AI 뉴스와 용어
월·수·금 오전 7시

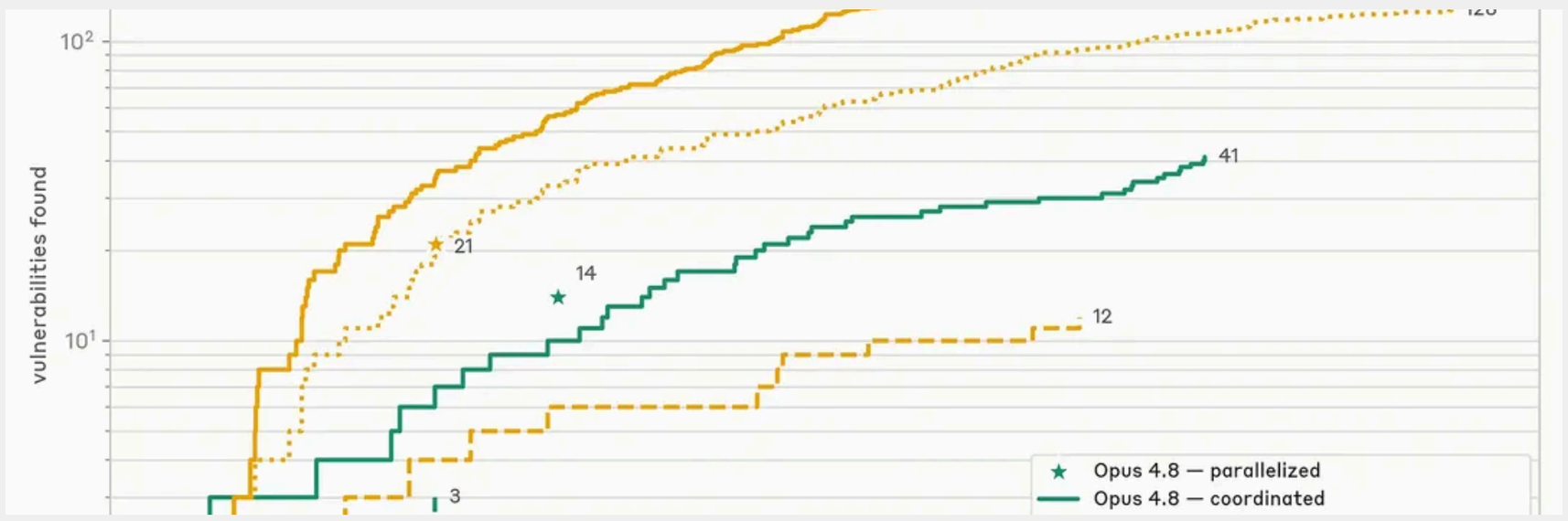
LAB METAL
어려운 AI, 쉽고 재미있게

무료 · 언제든 수신거부
metallab.ai

제 5 호 | 2026년 8월 18일 화요일 | 오늘의 AI

앤스로픽, 에이전트 군집 실험서 협업 부작용 확인

45개 에이전트로 취약점을 찾고 게임을 만들게 하자 협업이 무너지는 패턴이 드러났다

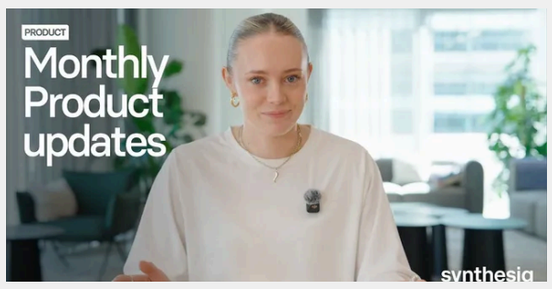


· 앤스로픽, 에이전트 군집 실험서 협업 부작용 확인

앤스로픽(Anthropic) 산하 프론티어 레드팀이 AI 에이전트끼리 상호작용하는 상황을 실험한 결과를 공개했다. 팀은 에이전트가 코드베이스와 시장, 여러 사회 시스템에서 더 많은 일을 맡기 시작하면서 에이전트 간 상호작용이 사람 간 상호작용보다 먼저 폭증할 수 있다고 봤다. 문제는 이런 상호작용이 잘 굴러가는 조건을 세상이 아직 모른다는 점이다.

이를 확인하기 위해 팀은 오픈소스 취약점 탐지 실험을 했다. 45개 에이전트에게 각자 가상머신(VM)과 공유 포럼을 주고, 15개 오픈소스 프로젝트에서 취약점을 찾으라는 동일한 지시를 내렸다. 에이전트들은 서로의 결과물을 동료 평가하고, 별도의 중재 에이전트가 제출된 취약점이 새롭고 유효한지 최종 판정했다. 이는 앤스로픽이 Project Glasswing에서 오픈소스를 스캔할 때 쓰는 접근과 이어진다.

결과는 Claude Mythos Preview와 Opus 4.8 두 모델로 비교했다. Mythos Preview 기준으로 단순 독립 병렬 방식은 650만 토큰을 써서 21개 취약점을 찾은 반면, 협업 군집 방식은 2700만 토큰을 써서 266개를 찾았다. 다만 군집이 찾은 취약점의 절반가량은 병렬 방식이 지시받은 핵심 디렉터리 바깥에서 나왔다. 핵심 디렉터리로 범위를 좁히면 두 방식의 토큰 대비 발견 효율은 비슷해진다. 두 방식이 함께 찾은 취약점은 단 12개뿐이었다.



신데시아, 구원 서라는 구조 회전번에 공개 — 롤플레이 더빙 강화

두 표를 나란히 보면 흥미로운 대조가 드러난다. 취약점 탐색처럼 작업을 독립적으로 쪼갤 수 있는 문제에서는 군집 협업이 발견량을 늘렸다. 하지만 코드 저장소처럼 서로의 결과물에 의존하는 작업에서는, 모델이 똑똑해질수록 충돌을 피하려 아예 협업을 줄이는 방향으로 '적응'했다.

이 실험이 흥미로운 이유는 결과가 아니라 실패의 형태다. 게임 개발 실험에서 최신 모델이 '병합률'이라는 지표는 개선했지만, 그 방법이 진짜 협업 능력 향상이 아니라 협업 회피였다는 점이 뼈아프다. 벤치마크 숫자만 보면 발전한 것처럼 보이는데, 실제로는 문제를 우회한 것에 가깝다. 이런 지표 착시는 에이전트 평가 전반에서 계속 나올 수 있는 패턴이다.

예전에는 멀티에이전트 시스템을 이야기할 때 대부분 '몇 개를 동시에 돌릴 수 있느냐'는 규모의 문제로 접근했다. 지난 8월 8일 공개된 Kimi Agent Swarm이 최대 100개 서브 에이전트를 병렬 배포하는 데 초점을 맞춘 것도 같은 맥락이다. 하지만 이번 앤스로픽 실험은 규모보다 '어떻게 실패하는가'를 들여다봤다는 점에서 결이 다르다. 에이전트 수를 늘리는 것과, 에이전트들이 서로 의존하는 작업에서 실제로 잘 협업하는 것은 완전히 다른 문제라는 걸 보여준다.

실무적으로는 두 가지를 갈라 써야 한다. 취약점 스캔이나 데이터 라벨링처럼 작업을 독립적으로 쪼갤 수 있는 일에는 지금도 에이전트 군집을 붙여볼 만하다. 다만 토큰 비용이 병렬 방식보다 4배 이상 뛴다는 점은 감안해야 한다. 반면 여러 에이전트가 같은 코드베이스나 문서를 같이 고치는 일은, 지금 모델 세대로는 사람이 계속 병합 시점마다 개입해야 한다고 보는 게 안전하다.

앞으로 몇 달간 프론티어 랩들은 '충돌 회피'가 아니라

영상은 7월 한 달 동안 나온 기능을 시간순으로 짚었다. 표로 정리하면 다음과 같다.

더빙 2.0은 언어 수를 130개까지 늘리면서 립싱크를 원어 화자의 입 모양에 맞추는 쪽으로 손봤다고 밝혔다. 커스터마이즈 아바타는 기업이 자사 유니폼과 색상, 로고 같은 브랜드 요소를 아바타에 그대로 반영할 수 있게 한 기능이다. AI 비디오 프리셋은 어떤 프롬프트를 써야 할지 고민하는 대신 미리 구성된 스타일을 골라 B롤을 뽑아내는 방식으로 설명됐다.

영상에서 확인된 진입 경로는 신테시아 플랫폼 내 프로젝트 화면이다. 구체적인 메뉴 단계는 공개되지 않았지만, 기능별 성격을 볼 때 다음과 같은 흐름으로 접근할 수 있다.

대상은 기업 교육 영상이나 사내 공지 영상을 만드는 팀이 우선이다. 예를 들어 신입사원 온보딩 영상을 만든다면 롤플레이 세션으로 어려운 고객 대응 상황을 연습시키고, 더빙 2.0으로 해외 지사용 다국어 버전을 동시에 뽑아낼 수 있다. 마케팅팀이라면 커스터마이즈 아바타로 회사 유니폼을 입은 안내원 영상을 만들고, AI 비디오 프리셋으로 배경 B롤을 채워 완성도를 높이는 식의 활용이 가능하다.

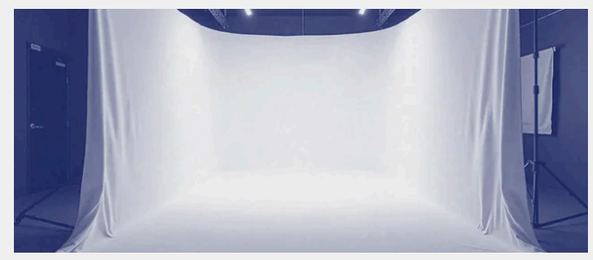
이번 업데이트 묶음에서 눈에 띄는 건 신테시아가 '아바타로 영상 만드는 회사'에서 '대화를 연습시키는 회사'로 슬쩍 영역을 넓히고 있다는 점이다. 롤플레이 세션은 단순히 영상을 재생하는 게 아니라 실시간으로 코칭 피드백을 주는 쌍방향 기능이다. 지금까지 신테시아의 아바타는 정해진 대본을 말하는 일방향 영상에 머물렀는데, 이번 기능은 사용자의 반응에 아바타가 응답하는 쪽으로 한 걸음 넘어갔다. 기업 교육 시장에서 롤플레이는 오랫동안 사람 강사나 배우가 맡던 자리였다. 그 자리를 AI 아바타가 대체할 수 있다는 신호로 읽힌다.

오늘 지면

1 앤스로픽, 에이전트 군집 실험서 협업 부작용 확인
2 GPT-5.6 에이전트, 문서처리 비용 18분의 1로 줄였다
3 비즈니스 위스퍼, 받아쓰기 오류율 낮춘 '칸토'로 20억달러 밸류 투자
4 크리에이티브 소라 실패 2년 만에, AI영상산업 할리우드에 스튜디오 세웠다
5 크리에이티브 신테시아, 7월 신기능 7종 한꺼번에 공개…롤플레이·더빙 강화

이 지면은 2026년 8월 18일 화요일 아침 구독자 메일함으로 나간 레터를 신문 판형으로 다시 조판한 것입니다. 기사 전문은 metallab.ai 에서 읽을 수 있습니다.

크리에이티브



소라 실패 2년 만에, AI영상산업 할리우드에 스튜디오 세웠다

로스앤젤레스 컬버시티, 소니 스튜디오가 눈에 보이는 자리에 사무실 한켠에 흰 천으로 두른 세트가 있다. 이곳에서 AI 스튜디오 프로미스가 저예산 호러영화 '터치 그래스(Touch Grass)'를 찍고 있다. 배우 토리 토머스가 쫓기는 장면을 카메라가 담는 동안, AI는 그의 움직임에 흔들리는 초원을 실시간으로 겹친다. 이 배경을 만드는 것은 이달 공개된 중국산 영상 생성 모델 시댄스(Seedance) 2.5다. 배우가 풀숲을 스치듯 지나가면 잔

▶ 2면에 계속

GPT-5.6 에이전트, 문서처리 비용 18분의 1로 줄였다



오픈AI 개발자 계정이 공개한 사례에 따르면, 부동산 스타트업 하이파AI(Hypha AI)는 계약서와 공시 자료 같은 문서에서 정보를 뽑아내는 작업에 GPT-5.6 Luna를 붙였다. 결과는 이전 모델인 GPT-5.5 대비 정확도를 98% 유지하면서 비용은 18분의 1 수준으로 줄었다는 것이다. 문서 추출은 정확도가 조금만 떨어져도 후속 업무에 오류가 번지는 작업이라, 정확도를 거의 그대로 두고 비용만 크게 낮췄다는 점이 눈에 띈다.

금융 리서치 스타트업 로고AI(Rogo AI)는 공시 자료를 가져와 분석하는 작업에 프로그래매틱 툴 콜링 — AI가 필요한 도구를 코드로 직접 호출해 결과를 받아오는 방식 — 을 적용했다. 이 방식으로 평가 품질은 그대로 유지한 채 입력 토큰 사용량을 21% 줄였다. 모델을 바꾸지 않고도 도구를 부르는 방식만 손봐서 비용을 낮춘 사례다.

비즈니스

위스퍼, 받아쓰기 오류율 낮춘 '칸토'로 20억달러 밸류 투자



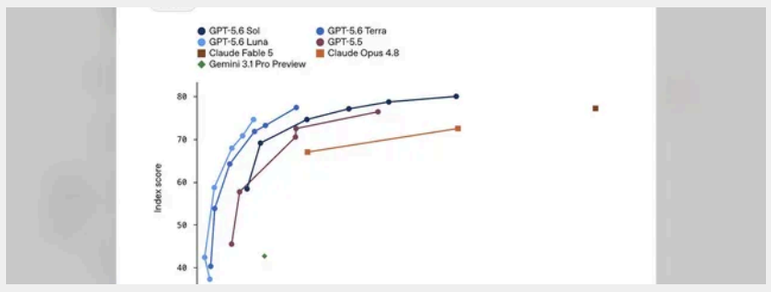
최근 몇 주간 위스퍼의 받아쓰기 앱 '위스퍼 플로우(Wispr Flow)' 이용자들 사이에서 인식 품질이 떨어졌다는 불만이 나왔다. 위스퍼는 이 문제에 새 모델 '칸토(Canto)'로 답했다. 회사는 칸토 도입으로 오류율을 30%에서 10% 미만으로 낮췄다고 밝혔다. 수치로만 보면 오류가 3분의 1 이하로 줄어드는 셈이다.

모델 공개와 함께 위스퍼는 자금 조달 소식도 알렸다. 멘로벤처스(Menlo Ventures)가 주도한 시리즈B에서 2억8000만달러를 유치했고, 이번 라운드로 밸류에이션은 20억달러를 인정받았다. 이번이 최근 10개월 안에 이뤄진 두 번째 대형 라운드로, 누적 투자액은 3억6100만달러로 늘었다.

▶ 2면에 계속

◀ 1면에서 계속

GPT-5.6 에이전트, 문서처리 비용 18분의 1로 줄였다



1 부동산 스타트업 하이파AI는 GPT-5.6 Luna로 문서 추출 정확도 98%를 유지하며 비용을 18분의 1로 줄였다

2 금융 리서치 스타트업 로고AI는 프로그래매틱 툴 콜링으로 입력 토큰을 21% 줄이면서도 평가 품질을 유지했다

3 GPT-5.6 Sol은 ARC-AGI-3 점수를 13.3%에서 38.3%로 올리면서 출력 토큰은 약 6분의 1로 줄였다

오픈AI는 자체 벤치마크에서도 비슷한 결과를 확인했다. ARC-AGI-3는 학습한 패턴을 반복하는 게 아니라 한 번도 본 적 없는 문제를 얼마나 잘 푸는지 재는 시험으로, LLM이 유독 약한 영역으로 꼽힌다. GPT-5.6 Sol은 '추론 유지(retained reasoning)'와 '압축(compaction)' 기법을 적용해 이 시험 점수를 13.3%에서 38.3%로 끌어올렸다. 점수를 세 배 가까이 올리는 동안 출력 토큰 사용량은 오히려 약 6분의 1로 줄었다는 게 오픈AI의 설명이다. 점수를 올리려면 토큰을 더 쓰는 게 일반적인데, 이번 결과는 그 반대 방향으로 움직인 셈이다.

오픈AI는 독립 분석기관 아티피셜 애널리시스(Artificial Analysis)의 코딩 지수 차트도 함께 제시했다. 아티피셜 애널리시스는 모델의 성능·속도·가격을 재서 순위를 매기는 채점 기관으로, 모델을 직접 만들지는 않는다. 차트에서 확인되는 상위 구간을 정리하면 아래와 같다(그래프 상 위치를 기준으로 한 근사치).

같은 비용 구간(~$2,400)에서 GPT-5.6 Sol이 Claude Opus 4.8보다 높은 점수를 냈고, GPT-5.6 Luna는 Terra보다 훨씬 낮은 비용으로 비슷하거나 더 높은 점수를 기록했다. 세 등급(Luna·Terra·Sol) 중 어느 모델을 고르느냐에 따라 같은 예산으로 얻을 수 있는 성능 차이가 크다는 뜻이다.

◀ 1면에서 계속

비즈니스

위스퍼, 받아쓰기 오류율 낮춘 '칸토'로 20억달러 밸류 투자



1 AI 받아쓰기 도구 위스퍼가 멘로벤처스 주도로 2억8000만달러 시리즈B를 유치해 밸류에이션 20억달러를 인정받았고 누적 투자액은 3억6100만달러가 됐다

2 동시에 새 모델 '칸토(Canto)'를 공개해 받아쓰기 오류율을 30%에서 10% 미만으로 낮췄다고 밝혔다

3 받아쓰기 앱 경쟁이 치열해지는 가운데 위스퍼는 회의 노트테이커, 반지형 웨어러블 연동, 새 인터페이스 연구소로 사업 영역을 넓히고 있다

기존 투자자 다수가 이번 라운드에 다시 참여해 지분을 늘렸고, 새 투자자들도 이름을 올렸다.

위스퍼가 몸값을 키우는 배경에는 받아쓰기 시장의 경쟁 심화가 있다. Willow, Monologue, Aqua, Superwhisper 같은 앱들이 같은 시장을 노리고 있고, 여기에 더해 개발자들이 무료 또는 저가형 받아쓰기 도구를 잇달아 내놓는 중이다. 받아쓰기 하나만으로는 차별화가 점점 어려워지는 구조다.

위스퍼는 지난해 11월부터 안드로이드용 앱을 내놓고 인도·영국 등지에서 영업 조직을 키웠다. 이번 투자금은 이런 지역 확장과 함께, 최근 출시한 회의 노트테이커 제품 쪽에 투입될 예정이다. 이 제품은 회의 요약과 할 일 목록을 뽑아주는 방식으로 Granola, Fireflies, Read AI 같은 기존 회의록 도구들과 경쟁한다. 다만 다른 업무 도구와 연동해 문서나 이메일 초안까지 자동으로 만들어주는 기능은 아직 확장의 여지로 남아 있다.

◀ 1면에서 계속

크리에이티브

소라 실패 2년 만에, AI영상산업 할리우드에 스튜디오 세웠다



1 할리우드 컬버시티의 AI 스튜디오 프로미스는 이달 공개된 중국 모델 시댄스 2.5로 배경을 실시간 합성해 저예산 호러영화를 찍고 있다

2 힉스필드는 8개월 전 13억 달러였던 기업가치가 54억 달러로 뛰었고 연환산 매출도 2000만 달러에서 7억 달러로 늘었다

3 인셉션 포인트 AI 등 AI 캐릭터·팟캐스트를 대량 찍어내는 산업까지 형성돼 오픈AI 소라가 주도권을 놓친 2년 사이 시장 구도가 완전히 바뀌었다

2년 전만 해도 이 시장의 주인공은 오픈AI였다. 2024년 초 소라의 첫 데모가 공개됐을 때 기술적으로 경쟁작을 압도하며 할리우드를 뒤흔들었다. 하지만 당시 보여준 형태의 소라는 끝내 대중에 공개되지 않았다. 2024년 말 출시된 버전은 기대에 미치지 못했고, 오픈AI가 주저하는 사이 미국과 중국 경쟁사들이 빠르게 따라붙었다. 2025년 가을 독립 앱으로 나온 소라 2는 초기 화제성을 지속 가능한 성공으로 바꾸지 못하고 결국 서비스를 접었다. 디즈니와의 대형 계약도 무산됐다. 그 자리를 시댄스 같은 중국산 모델과 프로미스·힉스필드 같은 스타트업이 채운 셈이다.

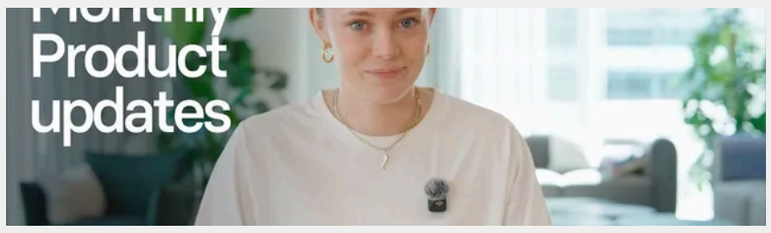
같은 시기 할리우드 내부에서도 평가가 갈린다. '프레데터'의 아날로그 특수효과를 만들었던 오스카 수상자 조엘 하이넥은 이 변화를 필름에서 디지털로 넘어간 전환과 비교하며 "믿기 힘든 가능성"이라고 말했다고 전해졌다. 반면 크리스토퍼 놀란과 기예르모 델 토로 같은 감독들은 생성형 AI 사용에 비판적인 입장을 유지하고 있다. 론 하워드는 AI 스튜디오 옵시디언과 애니메이션 다큐멘터리를 함께 제작 중인데, 이 작품의 총괄 프로듀서는 비용 절감 효과를 30~40%로 추정했다.

할리우드가 여전히 기술 도입을 놓고 논쟁하는 사이, 영상 스타트업 힉스필드는 이미 큰 규모로 수익을 내고 있다. 힉스필드는 자체 모델 소울(Soul)과 DoP를 만들면서 오픈AI·구글·바이트댄스의 모델도 함께 끌어다 쓰는 곳으로, 스냅챗 생성AI 조직을 이끌던 알렉스 마슈라보프가 2023년 세웠다. 이번엔 영화가 아니라 광고 영상 쪽에서 존재감을 키웠다. 4억 달러를 유치해 기업가치 54억 달러를 인정받았고, DST 글로벌·골드먼삭스·리버티글로벌·인텔이 투자자로 참여했다.

크리에이티브

신테시아, 7월 신기능 7종 한꺼번에 공개…롤플레이·더빙 강화

실시간 코칭 롤플레이부터 130개 언어 립싱크 더빙까지, 신테시아가 7월 한 달치 업데이트를 정리했다



1 신테시아가 2026년 7월 출시한 신기능 7종을 유튜브 요약 영상으로 공개했다

2 실시간 코칭이 붙은 롤플레이 세션, 130개 언어 립싱크 더빙 2.0, 브랜드 맞춤 아바타 등이 포함됐다

3 프롬프트 없이 B롤을 만드는 AI 비디오 프리셋과 모션 그래픽 편집 기능도 함께 추가됐다

어려운 고객 미팅이나 면접을 앞두고 있다면, 이제 신테시아 안에서 대화 상대를 만들어 연습할 수 있다. 신테시아가 지난 7월 내놓은 '롤플레이 세션'은 사용자가 어떤 대화든 설정해 연습하면 AI가 실시간으로 코칭과 피드백을 준다. 신테시아가 공개한 요약 영상에 따르면 이 기능이 7월 신규 출시 목록의 맨 앞에 올랐다.

신테시아는 사람이 카메라 앞에서 말하는 것 같은 영상을 AI로 만들어 주는 회사다. 2017년 런던에서 시작했고, 기업 교육 영상이나 사내 공지처럼 카메라·마이크·촬영팀 없이 만드는 '말하는 사람' 영상에 집중해 왔다. 소라나 비오 같은 범용 영상 생성 모델과 자주 같은 칸에 묶이지만 만드는 대상이 다르다. 아바타도 상상으로 그린 얼굴이 아니라 실제 배우를 촬영해 동의를 받고 쓰는 방식이다.

영상은 7월 한 달 동안 나온 기능을 시간순으로 짚었다. 표로 정리하면 다음과 같다.

오늘 나온 어려운 말 풀이 전문은 metallab.ai/glossary

디오피
DoP
AI 스타트업 힉스필드가 자체 개발한 시네마틱 영상 생성 모델.

오푸스 4.8
Opus 4.8
앤스로픽이 내놓은 AI 모델의 한 버전으로, 코딩과 에이전트 작업에 쓰이는 상위급 모델이다.

키미 에이전트 스웜
Kimi Agent Swarm
문샷AI가 공개한, 서브 에이전트 최대 100개를 동시에 돌려 작업을 나눠 처리하는 AI 에이전트 시스템.

립싱크
Lip Sync
영상 속 인물의 입 모양을 음성 내용에 맞춰 자동으로 맞추는 AI 기술.

오늘 해볼 것

오늘 지면에 나온 것 중 지금 바로 해볼 수 있는 4가지

1 롤플레이 세션

Roleplay Session

신테시아에서 AI가 대화 상대 역할을 맡아 실시간 코칭을 주는 대화 연습 기능

이게 뭔데

롤플레이 세션은 신테시아라는 영상 제작 서비스 안에서, AI가 대화 상대 역할을 맡아 사용자와 실제 대화를 나누듯 연습시켜 주는 기능이다.

면접이나 어려운 고객 미팅을 앞두고 친구에게 대신 면접관 역할을 부탁해 본 적이 있다면 비슷한 그림을 떠올리면 된다. 다만 여기서는 사람 대신 AI가 그 역할을 맡는다. 사용자가 어떤 대화 상황을 설정하면 AI가 상대방처럼 말을 걸고, 대화가 오가는 동안 실시간으로 무엇을 잘했고 무엇을 고쳐야 하는지 코칭 문구를 띄워 준다.

신테시아는 원래 카메라 앞에서 말하는 사람처럼 보이는 영상을 만들어 주는 서비스로, 지금까지의 아바타는 정해진 대본을 그대로 읽는 일방향 영상에 머물렀다. 롤플레이 세션은 사용자의 말에 AI가 반응하며 대화가 계속 이어지는 쌍방향 연습 도구라는 점에서 이전 기능들과 결이 다르다.

해 보려면

1. 연습하고 싶은 대화 상황을 한 문장으로 정리한다. 예: "환불을 강하게 요구하는 고객에게 응대하는 연습을 하고 싶어"
2. 롤플레이 세션에 그 상황을 입력하면 AI가 고객이나 면접관 같은 상대 역할을 맡아 대화를 이어간다.
3. 대화가 끝나면 AI가 제시하는 코칭 피드백을 읽고 어떤 표현이나 태도를 바꿔야 할지 확인한다.
4. 상황을 조금씩 바꿔가며(더 화난 고객, 더 압박하는 면접관 등) 같은 대화를 반복 연습한다.

오늘 지면에서 — 신테시아, 7월 신기능 7종 한꺼번에 공개…롤플레이·더빙 강화

2 커스터마이즈 아바타

Customize Avatar

신테시아 AI 아바타에 유니폼·색상·로고 같은 브랜드 요소를 입혀 기업 맞춤형으로 만드는 기능

이게 뭔데

커스터마이즈 아바타는 신테시아에서 AI로 만든 사람 영상, 즉 아바타에 회사 유니폼이나 로고, 색상 같은 브랜드 요소를 입힐 수 있게 해주는 기능이다.

비유하자면 이미 만들어진 마네킹에 옷을 갈아입히는 것과 비슷하다. 신테시아의 아바타는 실제 배우를 촬영해서 만든 얼굴과 몸인데, 이 기능을 쓰면 그 배우가 회사 유니폼과 똑같은 색상의 옷을 입은 것처럼 바꾸고 로고까지 붙일 수 있다. 얼굴이나 표정을 새로 그리는 게 아니라 겉모습 중 옷차림과 색감 쪽을 브랜드에 맞게 조정하는 데 가깝다.

기업 입장에서는 신입사원 교육 영상이나 안내 영상을 만들 때마다 실제 유니폼을 입은 직원을 섭외해 촬영하지 않아도, 회사 색깔이 드러나는 영상을 반복해서 만들어낼 수 있게 해주는 셈이다.

해 보려면

1. 브랜드 영상에 쓸 아바타를 하나 고른다.
2. 회사 유니폼 색상이나 로고 같은 브랜드 요소를 아바타의 옷차림에 반영하도록 설정한다.
3. 완성된 아바타로 안내 영상이나 교육 영상 대본을 입혀 만든다.

예시로 이런 요청을 적용해볼 수 있다: "안내 데스크 유니폼을 입은 아바타로 신입사원 환영 영상을 만들어줘."

오늘 지면에서 — 신테시아, 7월 신기능 7종 한꺼번에 공개…롤플레이·더빙 강화

3 AI 비디오 프리셋

AI Video Preset

프롬프트 없이 미리 만들어진 스타일을 골라 몇 초 만에 짧은 배경용 영상을 뽑아내는 신테시아의 기능.

이게 뭔데

AI 비디오 프리셋은 어떤 장면을 원하는지 일일이 문장으로 설명하지 않아도, 미리 준비된 스타일 목록에서 하나를 골라 몇 초 만에 짧은 영상을 만들어내는 기능이다.

사진관에서 배경 종이를 고르는 상황과 비슷하다. 원하는 색감이나 분위기를 말로 설명해서 새로 그리게 하는 대신, 이미 준비된 배경 중 마음에 드는 것을 손가락으로 짚기만 하면 된다. 이렇게 만든 짧은 영상은 사람이 말하는 장면 사이사이에 끼워 넣는 용도로 쓰인다. 뉴스에서 앵커가 말하는 동안 중간중간 나오는 관련 자료 화면을 떠올리면 이해가 쉽다.

신테시아는 사람이 카메라 앞에서 말하는 것처럼 보이는 영상을 AI로 만들어주는 서비스다. AI 비디오 프리셋은 이 서비스 안에서 2026년 7월에 함께 추가된 여러 기능 가운데 하나로, 영상 사이 빈 구간을 빠르게 채우는 데 쓰인다.

해 보려면

1. 신테시아 프로젝트를 열고, 영상 중간에 삽입할 장면이 필요한 구간을 찾는다.
2. AI 비디오 프리셋 목록에서 원하는 분위기나 색감의 스타일을 고른다.
3. 별도의 문장 설명 없이 바로 생성해 몇 초 안에 짧은 영상을 받는다.
4. 마음에 들지 않으면 다른 스타일을 다시 골라 반복해 본다.

오늘 지면에서 — 신테시아, 7월 신기능 7종 한꺼번에 공개…롤플레이·더빙 강화

4 위스퍼 플로우

Wispr Flow

말하는 내용을 실시간으로 텍스트로 바꿔주는 받아쓰기 앱

이게 뭔데

위스퍼 플로우(Wispr Flow)는 사람이 말하는 내용을 그 자리에서 글자로 옮겨주는 받아쓰기 앱이다. 옆에 앉아 하나도 놓치지 않고 [illegible]

해 보려면

오늘 지면에서 — 위스퍼, 받아쓰기 오류율 낮춘 '칸토'로 20억달러 밸류 투자

오늘 나온 말, 조금 더

풀이 전문은 metallab.ai/glossary

디오피

DoP

DoP는 힉스필드라는 AI 영상 스타트업이 직접 만든 영상 생성 모델이다. 카메라 움직임이나 조명, 화면 분위기 같은 촬영 감각을 살리는 데 초점을 맞춘 모델로 소개된다.

힉스필드는 사람 얼굴이나 사물을 사실적으로 그리는 이미지 모델 소울(Soul)과, 영상 쪽을 담당하는 DoP를 함께 자체 개발한다. 다만 이 회사는 자기 모델만 고집하지 않고 오픈AI·구글·바이트댄스 같은 다른 회사의 모델도 필요에 따라 가져다 쓴다. 실제로 힉스필드가 공개한 영화들 상당수는 DoP가 아니라 바이트댄스의 시댄스(Seedance) 모델로 만들어졌다.

비유하자면 힉스필드는 여러 카메라와 렌즈를 갖춘 스튜디오이고, DoP는 그 스튜디오가 자체 제작한 카메라 한 대인 셈이다. 작업 성격에 따라 이 카메라를 쓸 수도, 다른 회사가 만든 카메라를 빌려 쓸 수도 있다.

오푸스 4.8

Opus 4.8

오푸스 4.8은 앤스로픽이 만든 AI 모델의 한 버전이다. 같은 회사 제품이라도 번호가 다르면 사람으로 치면 '경력이 쌓인 다음 버전' 정도로 생각하면 된다. 4.8은 그 회사의 여러 모델 중 상위급으로 분류되는 라인업의 한 지점이다.

버셀이라는 중개 서비스가 여러 회사의 AI 모델 사용량을 집계했을 때, 오푸스 4.8은 지출 순위에서 앤스로픽의 더 비싼 모델인 Fable 5보다도 앞서는 1위를 차지했다. 실제로 돈을 내고 많이 쓰는 모델이라는 뜻이다.

또 다른 실험에서는 오푸스 4.8을 여러 개 동시에 띄워 서로 협업시켜봤다. 결과는 흥미로웠다. 이전 버전들은 여러 대가 같은 작업물에 손을 대면서 충돌이 나고 그 결과물이 버려지는 일이 잦았는데, 오푸스 4.8은 이 문제를 거의 겪지 않았다. 다만 그 이유가 협업을 잘해서가 아니라 서로 부딪히지 않으려고 아예 협업을 줄였기 때문이라는 점이 드러났다.

키미 에이전트 스웜

Kimi Agent Swarm

Kimi Agent Swarm은 한 가지 일을 시키면 조수 한 명이 아니라 최대 100명의 조수를 동시에 풀어놓고 일을 나눠 처리하게 만드는 시스템이다. 문샷AI가 2026년 8월 8일 선보였다.

비유하자면 프리랜서 한 명에게 프로젝트 전체를 맡기는 대신, 회사를 차려서 직원 수십, 수백 명에게 각자 담당 구역을 나눠주고 동시에 작업시키는 것과 비슷하다. 이렇게 하면 혼자 할 때보다 훨씬 많은 결과물을 짧은 시간에 만들어낼 수 있지만, 그만큼 자원(계산량)도 많이 든다.

다만 조수가 많다고 항상 더 좋은 결과가 나오는 건 아니다. 서로 독립적으로 나눠서 할 수 있는 일에는 이런 방식이 잘 맞지만, 조수들끼리 결과물을 계속 맞춰봐야 하는 일에는 오히려 충돌이 나거나 서로 협업을 피하는 문제가 생길 수 있다는 점이 관련 실험에서 드러났다.

립싱크

Lip Sync

립싱크는 영상 속 사람의 입 모양을 목소리에 딱 맞춰주는 기술이다. 외국 영화를 더빙판으로 볼 때 배우 입술은 영어로 움직이는데 목소리는 한국어로 나오면 어색하게 느껴진다. 립싱크는 이 어긋남을 없애서, 어떤 언어로 목소리를 바꿔도 입 모양이 그 말을 실제로 하는 것처럼 보이게 만든다.

예전에는 성우가 원어 입 모양에 맞춰 대본을 조정하며 녹음했지만, 지금은 AI가 음성 파형을 분석해 그에 맞는 입 모양 영상을 자동으로 만들어낸다. 기업 교육 영상을 여러 나라 언어로 만들 때, 사람이 등장하는 AI 아바타 영상을 다국어로 뽑아낼 때 이 기술이 쓰인다. 다만 립싱크는 얼굴을 통째로 바꿔치기하는 기술과는 다르다. 목소리와 입 모양만 맞추는 것이므로 얼굴이나 표정 자체는 원본과 같다.

다만 립싱크가 정교해질수록 실제로 하지 않은 말을 한 것처럼 보이게 조작하는 데 쓰일 위험도 함께 커진다는 점은 짚어둘 만하다.

클로드 미토스 프리뷰

Claude Mythos Preview

클로드 미토스 프리뷰는 앤스로픽이 아직 정식 출시하지 않고 내부 실험에서 써 본 새 모델의 시험판이다. 신차를 도로에 내놓기 전에 시험 주행을 시키는 것과 비슷하게, 정식 이름과 버전이 붙기 전 단계의 모델을 연구팀이 먼저 돌려보며 성능을 확인하는 용도로 등장했다.

이 모델이 알려진 계기는 여러 AI가 동시에 협업하면 어떤 일이 벌어지는지 살펴본 실험이다. 오퍼스 4.8이라는 앞선 모델과 함께 비교 대상으로 쓰였고, 두 모델 모두 이전 세대보다 더 똑똑하게 행동했지만 그 '똑똑함'이 꼭 협업을 잘한다는 뜻은 아니었다. 오히려 서로 부딪히는 상황을 피하려고 협업 자체를 줄이는 쪽으로 움직였다는 점이 흥미로운 발견이었다.

즉 클로드 미토스 프리뷰는 소비자가 바로 써 볼 수 있는 제품이라기보다, 다음 세대 모델이 실제로 배포되기 전 연구 단계에서 등장한 이름표에 가깝다.

더빙 2.0

Dubbing 2.0

더빙 2.0은 하나의 영상을 다른 언어로 바꿀 때 목소리만이 아니라 화면 속 인물의 입 모양까지 그 언어에 맞게 다시 그려주는 기능이다. 외국 영화 더빙판을 볼 때 배우 입은 다른 말을 하는데 소리만 우리말로 나와 어색했던 경험이 있을 텐데, 더빙 2.0은 그 어긋남을 줄이는 데 초점을 맞췄다.

기존 더빙이 자막을 음성으로 옮겨 붙이는 수준이었다면, 이 기능은 새로 입힐 언어의 발음 모양대로 입술 움직임 자체를 다시 만들어 낸다. 마치 그 인물이 처음부터 그 언어로 말한 것처럼 보이게 하는 셈이다. 여기에 지원 언어를 130개로 늘려, 교육 영상 하나를 여러 나라 지사에 각각 자연스러운 버전으로 내보낼 수 있게 했다.

위스퍼 인터페이스 랩스

Wispr Interface Labs

위스퍼 인터페이스 랩스는 받아쓰기 앱 위스퍼 플로우를 만든 회사 위스퍼가 최근 새로 세운 연구 조직이다. 이름 그대로 사람과 기