 시뮬레이션, 검증까지 보안 운영 전 과정을 AI-Native Operations로 자동화 및 지능화 • 보안 운영의 속도와 규모 향상
Access/Exposure Management	• 지속적인 표면 관리(Exposure Management)를 통해 공격에 노출될 수 있는 자산과 경로 최소화 • 신뢰를 기반으로 사용자·디바이스·애플리케이션의 접근을 지속적으로 검증(Access Management)해 잠재적 위험 감소
Unified Security	• 기존 보호 기능을 통합한 보안 플랫폼과 AI를 유기적으로 연계해 시너지 강화
Security Data Fabric	• 안랩의 자체 센서와 써드 파티 등 다양한 소스로부터 보안 데이터 수집·연결

안랩은 개별 보안 제품에 AI 기능을 추가하는 수준을 넘어, 제품과 데이터, 보안 운영 전반을 처음부터 AI 중심으로 연결하고 재설계하는 'AI-네이티브 보안'을 구현하고 있다. 이를 위해 AhnLab AI PLUS를 기반으로 엔드포인트·네트워크·클라우드·OT·위협 인텔리전스·보안관제 등 다양한 보안 영역의 데이터와 기능을 하나의 보안 맥락으로 연결한다.

각 제품과 서비스에서 발생한 경보·이벤트·텔레메트리는 공통 데이터 구조에서 활용된다. AI는 개별 신호의 연관성과 공격 흐름, 영향 범위를 분석해 여러 제품에서 발생한 신호를 하나의 보안 사건으로 재구성하고 대응 우선순위를 제시한다. 보안 운영자는 각 제품의 경보를 개별적으로 확인하는 방식에서 벗어나, 통합된 사건과 맥락을 바탕으로 중요한 위험과 필요한 조사·대응 조치를 확인할 수 있다.



[그림 5] 안랩의 AI 네이티브 센서 - 맥락 기반 위험 우선순위 식별

[그림 5]의 정보를 통해 좀 더 쉽게 알아보자. 안랩의 AI 네이티브 센서 20개는 최근 한 달 동안 총 1,969,908개의 알림(alerts)을 생성했다. 이 알림들을 건 당 수동 확인 2분을 기준으로 계산하면, 분석하는데 65,664 시간이라는 물리적으로 처리 불가능한 수치가 나온다. 안랩의 AI는 약 2백만 건의 알림을 자동 분석하여 실제 조사(investigation)가 필요한 알림 16건을 분류하고 실제 위험 13건을 식별해낸다. 이렇게 위험을 발견해 내는데 걸리는 시간은 불과 4분 23초다. 사용자가 AI가 식별한 위험을 클릭하면 상세 분석 내용과 AI가 제시하는 대응 방안을 보고 손쉽게 조치를 취할 수 있다.

안랩은 AI의 판단과 행동, 데이터 접근 과정을 기록·점검하며 주요 조치는 사용자의 승인을 거쳐 실행한다. AI에 자율성을 부여하되 권한·정책·감사·승인을 통해 관리하는 '통제된 자율성(Controlled Autonomy)' 원칙을 적용하는 것이다. 보안 운영의 속도와 효율을 높이면서도 안전성과 신뢰성을 확보하는 AI-네이티브 보안 운영 체계를 구현한다.

안랩은 에이전틱 AI 보안 뿐만 아니라 AI 보안 & 거버넌스 역량도 함께 제공한다. 안랩의 자회사 안랩클라우드메이트가 제공하는 **시큐어브리지(SecureBridge)**는 프롬프트에 담긴 기밀·민감 정보를 선별적으로 탐지 및 차단하여 LLM에 학습되지 않도록 하고, 안전한 프롬프트는 그대로 전송하여 데이터 유출 걱정 없이 AI를 자유롭게 활용할 수 있도록 한다.

실질적으로 체감하는 AhnLab AI PLUS의 효과

사용자 입장에서 실질적으로 체감할 수 있는 효과는 크게 다음 세 가지다.

	See What Matters	수많은 경보 중 중요한 위협과 위험을 우선적으로 식별
	Understand Faster	왜 위협한지, 어떤 자산과 사용자가 영향을 받았는지 빠르게 이해
	Act with Confidence	필요한 조사와 대응 방법을 제시하고 실제 대응까지 신속히 연결

안랩은 AhnLab AI PLUS를 활용해 이벤트 처리, 분석, 보안 관제 등 보안 운영의 다양한 영역에서 아래와 같은 효과를 누리고 있다.

3x

연간 이벤트 처리 규모

60%↓

분석 평균 소요 시간

2.5x

보안 관제 효율성

AhnLab AI PLUS의 목표는 고객에게 더 많은 경보나 정보를 보여주는 것이 아닌, 중요한 것을 빠르게 찾고, 왜 중요한지 이해하고, 무엇을 해야 할지 결정하는 시간을 줄이는 것이다. 그 과정에서, 지능적인 AI와 함께 보안 운영을 혁신할 수 있다.

FAQ

#1. AhnLab AI PLUS는 무엇인가?

안랩의 기술·데이터·제품·서비스를 AI 중심으로 연결해 보안 운영을 지능화하는 안랩의 AI 보안 플랫폼이자 전략이다.

#2. 왜 AhnLab AI PLUS인가?

AI 시대의 위협은 개별 제품이 아니라 연결된 보안 운영으로 대응해야 하기 때문이다. AhnLab AI PLUS는 안랩이 30년 이상 축적해온 보안 데이터와 인텔리전스를 학습하여 차별화된 AI 보안 역량을 제공한다.

#3. 사용자 입장에서 무엇이 달라지는가?

AI가 중요한 위협을 빠르게 이해하고 판단해 조사와 대응까지 연결함으로써, 사용자는 더 빠르고 효율적으로 보안을 운영할 수 있다. 이를 통해, 수동 작업에서 발생하는 '휴먼 에러(human error)'를 최소화하는 동시에 업무 생산성을 높이고 중요한 의사결정과 업무에 집중할 수 있다.

#4. 보안 솔루션에는 어떻게 적용되나?

안랩은 [EDR](#), [XDR](#) 등의 솔루션에 'AI Insight'와 같은 AI 이벤트 분석 기능을 탑재해 제공하고 있다. 그리고, [그림 5]의 AI 기반 Triage-Investigation-Response 기능을 곧 제공할 계획이다. 하루 수만 건 이상 생성되는 알림(alert)을 AI가 자동으로 분류 및 조사하고, 리스크 점수를 기반으로 정탐/오탐을 자동 식별한다. 사용자가 실제 위협에만 집중할 수 있도록 우선 순위를 최적화하여 소요 시간과 피로도를 크게 줄일 것으로 기대된다.

결론

에이전틱 AI를 한 문장으로 정의하면 '스스로 계획하고, 도구를 사용하여 행동해 목표를 달성하는 자율 실행형 AI'다. 보안 담당자들에게는 '강력한 자동화 기회'이자 '새로운 공격 표면'이다. AI 에이전트를 잘 사용하는 것도 중요하지만, 어떻게 통제하고, 어떤 기준으로 검증할 것인가에 대한 방안도 마련해야 한다.

에이전틱 AI는 멀리 있는 미래 기술이 아니라 실제 도입이 시작된 현재에 가깝다. 이 글을 읽고 가장 먼저 생각해봐야 할 질문은 다음과 같다.

“우리 조직은 어떤 AI 에이전트를 어떻게 쓸 것인가? 그리고 잘 사용할 준비가 되어 있나?”

이 질문이 의미하는 기술의 진화를 이해한다면, 급변하는 AI 시대에도 성공하는 보안 담당자가 될 수 있다.



AhnLab