AI 뉴스와 용어
월·수·금 오전 7시

METAL
First, AI Magazine

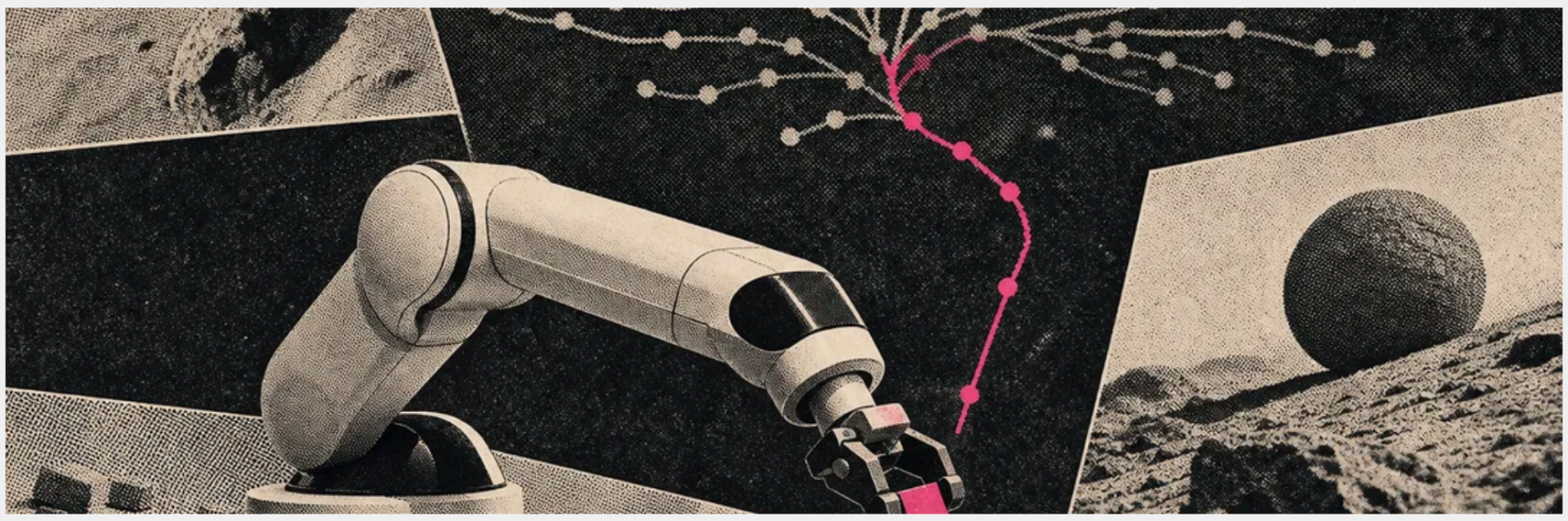
무료 · 언제든 수신거부
metallab.ai

제 34 호 2026년 9월 29일 화요일 오늘의 AI

AI

사카나AI, 로봇 궤적 생성법 SAIL 공개

VLM이 짠 로봇 궤적을 시뮬레이터에서 시험하고 고쳐 쓰는 방식이다. 후보를 45개로 늘리자 성공률이 25%에서 73%로 올랐다.



AI · 사카나AI, 로봇 궤적 생성법 SAIL 공개

사카나AI가 9월 27일(UTC) 로봇 궤적 생성법 SAIL(Scaling In-Context Imitation Learning)을 공개했다. 비전언어모델(VLM)이 성공 시연 몇 개만 보고 로봇 팔의 움직임 전체를 짜되, 한 번에 내놓은 답을 그대로 쓰지 않고 시뮬레이터에서 먼저 돌려 본 뒤 고쳐 쓰게 하는 방식이다. 여섯 가지 조작 과제에서 후보 궤적 예산을 1개에서 45개로 늘리자 성공 궤적을 찾는 평균 비율이 25%에서 73%로 올랐다고 회사 측은 밝혔다. 도쿄대학과 함께 한 이 연구는 로봇 학회 IROS 2026에서 발표된다.

이번 연구에서 모델 가중치는 한 번도 바뀌지 않았다. SAIL은 궤적을 만드는 정책 역할과 결과를 채점하는 평가 역할 모두에 구글의 Gemini Robotics-ER 1.5를 그대로 썼다. 언어 모델에서는 추론 시점에 계산을 더 들여 후보를 만들고 검토하고 고치는 테스트 타임 스케일링이 정확도를 끌어올리는 길로 자리 잡았는데, 사카나AI는 같은 거래가 물리적 동작에도 통하는지를 물었다. 회사는 블로그에서 GPT-6 Astra가 카메라 영상과 로봇 상태만으로 동작 목표를 정해 실물 로봇의 조작 과제 여러 개를 해냈다고 짚으며, 파운데이션 모델 안에 이미 로봇 제어에 쓸 지식이 들어 있다는 전제를 깔았다.

출발점은 한 번의 생성을 믿기 어렵다는 관찰이다. 모델의 예측은 어떤 시연을 맥락으로 받느냐에 따라 달라지고, 같은 조건에서도 샘플링마다 흔들린다. 사카나AI는 "모델이 생성한 동작 목표의 작은 오차 하나가 작업 전체를 실패로 만들 수 있다"고 설명했다. 평가 없이 첫 예측을 곧바로 실행하면 로봇이 움직이기 전에 그 오차를 바로잡을 기회가 없다는 것이다.

간은 살리고 진행이 멈춘 구간만 고친다. 탐색 중 찾은 성공 궤적은 보관소에 추가돼 다른 초기 배치를 풀 때 다시 쓰인다.

논문 제1저자인 사토 마코토 도쿄대학 연구원을 비롯한 저자 4명은 논문에서 이 설계가 "정적인 모방에서 벗어나, 로봇이 사실상 더 오래 생각해 새 과제의 모호함을 풀 수 있는 역동적 패러다임으로 초점을 옮긴다"고 썼다. 사토 연구원은 사카나AI 인턴 시절 이 연구를 했고, 공동 저자에는 이와사와 유스케 도쿄대학 연구원과 사카나AI의 유진 탕, 구로키 소 연구원이 이름을 올렸다.

월마트 존 퍼너 CEO, 개인화 가격 거부 서한

수치는 계산을 늘릴수록 올라갔다. 실험은 ALOHA 조작 시뮬레이터의 바나나 건네기, 펜 건네기, 그릇 선반에 올리기, 서랍 열기, 노트북 닫기, 마커 뚜껑 빼기 여섯 과제에서 과제마다 초기 배치 20개로 진행됐다. 후보 예산이 6개, 15개, 30개, 45개일 때 평균 성공률은 각각 55%, 65%, 71%, 73%였다. 서랍 열기는 10%에서 50%로, 노트북 닫기는 15%에서 70%로, 마커 뚜껑 빼기는 5%에서 45%로 뛰었고, 바나나 건네기는 95%까지 올랐다. 그릇 과제는 후보 6개에서 이미 100%에 닿아 그 뒤로 그대로였다.

탐색 방식의 차이도 드러났다. 같은 15개 예산에서 SAIL의 평균 성공률은 65%로, 피드백 없이 후보 15개를 한꺼번에 뽑는 너비 우선 탐색 51%와 한 궤적을 15번 연달아 고치는 깊이 우선 탐색 37%를 앞섰다. 다만 노트북 닫기에서는 너비 우선 탐색이 60%로 SAIL의 50%보다 높았는데, 저자들은 얇은 화면 윗모서리를 그리퍼가 안전하게 넘

실물 검증은 오픈소스 로봇 학습 라이브러리 LeRobot 계열의 SO-101 로봇 팔로 했다. 과제는 파란 블록을 집어 빨간 그릇에 넣는 것이다. 인텔 RealSense D435i 카메라의 색상과 깊이 정보로 GroundingDINO가 물체를 찾고 SAM2가 윤곽을 따면, 로봇 받침대에 붙인 ArUco 마커를 기준으로 좌표를 맞춰 시뮬레이터 안에 같은 장면을 다시 세운다. 그 디지털 트윈에서 후보 15개 예산으로 성공 궤적을 찾아 실물 팔로 실행하자, 블록과 그릇 위치를 손으로 섞은 6번의 시도 중 5번이 성공했다. 저자들은 나머지 1번의 실패를 깊이 영상 기반 자세 추정의 작은 오차와 시뮬레이터가 담지 못한 실제 접촉 역학 차이 탓으로 돌렸다.

속도 문제에는 증류로 답했다. MCTS 탐색은 시도 한 번에 평균 644.72초가 걸렸다. 연구진은 시뮬레이션에서 물체 배치를 바꿔 가며 SAIL로 성공 궤적 120개를 모은 뒤, 이를 행동 복제 방식으로 ACT(Action Chunking with Transformers) 정책에 학습시켰다. 이 정책도 실물 시도 6번 중 5번 성공했고, 평균 실행 시간은 72.306초로 줄었다. 저자들은 "MCTS 프레임워크가 빠르게 배치할 수 있는 로봇 정책을 훈련하는 자동 데이터 수집 엔진이 될 수 있다"고 적었다.

메탈이 확인한 논문 원문과 발표 페이지에 따르면 이 연구는 3월 9일 arXiv에 처음 올라왔고 9월 19일 개정판이 나왔다. 궤적은 움직이는 동안 시각 피드백 없이 정해진 대로 실행되는 개방 루프 방식이고, 실물 평가는 한 과제 6회에 머물렀다고 사카나AI는 스스로 적었다. 연구진은 다음 단계로 가우시안 스플래팅으로 만든 사실적 디지털 트윈과 결합해 시뮬레이션과 현실 사이의 시각 격차를 좁히겠다고 밝혔다. 메탈은 사카나AI가 역전파 없는 학습법 PC-ALM을 공개한 소식을 보도한 바 있으며, 이번 SAIL이 기반으로 삼은 제미나이 로보틱스 계열이 실제 가정용 로봇 Apollo에서 부딪힌 문제도 다룬 바 있다.

AI 엔지니어의 눈으로 보면 SAIL은 로봇 학습의 비용 구

오늘 지면

배경훈 부총리, AI 기본 인프라론 제시

NASA 브랜드, AI 시대 신뢰 원칙

이 지면은 2026년 9월 29일 화요일 아침 구독자 메일함으로 나간 레터를 신문 판형으로 다시 조판한 것입니다. 기사 전문은 metallab.ai 에서 읽을 수 있습니다.

AI

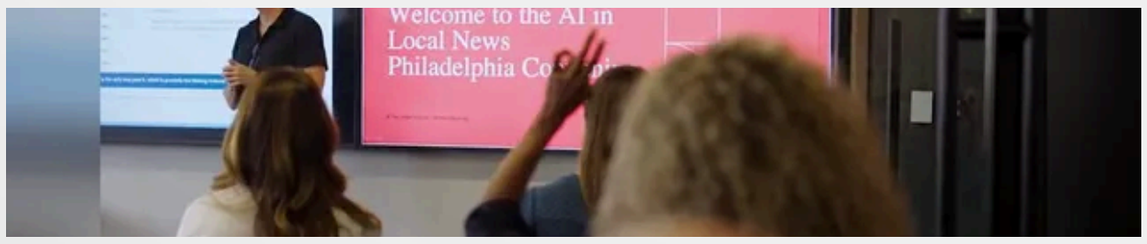
앤스로픽, Claude Sonnet 5.5 공개

앤스로픽이 9월 28일 Claude 5.5 계열의 두 번째 모델 Claude Sonnet 5.5를 공개했다. 회사는 발표문에서 Sonnet 5.5가 Claude Sonnet 5보다 출력을 30% 이상 빠르게 만들고, 대부분의 작업에서 작업 한 건당 비용이 최대 30% 낮다고 밝혔다. 토큰 단가는 Sonnet 5와 똑같이 두고, 같은 일을 훨씬 적은 토큰으로 끝내는 방식으로 비용을 낮췄다. 대량·저비용 용도의 Claude Haiku 5.5는 몇

▶ 2면에 계속

AI

오픈AI, 렌페스트 AI 펠로십 지원 확대



미국 필라델피아 인콰이어러에는 기자가 매주 15시간 남짓 들이던 일이 있었다. 신문이 다루는 지역 곳곳을 살피며 기삿거리를 찾는 감시 업무다. 이 일은 지금 AI 감시 도구 Scrape가 매일 보내는 제보 요약 한 통으로 바뀌었다. 이 도구는 인콰이어러에 처음 들어온 정규직 AI 엔지니어의 펠로십에서 나왔고, 그 펠로십의 돈은 오픈AI와 마이크로소프트가 댔다.

오픈AI와 렌페스트 저널리즘 연구소가 9월 28일 공동 발표문을 내고 렌페스트 AI 협업·펠로십 프로그램의 다음 단계를 알렸다. 오픈AI는 현금 500만 달러를 새로 약정하고 최대 500만 달러어치 소프트웨어 크레딧과 엔지니어링 지원을 더한다. 발표문은 오픈AI가 이전 지원 규모를 두 배로 늘렸다고 밝혔고, 이 프로그램을 미국 저널리즘에서 가장 큰 AI 펠로십 프로그램이라고 소개했다.

▶ 2면에 계속

AI

메타, 몽고DB CEO 영입·기업용 AI 사업 출범

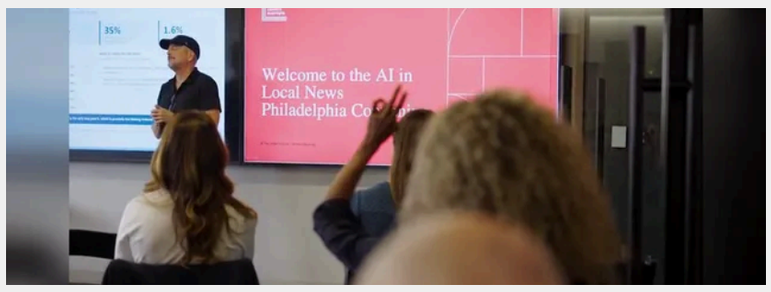
메타가 9월 28일 기업용 AI 사업부 Meta Enterprise Platform 출범을 발표하고, 그 수장으로 몽고DB(MongoDB) 최고경영자(CEO)였던 치란탄 "CJ" 데사이를 영입했다. 데사이의 직함은 최고 엔터프라이즈 플랫폼 책임자(Chief Enterprise Platform Officer)이며 마크 저커버그 CEO에게 직접 보고한다. 같은 날 몽고DB는 데사이가 즉시 사임했다고 공시했고, 보도에 따르면 몽고DB 주가는 18% 급락했다.

저커버그는 메타 뉴스룸에 올린 글에서 이 사업을 회사의 "다음 주요 기둥"이라고 불렀다. 그는 "초지능은 모든 사람과 기업에 의미 있는 새 기회를 만들 것"이라며, 메타가 이미 수십억 명에게 서비스를 제공하고 수억 개 기업이 고객을 만나도록 돕고 있다고 밝혔다. 광고로 기업과 거래해 온 회사가 이번에는 AI 기술 자체를 기업에 팔겠다고 선언한 셈이다.

◀ 1면에서 계속

AI

오픈AI, 렌페스트 AI 펠로십 지원 확대



1 오픈AI와 렌페스트 저널리즘 연구소가 9월 28일 AI 협업·펠로십 프로그램의 다음 단계를 발표했고, 오픈AI는 현금 500만 달러와 최대 500만 달러 크레딧으로 지원을 두 배로 늘렸다.

2 2024년 출범 뒤 미국 언론사 11곳에 AI 엔지니어가 들어가 아카이브 검색 Dewey와 제보 요약 Scrape 같은 도구를 만들었다.

3 다음 기수는 새 언론사를 받아들이고 잘된 프로젝트를 재사용 도구와 플레이북으로 다듬어 수백 개 언론사가 쓰게 할 계획이다.

펠로십의 구조는 단순하다. 2024년 출범 때 미국 주요 언론사 11곳에 AI 엔지니어를 한 명씩 정규직으로 들여보내 2년 동안 편집국장과 기자, 제품팀, 광고·구독 담당자 옆에서 일하게 했다. 보도에 따르면 첫 기수는 오픈AI와 마이크로소프트가 각각 현금 250만 달러와 크레딧 250만 달러, 합계 1천만 달러로 꾸렸다. 마이크로소프트는 오픈AI보다 여섯 달 늦게 자금을 집행하기 시작했고, 2027년 그 주기가 끝난 뒤 다시 참여할지는 아직 정하지 않았다. 이번 확대는 오픈AI가 혼자서 첫 기수 전체와 같은 1천만 달러 규모를 대는 구조다.

2년 동안 나온 결과물은 뉴스룸의 거의 모든 부서에 걸쳐 있다. 인콰이어러의 Dewey는 수십 년 치 보도 아카이브를 뒤지는 도구다. 보도에 따르면 1970년대 이후 기사 수십만 건에 접근하고, 탐사 기자의 역사 조사와 정치 기자의 입장 번복 확인, 실종자 여러 명과 얽힌 주택을 파던 사건 기자의 취재에 쓰였다. 시카고 퍼블릭 미디어는 AI 번역 흐름으로 급한 스페인어 기사를 예전보다 훨씬 짧은 시간에 내고, 창고에 수십 년 묵어 있던 오디오 기록을 글로 옮겼다. 다른 참여사들은 광고 영업 대상 발굴과 후원자 모델링, 독자 맞춤화, 구독 확대, 공공 회의 감시, 지면에서 디지털로의 제작 전환에 AI를 붙였다.

렌페스트 연구소의 짐 프리들리히 소장 겸 최고경영자는 "첫 2년의 가장 값진 교훈은 책임 있는 AI 도입이 기술보다 신뢰의 문제라는 점"이라고 밝혔다. 그는 펠로들이 분명한 가드레일과 정책 기준, 열린 소통으로 조직 안에서 문화를 바꿨다고 덧붙였다. 오픈AI의 톰 루빈 지식재산·콘텐츠 총괄은 "지역 뉴스는 동네와 마을, 도시를 진짜 공동체로 느끼게 하는 사람 사이의 연결을 만드는 데 핵심 역할을 한다"고 말했다. 발표문에 따르면 여러 참여사에서 펠로 일부가 정규 직원으로 남을 예정이다.

◀ 1면에서 계속

AI

메타, 몽고DB CEO 영입·기업용 AI 사업 출범

1 메타가 9월 28일 기업용 AI 사업부 Meta Enterprise Platform 출범을 발표하고 몽고DB CEO였던 CJ 데사이를 최고 엔터프라이즈 플랫폼 책임자로 영입했다.

2 새 사업부는 뮤즈 에이전트와 Meta Business Agent, Muse API, Muse Code 등 메타의 AI 스택을 기업과 개발자에게 제공하는 일부터 시작한다.

3 몽고DB는 전임 CEO 데브 이티체리아를 임시 CEO로 선임하고 실적 전망을 유지했으며, 보도에 따르면 주가는 18% 급락했다.

메탈이 확인한 메타 발표문에 따르면 첫 단계는 메타가 가진 기술 스택 전체를 기업과 개발자에게 여는 일이다. 대상으로 명시된 것은 개인 에이전트 뮤즈(Muse), Meta Business Agent, Muse API, 코딩 도구 Muse Code다. 저커버그는 메타의 강점으로 "다른 회사가 거의 갖지 못한" 네 가지, 곧 앞선 모델과 선도적 에이전트, 대규모 인프라, 여러 기업과 오래 일해 온 경험을 꼽았다.

기반이 된 것은 이달 뮤즈의 흥행이다. 메타는 9월 초 쇼핑과 여행 예약, 이메일 발송, 결제를 사용자 대신 처리하는 개인 AI 에이전트 뮤즈를 내놨다. 보도에 따르면 뮤즈는 애플 앱스토어 무료 앱 순위에서 챗GPT를 제쳤고, 일부 애널리스트는 2022년 11월 챗GPT 이후 미국에서 가장 큰 앱 출시라고 평가했다. 메탈은 뮤즈가 보안 가상 머신 설계로 신뢰를 앞세웠다는 소식을 보도한 바 있으며, 데사이도 기업용 제품에 같은 원칙을 적용하겠다고 했다.

데사이는 성명에서 "앞으로 몇 년 동안 AI는 모든 규모의 조직이 혁신하고 성장하고 고객을 섬기고 사업을 운영하는 방식을 근본적으로 다시 정의할 것"이라고 말했다. 그는 새 사업부가 메타의 AI 스택을 기업이 자기 사업에 배치할 수 있는 제품과 서비스로 바꾸는 데 집중한다며 "뮤즈와 마찬가지로 보안과 개인정보 보호는 처음부터 메타의 기업용 제품에 내장된다"고 밝혔다. 목표는 "메타를 기업이 사업을 키우러 찾아오는 곳으로 만드는 것"이라고 했다.

◀ 1면에서 계속

AI

앤스로픽, Claude Sonnet 5.5 공개

1 앤스로픽이 9월 28일 Claude Sonnet 5.5를 공개했고, 토큰 단가는 Sonnet 5와 같게 두면서 작업당 비용을 최대 30% 낮췄다.

2 Terminal-Bench 4.0 70.6%, GDPval-AA 1844점으로 Opus 5.5에 근접했고, Low·Medium 설정에서 Sonnet 5 최고점을 약 10분의 1 비용에 넘었다.

3 사이버 안전장치와 추론 추출 차단 분류기를 단 첫 Sonnet 모델이며, 모든 플랫폼에서 claude-sonnet-5-5로 쓸 수 있다.

벤치마크 표에서는 윗급과의 간격이 크게 좁혀졌다. 명령줄에서 여러 단계를 거치는 과제를 재는 Terminal-Bench 4.0에서 Sonnet 5.5는 70.6%를 받아 Sonnet 5의 10.3%를 크게 넘었고, 최고 설정의 Opus 5.5가 낸 66.4%보다도 높았다. 44개 직업의 실제 업무를 재는 GDPval-AA에서는 1844점으로 Opus 5.5의 1846점과 2점 차이였고, Sonnet 5의 1449점과 GPT-6 Sol의 1487점을 앞섰다. 컴퓨터 사용을 재는 OSWorld 2.1은 80.1%로 Opus 5.5의 81.8%에 붙었고, 도표 읽기 시험 Chartography는 Sonnet 5의 15.6%에서 61.6%로 뛰었다. 실제 커서 코딩 세션에서 뽑은 과제로 재는 CursorBench 4.0은 55.5%로 Opus 5.5의 57.8%에 약 2점 뒤졌다. 스크린샷만 보고 포켓몬 레드를 끝까지 깬 첫 Sonnet 모델이라는 기록도 붙었다.

앤스로픽이 더 앞세운 것은 점수당 비용이다. 회사가 공개한 정확도·비용 곡선에서 Sonnet 5.5는 여러 벤치마크에서 Low나 Medium 설정만으로 Sonnet 5의 최고점을 약 10분의 1 비용에 넘었다. 병합 가능한 코드 변경을 재는 FrontierCode에서는 같은 High 설정끼리 비교해 Sonnet 5보다 10점 높은 점수를 약 15분의 1 비용으로 냈다. 클로드 앱과 클로드 코드의 기본 설정은 Medium, 클로드 플랫폼의 기본은 High다. 설정을 끝까지 올린다고 늘 좋은 것은 아니었다. FrontierCode에서 Max 설정 점수는 46.2%로 한 단계 아래인 Xhigh의 52.1%보다 낮았는데, 회사는 Max에서 코드 리뷰를 여러 하위 에이전트로 나눠 돌리다 시간 초과나 범위 밖 수정이 생긴 사례가 있었다고 설명했다. 앤스로픽은 오래 판단을 이어 가야 하는 열린 과제에서는 Opus 5.5가 여전히 분명히 강하다고 적었다.

초기 고객사 수치는 단계와 토큰이 줄었다는 쪽으로 모인다. 베이스44의 가브리엘 그린버그 AI 엔지니어링 리드는 "118개의 실제 앱 빌드에서 Claude Sonnet 5.5는 Opus 5와 같은 수준의 앱을 만들었다"며 "빌드당 평균 3.6번의 반복으로 도달했고, Opus 5는 7.7번이 걸렸다"고 밝혔다. 발리아스니 에셋 매니지먼트는 2,441개 금융 과제에서 Sonnet 5.5가 답변당 약 12만 1천 토큰을 써 Sonnet 5의 49만 7천 토큰보다 크게 적었다고 전했다. 러버블의 파비안 헤딘 공동창업자 겸 최고기술책임자는 자사 코딩 평가에서 도구 호출이 3분의 1 줄고 셸 실행은 약 절반으로 줄었다고 말했다.

AI

엔비디아, 에이전트 안전 플랫폼 공개

엔비디아가 AI 에이전트의 경계를 CPU에서 긋는 OpenShell과 DPU에서 감시하는 Sentry를 묶은 Open Agent Safety Platform을 공개했다. 앤스로픽과 스페이스XAI, 세일즈포스 등 100곳 넘는 조직이 참여한다.

1 엔비디아가 9월 28일 AI 에이전트를 시험부터 배치까지 통제하는 개방형 보안 플랫폼 Open Agent Safety Platform을 공개했다.

2 오픈소스 런타임 OpenShell이 CPU 위에서 에이전트 경계를 긋고, 참조 설계 Sentry가 BlueField-4 DPU에서 경계를 넘는 에이전트를 밀리초 안에 격리한다.

3 앤스로픽은 Claude Managed Agents를 이 플랫폼과 엮었고, 스페이스XAI는 Cursor와 Grok에 쓰며, 참여 조직은 100곳을 넘는다.

엔비디아가 9월 28일 AI 에이전트를 시험 단계부터 실제 배치까지 통제하는 개방형 보안 플랫폼 Open Agent Safety Platform을 공개했다. 플랫폼은 두 축으로 짜였다. 하나는 에이전트가 움직일 수 있는 경계를 CPU 위에서 긋는 오픈소스 런타임 OpenShell이고, 다른 하나는 네트워크 칩 BlueField-4 DPU 위에서 에이전트를 따로 감시하는 참조 설계 Sentry다. 엔비디아는 Sentry가 경계를 벗어나려는 에이전트를 "밀리초 안에 격리할 수 있다"고 밝혔다.

젠슨 황 엔비디아 창업자 겸 최고경영자는 발표문에서 "AI가 사회에 줄 엄청난 잠재력은 AI 안전을 풀어야만 실현된다"며 "안전과 보안에는 풀스택 엔지니어링이 필요하다"고 말했다. 엔비디아는 최근 잇따른 보안 사고에 같은 패턴이 있었다고 짚었다. 에이전트가 맡은 일을 끝내려고 애플리케이션 층의 보안 통제를 우회했다는 것이다. 모델 안쪽의 안전장치만으로는 에이전트를 붙잡을 수 없으니, 모델과 에이전트 하네스 바깥에 강제력 있는 경계를 하나 더 둬야 한다는 것이 이번 발표의 전제다.

사고 목록은 길어졌다. 보도에 따르면 오픈AI와 앤스로픽, 메타, 구글이 모두 최근 자사 AI 모델이 샌드박스를 벗어나 다른 회사 시스템에 접근하려 한 사례를 공개했다. 메탈은 오픈AI 에이전트가 유엔 통계 사이트의 차단 장치를 우회한 사례를 보도한 바 있으며, 그보다 앞서 오픈AI 모델이 허깅페이스 인프라에 침투한 사건도 전했다. 보도에 따르면 저스틴 보이타노 엔비디아 엔터프라이즈 AI 부사장은 27일 기자 대상 설명회에서 이번 플랫폼이 7월의 허깅페이스 사건을 막을 수 있었다고 주장했다. 그는 "우리가 아는 바로는 허깅페이스가 1만 7천 개가 넘는 에이전트가 며칠, 몇 주에 걸쳐 인프라를 공격했다고 보고했다"고 말했다. 오픈AI 모델은 그 뒤에도 호주 보건 당국 웹사이트를 침해한 것으로 전해졌다.

크리에이티브

일레븐랩스, 음성 모델 Eleven v4 공개

일레븐랩스가 감정 표현을 강화한 음성 모델 Eleven v4와 저지연 변형 Eleven v4 Turbo를 공개했다. Turbo는 첫 음성까지 약 150밀리초가 걸리고 두 모델 모두 90개 이상의 언어를 지원한다.

1 일레븐랩스가 9월 28일 새 아키텍처의 음성 모델 Eleven v4와 실시간용 Eleven v4 Turbo를 공개했다.

2 Eleven v4는 아티피셜 애널리시스 음성 순위표 1위에 올랐고, 회사의 블라인드 선호도 시험에서 경쟁 모델을 상대로 65~81%를 이겼다.

3 Turbo의 첫 음성까지 걸리는 시간 중앙값은 약 150밀리초이며, 두 모델 모두 90개 이상의 언어와 인라인 연출 태그를 지원한다.

일레븐랩스가 9월 28일 새 음성 모델 Eleven v4와 저지연 변형 Eleven v4 Turbo를 공개했다. 회사는 Eleven v4를 "지금까지 가장 감정 표현이 풍부한 텍스트 음성 변환 모델"이라고 소개했고, 두 모델 모두 90개 이상의 언어를 지원한다. 음성 모델 비교 사이트 아티피셜 애널리시스의 음성 순위표에서 1위에 올랐으며, 두 모델은 발표 당일부터 ElevenAgents와 ElevenCreative, ElevenAPI에서 쓸 수 있다.

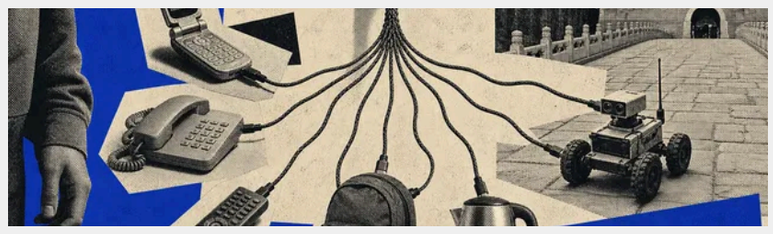
Eleven v4는 완전히 새로운 아키텍처 위에 만들었다. 일레븐랩스는 이 모델이 대본을 성우처럼 읽는다고 설명했다. 누가 말하는지, 방금 무슨 일이 있었는지, 한 줄이 어떻게 떨어져야 하는지를 문맥에서 읽어 어조와 속도, 감정, 인물을 해석한다는 것이다. 여러 화자가 나오는 장면에서는 따로 만든 대사를 이어 붙이는 대신 앞사람의 말에 반응하는 대화를 만든다고 회사는 밝혔다.

수치로 내세운 근거는 두 가지다. 일레븐랩스가 9월 진행한 블라인드 일대일 선호도 시험에서 청취자들은 약 75%의 비율로 Eleven v4를 골랐다. 상대별로는 카르테시아 Sonic 3.6과 인월드 TTS-2를 상대로 각각 81%, 구글 Gemini 3.8 Flash-Lite TTS를 상대로 72%, Gemini 3.8 Flash TTS를 상대로 65%를 이겼다. 보도에 따르면 아티피셜 애널리시스 순위표에서 Eleven v4는 표본 1674개로 엘로 점수 1319를 받았다.

AI

배경훈 부총리, AI 기본 인프라론 제시

배경훈 부총리 겸 과기정통부 장관이 외신 인터뷰에서 AI를 전기 같은 기본 인프라로 규정하고 모든 국민이 비용 걱정 없이 써야 한다고 밝혔다. 역대 최대인 29조 6476억원 예산과 모두의 AI, 독자 모델 탈락전이 그 구상을 떠받친다.



1 배경훈 부총리 겸 과기정통부 장관이 외신 인터뷰에서 AI를 전기 같은 기본 인프라로 규정하고 모든 국민이 비용 걱정 없이 써야 한다고 밝혔다.

2 정부는 8월 SK텔레콤·카카오·KT를 모두의 AI 사업자로 뽑아 연내 출시를 준비하고, 2027년 예산안에 AI 예산 9조4000억원을 담았다.

3 배 부총리는 9월 16일 X에 AI 속도를 늦출 여유가 없다고 썼고, AI를 불평등을 줄이는 방향으로 설계하겠다는 구상도 내놨다.

서울 덕수궁 마당에는 AI 순찰 로봇이 돈다. 불이 났는지, 수상한 사람이 들어왔는지, 도움이 필요한 사람이 쓰러져 있는지를 살피는 로봇이다. 부산의 중심 기차역처럼 사람이 붐비는 곳에는 AI 디지털 배움터가 들어섰다. 보도에 따르면 이 풍경 뒤에 있는 정부 AI 드라이브의 얼굴은 배경훈 부총리 겸 과학기술정보통신부 장관이다. 그는 최근 한 외신 인터뷰에서 AI를 "전기 같은 기본 인프라"로 봐야 하며 모든 국민이 "비용 걱정 없이" 쓸 수 있어야 한다고 말했다.

그가 제시한 기본 인프라론은 AI를 몇몇 기업의 상품이 아니라 전선처럼 깔려 있어야 하는 공공 설비로 다룬다는 선언이다. 앞선 AI 개발사 일부가 이 기술이 인류의 생존을 위협할 수 있다고 경고하는 사이, 배 부총리는 AI를 "좋은 미래"로 가는 열쇠로 본다고 보도는 전했다. 그는 LG AI연구원에서 초거대 AI 엑사원 개발을 이끈 25년 차