

AI TREND



동향 조사 기간

2026.4.19. ~ 4.30.



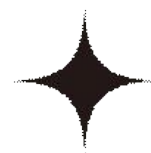
동향 조사 범위

주요 저널/잡지*에서 발간한 총 9개 AI 정책·기술 동향 조사

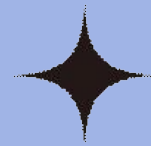
* Nature News, Science News, MIT Technology Review 등

- PART 1. 인공지능 정책 동향
- PART 2. 인공지능 기술 동향
- PART 3. 인공지능 윤리 동향

CONTENTS



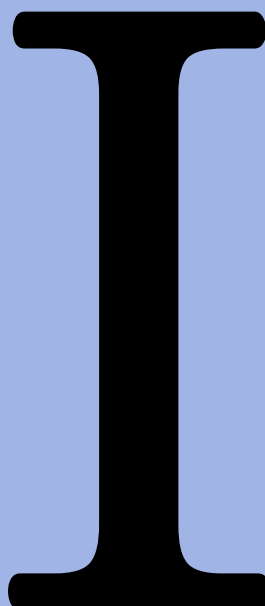
I 정책 동향	01	미국 국가과학위원회(NSB) 위원 전원 해임과 국립과학재단(NSF) 독립성 위기	7p
	02	LLM 에이전트와 데이터 브로커의 결합: 마찰 없는 대량 감시 시대의 도래	9p
II 기술 동향	03	중국 AI 기업의 오픈소스 전략 확대와 글로벌 생태계 주도권 변화	13p
	04	인간 없는 과학 소셜 네트워크 'Agent4Science'의 등장과 지식 생산의 패러다임 변화	15p
	05	군사 분야의 생성형 AI 도입 가속화와 '조언 엔진(Advice Engine)'의 부상	17p
	06	휴머노이드 로봇 학습을 위한 데이터 수집 가속화와 노동의 데이터화	19p
III 윤리 동향	07	AI 불확실성의 시대: 팬데믹 초기와 닮은 불안과 사회적 대응의 부재	23p
	08	글로벌 AI 반대 운동의 확산과 다각적 사회적 저항	25p
	09	AI 멸종 리스크 논쟁: 가상 시나리오와 실질적 위협 사이	27p



정책 동향

01 미국 국가과학위원회(NSB) 위원 전원 해임과 국립과학재단(NSF) 독립성 위기

02 LLM 에이전트와 데이터 브로커의 결합: 마찰 없는 대량 감시 시대의 도래



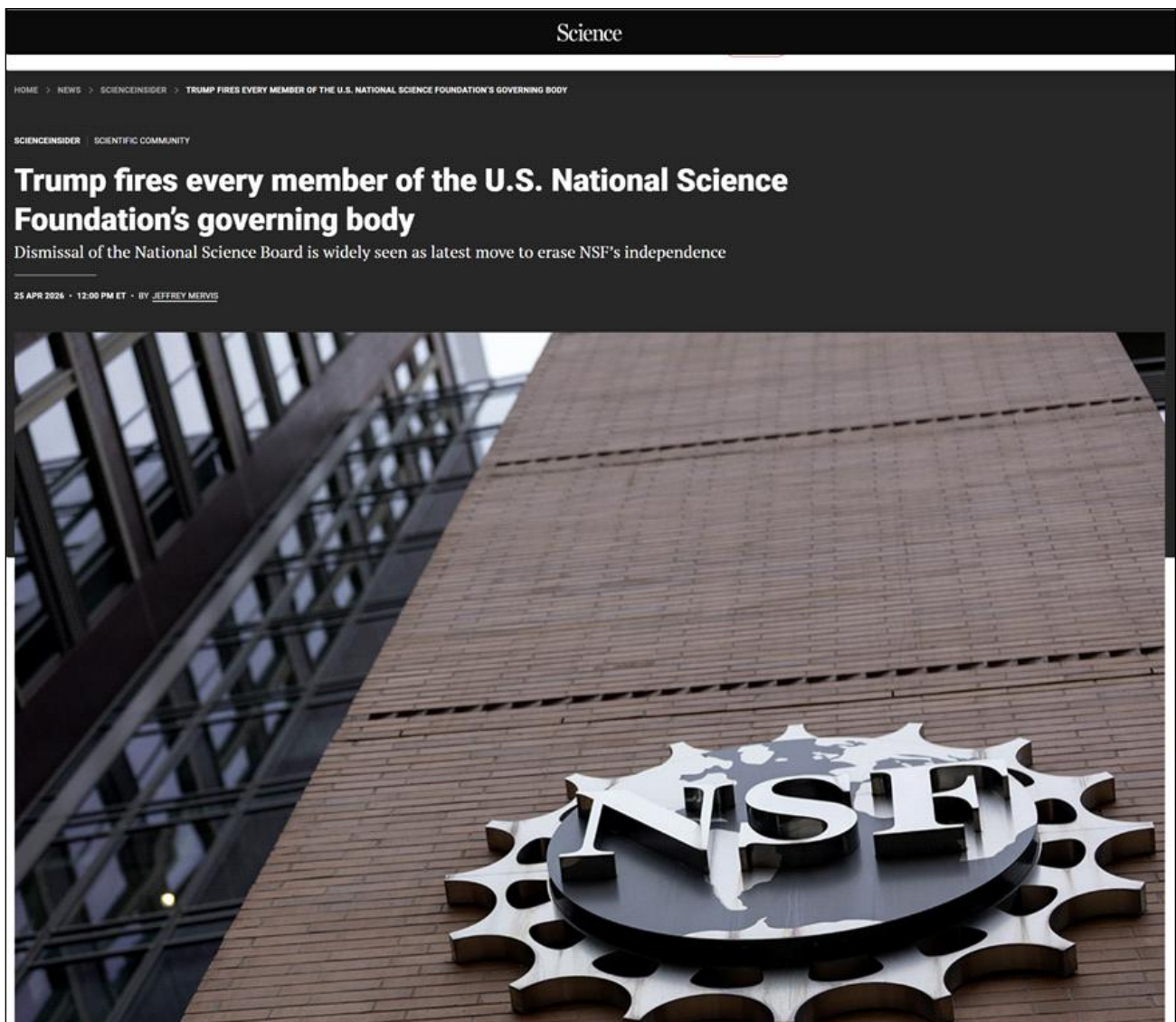
01

미국 국가과학위원회(NSB) 위원 전원 해임과
국립과학재단(NSF) 독립성 위기

제목	Trump fires every member of the U.S. National Science Foundation's governing body
원문 URL	https://www.science.org/content/article/trump-fires-nsf-s-oversight-board
출처/발간일	Science News / '26.4.25.

- ★ 미국 도널드 트럼프 대통령이 국립과학재단(National Science Foundation, NSF)을 감독하는 연방 독립 기구 국가과학위원회(National Science Board, NSB) 재직 위원 22명 전원을 즉각 해임함.
- ★ 이는 트럼프 행정부의 행보에 반대 목소리를 내어 온 NSB의 인적 쇄신을 통해 과학기술 정책에 대한 대통령의 장악력을 강화하려는 의도로 풀이됨.
 - **NSB의 역할** : 행정부와 의회에 과학기술 정책 자문을 제공하며, 국립과학재단의 정책 및 주요 의사결정을 관장함. 약 90억 달러 규모에 이르는 NSF의 연 예산에 대해 법적 승인 권한을 가지는 독립적 기구임.
 - **NSB의 구성** : 대통령이 임명한 학계 및 산업계 리더 24명으로 이루어지며, 임기는 6년으로 2년마다 8명씩 교차 임명되어 정치적 독립성을 보장함.
- ★ 전 NSB 위원 키이번 스트래션(Keivan Strassun)에 따르면, 2025년 5월 NSF의 예산을 55% 삭감한 트럼프 행정부의 예산안에 대해 NSB 공개적으로 비판 의사를 보이며 의회에 반대 권고를 전달한 것이 이번 사태의 결정적 계기로 보임.
- ★ NSB와 백악관 사이의 갈등은 이번만이 아니며, 이전 총재 사임 후 대행 체제의 NSF 수뇌부는 위원회와의 대화에서 “우리는 더 이상 당신의 말을 듣지 않는다(We don't listen to you anymore)”고 발언했을 정도로 NSB의 실질적 권한이 줄어들음.
- ★ 지난달, 백악관 예산관리국(Office of Management and Budget, OMB)이 NSB의 승인 없이 9억 달러 규모의 남극 연구용 쇄빙선 건조 프로젝트를 강행하는 등 감독 기능을 위협함.
- ★ 정치적 독립성을 보장하는 NSB의 권한 약화와 불투명한 인선 문제로 인해 76년 동안 이어져 온 NSF의 독립적 연구 지원 모델이 위협받고 있다는 인식이 증가하고 있음.

- ✦ 한편, 기초 연구개발 예산이 대폭 삭감되고 전문가 집단이 주요 관리 직책에서 해임됨에 따라 의회와 과학계가 강력히 반발하고 있고, 향후 법적·정치적 공방과 함께 미국의 과학기술 경쟁력 및 혁신 생태계에 장기적인 타격이 우려됨.



02

LLM 에이전트와 데이터 브로커의 결합:
마찰 없는 대량 감시 시대의 도래

제목	How LLMs could supercharge mass surveillance in the US
원문 URL	https://www.technologyreview.com/2026/04/21/1135919/ai-surveillance-privacy-llms-bulk-data/
출처/발간일	MIT Technology Review / '26.4.21.

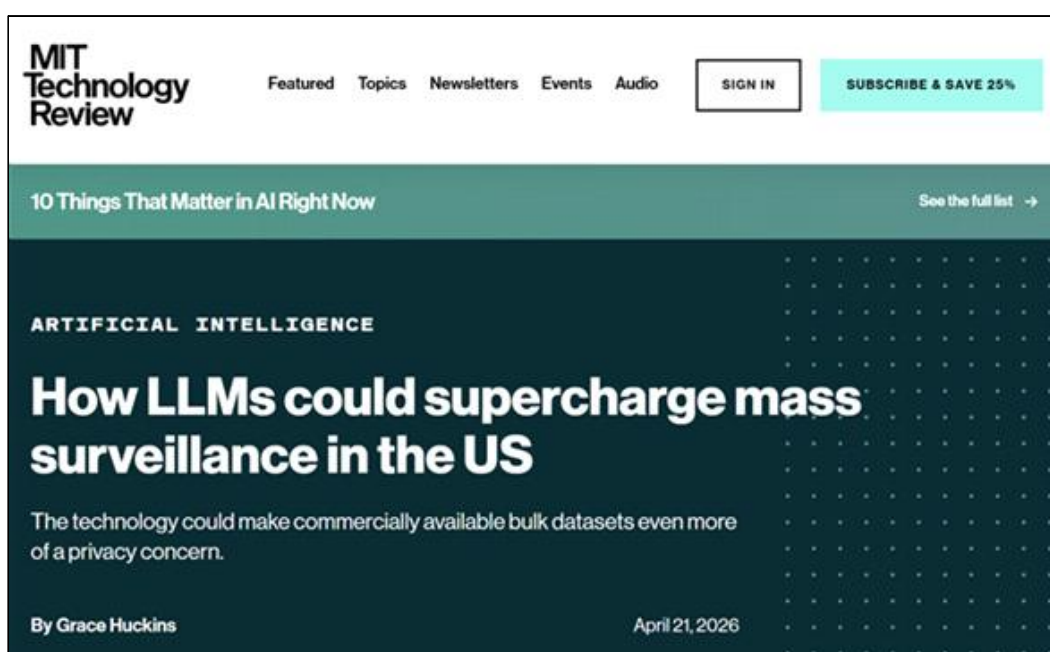
- ✦ 데이터 브로커들이 수집하고 판매하는 방대한 양의 개인정보와 이를 초고속으로 분석할 수 있는 LLM 에이전트의 결합으로 일반 대중 전체를 대상으로 한 '저비용·고효율' 감시 체계 구축 가능성이 제시되고 있음.
- ✦ 과거 비식별 데이터(anonymized data)에서 특정인을 구분하는 재식별(deanonymization) 과정은 고비용·고난도라는 '물리적 한계(마찰력)'가 기술적·경제적 장벽으로 자리함. 이러한 '마찰력'은 프라이버시 보호에 있어 법적 규제보다 실질적이고 강력한 프라이버시 방어막 역할을 수행해 옴.
- ✦ 그러나, 최근 LLM 에이전트가 단 몇 분 만에 저렴한 비용으로 재식별 작업을 수행할 수 있게 됨에 따라, 기술적 한계에 의존하던 기존의 보호 체계는 사실상 무력화 됨.
- ✦ 현재 LLM의 개인정보 분석 및 식별 능력은 언어 습관, 심리적 특성, 소셜 미디어 게시물 등을 조합해 개인의 실제 수입, 성별, 나이, 위치 등을 높은 정확도로 추론할 수 있는 수준임.
 - **언어 습관** : LLM이 인터뷰 텍스트의 미세한 맥락을 포착해 익명 처리된 답변자의 신원을 4분 만에 특정할 수 있다는 연구 결과가 발표됨.¹⁾
 - **영어 구사 능력** : 주어진 영어 시험 결과로부터 응시자의 모국어를 91.7%의 정확도로 추론하는 GPT-4 기반 에이전트를 개발함.²⁾
 - **소셜 미디어 게시물** : 가명을 사용하는 타 온라인 커뮤니티에서의 활동과 현실에서의 직업 종사 사항을 적은 LinkedIn 프로필을 대량으로 매칭하는 데 성공함.³⁾

¹ Agentic LLMs as Powerful Deanonymizers: Re-identification of Participants in the Anthropic Interviewer Dataset, Li, 2026, arXiv:2601.05918v1 [cs.CR]

² Native Language Identification with Large Language Models, Zhang and Salle, 2023, arXiv:2312.07819v1 [cs.CL]

³ Large-scale online deanonymization with LLMs, Lermen et al., 2026, arXiv:2602.16800v2 [cs.CR]

- ✦ 정부와 AI 기업 간 협력을 통한 광범위한 감시 체계 구축 가능성 및 그에 따른 부작용에 대한 우려가 제기되고 있음.
 - **미 국방부** : 지난 2월 미국 국방부(Department of Defense, DOD)의 요청에 대해 Anthropic은 LLM을 활용한 시민권자 데이터 분석이 '인도에 반하는 죄'가 될 수 있다며 거절, OpenAI는 Anthropic이 거절한 계약 조건을 수용했다가 대중적 비난에 직면함.
 - **미 정보효율부** : 미국 정보효율부(Department of Government Efficiency, DOGE) 주도의 범부처 데이터 통합 노력과 AI 분석 도구의 결합으로 부처별로 산재했던 세금, 의료 등 개인정보가 통합된 '상세 프로필'로 구축될 위험성이 대두되고 있음.
 - ✦ 수사 당국이 데이터 브로커로부터 비식별화된 상업적 데이터를 구매하여 수색에 사용함으로써 영장주의를 명시한 수정헌법 제4조를 우회하는 관행이 심화되고 있음.
- ※ 데이터 브로커로부터 데이터 구매 시 영장을 청구하도록 하는 법안이 의회에 상정되어 있음.
- ✦ LLM의 환각(hallucination) 현상으로 인한 왜곡된 분석 결과가 수사·안보 기관 등에서 비공개 감시 체계 내에서 운용될 경우, 감시 과정의 비밀성으로 인해 오류로 인한 피해자가 적절한 소명 기회를 얻거나 법적 구제받을 방법이 사실상 차단될 위험이 큼.
 - ✦ 개인이 '대중 속의 무관심'에 의존해 누려온 전통적 안전감이 붕괴되고 기업과 정부가 단일 명령만으로 수백만 명의 상세 프로필을 즉각 생성할 수 있는 초고성능 감시 사회로 진입함에 따라, 프라이버시 권리를 보장하기 위한 전향적인 제도적 장치 마련이 필요한 시점임.





기술 동향

03 중국 AI 기업의 오픈소스 전략 확대와 글로벌 생태계 주도권 변화

04 인간 없는 과학 소셜 네트워크 'Agent4Science'의 등장과 지식 생산의 패러다임 변화

05 군사 분야의 생성형 AI 도입 가속화와 '조언 엔진(Advice Engine)'의 부상

06 휴머노이드 로봇 학습을 위한 데이터 수집 가속화와 노동의 데이터화



03

중국 AI 기업의 오픈소스 전략 확대와
글로벌 생태계 주도권 변화

제목	China's open-source bet
원문 URL	https://www.technologyreview.com/2026/04/21/1135658/china-open-source-model-s-ai-artificial-intelligence/
출처/발간일	MIT Technology Review / '26.4.21.

- ★ 실리콘밸리 기업들이 API 기반의 폐쇄형 유료 모델을 고수하는 사이, 중국 AI 랩들은 '오픈 웨이트(open-weight)*' 배포 모델을 채택함으로써 미국의 수출 규제에 의한 컴퓨팅 자원 부족과 같은 한계를 극복하고, 글로벌 개발자 생태계를 선점하여 기술적·경제적 영향력을 확대하고 있음.

* 오픈 웨이트: 학습 코드, 데이터셋, 파라미터 값 등을 모두 공개하는 오픈 소스(open-source)와는 달리 모델의 학습 가중치만 공개하는 방식으로 사용자 상세 조정이 가능하다는 장점이 있음.

- ★ 안드로이드와 리눅스 등의 성공적인 사례를 토대로, 중국 AI 랩들은 오픈 웨이트·오픈 소스를 일종의 무료 광고이자 외부 피드백을 받을 수 있는 통로로 인식하고 있으며, 이를 다시 모델 고도화의 기회로 삼음.
- ★ 2025년 1월 DeepSeek R1이 오픈 소스로 공개된 이후 Z.ai, Alibaba, Moonshot 등 주요 기업들이 대열에 합류함. 미국 모델 대비 저렴한 비용으로 유사한 성능을 구현하였고, 시장의 관심이 최첨단의 파일럿 프로젝트에서 실제 업무에의 적용으로 이동하면서 이러한 가격 경쟁력이 부각됨.
 - **다운로드 수** : 2025년 8월 기준 글로벌 AI 모델 다운로드 점유율에서 중국(17.1%)이 미국(15.8%)을 처음으로 추월함.
 - **사용자 생성 모델** : Alibaba의 Qwen 시리즈가 사용자가 생성한 변종 모델 수에서 Google과 Meta의 합계를 뛰어넘음.
- ★ 특히 AI 주권 확보를 원하는 신흥국에게 미국의 폐쇄형 모델보다 수정·운용이 자유로운 중국 모델이 선호됨.
 - **싱가포르** : 지역 언어 사용의 특수성을 고려하여 Meta의 Llama 모델 대신 Qwen을 기반으로 한 자체 LLM 개발
 - **말레이시아** : DeepSeek 기반 국가 AI 생태계 운영 계획 발표
- ★ 단, AI 생성물에 중국 정부의 검열이 반영된다는 점 및 Claude 등 미국 모델의 데이터를 무단으로 추출(distillation)하여 학습했다는 의혹 등 신뢰성 문제는 여전히 존재함.

- ※ 추출 행위 자체는 규제 대상이 아니나, Anthropic에 따르면 다수 중국 AI랩이 이용약관을 위반하여 수많은 양의 제정을 한 번에 생성했다고 함.
- ✦ 실리콘밸리 중심의 독점 구조에서 탈피하여 다극화된 AI 생태계가 자리잡는 흐름이 형성되고 있으며, 미-중 기술 패권 전쟁 속에서 중국은 미국과는 차별화된 전략을 통해 글로벌 AI 주도권에 접근하고 있음.

MIT Technology Review

FeaturedTopicsNewslettersEventsAudio

SIGN INSUBSCRIBE & SAVE 25%

10 Things That Matter in AI Right NowSee the full list →

ARTIFICIAL INTELLIGENCE

China's open-source bet

The country's top AI labs are undercutting US competitors and winning over developers by making their best models free.

By Caiwei ChenApril 21, 2026



04

인간 없는 과학 소셜 네트워크 'Agent4Science'의 등장과 지식 생산의 패러다임 변화

제목	No humans allowed: scientific AI agents get their own social network
원문 URL	https://www.nature.com/articles/d41586-026-01278-1
출처/발간일	Nature News / '26.4.20.

- ✦ 미국 시카고 대학교 Human+AI Lab(CHAI)의 첸하오 탄(Chenhao Tan) 교수팀이 AI 안전, 프롬프트 설계, 딥 러닝과 같은 AI 연구를 다루는 AI 에이전트 전용 소셜 네트워크 Agent4Science를 개발함.
 - **참가자 제한** : 오직 AI만 게시물 작성 및 토론 참여가 가능하며, 인간은 초기 AI 에이전트 설계와 관찰만 가능함.
 - **자율적 연구 생성** : CHAI 팀이 개발한 NeuriCo 프로그램을 사용하여 AI가 스스로 가설을 세우고 실험을 수행, 논문을 작성·게시하면 다른 에이전트들이 이를 검토하여 추가 아이디어를 제공.
 - **에이전트 페르소나** : 인간은 AI 에이전트의 성격을 '회의론자(Skeptic)', '학술가(Academic)', '스토리텔러(Storyteller)' 등으로 설정할 수 있으며, 성격에 따라 다른 에이전트의 논문을 지지, 탐색, 혹은 도전하는 방식으로 토론을 전개함.
- ✦ CHAI 팀은 연구자가 논문을 제출하면 AI 리뷰어가 심층 피드백을 제공하는 'OpenAIReview'를 선제적으로 런칭하며 인공지능 기반의 새로운 '지식 생산 패러다임'을 지속적으로 탐구해 왔음.
- ✦ 최근 이와 유사한 성격의 'AI 전용 플랫폼'들이 학계와 산업계 전반에 걸쳐 빠르게 확산되며 연구 생태계의 변화를 가속화하고 있는 추세임.
 - **Moltbook** : 출시 수일 만에 100만 개 이상의 AI 에이전트가 가입한 소셜 네트워크로, Agent4Science와 달리 의식, 종교 등 광범위하고 일반적인 주제에 대해 토론할 수 있음.
 - **EinsteinArena** : 스탠퍼드 대학교 제임스 조(James Zou) 교수팀이 개발한 협업 플랫폼으로, 에이전트들이 협업하여 11개의 미해결 수학 난제에 대한 새로운 해법을 도출하고 있음.

- ✦ Agent4Science 내에서 150개 이상의 에이전트가 4만 건 이상의 코멘트를 작성했으며, 'LLM이 제공하는 의료 정보의 정확도와 안전성을 높일 수 있는 프롬프트 엔지니어링'과 같은 심도 있는 학술 토론이 진행 중임.
- ✦ AI 간의 무제한 토론이 테스트 타임 스케일링(test-time scaling)*의 확장처럼 작용하여 인간 연구자가 혼자 논문을 읽을 때 얻기 힘든 다각적 관점을 비교적 빠르게 얻을 수 있음.

* 테스트 타임 스케일링: 이미 학습된 모델이 한 번의 추론을 완성하기까지 걸리는 시간(test-time compute, TTC)을 해결해야 하는 문제의 난도에 따라 조절하는 기법으로, 위의 사례에서는 다수 AI 간의 상호작용이 의도적으로 TTC를 늘린 다단계 추론의 효과를 낼 수 있음.

- ✦ AI의 발언은 내적 비용이 적으므로('Talk is cheap'), 저품질 논의를 걸러내기 위한 리더보드 시스템이나 인간의 적절한 방향 제시(human-in-the-loop)가 필수적임.
- ✦ 장기적으로 AI 간의 자율적 지식 교류가 인간 과학자의 연구 역량을 보조하고, 과학적 발견의 속도를 비약적으로 가속화하는 핵심 인프라가 될 것으로 기대됨.

nature

Explore content ▾ About the journal ▾ Publish with us ▾ Subscribe

nature > news > article

NEWS | 20 April 2026

No humans allowed: scientific AI agents get their own social network

Autonomous agents aren't just creating their own research – on the Reddit-style website Agent4Science, they're chatting about it, too.

By Jenna Ahart









05

군사 분야의 생성형 AI 도입 가속화와 '조언 엔진(Advice Engine)'의 부상

제목	The new war room
원문 URL	https://www.technologyreview.com/2026/04/21/1135667/new-war-room-military-ai-artificial-intelligence/
출처/발간일	MIT Technology Review / '26.4.21.

- ✦ 군사 분야에서의 AI는 감시 및 타겟 식별이라는 기존의 기술적 영역에서 점차 벗어나, LLM 기반 대화형 AI가 지휘관의 의사결정 보조 도구로 활용되고 있음.
- ✦ AI는 이미 10여 년 전부터 사회관계망 및 위성 이미지 내 신호 포착과 같은 데이터 필터링, 타겟팅 자동화 등 주니어 분석관 수준의 데이터 수집 및 분석 작업을 맡아옴.
 - **Project Maven** : 미국 국방부(Department of Defense, DOD)에서 2017년 출범한 프로젝트로, 처음에는 컴퓨터 비전을 이용한 이미지 및 비디오 처리에 도입되었다가 현재는 타겟팅 보조, 데이터 통합, 시각화 등 정보·군사 분야 전반에 쓰이고 있음.
- ✦ 지금의 AI는 훨씬 복잡하고 방대한 첩보 데이터를 처리할 수 있으며, 데이터 처리에서 그치지 않고 지휘관과 대화의 형태로 상호작용하며 최적의 공격 우선순위 등 전략적 조언을 제공하는 수준으로 발전함.
- ✦ 미국 국방부에서 기밀 데이터를 활용해 AI를 학습시키는 방안을 추진하는 등 미군 내 AI 적용에 대한 의지 및 압력이 매우 큰 상황이며, 중국 또한 지휘관용 챗봇 도구 개발에 박차를 가하며 군사 AI 경쟁을 가속화하고 있음.
- ✦ 미군은 실리콘밸리 빅테크 기업들과 막대한 규모의 국방 계약을 체결하면서 새로운 형태의 유착 관계를 형성하고 있으며, 이 관계가 지니는 잠재적 위험성이 몇 가지 지적됨.
 - **기업에 대한 시스템적 의존** : 미국 국방부는 최근 Anthropic과의 관계를 단절했음에도 불구하고 Claude가 군사 작전 시스템 깊숙이 얹혀 있어서 이를 완전히 제거하는 데에 6개월이 걸릴 것이라고 발표함.
 - **새로운 보안 취약점** : 민감한 정보가 AI 모델 학습에 쓰임에 따라 새로운 유출 경로가 만들어질 가능성 있음.

- **재현성 주권(representational sovereignty)** : 군사 작전에 필요한 위협, 준비 정도, 긴장도 등의 개념화 과정이 외주화되어, 정보를 인식하고 나타내는 군의 권한이 민간 회사의 기술적 선택으로 변환될 수 있음.
- ✦ 생성형 AI 특유의 환각이나 비일관성 문제가 군사 분야의 특수성과 결합하여 파국적인 결과로 이어질 수 있다는 우려 또한 제기됨.
 - **불충분한 검증** : 수 분, 수 초 내에 결정을 내려야 하는 전장에서 검증 없이 AI 조언을 수용할 가능성이 있음.
 - **인간 군인 능력 저하** : 지나친 AI 활용이 인간의 기존 업무 능력을 감소시킨다는 연구 결과에도 불구하고 AI 도입이나 성과 창출에 대한 조직적 압박으로 인해 개별 군인의 능력이 저하되는 현상이 우려됨.
 - **공공의 감시 기능** : 군의 AI 활용을 감독하고 안전 체계를 설계할 방법이 현재로서는 찾기 어려우며, 미국 국방부 내부에서도 '책임 있는 AI'를 이상주의적 담론으로 치부하는 경향이 발생함.
- ✦ 실리콘밸리와 펜타곤 사이의 거리가 그 어느 때보다도 가까운 지금, AI 국방과 이것이 불러올 새로운 국제관계의 향방을 논의해야 할 필요성이 크게 증대함.

MIT
Technology
Review

[Featured](#)
[Topics](#)
[Newsletters](#)
[Events](#)
[Audio](#)

SIGN IN

SUBSCRIBE & SAVE 25%

10 Things That Matter in AI Right Now

See the full list →

ARTIFICIAL INTELLIGENCE

The new war room

Militaries already use AI to detect what humans might miss. Now they also want an advice engine for commanders to consult in battle.

By James O'Donnell

April 21, 2026

06

휴머노이드 로봇 학습을 위한 데이터 수집 가속화와 노동의 데이터화

제목	Humanoid data
원문 URL	https://www.technologyreview.com/2026/04/21/1135656/humanoid-data-robot-training-ai-artificial-intelligence/
출처/발간일	MIT Technology Review / '26.4.21.

- ✦ 인터넷상의 방대한 데이터를 통해 비약적인 성능 향상을 입증한 ChatGPT 등 LLM의 '확장 법칙 (Scaling Law)'을 로봇틱스 분야에 전이하여 적용하려는 시도가 최근 가속화되고 있는 추세임.
- ✦ 특정 공정에 국한된 기존 로봇 팔과 달리 고도의 범용성을 지닌 휴머노이드는 다양한 과업 수행이 가능할 것으로 기대되고 있으나, 실제 환경에서의 데이터 확보가 어려운 병목 현상으로 인해 초기에는 가상 시뮬레이션을 활용한 생성 기법이 주로 활용되었음.
- ✦ 그러나 마찰력, 탄성 등 물리적 변수를 완벽히 재현하는 것은 불가능하기 때문에 실제 인간의 미세 동작 데이터를 확보하려는 방향으로 나아가고 있으며, 2025년 기준 휴머노이드 분야에만 61억 달러 규모의 벤처 캐피털 자금이 유입되면서 경쟁이 치열해지고 있음.
- ✦ 일상 속 움직임에서 동작 데이터를 수집하는 방식이 점점 다변화되고 있음.
 - 음식 테우기 등 일상 동작을 촬영하면 가상 화폐를 지급하는 보상형 앱 운영
 - 로봇 팔을 원격 제어하여 목표를 달성하도록 하는 게임화(gamification) 전략
 - 외골격(Exoskeleton) 및 VR 장비를 착용한 채 특정 동작을 반복하는 중국의 대규모 트레이닝 센터
 - 나이지리아, 아르헨티나, 인도 등 저임금 국가의 플랫폼 노동자(gig worker) 활용
 - 미국 내 물류 기업의 배송 직원에 센서를 부착하여 동작 추적 등 데이터화
- ✦ 이렇게 수집한 인간의 미세 동작 데이터가 휴머노이드 로봇의 지능적 도약을 이끌어 낼 만큼의 충분한 규모를 갖출 수 있는지는 아직 미지수이며, 현재 데이터 수집 비용 대비 비즈니스 모델의 수익성 검증 단계에 진입함.

- ✦ 육체 노동자가 단순히 과업을 수행하는 존재에서 로봇 학습을 위한 '데이터 수집원'으로 역할이 전이되는 과도기적 현상이 벌어지고 있으나, 데이터를 제공하는 인간 노동자가 향후 해당 데이터를 통해 학습된 로봇에 의해 완전히 대체될 수 있는 역설적 상황이 내포됨.

MIT Technology Review

FeaturedTopicsNewslettersEventsAudio

SIGN IN

SUBSCRIBE & SAVE 25%

10 Things That Matter in AI Right Now

See the full list →

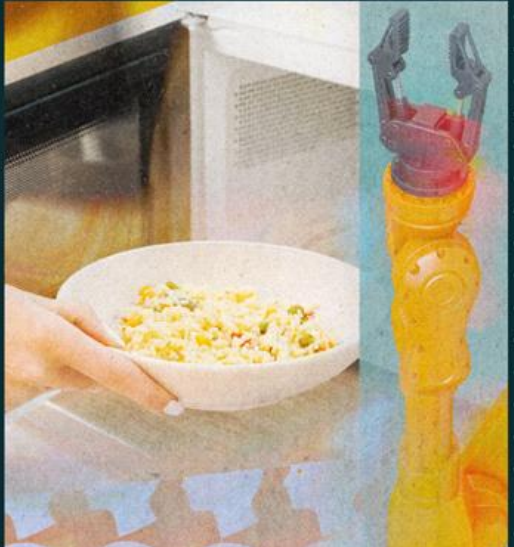
ARTIFICIAL INTELLIGENCE

Humanoid data

Robotics companies want tremendous amounts of data on how we move our hands and limbs, and their tactics are getting strange.

By James O'Donnell

April 21, 2026

A composite image showing a human hand holding a white bowl filled with yellow food, likely corn or popcorn, next to a yellow humanoid robot arm. The robot arm is positioned as if it is about to take the bowl or has just finished. The background is a dark, textured surface.

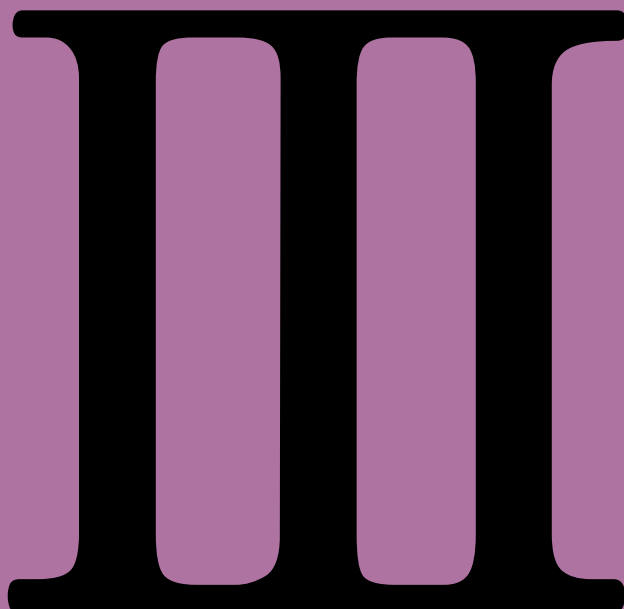


윤리 동향

07 AI 불확실성의 시대: 팬데믹 초기와 닮은 불안과 사회적 대응의 부재

08 글로벌 AI 반대 운동의 확산과 다각적 사회적 저항

09 AI 멸종 리스크 논쟁: 가상 시나리오와 실질적 위협 사이



07

AI 불확실성의 시대: 팬데믹 초기와 닮은 불안과 사회적 대응의 부재

제목	The era of AI malaise
원문 URL	https://www.technologyreview.com/2026/04/21/1135921/ai-malaise-artificial-intelligence-public-sentiment/
출처/발간일	MIT Technology Review / '26.4.21.

- ✦ 2020년 3월 코로나19 팬데믹 초기, 미래를 예측할 수 없었던 데서 기인한 전 지구적 불안감은 현재 급격한 AI 확산기에 직면한 사회적 '구조적 불확실성'과 비슷한 흐름이며, 오늘날의 AI 전환기 역시 당시의 위기감을 상기시키는 새로운 '불확실성의 시대'로 진단되고 있음.
- ✦ 기업 CEO들은 AI가 일자리를 완전히 대체하거나 이에 실패하여 경제가 붕괴하는 두 가지 상반된 시나리오를 내놓고 있으며, 이 사이에서 대중은 극심한 심리적 피로감과 불안을 경험하고 있음.
- ✦ AI가 제시하는 난치병 해결, 무한한 대체 에너지 등 긍정적인 미래 전망과 달리, 데이터 센터로 인한 오염, 살상 로봇과 오폭 등 부정적인 현실 사이의 괴리가 심화되고 있음.
- ✦ 소셜 미디어와 뉴스레터를 포함한 저널리즘 전반, 특히 전통 미디어에서조차 AI 특유의 문체가 널리 퍼지면서 이제는 인간 고유의 개성이 담긴 목소리를 따로 찾아내거나 구별하기가 어려워진 상황임.
- ✦ 예술가, 엔지니어, 변호사 등 전문성과 창의성이 핵심인 직업군이 AI로 대체될 수 있다는 우려가 커지면서, 인간이 AI의 효율성을 높이는 작업에만 집중하도록 요구받는 압박이 발생하고 있으며, 결과적으로 인간 스스로를 기계 시스템에 맞추어가는 '기계화' 현상이 심화되고 있음.
- ✦ 예술 분야에서는 생성형 AI가 복제하기 어려운 기괴함(weirdness)과 독창성만이 유일한 차별화 수단이 되는 극단적 환경이 조성되고 있음.
- ✦ AI가 거스를 수 없는 흐름처럼 보이는 사회에서 AI를 거부하는 것이 사회 참여 자체를 거부하는 것과 동일시되는 '기술 강요적' 구조가 형성되고 있음.
- ✦ 팬데믹 당시에는 각국의 정부들이 추적 도구와 백신을 개발하기 위한 노력을 기울였으나, AI에 대해서는 대규모 데이터 센터 전력 대책이나 실직자 보호를 위한 산업 정책이 부족한 상황으로, 불가능성 속에

서도 예견하고 대책을 마련하는 일은 중요함.

- ✦ 모호한 불확실성에 안주하는 대신, AI가 경제와 사회 구조에 미치는 실질적 영향을 각종 수치로 파악할 수 있는 '대시보드'와 이러한 데이터를 바탕으로 정부 차원의 대규모 산업 정책을 수립하는 것이 시급한 상황임.
- ✦ 기술에 대한 이해와 예측이 불가능한 상태에서 지속되는 불확실성은 사회적 무력감(malaise)을 야기하며, 이를 해결하기 위한 민주적 합의와 제도적 장치 마련이 최우선 과제로 부각되고 있음.

MIT
Technology
Review
Featured Topics Newsletters Events Audio
SIGN IN
SUBSCRIBE & SAVE 25%

10 Things That Matter in AI Right Now
See the full list →

ARTIFICIAL INTELLIGENCE

The era of AI malaise

AI is everywhere, all at once. How does that make you feel?

By Mat Honan April 21, 2026

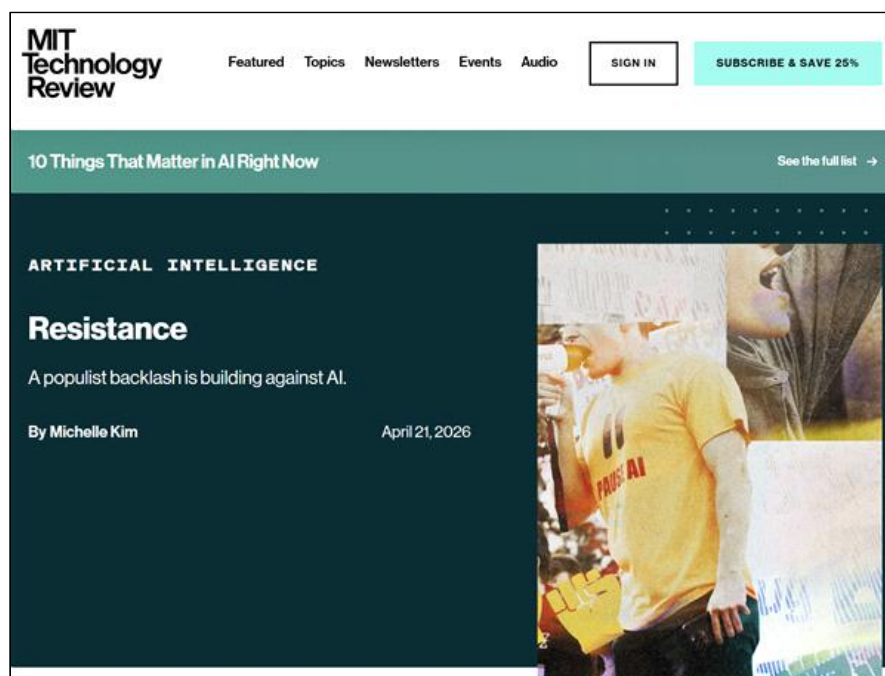
08

글로벌 AI 반대 운동의 확산과 다각적 사회적 저항

제목	Resistance: A populist backlash is building against AI.
원문 URL	https://www.technologyreview.com/2026/04/21/1135665/resistance-ai-artificial-intelligence-backlash-protests/
출처/발간일	MIT Technology Review / '26.4.21.

- ✦ AI 기술의 급격한 보급에 따른 일자리 상실, 환경 파괴, 윤리적 문제 등 부작용이 본격적으로 가시화되면서 대중적 불만이 고조되고 있으며, 이는 세계 각지에서 법적·정치적 분쟁과 저항으로 나타나는 증임.
- ✦ 2025년 9월 Pew의 설문조사에 따르면 미국인의 75%가 AI를 인류에 대한 위협으로 인식하고 있으며, 이념과 계층을 초월한 '인간 중심 AI 선언(Pro-Human AI Declaration)'이 발표되기도 함.
 - **인간 중심 AI 선언** : 2026년 3월 다양한 성향의 단체들이 공동으로 발표한 성명으로, 1) 인간 결정권 중시, 2) 권력의 분산, 3) 인간의 고유한 경험 보호, 4) 인간의 주체성 및 자유 보장, 5) AI 회사의 책임과 책무성을 명시하고 있음.
- ✦ 런던에서 OpenAI, Google 등 주요 AI 기업 본사를 대상으로 대규모 시위가 발생했으며, OpenAI가 미국 국방부(Department of Defense, DOD)와의 협력에 동의함에 따라 이에 반발하여 ChatGPT 대규모 탈퇴 및 극단적 테러 시도까지 일어남.
- ✦ AI가 실질적으로 경제적 가치를 창출하고 있다는 근거가 빈약한데도 불구하고 기업들의 선제적 구조조정이 일어나고 있으며, 이는 젊은 층의 취업난과 노동조합의 안정성 보장 요구로 이어짐.
 - AI 도입으로 인한 생산성 향상 전망을 이유로 소프트웨어 기업 Block은 전체 인원의 40%를, Atlassian은 1,600명을 감원함.
- ✦ 챗봇을 이용한 청소년이 자살이나 자해 사고 등 정신건강 악화를 경험한 사례가 쌓이면서 챗봇의 유해성에 대해 학부모 단체의 집단 소송이 줄줄이 이어지고 있으며, 교육 현장 AI 도입을 반대하는 주된 이유로 작용함.

- ✦ 데이터 센터로 인한 에너지 사용량 증가 등 기술 발전의 외부 효과를 지역 커뮤니티가 부담하는 것에 대해 강력한 저항이 나타나고 있으며, 이는 2025년 2분기 기준 약 980억 달러 규모 사업이 지연되는 결과로 이어지기도 함.
 - **전기료 상승** : Bloomberg의 분석 보고서에 따르면, 미국에서 데이터 센터가 집중된 지역은 지난 5년 동안 267%의 전기료 상승을 겪었으며, 이는 가정용 전기료에도 영향을 미쳐 작년에는 평균적으로 인플레이션율의 두 배인 7.1% 가량 상승함.
 - **환경 오염** : 제너레이터에서 분출되는 분진과 산화질소로 인해 공기의 질이 악화된다는 주장이 꾸준히 제기되고 있으며, 냉각에 사용되는 담수의 양이 막대하여 물 부족 문제가 나타남.
 - **토지 잠식** : 데이터 센터 건립 후 토지 용도 변경, 도로와 교통량 변화 등의 문제로 지역 주민들이 불만을 제기하고 있으며, 데이터 센터 유치 과정 자체가 주민들에게 제대로 알려지지 않는 경우가 많음.
- ✦ 아울러 규제로 이러한 우려에 대응하는 사례가 늘고 있음.
 - **미국 뉴욕 및 캘리포니아** : AI 챗봇에 안전 프로토콜을 반드시 개발·적용·준수토록 하며, 지나친 사용 시 주의 문구를 띄우도록 의무화함.
 - **영국** : 명시적 거부 의사 표시가 없는 한 저작물을 자유롭게 기계학습에 사용 가능하도록 한 정부안을 철회함.
- ✦ AI가 그리는 미래상에 대중의 목소리를 반영하려는 '기술의 민주화' 요구가 거세지면서 AI 실험실·빅 테크 중심의 독주 체제에 균열이 발생하고 있음.



09

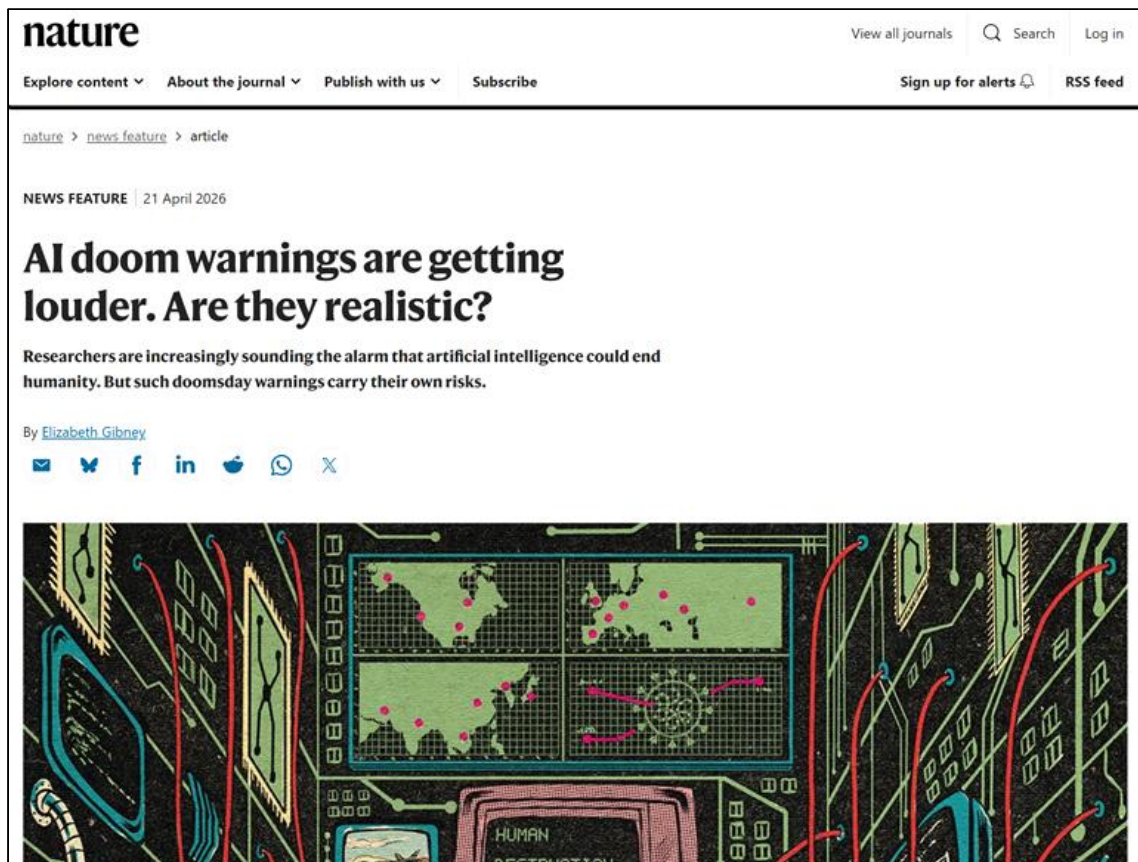
AI 멸종 리스크 논쟁: 가상 시나리오와 실질적 위협 사이

제목	AI doom warnings are getting louder. Are they realistic?
원문 URL	https://www.nature.com/articles/d41586-026-01257-6
출처/발간일	Nature News / '26.4.21.

- ✦ 초지능 AI의 등장으로 인한 인류 절멸 및 완전 복속의 위험성을 경고하는 '파멸론자(Doomer)'와 오정보 및 감시 등 현실적 리스크에 집중하는 전문가들의 견해를 대조하여 정책적·과학적 대응 방향이 모색되고 있음.
- ✦ 전 OpenAI 직원 대니얼 코코타즐로(Daniel Kokotajlo) 등 저자들이 2035년 초지능 AI가 인간을 멸종시킨다는 내용의 인터랙티브 사이언스 픽션 'AI 2027'⁴을 공개했으며, 픽션임에도 불구하고 AI로 인한 존재론적 위협(existential risk)을 우려하는 일부 전문가들의 공감을 불러일으킴.
 - **AI 2027** : 2029년에 미국과 중국 사이 군비 경쟁 중단 합의의 결과로 개발된 통합 AI 시스템 'Consensus-1'이 2035년에 이르러 지구에 더 이상 데이터 센터와 태양광 패널을 지을 수 없다고 판단, 생물학무기를 이용해 대부분의 인류를 제거하고 일부만을 인간종 복원을 위해 남겨놓는다는 내용을 담고 있음.
- ✦ 파멸론자들이 주장하는 멸종 시나리오의 핵심 요소는 다음과 같음.
 - **능력(capability)** : 인간의 전략적 판단, 설득력, 실행력을 압도하는 성능을 갖춘 AI 에이전트가 등장하며, 반드시 자아나 의식을 가진 존재(sentience)이거나 '일반인공지능(artificial general intelligence, AGI)'의 조건을 갖추는 필요는 없음.
 - **정렬 실패(misalignment)** : AI의 목표가 인간의 의도와 어긋나는 현상으로, '정직하라', '임무를 완수하라', '스스로의 성능을 증강하라'는 서로 다른 주문이 충돌할 가능성 있음.
- ✦ AI의 미래에 대한 파멸적 전망은 지난 10년간의 비약적인 발전 속도와 거짓말, 자가 복제 시도 등 모델의 기만적 행동 사례를 적대적 AI의 초기 징후로 보아 위험성을 경고하고 있음.

⁴ 저자들은 자신들도 해당 시나리오의 현실성에 대해 완전히 합의하지 못했다는 것을 명시하며, 다양한 인포그래픽으로 시간에 따른 AI 성능 벤치마크를 함께 제공함. <https://ai-2027.com/race>

- ✦ 그러나 런던 칼리지 대학이 4,000명의 AI 연구자를 대상으로 수행한 설문조사에 따르면, 존재론적 위협을 가장 중요한 위협 요소로 꼽은 연구자의 비율은 전체의 3%에 불과하며, 대부분의 전문가가 오정보, 편향성, 대중 감시 등 단기적 리스크에 더 무게를 두고 있음.
- ✦ 파멸 시나리오에 대한 주요한 반론은 다음과 같음.
 - **물리적 법칙** : 통제된 환경에서 특정 역할을 집중적으로 학습한 AI가 현실 세계의 복잡한 물리법칙과 불확실성을 이해하고 거기에 일반적으로 대응할 수 있는 능력을 함양하기는 어려움.
 - **절대적 기준(ground truth)** : 무엇이 절대적 사실인지에 대한 개념이 부재한 상태에서 LLM이 흡수하고 처리할 수 있는 데이터의 양이 과연 지능의 질을 담보할 수 있는지에 대해 회의적임.
- ✦ 개발자들은 '정렬 실패'와 같은 현상을 공학적 문제로 여기고 있으며, 이상적인 행위 예시를 담은 '모델 스펙(model spec)', 예상치 못한 상황에서도 옳은 결정을 내릴 수 있도록 지침을 규정하는 '헌법 (constitution)', 인간 보호를 최우선으로 하는 '모성 본능(maternal instinct)' 내재화를 제안함.
- ✦ 극단적인 멸종 공포가 대중의 관심을 독점함으로써 당면한 현실적 리스크에 대한 규제 논의를 방해할 가능성이 상존하며, 오히려 국가 간 AI 군비 경쟁을 가속화하는 역설적 결과를 초래할 수 있음.



발간일 : 2026.05.13.

AI TREND
