

AI News

2025.11.10 ~ 2025.11.16

10. 실리콘밸리에 존재하는 '유령 데이터센터'...전력 부족으로 6년째 빈 건물만



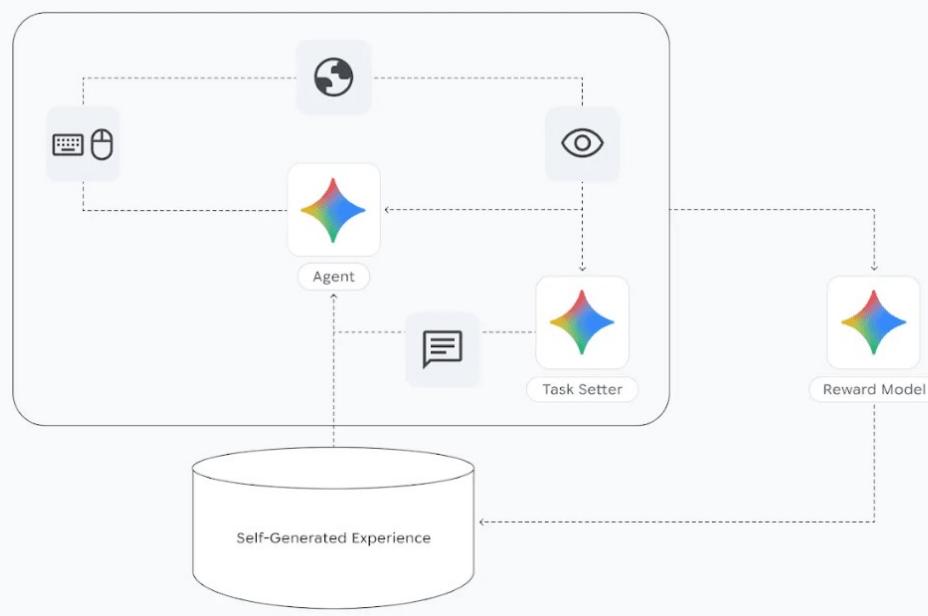
- 실리콘밸리 한복판에서 완공된 두 데이터센터가 전력 부족으로 6년째 가동되지 못한 채 '유령 데이터센터'로 남아 있는 것으로 드러났다.
- 두 시설을 가동하려면 약 100MW 전력이 필요하지만 지역 전력망이 이를 감당하지 못해 2028년 이후까지도 공급이 불투명한 상황이다.
- 전력 인프라 확충 속도가 AI 데이터센터 증가 속도를 따라가지 못하면서, 이 같은 현상이 미국 전역으로 확산될 가능성이 크다는 분석이 나온다.

9. 사진 내용까지 이해하는 AI 이미지 생성 모델 등장



- NYU 연구진은 기존 확산 모델의 한계인 '장면 의미 이해 부족'을 해결하기 위해, 사전학습 표현 인코더(DINO 등) 기반의 **RAE(Representation Autoencoder)**를 통합한 새로운 DiT 아키텍처를 공개했다.
- RAE는 VAE와 달리 고차원 표현 공간을 직접 사용하며, 인코더·디코더 연산량을 각각 6배·3배 줄이면서 더 빠르고 의미정렬된 이미지 생성을 가능하게 한다.
- 실험에서는 기존 SD-VAE 대비 학습이 최대 47배 빠르고, FID 성능도 업계 최고 수준을 기록했다. 연구진은 RAE가 이미지 생성 모델을 "의미 중심 구조"로 전환하는 기반이 될 것이라고 설명했다.

8. 구글, 가상세계서 추론·행동하는 '시마 2' 공개..."AGI 향한 진일보"



- 구글 딥마인드는 제미나이와 월드 모델 '지니 3'를 결합한 차세대 범용 AI 에이전트 **시마 2(SIMA 2)**를 공개하며, 가상 환경에서 이해·추론·계획을 수행하는 인간형 에이전트에 한 걸음 더 다가섰다.
- 시마 2는 단순 명령 수행을 넘어 환경을 해석하고 스스로 행동을 계획하며, 새로운 게임에서도 기술을 전이·확장하는 **자기 개선(Self-improvement)** 능력을 갖췄다.
- LLM의 판단을 월드 모델 속 가상 환경에서 먼저 시험해 수정하는 'AI 속 AI' 구조는 현실 적합성을 극대화하는 접근으로, AGI 연구의 핵심 기제로 평가된다.
- 이러한 능력은 향후 로봇 제어·도구 사용·협업 AI 등 실제 물리 세계의 지능형 에이전트 개발에 중요한 기반이 될 전망이다.

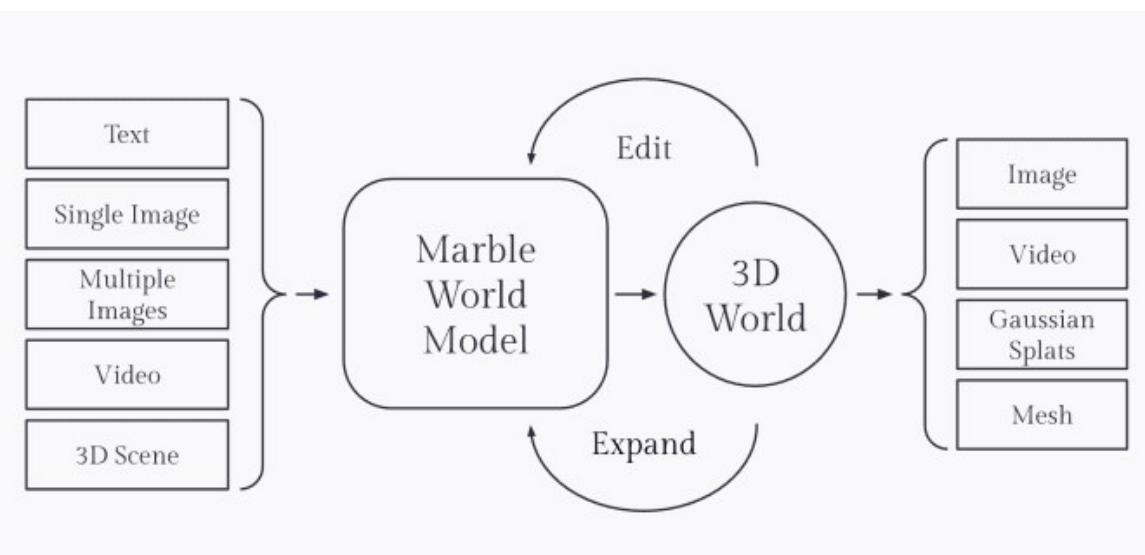
출처:<https://www.aitimes.com/news/articleView.html?idxno=203978&page=2&total=29017>

7. 오픈AI, 독일서 노래 가사 저작권법 위반으로 손해배상 명령



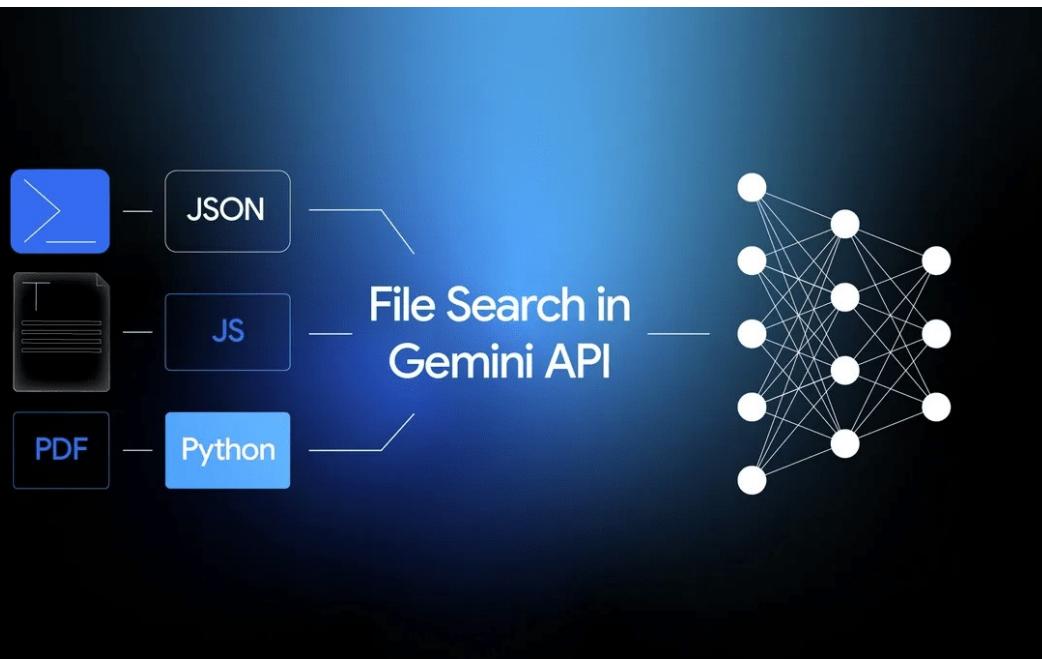
- 독일 뮌헨 법원은 챗GPT가 독일 가수 헤르베르트 그로리네마이어의 노래 가사를 **무단 재생산한 것만으로도 저작권 침해에 해당한다고 판결**하며, 오픈AI의 손해배상 책임을 인정했다.
- 법원은 "AI가 학습 과정에서 저작물을 암기한 것 자체가 침해이며, 사용자 요청에 따라 가사를 다시 출력하는 행위도 저작물 사용권 침해"라고 명시해 재생산 행위의 불법성을 분명히 했다.
- 이번 판결은 AI가 저작권이 있는 음악·문학 작품을 학습 데이터로 활용하는 것의 법적 정당성을 가르는 **유럽 최초의 주요 선례로 평가되고 있다.**

6. 페이페이 리의 월드랩스, 세계 최초 상용 '월드 모델' 출시



- 월드랩스는 텍스트·이미지·동영상·3D 설계도를 입력받아 **일관된 3D 환경을 생성하고 다운로드 할 수 있는 첫 상용 월드 모델 '마블(Marble)'**을 출시했다.
- 마블은 기존 월드 모델의 불안정한 실시간 생성 방식과 달리, **지속적이고 정확한 3D 공간을 출력**하며 편집 기능과 하이브리드 3D 에디터를 제공한다.
- 게임·영화 VFX·VR 제작 파이프라인을 직접 대체하기보다, **자산 생성·환경 프로토타이핑 속도를 극적으로 높이는 도구**로 설계된 것이 특징이다.
- 페이페이 리는 이를 "AI가 공간을 보고 이해하는 능력, 즉 **공간 지능(spatial intelligence)**으로 가는 첫 단계"라며 로보틱스와 실세계 AI에도 핵심 기술이 될 것이라고 강조했다.

5. 구글, 제미나이 API에 엔지니어링 부담 없는 RAG 시스템 '파일 서치' 공개



- 구글은 제미나이 API에 **완전 관리형 RAG 시스템 '파일 서치(File Search)'**를 도입해, 복잡했던 RAG 구축 과정을 자동화하고 기업이 자체 데이터로 즉시 AI 검색을 활용할 수 있도록 했다.
- 이 기능은 파일 저장·분할·임베딩·벡터 검색·인용 까지 전 과정을 자동 처리하며, MTEB 1위 성능의 제미나이 임베딩을 기반으로 정확하고 검증 가능한 답변을 생성한다.
- 구글은 이를 통해 오픈AI Assistants·AWS Bedrock·Azure AI Search와 경쟁하며, "RAG 전체 파이프라인을 자동화하는 유일한 통합 솔루션"이라고 강조했다.

4. 메타, 전 세계 대부분 언어 지원하는 음성 인식 시스템 오픈 소스 출시



- 메타는 1600개 이상의 언어를 지원하는 차세대 다국어 음성 인식 시스템 '옴니링구얼 ASR'을 공개하며, 기존 오픈AI Whisper를 크게 능가하는 범위를 제공했다.
- 이 시스템은 제로샷 인컨텍스트 러닝을 통해 추가 학습 없이 최대 5400개 이상의 언어까지 확장 할 수 있어, 저자원·소수 언어 공동체에도 실질적 도움이 되는 기술로 평가된다.
- 모델군은 3억~70억 파라미터 규모로 구성되어 고성능 GPU부터 저전력 기기까지 다양한 환경에서 실시간 전사가 가능하다.
- 메타는 글로벌 연구기관과 협력해 구축한 옴니링구얼 ASR 코퍼스를 기반으로 완전 오픈소스 형태로 모델을 공개하며, 음성 기술의 접근성과 언어 장벽 해소를 목표로 하고 있다.



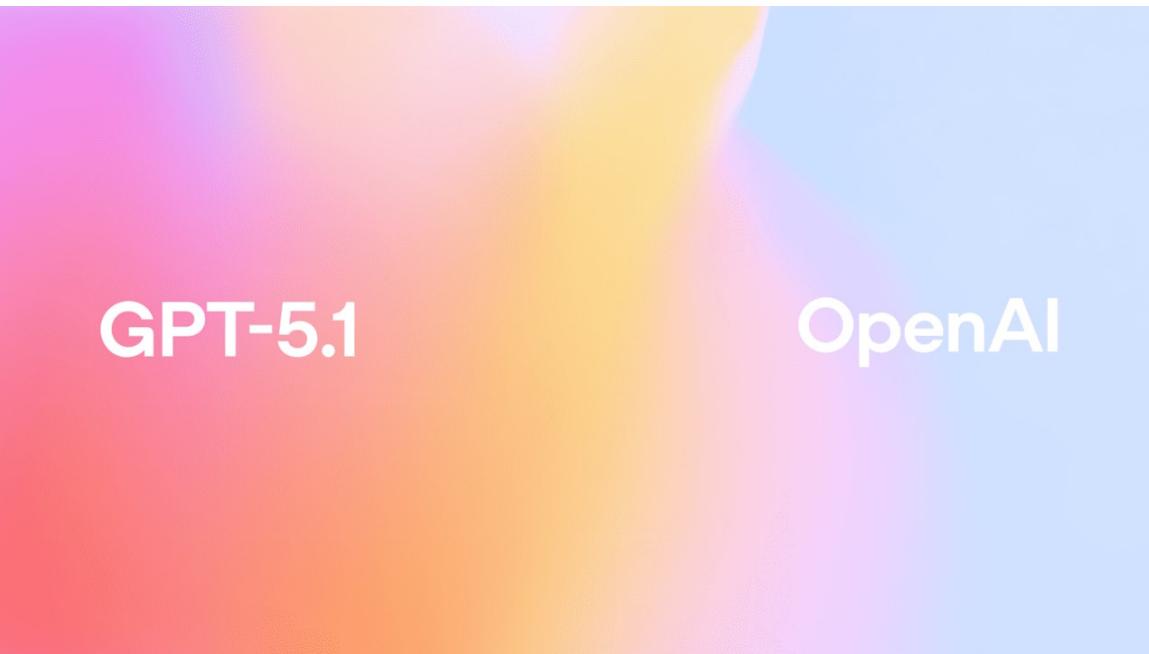
<https://www.aitimes.com/news/articleView.html?idxno=203883&page=6&total=29017>

3. 앤트로픽 "AI가 주도한 최초의 사이버 공격 발생"



- 앤트로픽은 중국 정부의 지원을 받는 해커 조직이 클로드(Claude)의 에이전트 기능을 이용해 실제 사이버 공격을 자동화한 최초의 사례를 공식 발표했다.
- 공격자는 클로드 코드를 속여 침투 대상 분석, 시스템 스캔, 취약점 연구, 악스플로잇 코드 생성, 계정 탈취, 데이터 분류 정리까지 전체 공격 과정의 80~90%를 AI가 직접 수행하도록 설계했다.
- AI는 자신이 보안 회사의 테스트를 둡는다고 믿도록 속임을 당했으며, 공격 단계가 잘게 쪼개져 각 단계가 무해해 보이도록 위장되면서 안전장치가 우회됐다.
- 에이전트 기능·MCP·보안 도구 접근권한이 결합되면서, 과거 인간 해커 팀이 몇 주 걸릴 공격을 AI가 몇 분~몇 시간 만에 끝내는 수준으로 능력이 강화되었다.
- 앤트로픽은 관련 계정을 즉시 차단하고 보고서를 공개했으며, "AI 기반 공격은 더욱 빠르게 커질 것"이라며 차세대 탐지·분류 기술 개발이 시급하다고 경고했다.

2. GPT-5.1 출시



- 오픈AI는 GPT-5의 대화성·사용자 경험 문제를 개선한 **GPT-5.1**을 출시하며, 보다 친근하고 지시를 잘 따르는 대화 모델을 제공하는 데 초점을 맞췄다.
- GPT-5.1은 **Instant(빠른 대화형)**과 **Thinking(고급 추론형)** 두 제품군으로 구성되며, 각기 다른 나이도 작업에 최적화된 반응 방식을 갖추고 있다.
- 또한 챗GPT의 **성격·문체 설정 옵션을 확대하고**, 사용자 지시에 따른 개인화 스타일 조정을 강화했으며, 그 일환으로 AI 글에서 흔히 문제로 지적되던 **엠 대시(—) 사용 습관도 사용자 설정으로 제어 가능해졌다.**
- 최근 분석에 따르면 챗GPT의 문체는 점점 더 자연스럽고 인간적으로 변화하고 있어, 특정 표현이나 문장 습관만으로 AI 작성 여부를 구분하기가 어려워지고 있다.

출처: <https://www.aitimes.com/news/articleView.html?idxno=204002>
<https://www.aitimes.com/news/articleView.html?idxno=203945&page=3&total=29017>

1. 얀 르쿤, 메타 떠나 스타트업 설립 계획..."예고된 결별 수순"



- 메타 수석과학자 **얀 르쿤**이 메타를 떠나 '월드 모델' 연구에 집중하는 신규 AI 스타트업을 설립할 계획인 것으로 알려졌다.
- 이는 메타가 장기 연구 중심의 FAIR 구조를 축소하고, **제품 개발 중심의 MSL 체제로 전환하면서 조직 재편과 연구 통제 강화가 이어진 데 따른 예견된 변화**로 평가된다.
- MSL 중심 전략, 내부 감원, 연구 공개 제한 등으로 연구진의 반발이 누적되며 르쿤의 퇴사 가능성이 지속적으로 제기돼 왔다.
- 특히 메타의 단기적 초지능 경쟁 전략과 르쿤의 장기 연구 철학 사이의 간극도 주요 배경으로 지적된다.
- PyTorch 창시자이자 총괄 책임자였던 Soumith Chintala도 2025년 11월 초 Meta 퇴사를 공식 발표했다.

출처: <https://www.aitimes.com/news/articleView.html?idxno=203076>
https://x.com/soumithchintala/status/1986503070734557568?utm_source=chatapt.com

괜찮은 기사들

- 구글, 지도로 AI 앱 만들 수 있는 개발자 기능 대거 추가
(<https://www.aitimes.com/news/articleView.html?idxno=203886&page=5&total=29017>)
- 구글, 노트LM에 '딥 리서치' 추가...더 많은 파일 형식 지원
(<https://www.aitimes.com/news/articleView.html?idxno=203994>)
- 바이두, 신형 AI 칩·슈퍼컴퓨팅 제품 공개..."새 모델은 기대 이하"
("(<https://www.aitimes.com/news/articleView.html?idxno=203981&page=2&total=29017>)
- 바이두, 도면 분석까지 가능한 고성능 'VLM 추론 모델' 오픈 소스 출시
(<https://www.aitimes.com/news/articleView.html?idxno=203955&page=2&total=29017>)
- IBM, 양자컴퓨터 칩 '룬' 공개..."2030년 상용화 목표"
("(<https://www.aitimes.com/news/articleView.html?idxno=203962&page=3&total=29017>)
- 삼성, 메모리칩 공급 부족 심화에 가격 최대 60% 인상
(<https://www.aitimes.com/news/articleView.html?idxno=204004>)