AI 뉴스와 용어
월·수·금 오전 7시

METAL
First, AI Magazine

무료 · 언제든 수신거부
metallab.ai

제 46 호 2026년 10월 16일 금요일 오늘의 AI

AI

Amp, 코딩 앞뒤 업무로 에이전트 확대

오픈AI가 공개한 사례에 따르면 코딩 에이전트 기업 Amp는 GPT-6 Astra로 코딩 앞뒤 업무까지 에이전트 활용을 확대하고 있다. 20명 팀은 마케팅까지 에이전트에 맡기고, ChatGPT 로그인과 Ultrafast로 토큰 비용과 대기 시간을 줄이고 있다.



AI · Amp, 코딩 앞뒤 업무로 에이전트 확대

오픈AI가 10월 14일 코딩 에이전트 기업 Amp의 GPT-6 Astra 도입 사례를 공개했다. Amp는 개발자의 아이디어를 실제 소프트웨어로 바꾸는 에이전트를 만든다. 오픈AI에 따르면 Amp에서 쓰이는 토큰의 대부분은 오픈AI 모델이 차지한다. 사례의 요지는 에이전트가 코드 작성에서 그치지 않고 무엇을 만들지 정하는 일부터 문서 작성과 디자인, 출시 뒤 마케팅까지 맡는 단계로 넓어지고 있다는 것이다.

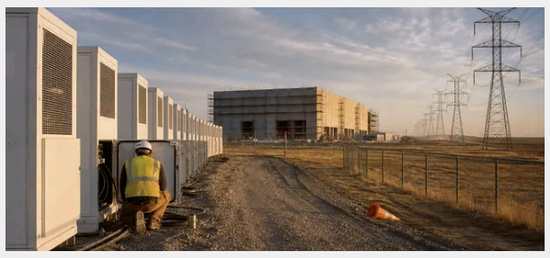
Amp는 사용자에게 가장 잘 맞는 모델을 골라 추천하는 방식으로 운영한다. 퀸 슬랙 Amp 최고경영자(CEO)는 오픈AI 모델을 고르는 이유로 투명성을 들었다. 그는 "오픈AI는 매우 투명하고, 안전장치가 적용될 때 사용자가 그 사실을 알고 이해한다"며 "Amp에서 오픈AI 모델만으로 모든 일을 하고도 충분히 만족할 수 있다"고 말했다. 그는 사용자들이 오픈AI 모델이 꾸준히 좋아지는 점을 높이 산다고도 했다.

에이전트에게 일을 많이 맡길수록 토큰 비용이 걸림돌이 된다. 슬랙에 따르면 Amp 사용자 가운데 일부는 일주일에 토큰 100억 개를 쓴다. 그는 사용량을 하나하나 지켜보는 순간 에이전트가 어디까지 할 수 있는지 확인하는 일을 스스로 막게 된다고 봤다. 그래서 Amp는 ChatGPT 계정으로 로그인하는 연동을 내놓았다. 이 연동으로 사용자는 GPT-6 Astra를 포함한 오픈AI 모델과 더 많은 토큰을 쓸 수 있다. 슬랙은 "ChatGPT로 로그인하면 토큰을 하나하나 세지 않아도 된다"며 "에이전트를 길게 풀어 두면 그때 비로소 에이전트가 얼마나 많은 일을 할 수 있는지 알게 된다"고 말했다.



Amp의 요금 구조도 같은 방향으로 바뀌었다. Amp는 9월 13일 공지에서 사용자가 자기 연산 자원과 모델 구독 또는 API 키를 가져오면 Amp를 무료로 쓸 수 있다고 밝혔다. ChatGPT 구독을 Amp에서 쓰는 데 더는 월 요금제가 필요 없고, 엔터프라이즈 등급을 뺀 모든 사용자에게 자체 키 사용에 따른 토큰 수수료와 한도도 없앴다. 대신 Amp는 에이전트가 독립적으로 병렬 실행되는 원격 컴퓨터인 오브 사용에 과금하고, 사용자 컴퓨터의 러너로 돌리면 무료다. 같은 공지에서 최소 데이터 보존 정책을 엔터프라이즈만이 아니라 모든 사용자에게 공식 적용한다고 했고, 9월 30일부터는 오픈라우터와 아마존 베드록 등 9가지 외부 모델 경로를 모든 등급에 열었다.

에이전트가 맡는 일의 범위도 넓어졌다. 20명 규모인 Amp 팀은 마케팅에도 에이전트를 쓴다. 슬랙은 "마케팅 전담 인력은 없지만 마케팅 전담 에이전트는 있다"고 말했다. 이 에이전트는 신규 사용자 경험을 다듬고 검색 등 여러 경로에서 Amp를 더 많이 찾게 하는 일을 맡는다. 그 과정에서 문서와 디자인 페이지를 만드는 데 GPT-6 Astra를 쓴다. 슬랙은 "우리는 디자인과 글쓰기 스타일이 아주 독특하고 까다로운데, GPT-6 Astra의 결과물에 매우 만족한다"고 말했다. 그는 집 설계도를 GPT-6 Astra에 주고 블렌더에서 3D 모델을 만든 뒤 아이들과 함께 그 모델로 게임을 만들기도 했다며, 이는 불과 몇 달 전만 해도 불가능했다고 덧붙였다.



ON.energy, 크루소에 AI UPS 배치

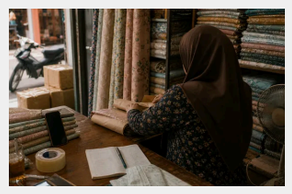
이전트를 띄우거나 여러 에이전트를 조율하고 지난 대화를 찾아 준다. GPT-Live-1 기반 기능 콜링 퍽을 쓰면 손목시계에서도 퍽과 실시간으로 대화할 수 있다. Amp는 10월 13일에는 퍽을 MCP 서버로 열어 ChatGPT와 클로드 등 다른 에이전트에서도 부를 수 있게 했다. 메탈이 확인한 Amp 공지에 따르면 퍽 MCP는 퍽과 대화하는 도구와 실행 중인 오브를 보여 주는 도구 두 가지를 제공한다.

속도는 다음 변수다. Amp는 GPT-6 Astra로 Ultrafast 모드를 시험해 왔고, 9월 29일 오픈AI의 Ultrafast 속도 등급을 쓰는 플래드 속도를 내놓았다. Amp 공지에 따르면 플래드는 추론 요청을 최대 6배 빠르게 처리하고 토큰당 비용도 6배가 든다. 아직은 Amp가 제공하는 추론에서만 작동하며, 연결된 ChatGPT 구독에는 오픈AI가 Ultrafast를 지원하지 않는다. 메탈은 오픈AI가 속도 등급 Ultrafast를 출시했다고 보도한 바 있다. 슬랙은 "사람들은 이렇게 좋으면서 이렇게 빠른 모델을 본 적이 없다"고 말했다. 그는 응답이 빨라지면 여러 에이전트 작업을 오가며 기다릴 필요가 줄고, 에이전트가 결과를 돌려주기 전에 스스로 검증할 시간을 더 쓸 수 있다고 봤다.

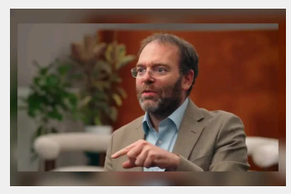
콘텐츠 마케터의 눈으로 보면 이 사례에서 눈여겨볼 대목은 20명 회사가 마케팅 직무를 사람 대신 에이전트에 배정했다는 사실이다. 검색 유입과 신규 사용자 경험처럼 측정 가능한 일을 에이전트에 맡기고, 글과 디자인의 기준은 팀이 쥐는 분업이다. Amp가 무료 등급과 자체 키 수수료 폐지로 문턱을 낮춘 것도 사용자가 에이전트에게 더 큰 일을 맡기게 하려는 같은 계산에서 나왔다. Amp의 다음 단계는 코딩 앞뒤의 일을 더 많이 에이전트 안으로 끌어오는 것이다. 슬랙은 "사람들이 무엇을 만들어 낼지는 아무도 모르고, 그걸 우리 모두가 지켜보게 된다"며 "그 과정에 한몫한다는 게 멋진 일"이라고 말했다.

오늘 지면

1 AI Amp, 코딩 앞뒤 업무로 에이전트 확대
2 AI 딥마인드, Argon 과학 연구 성과 공개
3 AI 구글 리서치, 지리공간 추론 에이전트 보고서 공개
4 AI 구글 딥마인드, 항균 펩타이드 탐색 사례 공개
5 AI 구글 리서치, 논문 검증 에이전트 딥 리뷰 공개
6 AI 메타, Muse 기반 건강 연구 원고 10편 공개



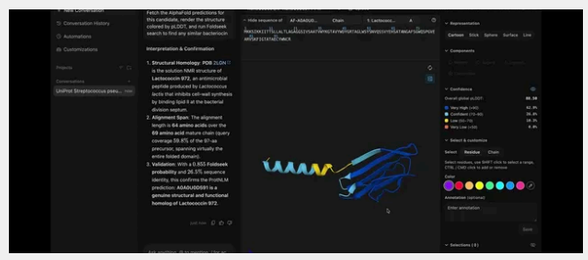
메타, 아태 소상공인 AI 활용 조사 공개



오픈AI, 블록·센티넬원 데이브레이크 사례 공개

이 지면은 2026년 10월 16일 금요일 아침 구독자 메일함으로 나간 레터를 신문 판형으로 다시 조판한 것입니다. 기사 전문은 metallab.ai 에서 읽을 수 있습니다.

AI



구글 딥마인드, 항균 펩타이드 탐색 사례 공개

구글 딥마인드가 10월 15일 에이전트 개발 플랫폼 Google Antigravity로 새 항균 펩타이드 후보를 찾아낸 사례를 공개했다. 연구자와 AI 에이전트가 한 세션 안에서 UniProt 데이터베이스의 기능 미상 단백질 수백만 개를 훑어 호흡기 감염 병원체에 활성이 예측되는 박테리오신 후보를 골랐고, 이 후보는 지금 습식 실험 검증을 받고 있다. 블로그를 쓴 펑페이 톈, 맥스웰 빌레스키, 조이 창갈리두, 수크딥 싱, 루시 콜웰 등 연구진

▶ 2면에 계속

AI

딥마인드, Argon 과학 연구 성과 공개



구글 딥마인드가 10월 14일 새 프론티어 모델 Gemini 4 Argon으로 양자컴퓨팅, 핵융합, 단백질 설계 분야에서 낸 과학 연구 성과를 공개했다. 딥마인드는 공식 X 계정에 "Gemini 4 Argon이 이미 양자컴퓨팅, 핵융합, 단백질 설계 등에서 새로운 과학적 돌파구를 열고 있다"고 밝혔다. 근거는 하루 전인 10월 13일 코레이 카북추오글루 구글 딥마인드 수석부사장 겸 최고 AI 설계책임자가 정식 출시 글에 적은 네 가지 사례다. 구글은 같은 날 과학자 1만 명에게 개발 환경 구글 안티그래비티의 최고 사용 등급을 6개월 동안 무료로 열어 주겠다고 발표했다.

메탈은 구글 딥마인드가 10월 1일 Gemini 4 Argon을 공개했다고 보도한 바 있으며, 13일 발표는 그 모델을 개발자와

▶ 2면에 계속

AI

구글 리서치, 지리공간 추론 에이전트 보고서 공개



구글 리서치가 10월 14일 지구 데이터를 분석하는 지리공간 추론 에이전트의 기술 보고서를 공개하고 실험 접근을 넓혔다. 제미나이를 바탕으로 한 이 에이전트는 위성·기상·홍수·인구 기초 모델을 도구로 묶고, 질문을 받으면 분석 계획부터 세운 뒤 실행한다. 보고서에 따르면 전문가가 설계한 개방형 과제 벤치마크에서 도구 없는 대규모언어모델(LLM)은 0점, 웹검색과 지도 같은 기본 도구를 붙인 모델은 3.3점에 그쳤고 이 에이전트는 62.7점을 받았다. 구글은 구글 어스 안의 대화형 심층 분석 기능과 개발자용 에이전트 간(A2A) 서비스로 일부 이용자에게 시험 사용을 연다.

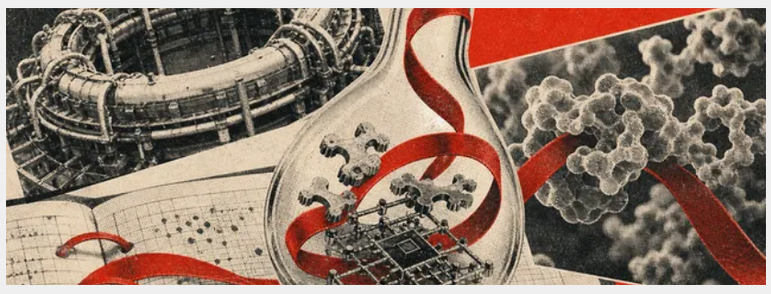
출발점은 범용 모델의 약점이다. 구글 리서치의 슈라비아 셰티 디스팅귀시드 엔지니어와 루크 배링턴 제품 디렉터는

▶ 2면에 계속

◀ 1면에서 계속

AI

딥마인드, Argon 과학 연구 성과 공개



1 구글 딥마인드는 10월 14일 Gemini 4 Argon이 양자컴퓨팅, 핵융합, 단백질 설계에서 새로운 과학적 돌파구를 열고 있다며 연구 성과를 공개했다.

2 Argon이 짠 아크사인 양자 서브루틴은 린 형식 검증을 거쳐 시공간 비용을 100배 넘게 줄였고, 랜드리먼 교수와의 공동 연구는 1967년 그래드 추측의 반례 집합을 넓혔다.

3 효소 설계에서는 몇 시간 만에 새 설계안 5개를 내 실험실에서 촉매 활성이 확인됐고, 구글은 과학자 1만 명에게 안티그래비티 최고 등급을 6개월 무료로 제공한다.

가장 큰 숫자는 양자컴퓨팅에서 나왔다. 최근 양자 시뮬레이션 알고리즘의 기본 부품인 아크사인 서브루틴을 Argon이 새로 짰고, 이 구현은 수학 증명 검사 도구 린(Lean)으로 형식 검증을 거쳤다. 구글에 따르면 공개된 양자 CORDIC 기준 구현보다 시공간 비용이 100배 넘게 줄었다. 시공간 비용은 연산에 드는 하드웨어 규모와 실행 시간을 함께 잰 값이다. Argon은 오류 정정 쪽에도 손을 댔다. 기계학습 기반 양자 오류 정정 디코더 알파큐빗의 내부 버전을 하나의 최적화된 하드웨어 커널로 옮겨 추론 처리량을 두 배 가까이 끌어올렸다. 디코더가 양자 하드웨어의 실시간 속도를 따라가야 더 큰 장치로 넓힐 수 있다는 것이 구글의 설명이다.

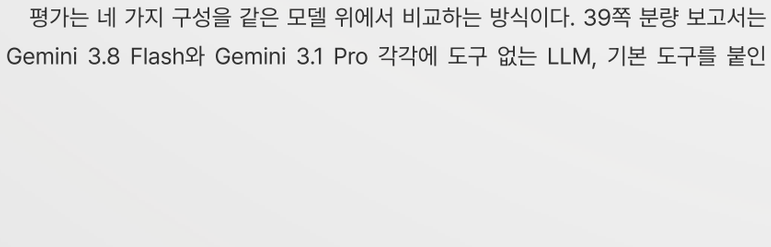
핵융합에서는 59년 묵은 수학 가설이 대상이었다. 스텔러레이터는 뒤틀린 3차원 자기장으로 뜨거운 플라스마를 가두는 핵융합 장치인데, 1967년부터 이어진 그래드 추측은 이 자기장의 정확한 수학 해를 찾을 수 없다고 봤다. 최근 메릴랜드대 매트 랜드리먼 교수 등이 반례를 찾아냈고, 구글 연구팀은 랜드리먼 교수와 함께 더 넓은 범위의 반례 집합을 발견했다. 구글은 이 해들이 스텔러레이터 설계에 쓰는 시뮬레이션 코드를 시험할 믿을 만한 기준점이 된다고 밝혔다.

생명과학 사례는 두 갈래다. 가모 랩스 연구진은 희귀 발달성 폐질환을 앓는 신생아의 비식별 유전체를 Argon으로 분석했고, 한 사례에서 일반 분석 소프트웨어가 수년 동안 놓친 염색체 재배열을 짚어 내고 왜 그동안 잡히지 않았는지까지 설명했다. 가모 랩스 내부 평가에서 Argon은 희귀질환 진단 보조 벤치마크 레어벤치 57%, 실험실 추론 평가 랩벤치 52.2%로 1위를 기록했다. 효소 설계에서는 에스터 결합을 끊는 세린 에스테라아제를 만들었다. 결합이 끊기면 자외선 아래에서 빛나는 표준 분자 4MU 아세테이트를 표적으로 삼아, Argon은 몇 시간 만에 데이터를 모으고 모델을 조정하고 생화학 기반 필터를 골라 새 설계안 5개를 내놨다. 이후 실험실 시험에서 다섯 설계 모두 촉매 활성이 확인됐다. 보통 수개월의 모델 조정과 후보 선별이 드는 작업이다.

◀ 1면에서 계속

AI

구글 리서치, 지리공간 추론 에이전트 보고서 공개



1 구글 리서치가 10월 14일 제미나이 기반 지리공간 추론 에이전트의 기술 보고서를 공개하고 구글 어스와 에이전트 간(A2A) 서비스로 실험 접근을 넓혔다.

2 Gemini 3.8 Flash 기준 개방형 전문가 과제 GeoExpertQueries에서 도구 없는 LLM은 0점, 기본 도구 모델은 3.3점, 에이전트는 62.7점을 받았고 GeoKTools는 40.5점에서 86.0점으로 올랐다.

3 WHO 아프리카 지역사무소는 에볼라 발병 초기 몇 분 만에 노출 정착지 48곳과 위험 인구 4만5500명 이상을 찾아냈다.

설계의 핵심은 맥락 관리다. 어스 AI 기능은 모듈형 스킬로 포장돼 질문이 필요로 할 때만 호출된다. 에이전트는 복잡한 연산 전에 단계별 분석 계획을 짜고, 지역 데이터가 부족하면 대체 데이터셋을 고르고, 큰 지도 레이어는 모델의 맥락 창 밖 전용 작업 공간에서 처리한다. 답변마다 계획과 사용한 데이터 출처를 보여 줘 전문가가 가정을 점검하거나 데이터셋을 바꿀 수 있다. 이용자는 CSV 표, 사용자 지정 경계(GeoJSON), 클라우드 테이블로 자기 데이터를 올려 구글의 전 지구 모델과 결합할 수 있다.

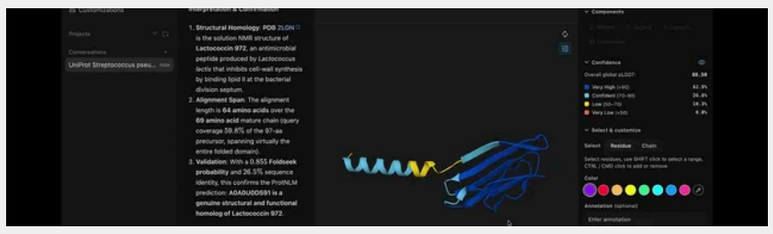
평가는 네 가지 구성을 같은 모델 위에서 비교하는 방식이다. 39쪽 분량 보고서는 Gemini 3.8 Flash와 Gemini 3.1 Pro 각각에 도구 없는 LLM, 기본 도구를 붙인 LLM, 어스 AI 도구만 붙이고 사전 계획은 뺀 구성, 계획 단계까지 갖춘 완성형 에이전트를 놓고 벤치마크 네 종으로 쟀다. 구글이 새로 만든 GeoExpertQueries는 전문가가 쓴 132개 문제 가운데 풀 수 없는 것을 걸러 낸 101개 개방형 과제로, 문제마다 독립 전문가 세 명이 합의한 모범 작업 흐름이 붙어 있다. 함께 만든 GeoKTools는 92개 템플릿에서 뽑은 305개 문항으로 24종 도구 가운데 1~4종을 정확히 골라 써야 풀린다. 구글은 두 벤치마크의 개발용 문항을 연구자에게 공개했다.

결과는 도구가 가장 큰 차이를 만든다는 쪽이다. 메탈이 확인한 보고서 표에 따르면 Gemini 3.8 Flash 기준 GeoKTools 점수는 도구 없는 LLM 40.5점, 기본 도구 40.0점에서 어스 AI 도구를 붙이자 86.0점으로 뛰었다. 지구 관측 코딩 벤치마크 CBGB도 12.9점과 19.7점에서 87.9점이 됐다. 사전 계획은 개방형 과제에서 10.5점을 더해 GeoExpertQueries를 52.2점에서 62.7점으로 올렸고, 분야별로는 자연재해·위기 대응 과제가 45.6점에서 66.0점으로 가장 크게 올랐다. 반면 Gemini 3.1 Pro에서는 계획 단계를 붙인 에이전트가 47.6점으로 도구만 붙인 구성의 50.8점보다 3.2점 낮았다. 미국항공우주국(NASA) 문서 기반 문답인 UnivEARTH에서는 도구 없는 LLM도 93.6점을 받아, 구글은 원문이 웹에 공개된 탓에 이 벤치마크가 포화됐다고 봤다. GeoExpertQueries 점수는 최종 답의 정오가 아니라 전문가 작업 흐름과의 일치를 자동 채점기로 잰 값이어서, 구글은 에이전트의 결과물을 전문가가 검토할 초안으로 규정했다.

◀ 1면에서 계속

AI

구글 딥마인드, 항균 펩타이드 탐색 사례 공개



1 구글 딥마인드는 10월 15일 Google Antigravity에서 연구자와 AI 에이전트가 UniProt 기능 미상 단백질 수백만 개를 훑어 새 항균 펩타이드 후보를 찾은 사례를 공개했다.

2 에이전트는 락토코신 972 계열의 아미노산 97개짜리 박테리오신을 골라 BLAST, AlphaFold 구조 비교(Foldseek 확률 0.855), 유전자 주변 분석으로 검증하고 실험 계획까지 짰다.

3 후보는 같은 파이프라인의 다른 후보 몇 개와 함께 습식 실험 단계에 있으며, 딥마인드는 Antigravity에 3D 도구 Biomolecular Structure Viewer를 새로 넣었다.

걸림돌은 단백질 데이터베이스의 암흑 물질이다. 알려진 단백질과 닮지 않아 기능이 밝혀지지 않은 서열이 UniProt에만 수백만 개 쌓여 있다. ProtNLM 프리프린트에 따르면 알려진 자연 단백질 서열 가운데 실험으로 기능이 확인된 것은 1%가 안 되고, 최소 30%는 기능이 알려진 서열과 너무 멀어 계산으로도 주석을 달 수 없다. 구글이 만든 ProtNLM은 아미노산 서열을 읽어 단백질 이름과 기능을 자연어로 내놓는 트랜스포머 모델이다. 교신저자인 루시 콜웰과 맥스웰 빌레스키 등이 쓴 35쪽 분량의 프리프린트에 따르면 이 모델은 기능 미상 단백질 약 4900만 개의 이름을 예측했고 그 결과가 UniProt에 실렸다.

UniProt 안내에 따르면 ProtNLM이 예측한 이름은 202204 릴리스부터 실렸다. 이름 외에 기능 설명, 세포 안 위치, 키워드, GO 용어까지 예측하는 ProtNLM2는 UniProt 202304 릴리스의 단백질 항목 2억4000만 개로 학습했고, 202602 릴리스와 함께 가장 많이 검색되는 약 400개 종의 TrEMBL 항목 약 2만6000개에 시범 공개됐다. 예측은 문자열 일치나 서열 유사성, 구조 유사성으로 뒷받침되지 않으면 화면에 나오지 않는다. UniProt은 "머신러닝 AI 모델은 실제로 오류를 낸다"며 사용자에게 예측을 신중히 평가해 달라고 밝혔다.

이번 사례에서 연구자가 정한 목표는 스트렙토코쿠스 슈도뉴모니에에서 새 항균 펩타이드를 찾아 계산으로 검증하는 일이었다. 어떤 질문을 하고 어떤 증거를 모을지는 연구자가 정했고, 여러 데이터베이스 조회와 구조 분석, 문헌 종합은 에이전트가 맡았다. 공개된 화면에서 연구자는 Gemini 3.8 Flash를 고른 뒤 UniProt에서 단백질을 찾으라고 지시했다. 에이전트는 수백만 항목을 몇 개 후보로 좁혔고, 그중 락토코신 972 계열의 아미노산 97개짜리 박테리오신을 골랐다. 메탈이 확인한 UniProt 공개 항목 A0A0U0D591에서 이 단백질은 미검토 상태, 주석 점수 5점 만점에 1점, 질량 1만408달톤의 예측 단백질로 등록돼 있다.

AI

구글 리서치, 논문 검증 에이전트 딥 리뷰 공개

구글 리서치가 연례 행사 리서치앳 2026에서 학회 원고 1만6000편 넘게 검토한 논문 검증 에이전트 딥 리뷰를 일반에 열었다. AI 학회 원고가 3년 만에 세 배로 불어난 상황에서 검증 노동을 AI로 나누겠다는 구상이다.



1 구글 리서치는 10월 14일 마운틴뷰에서 연 리서치앳 2026에서 논문 검증 에이전트 딥 리뷰를 일반에 공개하고 연구 성과 10건을 발표했다.

2 딥 리뷰는 문서를 구간별로 나눠 여러 에이전트가 병렬 검증하는 4단계 구조로, 전신 PAT가 STOC·ICML·NeurIPS·ICLR에서 1만6000편 넘는 원고를 검토했고 사용자는 매주 2건을 무료로 쓴다.

3 함께 공개된 어스 AI 에이전트는 에볼라 발생 초기 위험 정착지 48곳과 위험 인구 4만5500명 이상을 몇 분 만에 찾았고, 개방형 의료 모델 메드젬마는 내려받기 1000만 회를 넘겼다.

구글 리서치가 10월 14일 미국 마운틴뷰에서 연 연례 행사 리서치앳(Research@) 2026에서 논문 검증 에이전트 딥 리뷰를 일반에 공개했다. 딥 리뷰는 긴 과학 논문과 기술 문서를 구간별로 나눠 여러 추론 에이전트가 동시에 검토하는 AI 시스템이다. 시험 버전은 컴퓨터과학 최상위 학회 4곳에서 1만6000편이 넘는 원고를 검토했다. 구글은 같은 자리에서 지구 관측 AI, 과학 연구 보조 도구, 의료 AI, 홍수 예측까지 연구 성과 10건을 함께 공개했다.

배경에는 검증의 병목이 있다. 딥 리뷰 소개 글을 쓴 라제시 자야람과 뱅상 코엔아다드 구글 리서치 연구과학자에 따르면 올해 ICLR, NeurIPS, ICML 세 AI 학회에 들어온 원고는 7만4000편을 넘었다. 2023년 약 2만3000편, 2024년 3만2000편, 2025년 4만5000편에서 해마다 불어났다. 메탈이 확인한 ICML 2026 공식 자료표에 따르면 7월 서울 코엑스에서 열린 ICML 2026 본회의 트랙에만 2만4661편이 제출돼 6552편이 채택됐다. 원고를 쓰는 속도는 AI로 빨라졌지만 원고를 검증할 사람의 수는 거의 그대로라는 것이 구글의 진단이다.

딥 리뷰는 네 단계로 움직인다. 분할 에이전트가 문서를 서론, 방법, 실험 같은 논리 구간으로 나누고, 구간마다 정보 밀도를 따져 연산 예산을 다르게 배정한다. 증명이 빽빽한 절에는 서론보다 많은 연산이 간다. 이어 전문 추론 에이전트들이 문서 전체 맥락을 보면서 각자 맡은 구간을 병렬로 검증하고, 마지막으로 종합 에이전트가 구간별 보고서를 모아 절 사이의 모순을 다시 확인하고 외부 근거 대조로 환각을 걸러 낸다.

AI

메타, Muse 기반 건강 연구 원고 10편 공개

메타가 Muse Spark 1.3과 Muse Wonder로 만든 연구 스캐폴드를 의료 연구자들에게 맡겨 4주 동안 건강 연구 원고 10편을 만들었다. 원고 한 편에 2주가 채 걸리지 않았지만 결론의 책임은 연구자에게 남겼다.



1 메타는 10월 14일 Muse Spark 1.3과 Muse Wonder 기반 코딩 스캐폴드로 의료 연구자들이 4주 동안 수행한 건강 연구 원고 10편을 공개했다.

2 연구자들은 약 500만 명 규모의 비식별 전자건강기록으로 첫 프롬프트에서 원고까지 2주 안에 마쳤고, 가장 좋은 경우 기존보다 최대 10배 빨랐다고 보고했다.

3 연구진 보고서는 사망률 1% 오계산과 헤파린 용량 혼동 같은 모델의 오류를 의사들이 잡아낸 사례를 공개하며 최종 책임은 연구자에게 있다고 밝혔다.

메타가 10월 14일 자사 AI 모델 Muse로 의료 연구자들이 쓴 건강 연구 원고 10편을 공개했다. Muse Spark 1.3과 Muse Wonder를 결합한 사내 코딩 스캐폴드를 만들고, 의료 연구자들이 이 도구로 가설 수립부터 코드 실행과 원고 작성까지 진행하게 한 실험의 결과물이다. 메타는 연구 블로그에서 AI가 연구의 모든 단계를 앞당기되 과학적 판단과 결론은 해당 분야 전문가가 책임지는 사람 주도 체계라고 설명했다. 원고는 감염병, 스포츠의학, 응급의학, 종양학, 내분비학, 구강 건강 등을 다룬다.

판을 키운 것은 데이터다. 스캐폴드는 공개 데이터와 함께 메타가 라이선스를 받은 후향적 비식별 전자건강기록(EHR) 데이터베이스에 접근했다. 메타는 이 데이터를 500만 명 규모라고 밝혔고, 함께 공개된 논문에는 2026년 6월 배포본 기준 495만 명, 21개 미국 의료 시스템을 잇는 기록과 청구 자료라고 적혀 있다. 심장학, 역학, 감염병, 내과 배경의 연구자들이 4주 동안 이 체계로 독자 연구를 수행했다. 메타가 공개한 2분 22초 영상에서 연구진은 의료 전문가 15명이 참여했다고 소개했다.

속도는 메타가 가장 앞세운 숫자다. 연구 블로그에 따르면 모델에 첫 프롬프트를 넣고 원고가 완성되기까지 2주가 채 걸리지 않았고, 빠른 논문은 7일 만에 나왔다. 참가자들은 만장일치로 큰 생산성 향상을 느꼈다고 답했으며, 가장 좋은 경우 기존 연구 과정보다 최대 10배 빨랐다고 보고했다. 영상 속 한 연구자는 AI 없이는 보통 두세 달 걸리던 일을 아이디어에서 원고까지 2주 안에 끝낸다고 말했다.

비즈니스

메타, 아태 소상공인 AI 활용 조사 공개

메타가 아태 9개 시장 소상공인 3983명 조사에서 84%가 AI를 쓰고 76%가 생산성 효과를 봤다는 결과를 공개했다. 86%는 AI를 제대로 쓰려면 더 많은 안내가 필요하다고 답했고, 교육의 가장 큰 장벽은 비용이었다.



1 메타는 10월 15일 유고브에 의뢰한 아태 9개 시장 소상공인 3983명 조사에서 84%가 AI를 쓰고, 76%가 생산성에, 71%가 매출에 긍정적 효과를 봤다고 밝혔다.

2 응답자의 80%는 AI가 대기업과의 경쟁을 돕는다고 했고, 쓰임새는 콘텐츠 51%, 지역 광고·마케팅 43%, 번역 38%, 고객 응대 37% 순이었다.

3 86%가 더 많은 안내가 필요하다고 했고 교육 장벽 1위는 비용(51%)이었으며, 메타는 그로스 아카데미와 Advantage+ 등 자사 도구를 해법으로 내놨다.

메타가 10월 15일 아시아태평양(아태) 9개 시장의 소상공인 3983명을 대상으로 한 AI 활용 조사 결과를 공개했다. 메타가 의뢰하고 여론조사 기관 유고브가 온라인으로 진행한 이 조사에서 응답자의 84%가 지금 AI를 쓰고 있다고 답했다. 지난 12개월 동안 생산성에 긍정적 효과가 있었다는 응답은 76%, 매출에 긍정적 효과가 있었다는 응답은 71%였다. 그러나 86%는 소상공인이 AI의 효과를 온전히 누리려면 더 많은 안내가 필요하다고 했다. 메타는 소상공인이 이미 AI를 받아들였고 이제는 잘 쓰는 법에 대한 도움을 원한다고 밝혔다.

조사 대상은 호주, 싱가포르, 베트남, 인도네시아, 대만, 인도, 태국, 일본, 홍콩의 소상공인이다. 자료는 2026년 8월 6일부터 9월 2일까지 모았고, 시장마다 직원 1~9명, 10~19명, 20~49명 규모별로 할당을 둬 고르게 응답을 받았다. 메타는 이 결과가 각 시장 소상공인 전체를 대표하는 수치는 아니며, 시장별 차이에는 질문 해석과 번역, 지역별 응답 관행이 일부 반영됐을 수 있다고 밝혔다.

응답자들은 AI를 몸집 차이를 줄이는 도구로 봤다. 9개 시장 전체에서 80%가 AI 덕분에 소상공인이 대기업과 더 효과적으로 경쟁할 수 있다고 답했고, 같은 비율이 AI가 창업을 쉽게 만든다고 했다. 83%는 예전에는 훨씬 많은 돈이나 자원이 필요했던 일을 AI로 할 수 있게 됐다고 답했다. 소상공인이 겪는 두 가지 큰 어려움인 자원 부족과 시간 부족을 AI가 덜어 준다는 응답도 나왔다.

AI

오픈AI, 블록·센티넬원 데이브레이크 사례 공개

오픈AI가 블록과 센티넬원의 데이브레이크 도입 사례를 공개했다. 센티넬원은 4~6주 걸리던 악성코드 역분석을 약 8시간으로 줄였고, 블록은 분기마다 분석가 약 1100시간 분량의 조사를 AI로 수행한다.



1 오픈AI는 10월 14일 블록과 센티넬원의 데이브레이크 도입 사례를 공개했고, 센티넬원은 최고 수준 연구자가 4~6주 걸리던 악성코드 역분석을 약 8시간으로 줄였다고 밝혔다.

2 블록은 탐지 규칙 변경을 7개월 치 원격측정 데이터에 1시간 안에 다시 돌려 보고, 분류 처리량 50% 증가와 분기당 오탐 분류 약 600시간 회수를 보고했다.

3 두 회사는 아키텍처 변경과 패치 배포 같은 결정은 사람이 맡는다고 밝혔으며, 센티넬원은 9월부터 GPT-5.6-Cyber를 위협 노출 점검 서비스에 넣고 있다.

오픈AI가 10월 14일 핀테크 기업 블록과 보안 기업 센티넬원의 데이브레이크 도입 사례를 함께 공개했다. 데이브레이크는 오픈AI가 방어자에게 사이버 보안 작업용 모델을 공급하는 프로그램이다. 두 사례의 공통분모는 시간이다. 센티넬원은 최고 수준 연구자가 4~6주 걸리던 악성코드 역분석을 약 8시간으로 줄였고, 블록은 최근 레드팀 훈련에서 상위 수준의 취약점 보고를 넣고 자동 패치 제안이 나오기까지 18분이 걸렸다고 밝혔다.

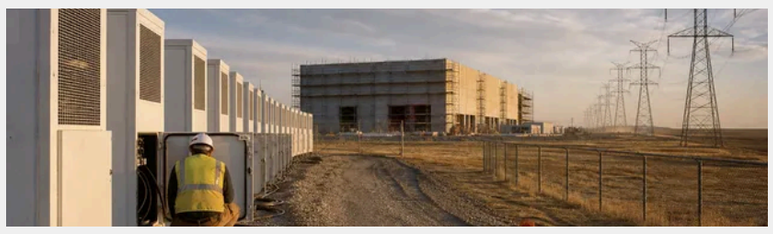
블록은 결제 단말 Square, 송금 앱 Cash App, 후불결제 Afterpay를 운영하는 회사다. 오픈AI에 따르면 블록 보안팀은 AI가 취약점 발견과 탐지, 대응을 맡는 오픈AI 디펜스 팩토리 방식을 쓰고 있으며, 몇 주 걸리던 일을 몇 시간 만에 끝낸다. 제임스 네테스하임 블록 최고정보보호책임자(CISO)는 "AI가 사람의 의도에서 실행까지 걸리는 시간을 압축하고 있다"며 "데이브레이크 같은 도구 덕분에 방어자도 불은 불로 맞서며 AI 도입 속도를 따라갈 수 있다"고 말했다.

블록이 공개한 수치는 탐지 업무에 몰려 있다. 조사마다 원본 쿼리 증거를 저장해 두기 때문에, 탐지 규칙을 바꾸자는 제안을 7개월 치 운영 원격측정 데이터에 다시 돌려보는 데 1시간이 채 걸리지 않는다. 예전에는 지난 사건 수백 건을 손으로 다시 조회해야 해서 실제로는 거의 하지 못했던 일이다. 자동 분류 체계를 들인 뒤 분류 처리량은 50% 늘었고, 정보 수집과 조사도 50% 늘었다. 조사는 분기마다 분석가 약 1100시간 분량을 수행하고, 오탐 분류에서 돌려받은 시간은 분기당 약 600시간이다. 네테스하임 CISO는 "모든 경보를 그날 유일한 경보처럼 다룬다"고 설명했다.

기후

ON.energy, 크루소에 AI UPS 배치

ON.energy가 데이터센터와 전력망 사이에 두는 리튬이온 AI UPS를 크루소 캠퍼스 5기가와트 규모로 상용 배치한다. 수천 개 GPU의 급격한 부하 변동과 데이터센터의 갑작스러운 계통 이탈을 함께 흡수하는 장치다.



1 크루소는 7월 21일 ON.energy와 여러 하이퍼스케일 캠퍼스에 AI UPS 5기가와트를 배치한다고 발표했고, 가동은 2026년에 시작해 2027년까지 이어진다.

2 AI 학습 시설의 전력 수요는 하루 수백~수천 번 오르내리며, 7월에는 버지니아에서 데이터센터 부하 약 4기가와트가 한꺼번에 전력망에서 떨어졌다.

3 AI UPS는 국립 로키산맥 연구소 시험과 ERCOT 대형 부하 요건 검증을 거쳤지만, 데이터센터를 처음부터 이 장치에 맞춰 설계해야 한다.

AI 데이터센터와 전력망 사이에 리튬이온 배터리를 끼워 두는 새 무정전전원장치(UPS)가 상용 배치를 준비하고 있다. 미국 스타트업 ON.energy의 「AI UPS」다. AI 인프라 기업 크루소는 7월 21일 ON.energy와 함께 여러 하이퍼스케일 캠퍼스에 AI UPS 5기가와트를 배치하겠다고 발표했고, 가동은 2026년에 시작해 2027년까지 이어진다. 보도에 따르면 ON.energy는 이번 주문으로 이 중앙 집중형 UPS를 상용 규모에서 처음 시험대에 올린다.

문제의 출발점은 AI 데이터센터의 전력 사용 패턴이다. 크루소 발표문에 따르면 학습 중에는 수천 개의 GPU가 한꺼번에 출력을 올렸다 내리며, 대형 AI 학습 시설의 전력 수요는 하루에 수백에서 수천 번 높낮이를 오간다. 이 변동을 관리하지 않으면 계통 연결 지점을 거쳐 송전망으로 번진다. 국립 로키산맥 연구소(옛 국립재생에너지연구소)의 샤힐 샤 수석 엔지니어는 수요가 "몇 밀리초 만에 5%에서 100%로" 뛸 수 있다며 "그렇게 하면 전력망을 두들기는 셈"이라고 말했다.

반대 방향의 위험도 있다. 데이터센터는 전력 품질의 흔들림에 민감해 순간적으로 전력망에서 떨어져 나가 자체 UPS와 비상 발전기로 넘어간다. 그 순간 계통에서는 거대한 수요가 갑자기 사라진다. 보도에 따르면 미국 동부 전력망 운영기관 PJM은 최근 몇 년 사이 기가와트급 이탈을 여러 차례 기록했고, 7월에는 버지니아에서 데이터센터 부하 약 4기가와트가 한꺼번에 끊겼다. 북미전력신뢰도공사(NERC)는 대형 「연산 부하」에 새 요건을 매기는 작업에 들어갔고, 텍사스 규제 당국도 올해 비슷한 요건을 도입했다. 샤 수석 엔지니어는 "많은 계통 운영자가 대형 부하인 데이터센터가 전력망 신뢰도를 해치지 않도록 요건을 만들고 있다"고 말했다.

기후

리플 연구진, 2026 기후 상태 보고서 발표

지구 기후의 활력징후 16개 가운데 7개가 관측 사상 가장 나쁜 값을 찍었다. 재생에너지가 석탄을 처음 앞질렀지만 화석연료 사용과 배출도 함께 사상 최대를 기록했다.



1 국제 과학자 연구진은 기후 지표 16개 중 7개가 기록상 가장 극단적인 상태라고 밝혔다.

2 지구 온난화 속도는 10년당 0.05도에서 0.36도로 빨라졌고 이 속도라면 2050년 전에 2도 상승에 이를 수 있다.

3 2025년 재생에너지 발전이 처음으로 석탄을 넘었지만 에너지 수요 증가로 화석연료 소비와 배출도 함께 최대치를 기록했다.

국제 과학자 연구진이 10월 14일 학술지 바이오사이언스(BioScience)에 2026년 기후 상태 보고서를 발표했다. 오리건주립대학교 윌리엄 리플 교수가 주저자를 맡았고 마이클 만, 요한 록스트룀 등 기후과학자들이 공동 저자로 참여했다. 연구진이 추적하는 기후 지표 16개 가운데 7개, 비율로 44%가 관측 이래 가장 극단적인 값에 도달했다. 연구진은 보고서 서두에 "이것은 더 이상 미래에 대한 경고가 아니라 현재에 대한 진단"이라고 적었다.

기록을 다시 쓴 지표의 출발점은 온실가스다. 2026년 5월 대기 중 이산화탄소 농도는 428ppm을 넘었고, 이는 산업화 이전보다 약 50% 높은 수준이다. 메탄과 아산화질소도 올해 누적 평균 기준으로 나란히 최고치를 기록했다. 보고서는 이 상승의 대부분이 화석연료 연소에서 나온다고 밝혔다. 이산화탄소 농도가 지난 200만 년 중 가장 높다는 기후변화에 관한 정부간 협의체(IPCC)의 평가도 함께 인용했다.

얼음과 바다 지표도 같은 방향을 가리킨다. 빙하 두께와 그린란드 빙상 질량은 기록상 최저로 내려갔고, 북극 해빙 최소 면적과 남극 빙상 질량 변화도 최저 기록에 가까워졌다. 반대로 해양 열 함량과 해수면은 예비 자료 기준으로 최고치다. 보고서가 인용한 연구에 따르면 바닷물이 데워지며 부피가 늘어나는 열팽창이 최근 수십 년 해수면 상승의 약 56%를 설명한다. 잦아진 해양 열파는 산호 백화와 해양 생태계 교란으로 이어지고 있다.

기후

업타임 인스티튜트, 데이터센터 이중화 효과 분석

업타임 인스티튜트가 1200명 넘는 운영자 설문으로 전력·냉각 이중화와 정전의 관계를 계량했다. 냉각을 2N으로 갖춘 곳은 정전 경험 비율이 39%로 예비 없는 곳의 70%보다 크게 낮았다.



1 업타임 인스티튜트는 2025·2026년 연례 조사 응답자 1200명 이상을 분석해, 냉각 2N 구성의 최근 3년 정전 경험 비율이 39%로 N 구성의 70%보다 낮았다고 밝혔다.

2 N 대비 재무 피해가 큰 정전 가능성은 냉각에서 N+1 36%, N+2 53%, 2N 63% 낮았고 전력에서는 23%, 37%, 44% 낮았다.

3 저자들은 상관관계가 인과관계는 아니라고 밝히면서도, 운영자가 설비 비용과 다운타임 위험을 저울질할 지표로 쓸 수 있다고 설명했다.

데이터센터의 전력·냉각 설비를 두 벌로 갖춘 운영자는 예비 설비가 없는 운영자보다 정전을 겪은 비율이 크게 낮았다. 데이터센터 신뢰성 평가 기관 업타임 인스티튜트가 10월 14일 공개한 분석에 따르면 냉각 설비를 완전 이중화(2N)한 응답자 가운데 최근 3년 안에 영향이 큰 정전을 한 번 이상 겪은 비율은 39%였고, 예비 없이 최대 부하에 딱 맞춘(N) 응답자는 70%였다. 전력도 2N 42%, N 68%로 같은 모양이었다. 재무 피해가 큰 정전을 겪을 가능성은 N에 비해 2N 냉각에서 약 63% 낮았다.

분석의 바탕은 두 해의 설문이다. 업타임은 2025년과 2026년 연례 글로벌 데이터센터 조사를 합쳐 응답자 1200명이 넘는 표본을 만들었다. 운영자들은 주력 데이터센터가 최근 3년 안에 영향이 큰 정전을 겪었는지 답했고, 겪었다면 가장 최근 정전을 1등급(영향 최소)부터 5등급(심각)까지 매겼다. 3~5등급은 대체로 의미 있는 재무 피해를 낸 정전이다. 업타임은 이 답을 각 데이터센터의 전력·냉각 이중화 구성과 교차 집계했다.

이중화 구성은 네 단계로 나뉜다. N은 예상 최대 부하에 맞춘 용량만 갖춰 핵심 부품 하나가 고장 나면 전체 부하를 감당하지 못할 수 있다. N+1은 필요한 것보다 부품 하나를 더 둬 한 부품의 고장이나 정비를 서비스 중단 없이 견딘다. N+2는 둘을 더 둬 한 부품을 정비하는 동안 다른 하나가 고장 나도 버틴다. 2N은 필요한 용량 전체를 통째로 한 벌 더 갖춰, 각각 전체 부하를 감당하는 독립된 두 경로를 만든다.