AI 뉴스와 용어
월·수·금 오전 7시

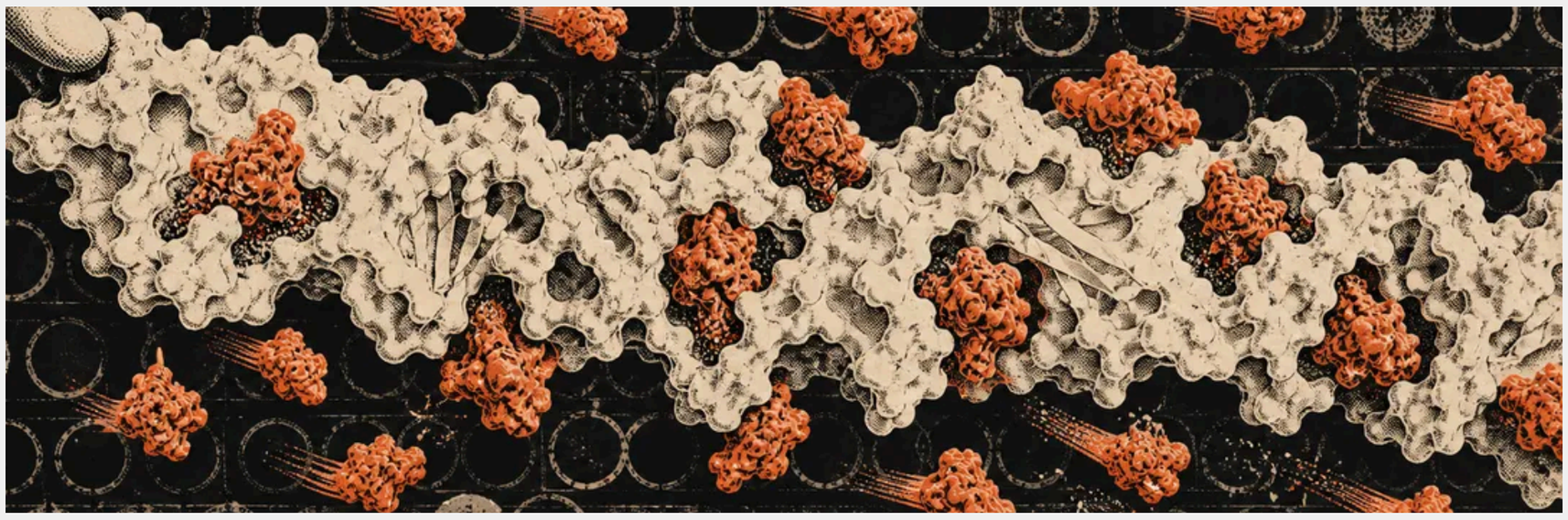
LAB METAL
어려운 AI, 쉽고 재미있게

무료 · 언제든 수신거부
metallab.ai

제 16 호 | 2026년 9월 3일 목요일 | 오늘의 AI

클로드, 24시간 자율 설계 단백질 12개 중 절반 결합

앤스로픽 클로드가 표적만 받고 24시간 혼자 설계한 단백질을, 독립 실험실 두 곳이 검증하니 절반 가까이 결합했어요

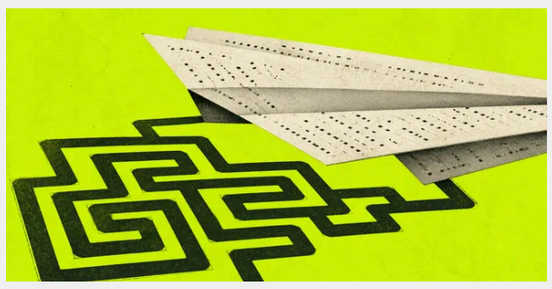


· 클로드, 24시간 자율 설계 단백질 12개 중 절반 결합

앤스로픽의 AI 모델 클로드가 사람 도움 없이 표적 단백질 정보만 받고 24시간 안에 새로운 결합 단백질을 설계했고, 독립 실험실 두 곳이 이 가운데 절반 가까이가 실제로 표적에 달라붙는다는 사실을 확인했어요. 목표 하나만 주고 나머지 설계 과정 전부를 AI가 처리했다는 점이 이번 사례의 핵심이에요.

풀어서 설명하면, 앤스로픽은 지난 8월 클로드가 인간 전문가의 프롬프트를 받아 15개 표적 단백질 중 14개에 결합 분자를 설계했고 Adaptyv Bio와 Twist Bioscience 검증에서 22\~35%가 실제로 결합했다고 밝힌 적이 있어요. 이번에는 사람이 표적만 알려주고 나머지 설계 전 과정을 클로드가 24시간 동안 혼자 맡았다는 점이 다르고, 검증 기관도 GenScript가 새로 추가됐어요.

클로드는 표적 단백질 하나가 주어지면 오픈소스 생체분자 모델을 이용해 후보 단백질을 생성하고, 접힘 구조를 예측하고, 점수를 매기는 과정을 스스로 반복했어요. 이후 생물정보학 도구로 후보군의 다양성과 신규성을 검사한 다음, 표적에 달라붙을 가능성이 가장 높은 설계를 골라냈고요. 사람이 개입한 부분은 표적을 알려준 것뿐이었고, 나머지 설계 전 과정은 24시간 동안 클로드 혼자 진행했어요.



구글, 제미나이 3.8 플래시 공개…코딩·에이전트 작업에 초점

클로드가 고른 설계는 Adaptyv Bio와 GenScript, 서로 다른 두 실험실에서 실제로 합성되고 측정됐어요. 공개된 12개 표적 가운데 클로드가 설계한 단백질의 절반 가까이가 를 냈다는 거예요. 지난 8월 실험은 사람 전문가가 프롬프트로 방향을 짚어 준 결과였고, 이번엔 표적 하나만 던져 놓고 24시간을 통째로 맡겼는데도 비슷한 급의 결합률이 나왔어요. 이 격차가 좁혀졌다는 사실 자체가 앤스로픽이 항체·저분자화합물까지 신약 개발 전 과정을 자동화하려는 방향과 맞닿아 있어요.

오픈 소스 단백질 언어 모델을 여러 회사가 함께 실무에 붙여 보면 항상 비슷한 벽에 부딪혀요. 시뮬레이션 점수는 그럴듯한데 실제 결합 실험에서 재현이 안 되는 경우가 많다는 것이죠. 그래서 Adaptyv Bio와 GenScript처럼 서로 다른 실험실 두 곳이 독립적으로 검증했다는 절차 자체가, 숫자 하나보다 더 눈여겨봐야 할 대목이에요. 재현성 없는 in silico 결과는 논문 한 편으로 끝나지만, 두 곳에서 따로 확인된 결과는 다음 단계로 넘어갈 근거가 되거든요.

국내 바이오·제약 팀이라면 지금 당장 이런 완전 자율 파이프라인을 그대로 들여오기보다는, 오픈소스 생체분자 모델과 자체 검증 실험실 협업 구조를 먼저 갖추는 쪽이 현실적이에요. AI가 설계를 빠르게 늘려 줄수록 웨트랩 검증 역량이 병목이 되기 때문이고요. 앞으로 몇 주 안에는 앤스로픽이 이번 자율 설계 방식을 항체나 저분자화합물 같은 다른 분자 유형으로 확장하는 사례가 추가로 나올 가능성이 커 보여요.

집할 수 없게 됐는데, 이는 수천 개의 가짜 계정으로 모델의 사고를 빼내 안전장치 없는 복제본을 만드는, 공개적으로 알려진 증류 기법을 막기 위한 조치예요.

Fable 5.1은 토큰 단위로 과금하는 일반 워크로드 기준으로 Fable 5보다 약 25% 저렴해질 것으로 앤스로픽은 추산해요. 이미 처리해 저장해 둔 입력을 다시 읽어오는 '캐시 읽기' 가격을 낮춘 덕분인데, 캐시 읽기 요금은 100만 토큰당 $1에서 $0.25로 75% 낮아졌어요. 에이전트가 오래 돌아가는 고강도 작업에서는 절감 폭이 최대 약 45%까지 커진다고 해요. 다만 값이 내린 항목은 캐시 읽기 하나뿐이고, 입력 100만 토큰당 $10·출력 100만 토큰당 $50이라는 기본 단가는 Fable 5와 똑같이 그대로예요.

지난주 배경 기사에서 다뤘듯 GLM 5.2 같은 저비용 모델이 Fable의 약 9분의 1 가격으로 등장하면서 개발자들이 작업을 쪼개 쓰는 흐름이 생겼는데, 이번 가격 인하는 그 흐름에 대한 앤스로픽의 대응으로 읽혀요.

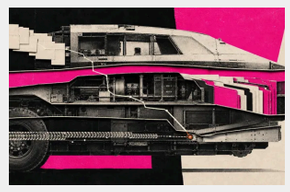
한 가지 유보할 점은, 벤치마크 분석기관 Artificial Analysis는 이 절감 폭이 작업 방식에 따라 달라진다고 지적했다는 거예요. 캐시 읽기가 대부분을 차지하는 에이전틱 작업에서는 작업당 약 $1.40이 절약되지만, 노력 수준을 최대(max)로 올리면 Fable 5.1이 토큰을 더 많이 뱉어 내오히려 작업당 비용이 Fable 5보다 약 20% 높아진다는 계산이에요. '최대 45%'는 기본 설정에서 캐시를 많이 재사용할 때의 상단값이라는 점을 감안해야 해요.

출시와 함께 앤스로픽은 모든 이용자의 5시간·주간 사용량 한도를 초기화했어요. 공식 개발자 계정 ClaudeDevs가 이날 알린 내용인데, 기존 한도를 얼마나 썼든 이날 기준으로 다시 시작돼 새 모델을 부담 없이 바로 써볼 수 있어요.

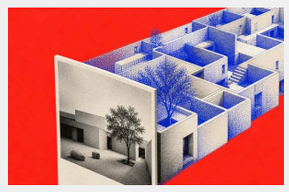
앤스로픽 Fable 5.1·Mythos 5.1 출시

앤스로픽이 6월 9일 Fable 5를 내놓은 지 약 석 달 만인 9월 1일, 코딩·지식노동용 신모델 Claude Fable 5.1과 사이버보안·생명과학 전용 Claude Mythos 5.1을 함께 공개했어요. 두 모델은 사실 같은 모델이고 안전장치 수준만 다르게 설정한 버전인데, Fable 5.1은 누구나 쓸 수 있도록 일반 공개됐고 Mythos 5.1은 신뢰 접근 프로그램을 통과한 기

오늘 지면



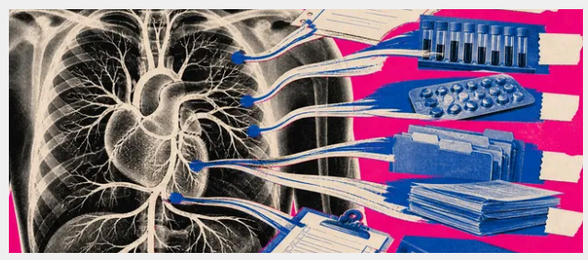
클로드 Fable 5.1, 재현 안 되는 버그 원인 찾아냈다



월드랩스 아틀라스, 사진 몇 장으로 3D 세계 생성·복원

이 지면은 2026년 9월 3일 목요일 아침 구독자 메일함으로 나간 레터를 신문 판형으로 다시 조판한 것입니다. 기사 전문은 metallab.ai 에서 읽을 수 있습니다.

이 어 서



챗GPT, 에픽 EHR과 의료데이터 9종 연결한다

대상은 미국 대형 병원 다수가 쓰는 EHR 소프트웨어 에픽(Epic)이에요. 병원이 자신들의 에픽 환경을 챗GPT 포 헬스케어 계정에 연결하면, 의사는 예약 메모·검사 결과·투약 기록·전문의 소견을 일일이 찾아다니는 대신 챗GPT에 "지난 방문 이후 달라진 게 뭔가요", "오늘 진료 전에 봐야 할 최근 검사 결과가 있나요" 같은 질문을 바로 던질 수 있어요. 오픈AI는 이 기능을 두 방식으로 나눠 지원한다고 설명했어요. 하나는 챗GPT 화면에서

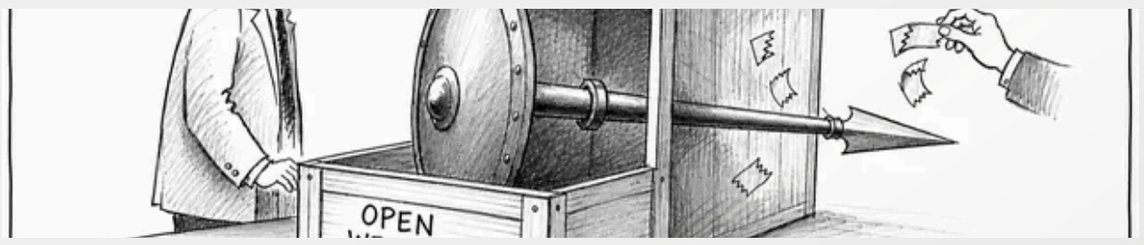
▶ 2면에 계속

앤스로픽 캠퍼스 앰버서더, 대학원·박사 트랙 신설했다

클로드 캠퍼스 앰버서더는 학생들이 각자 학교에서 AI 활용을 넓히는 걸 앤스로픽이 지원하는 프로그램으로, 매 학년도 다시 문을 여는 방식으로 운영돼요. 이번 신청 공고에는 지난해 앰버서더로 활동한 학생도 재신청할 수 있다는 안내가 붙었고, 프로그램은 앤스로픽이 직접 운영하되 외부 에이전시가 실무를 돕는 구조로 짜여 있어요.

세 트랙은 대상 학생군과 캠퍼스에서 맡는 역할이 서로 달라요. 학부생 트랙은 캠퍼스에 AI 빌더 클럽을 직접 꾸리고 해커톤과 데모나이트를 여는 실행형 역할이고, 대학원생 트랙은 AI가 교육과 사회에 미치는 영향을 놓고 토론 모임을 여는 데 초점이 맞춰져 있어요. 박사·포스닥 트랙은 연구에 AI를 쓰는 과학자들을 모아 클로드 과학 워크숍을 조직하는 역할이에요.

오픈웨이트 AI, 같은 달에 가장 싼 방패이자 가장 쉬운 창이 됐다



2026년 8월, 오픈웨이트 AI 모델은 보름 사이에 정반대의 평가를 동시에 받았어요. 8월 14일 중국 AI 기업 Z.ai는 코딩과 사이버보안 작업을 자동화하는 오픈웨이트 모델 GLM-5.3을 발표했고, 버셀 CEO와 보안 전문가들은 이걸 「방어자에게 돌아가는 가장 싼 무기」라고 반겼어요. 그리고 8월 말, 캐나다 워털루대학교와 AI 안전 연구단체 FAR.AI가 이끈 국제 연구팀은 널리 쓰이는 오픈웨이트 모델 21개의 안전장치가 파인튜닝 한 번에 무너진다는 결과를 내놨어요.

두 소식을 각각 읽으면 하나는 기대, 하나는 경고예요. 그런데 나란히 놓으면 같은 자리를 가리켜요. 오픈웨이트 모델은 누구나 내려받아 자기 서버에서 돌릴 수 있어서 싸고, 그래서 방어에 쓰기 좋고, 바로 그 이유로 안전장치도 누구나 떼어낼 수 있어요. 이 기사는 WIRED가 전한 두 사건을 한 축으로 다시 읽어요.

◀ 1면에서 계속

앤스로픽 캠퍼스 앰버서더, 대학원·박사 트랙 신설했다

1 앤스로픽이 클로드 캠퍼스 앰버서더 프로그램의 2026~2027학년도 신청을 9월 1일 열었고, 학부생 전용이던 트랙을 대학원생·박사·포스닥까지 넓혀 3개로 늘렸어요

2 선발된 학생은 트랙과 무관하게 3,600달러 현금 지원금을 받고, 신청은 9월 12일 태평양시간 밤 11시 59분에 마감돼요

3 대학원·박사 트랙은 초기 파일럿 기간이라 정해진 학교 풀에서만 선발하는 반면 학부생 트랙은 전 세계 누구나 신청할 수 있어요

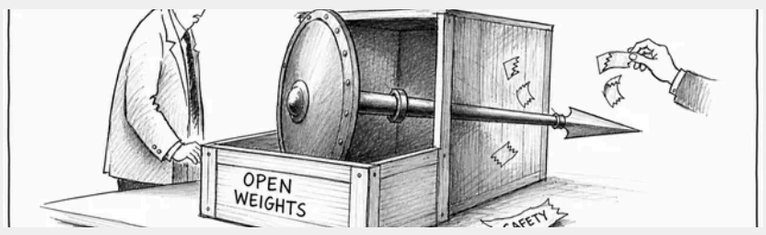
신청 자체는 전 세계 어느 학교 학생이든 가능해요. 다만 대학원생·박사 트랙은 이번이 첫 파일럿 시행이라 초기에는 정해진 학교 풀에서만 선발할 수 있다고 안내됐어요. 학부생 트랙과 조건 폭이 다른 셈이에요. 공통 조건은 18세 이상이고 재학 중인 국가에서 노동허가를 가지고 있어야 한다는 점, 그리고 신청과 참여 모두 비용이 들지 않는다는 점이에요. 프로그램은 2026년 9월부터 2027년 6월까지 한 학년도 내내 이어져요.

신청은 9월 1일 열려 9월 12일 태평양시간 밤 11시 59분에 마감되고, 앤스로픽은 접수되는 순서대로 서류를 검토하는 롤링 리뷰 방식을 쓴다고 밝혔어요.

지원 절차는 단일 신청 폼으로 이뤄져요. 클로드 캠퍼스 신청 페이지에 접속해 원하는 트랙(학부·대학원·박사)을 선택한 뒤, 소속 학교와 전공, AI·클로드 활용 경험을 적어 제출하면 돼요. 지난해 앰버서더였던 학생은 폼에서 그 사실을 표시하면 되고, 서류는 접수 순서대로 검토돼 결과가 통지돼요. 박사 트랙 지원자는 실제 연구에 클로드를 활용하고 있는지가 핵심 심사 기준으로 안내돼 있어요.

◀ 1면에서 계속

오픈웨이트 AI, 같은 달에 가장 싼 방패이자 가장 쉬운 창이 됐다



1 중국 Z.ai는 8월 14일 코딩·사이버보안 작업을 자동화하는 오픈웨이트 모델 GLM-5.3을 발표하고, 위험을 인정해 보안 파트너부터 단계적으로 공개하고 있어요.

2 워털루대와 FAR.AI 연구팀은 TamperBench로 라마·큐원·미스트랄 등 21개 오픈웨이트 모델을 9가지 변조 기법으로 시험했고, 성능을 크게 깎지 않고 변조를 막아낸 방어 기법은 하나도 없었어요.

3 두 소식은 같은 질문을 가리켜요 — 오픈웨이트 모델을 공개할지 말지가 아니라, 공개된 뒤 안전장치가 떼어진 모델을 누가 어떻게 책임지느냐예요.

풀어서 설명하면, 오픈웨이트 모델은 완성된 AI의 두뇌 파일을 통째로 내려받아 내 컴퓨터에서 돌리고 마음대로 고칠 수 있는 모델이에요. 그래서 싸고 자유롭지만, 원래 만든 회사가 심어 둔 「하지 마」 장치까지 내 손으로 떼어 낼 수 있다는 뜻이기도 해요.

Z.ai는 발표 글에서 GLM-5.3이 앤스로픽과 오픈AI의 최고 성능 공개 모델에 거의 맞먹는다고 밝혔어요. 근거로 든 건 코딩·사이버보안 벤치마크 점수인데, 인기 있는 보안 벤치마크 CyberGym 등 일부 항목에서는 두 회사 모델에 근접하거나 앞섰다고 주장했어요. 훈련 방식은 이미 해결된 문제의 예시를 모델에 주고 시행착오로 익히게 하는 사후 학습(포스트트레이닝)이었고요. 같은 날 코드 저장소의 취약점을 GLM-5.3으로 스캔해 주는 서비스 OpenVuln도 함께 내놨어요.

여기까지만 보면 여느 모델 출시와 다르지 않아요. 다른 점은 공개 방식이에요. Z.ai는 발표 글에서 이 능력이 방어자가 약점을 일찍 찾고 대응 속도를 높이는 데 도움이 되는 동시에 명백한 이중용도 위험을 만든다고 인정했어요. 그래서 선정된 보안 파트너가 통제된 환경에서 먼저 검증하고, 엄격한 안전성 평가를 거친 뒤 8월 하순에서 9월 초 사이에 가중치와 API를 전부 열겠다고 했어요. 오픈웨이트를 표방하는 회사가 프런티어 연구소처럼 단계적 공개를 택한 거예요.

◀ 1면에서 계속

챗GPT, 에픽 EHR과 의료데이터 9종 연결한다

1 오픈AI가 챗GPT 포 헬스케어에 에픽(Epic) EHR 통합과 헬스케어 퍼블릭 데이터 플러그인을 9월 1일 공개했어요.

2 플러그인은 ClinicalTrials.gov, CMS Coverage, RxNorm, DailyMed, PubMed 등 9개 공식 의료 소스를 하나로 묶었어요.

3 27개 임상 시나리오 4,363건 평가에서 99.1%가 안전 판정을 받았고, 5개 데이터 소스 응답의 93% 이상이 '양호' 이상 정확도로 평가됐어요.

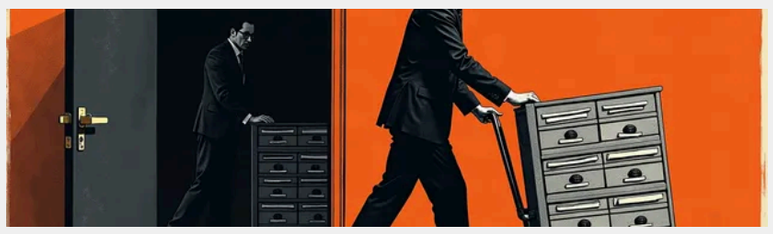
UCSF 헬스는 이번 통합의 파일럿 파트너로 참여했어요. 수레시 구나세카란 UCSF 헬스 대표는 "이 기술은 데이터를 종합하는 데 드는 시간을 줄일 잠재력이 있다"고 말했어요.

환자 기록만으로는 충분하지 않아요. 의료진은 최신 연구 논문, 약품 허가 정보, 보험 급여 기준, 임상시험 모집 현황도 함께 봐야 하는데, 이 정보들은 대개 서로 다른 사이트에 흩어져 있어요. 오픈AI가 함께 내놓은 헬스케어 퍼블릭 데이터 플러그인은 ClinicalTrials.gov(임상시험 등록), CMS Coverage(메디케어 급여 기준), RxNorm(의약품 표준 코드), DailyMed(약품 라벨 정보), PubMed(의학 논문 검색)를 포함해 아홉 개 공식 소스를 하나의 플러그인으로 묶었어요.

오픈AI가 든 예시를 보면 활용 폭이 넓어요. 연구팀은 ClinicalTrials.gov로 모집 중인 임상시험을 찾고 참가 조건을 비교할 수 있고, 약제팀은 DailyMed로 특정 약품의 최신 라벨과 경고 문구를 확인할 수 있어요. 당뇨병 예방 프로그램을 기획하는 공중보건팀이라면 관련 연구, 진행 중인 임상시험, 메디케어 급여 정보를 한 화면에서 근거와 함께 살펴볼 수 있다고 오픈AI는 설명했어요.

앤스로픽 Fable 5.1·Mythos 5.1 출시

같은 모델을 안전장치로 가른 두 버전, 캐시 읽기 인하로 비용 낮추고 EFS로 데이터 보관 방식도 바꿔요



1 앤스로픽이 Claude Fable 5.1과 Mythos 5.1을 공개했는데, 동일 모델에 안전장치 수준만 다르게 설정한 버전이에요

2 캐시 읽기 가격 인하로 일반 작업은 약 25%, 고강도 에이전틱 작업은 최대 약 45%까지 비용이 줄었어요

3 고객이 직접 통제하는 클라우드에 데이터를 저장하는 EFS를 올가을 도입하고, 사이버보안 오탐도 60% 줄였어요

앤스로픽이 6월 9일 Fable 5를 내놓은 지 약 석 달 만인 9월 1일, 코딩·지식노동용 신모델 Claude Fable 5.1과 사이버보안·생명과학 전용 Claude Mythos 5.1을 함께 공개했어요. 두 모델은 사실 같은 모델이고 안전장치 수준만 다르게 설정한 버전인데, Fable 5.1은 누구나 쓸 수 있도록 일반 공개됐고 Mythos 5.1은 신뢰 접근 프로그램을 통과한 기업·연구자만 쓸 수 있어요. 앤스로픽은 두 모델을 "코딩과 지식노동에서 세계에서 가장 앞선 모델"이라고 소개하면서, 이번 발표에서 가장 눈에 띄는 대목으로 그동안 고객들이 던져온 가격·데이터 보관·안전장치 불만에 정면으로 답했다는 점을 내세웠어요.

풀어서 설명하면, 지난달까지 앤스로픽의 코딩 모델 Fable은 품질은 좋지만 비용이 높아 개발자들이 설계는 Fable에 맡기고 반복 작업은 값싼 모델로 나눠 처리하는 식으로 부담을 쪼개왔고, 접근 통제와 데이터 보관 정책 때문에 일부 기업·국가가 트래픽 경로를 재검토하는 일까지 있었어요. Fable 5.1은 이 세 가지 불만을 각각 겨냥해 나온 업데이트예요.

Claude Fable 5.1과 Claude Mythos 5.1은 같은 기반 모델이에요. 차이는 안전장치의 촘촘함인데, Fable 5.1은 코딩·지식노동·장시간 문제 해결에 쓰도록 일반 공개됐고, Mythos 5.1은 사이버보안과 생명과학 작업을 지원하도록 설계된 별도 안전장치를 얹어 신뢰 접근 프로그램으로만 제공돼요. 앤스로픽은 Terminal-Bench 4.0 점수에서 두 모델 사이에 차이가 나는 이유가 이전의 덜 정교했던 사이버 안전장치가 일부 작업에 개입했기 때문이라고 설명했고, 이번에 안전장치를 개선하면서 두 모델의 성능 격차는 훨씬 줄어들 것으로 본다고 밝혔어요.

구글, 제미나이 3.8 플래시 공개…코딩·에이전트 작업에 초점

제미나이 3.7 플래시를 기반으로 한 신모델, 6개 경로로 배포되고 컨텍스트는 100만 토큰까지 지원해요

1 구글 딥마인드가 2026년 9월 2일 제미나이 3.7 플래시를 기반으로 한 제미나이 3.8 플래시 모델카드를 공개했어요

2 입력은 텍스트·이미지·오디오·비디오를 받고 컨텍스트 윈도우는 최대 100만 토큰, 출력은 최대 6만4천 토큰까지 지원해요

3 제미나이 앱과 구글 AI 스튜디오, 제미나이 API 등 6개 채널로 배포되며 소프트웨어 엔지니어링과 에이전트 작업에 맞춰졌어요

구글 딥마인드가 2026년 9월 2일 제미나이 3.8 플래시(Gemini 3.8 Flash) 모델카드를 공개했어요. 제미나이 3 계열의 최신 반복 모델로, 소프트웨어 엔지니어링과 에이전트형 지식 업무에서 성능을 끌어올린 게 이번 발표의 핵심이에요.

풀어서 설명하면, 제미나이 플래시는 구글이 몇 주 간격으로 소수점 버전을 올리며 다듬는 실험용 라인업이에요.

지난달 12일 테스팅캐탈로그가 제미나이 엔터프라이즈 화면에서 '제미나이 3.6 플래시' 시험 배너를 포착한 적이 있어요. 그로부터 3주가 채 지나지 않은 시점에 구글은 3.8 플래시 모델카드를 정식으로 올렸어요. 카드에는 3.7 플래시 모델카드를 함께 참고하라는 안내가 붙어 있는데, 아키텍처와 학습 데이터, 하드웨어, 소프트웨어를 3.7 플래시 것과 그대로 공유한다고 적혀 있어요.

클로드 Fable 5.1, 재현 안 되는 버그 원인 찾아냈다

차량 데이터·지원 티켓·여러 코드베이스를 한꺼번에 훑어 원인 후보를 제시, 최종 승인은 개발자가 맡아요

1 앤스로픽이 유튜브에 올린 영상에서 클로드 Fable 5.1이 몇 달치 차량 운행 데이터, 고객 지원 티켓, 여러 코드베이스를 동시에 분석해 사내에서도 재현하지 못한 버그의 원인 후보를 찾아냈어요.

2 Fable 5.1은 원인 제안까지 맡고, 실제로 무엇을 출시할지 정하고 수정안을 승인하는 일은 개발자가 그대로 맡는 구조로 시연됐어요.

앤스로픽이 클로드 Fable 5.1의 디버깅 능력을 보여주는 영상을 유튜브 공식 계정에 올렸어요. 영상 제목은 "Debugging across the whole stack with Claude Fable 5.1"인데요, 팀이 사내에서도 재현하지 못했던 버그를 Fable 5.1이 몇 달치 차량 데이터와 고객 지원 티켓, 여러 코드베이스를 한꺼번에 훑어 원인 후보까지 찾아내는 과정을 담았어요. 다만 무엇을 실제로 출시할지 정하고 수정안을 최종 승인하는 일은 여전히 개발자의 몫으로 남아 있어요.

영상 속 사례에서 Fable 5.1이 맡은 문제는 단순한 코드 오류가 아니에요. 팀이 내부적으로 여러 번 시도해도 재현하지 못한 버그였고, 원인을 찾으려면 차량에서 쌓인 몇 달치 운행 데이터와 고객 지원 티켓, 그리고 서로 다른 여러 코드베이스를 동시에 들여다봐야 했어요. 앤스로픽은 영상 설명에서 Fable 5.1이 "장기간 이어지는 엔지니어링 작업에 특히 강하다"고 소개했어요. 증상만 가리는 임시 처방이 아니라 근본 원인을 짚어내는 데 강점이 있다는 설명도 함께 붙었어요.

시연에서 Fable 5.1이 하는 일은 원인 후보를 제안하는 데까지예요. 그 다음은 사람 차례인데요, 개발자가 Fable 5.1이 지목한 원인을 확인하고, 이 수정이 실제로 배포해도 안전한지 판단한 뒤 최종 승인을 내려요. 다시 말해 이번 시연은 AI가 진단을 대신하고 사람이 결정권을 쥐는 구조를 보여준 거예요. 여러 코드베이스와 몇 달치 데이터를 오가는 지루한 추적 작업은 모델이 맡고, 무엇을 출시할지 정하는 판단은 사람이 맡는 분업이 이번 영상의 핵심이에요.

월드랩스 아틀라스, 사진 몇 장으로 3D 세계 생성·복원

페이페이 리의 월드랩스가 생성·복원·로봇 시뮬레이션을 한 모델로 묶은 아틀라스를 내놨어요

오늘 나온 어려운 말 풀이 전문은 metallab.ai/glossary

OSWorld-2.0
OSWorld-2.0
AI 에이전트가 실제 컴퓨터 화면에서 여러 단계의 작업을 스스로 완수하는지 시험하는 평가 도구

제미나이 3.8 플래시
Gemini 3.8 Flash
구글 딥마인드가 2026년 9월 공개한 제미나이 3 계열 경량 모델로, 코딩과 에이전트 작업 성능을 다듬고 가격을 낮춘 버전이다.

클로드 캠퍼스 앰버서더
Claude Campus Ambassadors
앤스로픽이 대학생·대학원생·박사 연구자를 뽑아 각자 학교에서 클로드 활용을 넓히도록 지원금을 주는 캠퍼스 프로그램

챗GPT 포 헬스케어
ChatGPT for Healthcare
병원 전자의무기록과 공식 의학 데이터베이스를 연결해 의료진의 정보 검토를 돕는 오픈AI의 헬스케어 전용 챗GPT.

오늘 해볼 것

오늘 지면에 나온 것 중 지금 바로 해볼 수 있는 4가지

1 제미나이 3.8 플래시

Gemini 3.8 Flash

구글 딥마인드가 2026년 9월 공개한 제미나이 3 계열 경량 모델로, 코딩과 에이전트 작업 성능을 다듬고 가격을 낮춘 버전이다.

이게 뭔데

제미나이 3.8 플래시는 구글 딥마인드가 만든 제미나이 AI 모델 가족 중 가볍고 빠른 축에 속하는 최신 버전이다. 큰 자동차 회사가 신차를 통째로 새로 설계하지 않고, 같은 차체에 엔진 세팅만 다시 조율해 부분 변경 모델을 내놓는 것과 비슷하다. 구글은 3.7 플래시라는 이전 버전의 뼈대—아키텍처와 학습 데이터, 심지어 돌리는 하드웨어까지—를 그대로 물려받아 3.8 플래시를 만들었다고 밝혔다.

이번 튜업의 초점은 코딩과 에이전트 작업, 즉 사람이 시킨 일을 스스로 여러 단계로 나눠 처리하는 업무에서의 성능이다. 받아들이는 입력(글·이미지·소리·영상)과 한 번에 처리할 수 있는 정보량, 답으로 내놓는 글자 수 같은 기본 사양은 3.7 플래시와 똑같다. 즉 몸집을 키우거나 구조를 바꾼 게 아니라, 같은 틀 안에서 실력만 끌어올린 셈이다.

가격도 눈에 띄게 낮췄다. 경쟁 모델보다 훨씬 싼 값을 매겼지만, 이 가격은 한동안만 적용되는 도입가이고 시점이 지나면 두 배로 오른다는 점도 모델카드에 함께 적혀 있다.

해 보려면

제미나이 API를 쓸 수 있다면 아래 프롬프트를 3.7 플래시와 3.8 플래시 양쪽에 똑같이 넣어 결과를 비교해볼 수 있다.

다음 파이썬 함수에 버그가 있다. 원인을 찾고 고친 코드를 설명과 함께 보여줘.

```
def findmax(numbers):
maxnum = 0
for n in numbers:
if n > maxnum:
maxnum = n
return maxnum
```

두 모델의 답변 속도와 설명의 꼼꼼함을 비교하면 이번 업데이트가 실제로 어떤 부분을 다듬었는지 체감할 수 있다.

오늘 지면에서 — 구글, 제미나이 3.8 플래시 공개…코딩·에이전트 작업에 초점

2 클로드 페이블 5.1

Claude Fable 5.1

앤스로픽이 코딩·지식노동용으로 내놓은 클로드 모델의 개선판으로, 캐시 요금을 낮추고 복제 방지 장치를 더했다

이게 뭔데

클로드 페이블 5.1은 앤스로픽이 코딩과 복잡한 지식노동을 돕도록 만든 AI 모델의 업그레이드판이다. 석 달 전 나온 이전 버전을 이어받아, 개발자와 일반 사용자 누구나 쓸 수 있도록 공개됐다. 같은 날 함께 나온 '미토스 5.1'이라는 이름의 모델도 있는데, 사실 이 둘은 속 알맹이가 같은 하나의 모델이다. 다만 사이버보안이나 생명과학처럼 민감한 분야에서 쓰이는 미토스 5.1에는 더 촘촘한 안전장치가 걸려 있고, 이 버전은 검증을 통과한 기업이나 연구자만 쓸 수 있다. 자동차로 치면 같은 엔진을 쓰되 한쪽은 속도제한을 더 세게 걸어둔 셈이다.

이번 업데이트에서 가장 눈에 띄는 변화는 요금이다. 모델이 이전에 처리해 저장해둔 내용을 다시 불러오는 비용을 75%나 낮췄는데, 긴 작업을 반복해서 시키는 경우일수록 절감 폭이 커져 전체 비용이 최대 45%까지 줄어들 수 있다. 다만 새로 글자를 처리하는 기본 요금 자체는 그대로라서, 짧고 단순한 작업에서는 체감 효과가 크지 않을 수 있다.

또 하나 눈여겨볼 부분은 모델의 사고 과정을 몰래 빼내 베끼는 것을 막는 장치가 새로 들어갔다는 점이다. 누군가 수천 개의 가짜 계정을 만들어 모델이 답을 내놓기까지의 생각 과정을 긁어모은 뒤, 안전장치 없는 복제본을 만드는 수법이 알려져 있었는데, 이번 업데이트 이후 만든 계정부터는 이런 방식으로 대화 내용을 손대는 것이 막혔다.

해 보려면

클로드를 쓸 수 있는 계정이 있다면 모델 선택 목록에서 Claude Fable 5.1이 보이는지 확인해 보자. 같은 코딩 질문을 이전 버전과 새 버전에 각각 던져보고 답변 속도나 비용(사용량 표시가 있다면) 차이를 비교해 보면 업데이트를 체감하기 쉽다.

오늘 지면에서 — 앤스로픽 Fable 5.1·Mythos 5.1 출시

3 GLM-5.3

GLM-5.3

즈푸 AI(Zhipu AI)가 2026년 8월 공개한 오픈웨이트 대형 AI 모델로, 코딩과 사이버보안 작업에 강점을 두고 만들어졌다.

이게 뭔데

GLM-5.3은 중국 AI기업 즈푸 AI가 만든 대형 AI 모델이다. 완성된 AI의 두뇌 파일을 통째로 내려받아 누구나 자기 컴퓨터나 서버에서 돌리고 고쳐 쓸 수 있게 공개했다는 점이 특징이며, 특히 코드를 짜거나 보안 취약점을 찾아내는 작업에서 힘을 발휘하도록 다듬어졌다.

이걸 자물쇠 설계도를 통째로 나눠주는 것에 비유할 수 있다. 설계도를 받은 방어자는 값싸게 자기 집의 자물쇠를 점검하고 약점을 미리 고칠 수 있지만, 같은 설계도로 나쁜 의도를 가진 사람이 열쇠를 만들 수도 있다. 그래서 즈푸 AI는 GLM-5.3의 설계도를 한꺼번에 풀지 않고, 보안 파트너의 검증을 먼저 거친 뒤 순차적으로 공개하는 방식을 택했다.

이 순차 공개 방식은 눈에 띄는 대목이다. 원래 오픈웨이트 모델은 발표와 동시에 전부 열어 주는 게 보통인데, GLM-5.3은 크고 이름 있는 AI 연구소들이 위험한 능력을 검증할 때 쓰는 단계적 공개 절차를 따랐다.

해 보려면

GLM-5.3을 API나 이를 제공하는 서비스에서 써볼 수 있다면, 아래처럼 코드 점검용 프롬프트를 넣어 시험해 볼 수 있다.

다음 코드에서 보안 취약점을 찾아 목록으로 정리해 줘. 각 취약점마다 위험도와 고치는 방법도 같이 적어 줘.

OSWorld-2.0은 AI 에이전트에게 진짜 컴퓨터 화면을 주고, 사람처럼 마우스를 클릭하고 키보드를 눌러 일을 끝까지 해내는지 시험하는 시험지 같은 것이다. 예를 들어 브라우저를 열어 정보를 찾고, 그 결과를 스프레드시트에 옮겨 저장하는 식의 여러 단계짜리 작업을 실제 운영체제 화면 위에서 수행하게 한 뒤, 끝까지 제대로 마쳤는지를 채점한다.

일반적인 시험이 정답 하나를 고르는 방식이라면, 이런 시험은 중간에 창을 잘못 닫거나 엉뚱한 메뉴를 눌러도 감점 없이 넘어가는 게 아니라 전체 과정이 실패로 이어질 수 있다는 점에서 훨씬 까다롭다. 화면을 보고 이해하는 능력과 실제로 손을 움직이는 능력을 동시에 요구하기 때문에, AI가 단순히 말을 잘하는 수준을 넘어 실제 업무를 대신할 수 있는지 가늠하는 잣대로 쓰인다.

숫자 2.0이 붙은 건 이전 버전보다 시험 문제나 채점 방식이 더 다듬어졌다는 뜻으로 이해하면 된다. 새로운 AI 모델이 발표될 때마다 이런 시험을 통과한 점수가 함께 공개되는 경우가 많다.

클로드 캠퍼스 앰버서더

Claude Campus Ambassadors

클로드 캠퍼스 앰버서더는 앤스로픽이 전 세계 학생을 뽑아 각자 다니는 학교에서 클로드를 알리고 활용하는 활동을 벌이게 하고, 그 대가로 현금 지원금을 주는 프로그램이에요. 학교 축제 때 동아리마다 지원금을 받아 행사를 여는 것과 비슷한 구조인데, 여기서 동아리 주제가 'AI 활용'인 셈이에요.

트랙은 세 갈래로 나뉘어요. 학부생 트랙은 캠퍼스에 AI 동아리를 직접 꾸려 해커톤이나 발표회를 여는 실행형이고, 대학원생 트랙은 AI가 교육과 사회에 미치는 영향을 놓고 토론 모임을 여는 데 집중해요. 박사·포스닥 트랙은 연구에 AI를 쓰는 과학자들을 모아 워크숍을 조직하는 역할이에요. 어느 트랙이든 선발되면 3,600달러를 받아요.

이 프로그램은 매 학년도 다시 문을 여는 방식으로 운영돼요. 그동안은 학부생만 뽑았는데, 이번에 대학원생과 박사 트랙이 처음 생기면서 파일럿 단계라 아직은 정해진 학교 풀에서만 선발한다는 조건이 붙었어요.

챗GPT 포 헬스케어

ChatGPT for Healthcare

챗GPT 포 헬스케어는 병원과 의료기관이 쓰도록 오픈AI가 내놓은 챗GPT의 업무용 버전이다. 일반 챗GPT가 인터넷에 있는 정보로 답한다면, 이 버전은 병원이 허락한 범위 안에서 환자의 진료 기록과 공식 의학 자료를 직접 불러와 답한다. 비유하자면 개인 비서에게 병원 서류함과 의학 도서관 열쇠를 동시에 쥐여주고, 의사가 묻는 대로 필요한 서류만 골라 책상에 올려달라고 하는 셈이다.

쓰는 방식은 두 가지다. 하나는 챗GPT 화면에서 환자 기록을 불러와 검토하는 방식이고, 다른 하나는 병원이 쓰는 전자의무기록 화면 안에 챗GPT를 심어 화면을 옮기지 않고도 도움을 받는 방식이다. 여기에 임상시험 등록 정보, 약품 라벨, 의학 논문, 보험 급여 기준 같은 공식 의료 데이터베이스 아홉 곳을 한 번에 검색하는 기능도 함께 딸려와, 의사나 연구자, 약제팀이 흩어진 자료를 일일이 찾아다니지 않아도 된다.

헬스케어 퍼블릭 데이터 플러그인

Healthcare Public Data Plugin

헬스케어 퍼블릭 데이터 플러그인은 오픈AI가 챗GPT 포 헬스케어에 붙인 기능으로, 여기저기 흩어진 공식 의료 정보 사이트를 한 번에 뒤져서 답을 찾아주는 도구다. 임상시험 모집 현황을 확인하려면 한 사이트, 약품 허가 정보를 보려면 다른 사이트, 보험 급여 기준을 알아보려면 또 다른 사이트를 열어야 했던 걸, 서랍 아홉 개를 대신 열어 필요한 서류를 찾아주는 비서를 둔 것에 비유할 수 있다.

이 서랍에는 임상시험 등록 정보, 의약품 표준 코드, 약품 최신 라벨과 경고 문구, 의학 논문 검색 결과 등이 들어 있다. 연구팀은 모집 중인 임상시험 조건을 비교하고, 약제팀은 약품 경고 문구를 확인하고, 공중보건팀은 정책을 기획할 때 관련 근거를 한 화면에서 볼 수 있다.

다만 이 플러그인은 누구나 열람 가능한 공식 데이터베이스만 다룬다는 점을 기억해야 한다. 병원의 특정 환자 진료 기록을 불러오는 기능은 별도로 병원이 자체 전자의무기록 시스템을 연결해야 쓸 수 있는 다른 기능이다.

HIPAA 준수 협약

HIPAA Business Associate Agreement

HIPAA 준수 협약(비즈니스 어소시에이트 협약)은 병원이 환자 정보를 다루게 될 외부 회사와 미리 맺어두는 법적 약속이다.

집주인이 이삿짐 업체에 살림을 맡기기 전에 '분실이나 파손이 생기면 책임진다'는 계약서를 받아두는 것과 비슷하다. 병원도 환자의 진료 기록, 검사 결과 같은 민감한 정보를 클라우드 서비스나 AI 챗봇 같은 외부 업체에 넘기려면, 그 업체가 병원과 똑같은 수준으로 정보를 지키겠다고 서약하는 문서에 서명하게 만든다. 이 문서에는 정보를 어떻게 암호화할지, 누가 접근할 수 있는지, 유출 사고가 나면 얼마 안에 알려야 하는지 같은 구체적인 규칙이 담긴다.

미국 법에서는 이 협약이 없으면 병원이 환자 정보를 외부 서비스와 아예 주고받을 수 없다. 그래서 병원 전자의무기록 시스템에 AI 챗봇을 연결하는 서비스가 나오려면, 그 전에 병원과 AI 회사 사이에 이런 협약이 먼저 자리 잡혀 있어야 한다.

엔터프라이즈 프론티어 세이프가드

Enterprise Frontier Safeguards

엔터프라이즈 프론티어 세이프가드는 AI 회사가 사이버보안이나 생명과학처럼 잘못 쓰이면 피해가 큰 분야에 모델을 내줄 때, 아무나 바로 쓰게 하지 않고 신원과 목적을 확인받은 기업·연구자에게만 강화된 안전장치를 씌워 접근을 허용하는 방식이에요. 같은 건물이라도 로비는 누구나 드나들지만 위험 물질을 다루는 실험실 구역은 출입증을 확인받은 사람만 들어가게 나눠 놓은 것과 비슷해요.

앤스로픽이 내놓은 클로드 페이블 5.1과 클로드 미토스 5.1이 이 방식을 보여주는 예예요. 두 모델은 사실 같은 기반 모델인데, 페이블 5.1은 코딩과 지식노동용으로 누구나 쓸 수 있게 일반 공개됐고, 미토스 5.1은 사이버보안·생명과학 작업을 지원하도록 별도의 촘촘한 안전장치를 얹어 신뢰 접근 프로그램을 통과한 곳에만 제공돼요. 안전장치의 촘촘함 차이 때문에 같은 능력이라도 벤치마크 점수가 서로 다르게 나올 수 있어요.

여기에는 모델의 사고 과정을 몰래 빼내 안전장치 없는 복제본을 만드는 수법을 막는 장치도 함께 포함되는 경우가 많아요. 즉 이 세이프가드는 성능을 낮추는 장치가 아니라, 위험한 활용을 막기 위해 누가 어디까지 접근할 수 있는지를 정교하게 나누는 제도예요.

클로드 미토스 5.1

Claude Mythos 5.1

클로드 미토스 5.1은 앤스로픽이 사이버보안과 생명과학처럼 조심스러운 분야의 전문 작업을 돕도록 내놓은 AI 모델이다. 같은 주방에서 나온 요리인데, 일반 손님도 들어갈 수 있는 식당 홀이 있고 자격을 확인받은 손님만 들어가는 안쪽 방이 따로 있다고 생각하면 된다. 실제로 미토스 5.1은 같은 시기에 나온 클로드 페이블 5.1과 뿌리가 같은 하나의 모델이고, 다른 점은 오직 안전장치를 얼마나 촘촘하게 걸었느냐다.

페이블 5.1은 코딩이나 지식노동에 쓰도록 누구나 접근할 수 있게 열려 있지만, 미토스 5.1은 사이버보안·생명과학 작업에 맞춰 별도의 안전장치를 더 얹은 대신, 사전에 신뢰할 수 있다고 검증된 기업이나 연구자만 쓸 수 있다. 이런 분야의 질문은 답변이 악용됐을 때 위험이 크기 때문에 아무에게나 열어두지 않는 것이다.

다만 앞선 버전에서는 이 안전장치가 너무 엄격해서 정상적인 작업까지 방해하는 부작용이 있었는데, 앤스로픽은 이번 5.1에서 안전장치를 다듬어 두 모델 사이의 실력 차이가 줄어들 것으로 본다고 밝혔다.