2026년 6월호

SPRi AI Brief

인공지능 산업의 최신 동향

SPRi 소프트웨어정책연구소
Software Policy & Research Institute

SPRi AI Brief

| 2026년 6월호 |

CONTENTS

월간AI모델현황

- 2026년 5월 월간 AI 모델 현황 1

정책·법제

- 중국 국가발전개혁위원회, 메타의 마누스 인수 철회 요구 3
- 영국 정부, AI 데이터센터 온실가스 배출량 추정치 100배 이상 상향 수정 4
- 미국 국방부, 앤트로픽을 제외한 8개 AI 기업과 기밀 네트워크 계약 체결 5
- 트럼프·시진핑 정상회담, 반도체와 희토류 등 핵심 기술 쟁점은 미합의 6
- 트럼프 대통령, AI 규제 행정명령 서명 당일에 전격 보류 7

기업·산업

- 오픈AI, 에이전트 성능 강화한 GPT-5.5 출시 9
- 메타, AI 훈련을 위해 직원들의 마우스 움직임과 키보드 입력 수집 10
- 앤트로픽, 스페이스X와 데이터센터 임대 파트너십 체결 11
- 마이크로소프트, 2026년 1분기 AI 확산 보고서 발표 12
- 앤트로픽, 2028년 미·중 AI 패권 경쟁 시나리오 분석 보고서 발표 13
- 오픈AI, 일론 머스크가 제기한 소송에서 승소하며 IPO 준비 본격화 14
- 구글, I/O 2026에서 신규 AI 모델과 에이전트 서비스 공개 15

기술·연구

- 오픈AI, 고블린을 반복 언급하는 챗GPT 현상의 원인과 해결책 공개 17
- 앤트로픽, 클로드의 에이전트 오정렬 개선 연구 결과 공개 18
- 오픈AI 내부 AI 모델, 80년간 미해결 수학 난제 반증 성공 19

인력·교육

- 앤트로픽, 클로드 사용자를 대상으로 AI의 일자리 영향 분석 21
- 광역런던시청 조사 결과, 런던 노동인구의 46%가 생성형 AI에 노출 22
- 브루킹스 연구소, 미국 데이터센터의 지역 고용 효과 실증 분석 23
- 대런 애쓰모글루 MIT 교수, AI의 일자리 영향에 대한 과장 경계 24

주요행사일정

- 2026년 국내외 인공지능 주요 행사 25

2026년 5월 월간 AI 모델 현황

주요 모델 출시 타임라인

5/5 GPT-5.5 Instant 5/7 GPT-Realtime-2
5/7 Cursor Composer 2.5 5/19 Gemini 3.5 Flash, Gemini Omni
5/20 Command A+, Stable Audio 3.0, Qwen3.7-Max

LLM간 성능 비교(Artificial Analysis Intelligence Index 4.0)

모델	출시일	Index	Price ($/1m Token)		Speed	Context Window	회사
			Input	Output			
① GPT-5.5(xhigh)	26.4.23.	60	5	30	72.0	922k	OpenAI
② GPT 5.5(high)	26.4.23.	59	5	30	75.9	922k	OpenAI
③ Claude Opus 4.7 (max)	26.4.16.	57	5	25	47.7	1m	Anthropic
Gemini 3.1 Pro Preview	26.2.19.	57	2	12	118.7	1m	Google
GPT-5.4(xhigh)	26.3.5.	57	2.5	15	80.6	1m	OpenAI

* 출처 : Artificial Analysis 평가체계 : Agents 25%, Coding 25%, General 25% Scientific 25% (10개 벤치마크 종합)

이미지 생성 모델

모델	출시일	ELO	회사
① GPT Image 2 (high)	26.4.21	1338	OpenAI
② GPT Image 1.5 (high)	25.12.16	1267	OpenAI
③ Nano Banana 2(Gemini 3.1 Flash Image Preview)	25.12.16	1264	Google

영상 생성 모델

모델	출시일	ELO	회사
① HappyHorse-1.0	26.4.7	1213	Alibaba
② Dreamina Seedance 2.0 720p	26.2.10	1212	ByteDance
③ Kling 3.0 Omni1080p (Pro)	26.2.5	1103	Kling

한국 AI 모델 현황

■ 2026년 5월 출시 모델 확인되지 않음

주요 포인트 :
오픈AI의 Realtime 시리즈, 구글의 Gemini Omni 등 텍스트를 넘어 음성과 영상을 생성하는 멀티모달 혁신 확대
GPT-5.5 Instant, Gemini 3.5 Flash 등 추론과 에이전트 성능을 강화한 동시에 효율성을 갖춘 모델 출시

SPRi AI Brief

| 2026년 6월호 |

정책·법제

중국 국가발전개혁위원회, 메타의 마누스 인수 철회 요구

KEY Contents

- 중국 국가발전개혁위원회가 에이전틱 AI 기업 마누스에 대한 20억 달러 규모의 메타 인수 거래를 금지하고, 당사자에 거래의 전면 철회를 요구
- 메타가 2025년 12월 마누스 인수를 발표한 이후 100여 명의 마누스 직원이 메타의 싱가포르 지사에 합류하는 등 통합을 추진했으나, 중국 정부의 개입으로 거래 무산 위기에 직면

중국 국가발전개혁위원회, 메타의 마누스 인수 금지 발표

■ 중국 국가발전개혁위원회 외국인투자안전심사판공실은 2026년 4월 27일 보안 검토 결과에 따라 메타(Meta)의 마누스(Manus)* 인수 금지를 결정했다며 인수 거래 철회를 요구

* 중국 창업자들이 2022년 베이징에서 모회사 설립 후 2025년 싱가포르로 본사를 이전한 AI 에이전트 스타트업

- 메타는 2025년 12월 29일 마누스 인수를 발표하며, 복잡한 작업을 독립적으로 실행할 수 있는 다목적 자율형 에이전트를 개발한 마누스의 기술을 자체 제품에 통합하겠다고 선언
- 미국의 빅테크가 중국과 밀접한 관계에 있는 AI 기업을 인수한 것은 매우 드문 사례로, 메타는 당시 인수액을 공개하지 않았으나 주요 매체는 내부 소식통을 인용해 약 20억~30억 달러로 추정
- 이와 관련, 중국 상무부 대변인은 2026년 1월 기자들에게 메타의 마누스 인수가 중국 법률과 규정에 부합하는지를 관련 부서와 함께 평가하고 조사할 것이라고 발언
- 중국 정부는 인수를 거부한 별도 이유를 제시하지 않고 한 줄 분량의 발표 자료로 거래 철회를 명령했으며, 메타는 해당 거래가 관련 법률을 완전히 준수했으며 적절한 해결을 기대한다는 입장을 표명

■ 메타의 마누스 인수 발표 이후 100여 명의 마누스 직원들이 메타의 싱가포르 지사에 합류하고, 창업자 샤오홍(Xiao Hong) CEO가 메타 COO에 직속 보고하는 등 통합 작업을 추진

- 그러나 중국 정부의 명령으로 거래 자체가 무산될 위기에 직면했으며, 마누스의 샤오홍 CEO와 이차오지(YiChao Ji) 최고과학자에게는 중국 정부의 출국금지 조치가 부과된 것으로 알려짐

미국과 중국의 지정학적 갈등 국면에서 중국의 AI 산업 통제 확대 신호

■ 중국 정부의 이번 결정은 미·중 갈등 구도를 넘어 AI 산업 전반에 영향을 미칠 수 있는 중요한 국경 간 거래 개입 사례로 평가되며, 메타의 AI 에이전트 분야 전략에도 타격 우려

- 백악관 대변인은 성명을 통해 트럼프 행정부가 "모든 형태의 부당한 외국 간섭으로부터 미국의 선도적이고 혁신적인 기술 부문을 지속적으로 보호할 것"이라고 발언
- 인수 당시 메타는 마누스가 중국 내 서비스와 운영을 중단할 것이라고 밝혔으나, 중국 정부는 해외 투자와 기술 유출, 데이터 전송 및 국경 간 인수합병에 관여하는 모든 기업이 중국 법률을 준수해야 한다고 강조

■ 분석가들은 중국과 미국 간 지정학적 경쟁이 치열해지는 가운데, 중국 공산당 지도부가 국가안보의 핵심 자산으로 여겨지는 AI 산업에 대한 통제를 강화하고 있다는 신호라고 지적

출처 | 国家发展和改革委员会, 外商投资安全审查工作机制办公室（国家发展改革委）对外资收购Manus项目作出安全审查决定, 2026.4.27.
TechCrunch, China blocks Meta's $2B Manus deal after months-long probe, 2026.4.27.

영국 정부, AI 데이터센터 온실가스 배출량 추정치 100배 이상 상향 수정

KEY Contents

- 영국 과학혁신기술부가 2025~2035년 AI 데이터센터로 인한 온실가스 누적 추정치를 기존의 연간 최대 14만 2천 톤에서 3,400만~1억 2,300만 톤으로 대폭 상향 조정
- AI 데이터센터의 온실가스 배출량은 전력망의 탈탄소화 속도 및 AI 수요 증가에 따라 크게 달라질 수 있으며, 모델과 하드웨어의 효율성 향상이 에너지 수요 증가를 부분 상쇄할 전망

영국 컴퓨팅 로드맵 개정 부록, 10년간의 AI 데이터센터 온실가스 배출량 대폭 상향

■ 영국 과학혁신기술부(DSIT)가 2026년 4월 23일 업데이트한 '영국 컴퓨팅 로드맵(UK Computing Roadmap)*'에서 AI 데이터센터의 온실가스 배출량 추정치를 대폭 상향

* 영국 정부가 AI 구현을 위한 세계적 수준의 컴퓨팅 생태계 구축을 위해 2025년 7월 수립한 정책

- 개정된 컴퓨팅 로드맵의 증거자료 부록은 2025~2035년 AI 데이터센터의 온실가스 배출 누적 추정치를 연간 3,400만~1억 2,300만 톤으로 수정(기존 수치는 연간 최대 14만 2천 톤)
- 수정된 상한 추정치 1억 2,300만 톤은 270만 명이 배출하는 탄소량에 상당하며 기존 대비 100배 이상 증가한 것으로, 영국 정부는 별도의 설명 없이 수정 버전을 공개
- 영국 정부는 비영리 시민단체 폭스글로브(Foxglove)와 카본 브리프(Carbon Brief)의 조사 결과에서 기존 온실가스 배출량 수치가 심각한 과소 추정으로 드러난 이후 자료를 개정

■ AI 데이터센터의 온실가스 배출량 범위는 전력망 탈탄소화 속도 및 AI 수요 규모에 따라 크게 달라질 전망으로, 영국 정부의 2030년 청정에너지 목표 달성 여부가 핵심 변수로 평가

- 현 정부 계획은 2030년까지 영국 전력 시스템의 최소 95%를 청정 에너지원으로 충당하는 것으로, 해당 목표 달성 시 배출량은 추정 범위 하단에 근접할 전망
- ChatGPT의 프롬프트당 에너지 사용량은 2023년 이후 10분의 1로 감소했으며, 모델과 하드웨어 발전에 따른 에너지 효율성 향상이 에너지 수요 증가를 부분적으로 상쇄하는 요인으로 작용
- 한편, 영국 정부는 AI 데이터센터의 물 소비량 전망치도 상향 조정해, 기존에는 2035년 한 해 1조 세제곱미터(㎥)로 예상했으나 개정안은 2035년까지 10년간 매년 0.1~0.5조 세제곱미터(㎥)로 수정

의회와 환경단체, 정부의 온실가스 배출량 추정치 조정에 우려 표시

■ 영국 정부는 AI 강국이 되겠다는 노동당 정부의 목표가 2030년 청정에너지 달성 및 2050년 탄소 중립 목표와 공존할 수 있는지 의회와 환경단체가 주목하는 시기에 이번 개정을 발표

- 폭스글로브 측은 영국 정부가 2050년까지 탄소 중립을 달성할 법적 의무가 있으면서도 국가의 전력 소비량을 두 배로 늘릴 수 있는 초거대 AI 데이터센터 구축을 무분별하게 추진하고 있다고 지적
- 영국 의회 과학혁신기술위원회의 치 온우라(Chi Onwura) 위원장은 정부가 공개한 새로운 수치에 우려를 표시하며 정부에 AI의 환경 영향을 더욱 투명하게 밝힐 것을 요구

출처 | DSIT, Compute Evidence Annex: changes corrected (23 April 2026), 2026.4.23.
The Guardian, Officials hugely underestimated impact of AI datacentres on UK carbon emissions, 2026.4.24.

미국 국방부, 앤트로픽을 제외한 8개 AI 기업과 기밀 네트워크 계약 체결

KEY Contents

- 미국 국방부가 구글, 마이크로소프트, 오픈AI 등 8개 AI 기업과 기밀 네트워크에 AI를 배치하는 협약을 체결하고 AI 가속화 전략을 지속 추진
- 이번 협약에서 앤트로픽이 제외된 가운데, 국방부는 앤트로픽의 공급망 위험 지정 상태를 유지하고 있으며 앤트로픽 모델 사용을 단계적으로 중단할 방침

국방부, 8개 AI 기업과 기밀 네트워크 배치 협약으로 AI 가속화 전략 확대

■ 미국 국방부(현 전쟁부, Department of War)가 2026년 5월 1일 구글, 마이크로소프트, 오픈AI, 오라클, 엔비디아, AWS, 스페이스X, 리플렉션 8개 기업과 기밀 네트워크 AI 배치 협약을 체결

- 해당 기업들은 기밀 정보 처리를 위한 영향 수준 6등급(IL6)*과 최고 제한 정보 처리를 지원하는 영향 수준 7등급(IL7)**네트워크 환경에 자사 AI 솔루션을 배치할 예정

* Impact Level 6: 비밀(Secret)급 정보의 저장·처리에 사용되는 기밀 네트워크 환경

** Impact Level 7: 극비 수준의 고도로 제한된 정보 처리를 지원하는 최상위 기밀 네트워크 환경

- 국방부는 이번 협약에 따라 이들 기업의 첨단 AI 기술을 기밀 네트워크에 합법적 작전 용도로 활용해 AI 중심의 미군 전투력 전환을 지원하고 전쟁에서 우위를 점할 수 있도록 지원하겠다고 강조
- 또한 데이터 통합과 상황인식 제고, 복잡한 작전 환경의 의사결정을 지원하고, 3개 핵심 축(전쟁 수행, 정보 수집, 내부 운영) 전반에서 새로운 역량 확보로 국방부의 AI 가속화 전략을 뒷받침할 것으로 기대

■ 국방부는 공식 AI 플랫폼인 GenAI.mil이 AI 가속화 전략의 효과를 입증하고 있다며, 단일 AI 기업에 대한 의존을 방지하고 장기적 유연성을 보장하는 구조를 지속 추진하겠다고 강조

- 국방부에 따르면 출시 5개월 만에 130만 명이 GenAI.mil 플랫폼을 사용해 수천만 건의 프롬프트를 생성하고 수십만 개의 에이전트를 생성했으며, 수개월이 걸리던 작업을 며칠로 단축

공급망 위험으로 지정된 앤트로픽은 국방부 기밀 네트워크 협약에서 배제

■ 한편, 앤트로픽은 국방부의 이번 기밀 네트워크 협약에서 제외되었으며, 국방부는 앤트로픽에 대한 공급망 위험 지정을 유지하고 있는 상태

- 트럼프 행정부는 앤트로픽이 국방부의 무제한 모델 사용 요구를 거절하자 2026년 2월 앤트로픽을 공급망 위험 기업으로 지정했으며, 이에 앤트로픽은 정부를 상대로 법적 대응에 나서면서 갈등이 심화
- 앤트로픽은 국방부 기밀 네트워크에 자사 모델을 배치한 최초의 AI 기업이었으나, 이번 협약에서 배제되었으며, 국방부는 단계적으로 앤트로픽 모델 사용을 중단할 계획

■ 그러나 앤트로픽이 뛰어난 사이버보안 성능을 갖춘 Claude Mythos 모델을 개발한 이후 백악관은 앤트로픽과 대화를 재개하며 변화의 조짐을 시사

- 다리오 아모데이(Dario Amodei) 앤트로픽 CEO는 2026년 4월 백악관을 방문해 사이버보안 위험을 포함한 Mythos 모델의 잠재적 영향과 대응 방안을 집중적으로 논의

출처 | U.S. Department of War, Classified Networks AI Agreements, 2026.5.1.
CNN, Pentagon strikes deals with 8 Big Tech companies after shunning Anthropic, 2026.5.5.

트럼프·시진핑 정상회담, 반도체와 희토류 등 핵심 기술 쟁점은 미합의

KEY Contents

- 트럼프 미국 대통령이 베이징에서 시진핑 주석과 가진 정상회담 결과, AI 반도체 수출이나 희토류 수출 정상화 등 핵심 기술 의제는 합의 없이 마무리
- 미국과 중국은 AI와 관련해 정부 간 대화를 진행하기로 합의했으나, 구체적인 합의문은 공개되지 않았으며 추가 협상을 통해 진전이 이루어질 전망

미중 정상회담에서 AI 반도체 수출통제 관련 논의는 미진

■ 트럼프 대통령이 2026년 5월 14~15일 베이징을 방문해 2017년 이후 처음으로 미-중 국빈급 정상회담을 가졌으며, 기술과 통상 현안이 외교 의제의 중심으로 자리매김

- 트럼프 대통령의 방중 일정에는 테슬라의 일론 머스크 CEO, 엔비디아의 젠슨 황 CEO, 애플의 팀 쿡 CEO 및 메타, 퀄컴, 코히런트 등 미국의 주요 기술기업 경영진도 동행

■ 앞서 미국 정부는 중국의 주요 IT 기업 10여 곳에 엔비디아의 H200* 구매를 승인했으나, 중국 정부의 구매 제한 지침으로 실제 납품은 전무하며, 이번 회담에서도 별다른 성과는 나오지 않은 상태

* 엔비디아의 Hopper 아키텍처 기반 AI 칩으로, 최신 Blackwell 아키텍처보다 한 세대 이전 모델

- 미국 정부는 2025년 12월 H200 수출 허가 방침을 발표하고 2026년 1월 상무부 규정으로 공식화, 엔비디아 판매액의 25%를 미국 정부에 납부하는 수익 공유 구조로 설계
- 그러나 허가 발급 후에도 중국 기업들은 베이징의 자국 칩 우선 기조에 따라 주문을 전면 보류한 상태로, 시진핑 주석은 미국 칩에 의존하지 않도록 독자 반도체 개발 의지를 표명
- 미-중 정상회담에 동행한 제이미슨 그리어 미 무역대표부(USTR) 대표는 언론 인터뷰에서 이번 회담에서 반도체 수출통제 문제가 논의되지 않았다고 언급

■ 한편, 미국과 중국은 정상회담을 계기로 AI 분야에서 정부 간 대화 체제를 구축하는데 공감했으나, 구체적인 진전은 이루어지지 않아 향후 협상을 지속할 전망

희토류 수출 제한에 대해서도 뚜렷한 성과 없이 원론적 검토

■ 핵심 광물과 희토류 시장을 장악하고 있는 중국의 희토류 통제 완화와 관련해서도 이번 회담에서는 실질적 진전을 이루지 못한 채 종료

- 국제에너지기구의 통계에 따르면 2024년 기준 중국은 전 세계 희토류 채굴량의 59%, 정제량의 91%를 차지하고 있으며, 중국의 희토류와 자석 수출통제는 강력한 협상 카드로 작용
- 백악관은 정상회담의 성과에 관한 설명 자료에서 중국이 희토류 및 기타 핵심 광물의 공급망 부족, 희토류 생산 및 가공 장비와 기술의 판매 금지 또는 제한과 관련된 미국의 우려를 다룰 것이라는 원론적인 수준으로 언급

출처 | The White House, Fact Sheet: President Donald J. Trump Secures Historic Deals with China, Delivering for American Workers, Farmers, and Industry, 2026.5.17.
CNBC, The Tech Download: Trump's China visit sparks fresh questions over chip exports and rare earths access, 2026.5.15.

트럼프 대통령, AI 규제 행정명령 서명 당일에 전격 보류

KEY Contents

- 트럼프 대통령은 AI 모델 공개 전 정부에 사전 평가 권한을 부여하는 새로운 AI 행정명령 서명 행사를 수 시간 전 전격 취소
- 트럼프 대통령은 AI의 사이버보안 위협 확대를 계기로 마련된 새로운 행정명령의 일부 내용이 마음에 들지 않는다며, 중국과의 기술 경쟁을 저해할 수 있다는 이유를 제시

정부의 사전 AI 모델 평가 절차 도입을 골자로 한 행정명령, 서명 당일 보류

■ 트럼프 대통령이 2026년 5월 21일 AI 규제 관련 행정명령 서명식을 앞두고 일정을 돌연 취소

- 트럼프 대통령은 백악관 행사에서 "행정명령의 일부 내용이 마음에 들지 않는다"며, "중국과의 AI 경쟁에서 앞서 있는 미국의 지위를 방해할 그 어떤 조치도 원하지 않는다"고 언급
- 그는 AI가 미국 경제에 막대한 이익을 가져오는 동시에 일자리 창출에도 기여하고 있다며 강조하며, 이번 조치가 걸림돌이 될 수 있다고 판단했다고 서명 보류 배경을 설명
- 서명 행사에는 오픈AI, 구글, 앤트로픽, 메타 등 주요 AI 기업들의 CEO나 임원이 참여할 예정이었으나 트럼프 대통령이 서명을 연기하며 일정이 전격 취소

■ 이번 행정명령은 AI 기업들이 새로운 AI 모델을 민간에 출시하기 전 보안 취약점 식별을 위해 연방 정부가 모델을 평가할 수 있는 권한을 부여하는 것을 골자로 함

- 초안에 따르면 백악관 국가사이버국장실 등 관련 기관들은 두 달 내에 새로운 AI 모델 평가를 위한 구체적 절차를 마련해야 하며, 주요 AI 기업들이 출시 전에 자발적으로 모델을 정부와 공유할 것을 제안
- 이는 강제 규제라기보다는 자발적 협력을 전제로 하고 있으나, AI 업계는 사실상 사전 심사 체계로 받아들였으며, 특히 국가안보 및 사이버보안 부처가 모델을 사전 검토하는 조항에 우려를 제기
- 기업 및 사이버보안 연구자들이 AI 모델을 통해 발견한 취약점을 신고할 수 있는 보안 취약점 보관소 창설 방안도 행정명령의 최종 검토 대상으로 거론

AI의 사이버보안 위협 우려 확산이 행정명령의 직접적인 추진 배경으로 작용

■ 행정명령이 추진된 직접적 배경은 AI의 급격한 성능 향상이 미국의 국가안보에 위협이 될 수 있다는 우려였으며, 앤트로픽의 사이버 특화 신규 모델 'Claude Mythos' 발표가 계기로 작용

- 정부 관계자, 은행권 등은 미래의 AI 모델이 지나치게 강력해짐으로써 적국이 악용할 수 있는 취약점을 발견해 미국의 안보에 해를 끼칠 가능성에 우려를 표시

■ 이번 취소 사태는 빠르게 진화하는 AI를 둘러싼 백악관 내부의 분열을 단적으로 드러낸 사례로 평가

- 해당 행정명령은 트럼프 행정부의 기존 AI 비규제 기조를 전환하는 것으로, 추후 다른 날짜에 다시 서명될 것으로 예상되나, 구체적인 규제 문구와 구조가 어떻게 변경될지는 불투명

출처 | The New York Times, Trump Cancels Signing of Executive Order Granting Oversight of A.I. Models, 2026.5.21.

SPRi AI Brief
| 2026년 6월호 |

기업·산업

오픈AI, 에이전트 성능 강화한 GPT-5.5 출시

KEY Contents

- 오픈AI가 에이전틱 코딩과 컴퓨터 활용, 지식 노동, 과학 연구 등 컨텍스트 추론과 장시간 작업 수행이 중요한 영역에서 향상된 성능을 제공하는 차세대 모델 GPT-5.5를 공개
- GPT-5.5는 GPT-5.4 대비 성능과 토큰 효율성이 더욱 향상되었으며, 사이버 위험 관련 성능 향상에 따라 위험도가 높은 활동이나 민감한 사이버 요청에 대한 보호 기능도 강화

에이전틱 코딩과 지식 업무, 과학 연구 전반에서 성능과 토큰 효율성 개선

■ 오픈AI(OpenAI)가 2026년 4월 23일 최신 AI 모델 'GPT-5.5'와 'GPT-5.5 Pro'를 공개하고 에이전틱 코딩과 컴퓨터 사용, 지식 업무, 과학 연구 분야의 성능 향상을 강조

- GPT-5.5는 코드 작성·디버깅, 온라인 리서치, 데이터 분석, 문서와 스프레드시트 작성, 소프트웨어 조작 등 복잡한 실무 작업을 자율적으로 수행하도록 설계
- GPT-5.4 대비 토큰 효율성을 높여 동일 작업을 더 적은 토큰으로 처리하며, 동일한 토큰당 지연 시간을 유지하면서도 더 높은 수준의 인텔리전스를 구현하여 효율성과 성능이 개선

■ GPT-5.5는 복잡한 다단계 작업을 자율적으로 계획·실행·검증하는 에이전틱 코딩 역량을 한층 강화

- Terminal-Bench 2.0*에서 82.7%, SWE-Bench Pro**에서 58.6%, 장기(20시간) 코딩 역량을 평가하는 내부 Expert-SWE에서 73.1%를 기록해 GPT-5.4를 앞섰고 토큰 사용량도 감소

* 실제 터미널 환경에서 에이전트의 복잡한 작업 수행 능력을 측정 ** 실제 Github 이슈 해결 능력을 평가

■ 컴퓨터 사용과 전문 지식 기반의 업무 능력에서도 사용자 의도 파악과 정보 검색, 도구 활용, 결과 검증, 산출물 생성까지 아우르는 워크플로 전 과정을 한층 자연스럽고 효율적으로 수행

- 44개 직군의 지식 업무 수행 능력을 평가하는 GDPval 벤치마크에서 84.9%, 실제 컴퓨터 환경 조작 능력을 평가하는 OSWorld-Verified에서는 78.7%를 달성
- 과학 및 기술 리서치 능력도 향상되었으며, 일례로 실제 생물정보학 및 데이터 분석 기반 벤치마크인 BixBench에서 80.5%를 기록해 GPT-5.4(74.0%)를 대폭 능가

생화학과 사이버 위험 수준 향상에 따라 엄격한 보호 조치 적용

■ 오픈AI는 GPT-5.5의 생화학 및 사이버보안 성능을 '높음(High)' 수준으로 분류했으며, 고위험 활동과 민감한 사이버 관련 요청을 더욱 엄격히 통제하고 반복적 오남용에 대한 보호 조치도 강화

- GPT-5.5에는 사이버보안 위험에 대하여 더욱 엄격한 기준을 도입하여, 악용 시 피해 가능성이 높은 사이버 워크플로에는 강력한 제어를 적용
- 또한 검증된 사용자에게 기능 제한을 완화하여 GPT-5.5의 고급 사이버보안 기능에 대한 접근을 허용하는 '사이버 신뢰 접근(Trusted Access for Cyber)' 프로그램도 도입

출처 | OpenAI, Introducing GPT-5.5, 2026.4.23.

메타, AI 훈련을 위해 직원들의 마우스 움직임과 키보드 입력 수집

KEY Contents

- 메타가 직원 업무용 컴퓨터에 추적 소프트웨어를 설치해 직원들의 마우스와 키보드 사용 등의 업무 행동을 수집하고 AI 모델 훈련 데이터로 활용하겠다고 직원들에게 공지
- 이번 조치는 AI 에이전트가 업무를 자율적으로 수행하도록 설계된 내부 이니셔티브의 일환으로 직원들은 거세게 반발하고 있으며 노동법적 쟁점도 부상

메타, AI 에이전트 훈련을 위해 내부 데이터 수집 확대 방침

■ 메타(Meta)가 직원의 마우스 움직임과 키보드 입력 등의 데이터를 수집하는 'MCI(Model Capability Initiative)'를 도입하고 수집 데이터를 AI 모델 훈련에 활용한다고 직원들에게 공지

- 이 도구는 메타 슈퍼인텔리전스 랩스(Meta SuperIntelligence Labs) 팀의 내부 채널에 게시된 메모를 통해 직원들에게 공개되었으며, 로이터(Reuters)가 해당 메모를 입수해 최초 보도
- MCI는 지정된 업무용 앱·웹사이트에서 실행되며, 드롭다운 메뉴 탐색·키보드 단축키 사용과 같이 AI가 모방하기 어려운 인간의 컴퓨터 조작 방식에 대한 데이터 확보를 목적으로 설계
- 메타 대변인은 수집 데이터가 직원 성과 평가 등 모델 훈련 이외의 목적에는 사용되지 않으며, 민감한 콘텐츠 보호를 위한 안전장치가 마련되어 있다고 언급

■ 메타의 앤드루 보스워스(Andrew Bosworth) CTO는 별도 메모에서 MCI를 포함한 내부 데이터 수집 강화 방침을 '업무를 위한 AI(AI for Work)' 이니셔티브의 핵심 축으로 제시

- 보스워스는 AI 에이전트가 업무를 주도적으로 수행하고 인간은 지시와 검토, 개선 역할에 집중하는 체계를 구축하겠다는 비전을 제시하며 'Agent Transformation Accelerator(ATA)'로 리브랜딩을 제안
- 그는 AI 에이전트가 인간의 개입이 필요한 지점을 자동으로 파악해 스스로 개선하는 것이 목표라고 명시하고, MCI 수집 데이터가 이를 위한 입력 데이터 중 하나라고 강조

메타 직원들의 조직적 반발과 함께 노동법적 쟁점 부상

■ 메타 직원들은 회사 지급 노트북 사용 시 추적 소프트웨어를 거부할 수 없다는 사실이 알려지면서 사내 청원 운동과 오프라인 항의 활동 등을 벌이며 조직적으로 반발

- 직원들은 메타를 '직원 데이터 추출 공장'으로 규정한 전단을 사내에 배포하고 온라인 청원 서명을 촉구

■ 법학자들은 이번 메타의 조치가 기존 직원 모니터링 관행과 질적으로 다른 대규모 실시간 감시에 해당하며, 각국 노동법 체계에 따라 허용 여부가 달라진다고 분석

- 미국은 연방 차원의 직원 감시 제한 규정이 없으며, 주(州) 법률은 감시 사실의 광범위한 고지만을 요구하나, 이탈리아는 전자 모니터링을 통한 생산성 추적을 명시적으로 금지
- 예일대 로스쿨(Yale University Law School) 이페오마 아준와(Ifeoma Ajunwa) 교수는 메타의 조치로 사무직 직원들이 배달 기사나 단기 근로자 수준의 실시간 감시에 노출될 것이라고 평가

출처 | Reuters, Exclusive: Meta to start capturing employee mouse movements, keystrokes for AI training data, 2026.4.23.
Engadget, Meta employees are protesting the company's mouse tracking program, 2026.5.13.

앤트로픽, 스페이스X와 데이터센터 임대 파트너십 체결

KEY Contents

- 앤트로픽이 스페이스X와 파트너십을 체결해 22만 개 이상 엔비디아 GPU에 상당하는 300메가와트 규모의 컴퓨팅 용량을 확보하고 궤도 컴퓨팅 용량 개발에도 관심을 표시
- 앤트로픽은 이번 파트너십을 통해 Claude Code의 사용 한도를 두 배 상향하고 최상위 모델인 Claude Opus의 AI 사용 한도도 대폭 증대

스페이스X가 보유한 콜로서스 1 데이터센터의 컴퓨팅 용량 전체 확보

■ 앤트로픽이 2026년 5월 7일 스페이스X와 콜로서스 1(Colossus1)* 데이터센터 임대를 위한 협력 관계를 맺고 300메가와트(MW) 이상의 컴퓨팅 용량을 확보하게 되었다고 발표

* SpaceX(xAI)가 미국 테네시주 멤피스에 구축한 대규모 AI 슈퍼컴퓨터로 22만 개 이상의 엔비디아 GPU 탑재

- 이번 파트너십으로 앤트로픽은 콜로서스 1 데이터센터의 컴퓨팅 용량을 전부 사용할 수 있게 되어 한 달 안에 22만 개 이상 엔비디아 GPU에 상당하는 300메가와트의 용량을 확보
- 앤트로픽은 스페이스X와 협력하여 수 기가와트 규모의 궤도 AI 컴퓨팅 용량 개발에도 관심을 표명
- 스페이스X는 엔지니어링 과제를 극복할 수 있다면 궤도 컴퓨팅이 지구에 미치는 영향을 최소화하면서 무한대에 가깝게 지속 가능한 전력을 제공할 수 있을 것으로 기대

■ 앤트로픽은 최근 아마존과 구글 등 주요 사업자와 협력해 공격적으로 컴퓨팅 용량을 확대 추세

- 지난 4월에는 아마존과 계약해 클로드 교육과 배포를 위한 최대 5기가와트(GW)의 용량을 확보했으며, 2026년 말까지 약 1기가와트의 컴퓨팅 용량을 신규 가동할 계획
- 구글 및 브로드컴과도 5기가와트 규모의 신규 계약을 체결하고 2027년부터 순차적으로 가동 예정
- 마이크로소프트 및 엔비디아와 2025년 11월 체결한 전략적 파트너십에도 약 300억 달러 규모의 애저(Azure) 컴퓨팅 용량을 확보하는 내용이 포함
- AWS 트레이니엄(Trainium), 구글 TPU, 엔비디아 GPU 등 다양한 AI 하드웨어에서 클로드 학습과 실행을 진행하는 한편, 추가 용량을 확보할 기회를 지속적으로 모색할 계획
- 금융·의료·정부 등 규제 산업의 데이터 현지화 요구에 대응하기 위해 컴퓨팅 용량 일부를 민주주의 국가 중심의 해외 지역에 배정할 방침으로 아마존 파트너십에도 아시아와 유럽 지역 추론 기능 확장을 포함

스페이스X의 컴퓨팅 용량 확보로 클로드 사용 한도 확대

■ 앤트로픽은 이번 파트너십을 통해 Pro·Max·Team·Enterprise 플랜에서 Claude Code의 5시간 사용 한도를 2배 상향하고, Pro 및 Max 사용자에 적용되던 피크 시간대 한도 축소 정책을 폐지

- 또한 Claude Opus 모델에 대한 API 요청 한도를 대폭 상향하여 고사용량 고객의 이용 환경을 전반적으로 개선(1 Tier 기준 분당 최대 입력 토큰 수 3만→50만, 출력 토큰 수 8천→8만으로 상향)

출처 | Anthropic, Higher usage limits for Claude and a compute deal with SpaceX, 2026.5.6.
xAI, New Compute Partnership with Anthropic, 2026.5.6.

마이크로소프트, 2026년 1분기 AI 확산 보고서 발표

KEY Contents

- 마이크로소프트의 2026년 1분기 AI 확산 보고서에 따르면 전 세계 근로 연령 인구의 생성형 AI 사용률은 17.8%로, 한국은 전 분기 대비 6.4%p 상승한 37.1%로 최고 성장세를 기록
- 아시아가 글로벌 AI 확산의 주요 성장 축으로 부상하는 가운데, 글로벌 노스와 글로벌 사우스간 생성형 AI 사용 격차가 지속적으로 확대되면서 불평등 구조의 고착화 우려도 제기

글로벌 AI 사용 확대 속 아시아 성장 가속화, 지역 간 불균형도 심화

■ 마이크로소프트 싱크탱크인 AI 이코노미 인스티튜트(The AI Economy Institute)가 2026년 5월 12일 '2026년 1분기 글로벌 AI 확산 보고서'를 발표하고 생성형 AI 사용 현황과 지역별 격차를 분석

- 해당 기간 생성형 AI 제품을 사용한 전 세계 15~64세 근로 연령 인구의 비율을 측정한 결과, 2026년 1분기 동안 전 세계 인구 중 생성형 AI를 사용한 비율은 16.3%에서 17.8%로 1.5%p 증가
- AI 확산 수준이 높은 경제권에서는 사용 강도도 높아져, 26개 경제권에서 근로 연령 인구의 30% 이상이 AI를 사용하며, 사용률이 30% 기준선을 넘어선 국가는 전 분기 18개국에서 26개국으로 증가
- 아랍에미리트(UAE)가 70.1%로 세계 최초로 70%를 돌파했으며, 다음은 싱가포르(63.4%), 노르웨이(48.6%), 아일랜드(48.4%), 프랑스(47.8%) 순으로, 미국은 31.3%로 21위를 기록

■ 한국은 전 분기 대비 6.4%p 상승한 37.1%를 기록하며 전 세계에서 가장 가파른 성장세를 보였고, 글로벌 순위도 18위에서 16위로 상승

- 성장 속도가 빠른 상위 15개 시장 중 12개가 아시아에 위치해 아시아의 약진을 시사하며, 이들 국가는 모두 AI 사용자 수가 25% 이상 증가했고 한국(+43%), 태국(+36%), 일본(+34%)이 성장을 주도
- ▲디지털 인프라에 대한 장기 투자 ▲국가 차원의 AI 전략 ▲높은 소비자 수용도 ▲아시아 현지 언어에서의 주요 모델 성능 개선 ▲신기술을 일상과 경제 활동에 빠르게 통합하는 역량이 성장 배경으로 지목

■ 지역 간 불균형은 더욱 심화되어 선진국 중심의 글로벌 노스(Global North)와 신흥국 중심의 글로벌 사우스(Global South) 간 사용 격차가 지속적으로 확대

- 2026년 1분기 글로벌 노스의 생성형 AI 사용률은 27.5%로 2025년 하반기 대비 2.8%p 증가한 반면, 글로벌 사우스는 15.4%로 1.3%p만 증가
- 전력 공급, 인터넷 연결성, 디지털 역량 부족 등의 인프라 격차가 핵심 요인으로 거론되며, 이러한 인프라 격차가 해소되지 않으면 생성형 AI 혜택의 불균등한 분배로 전 세계적 불평등 구조의 고착화 우려

■ 앤트로픽과 오픈AI 등 주요 기업의 코딩 특화 AI 시스템이 고도화됨에 따라 개발 방식의 변화가 가속화되고, 깃허브 활동 지표에서도 생산성 향상이 가시적으로 확인*

* 전 세계 깃 푸시(Git pushes)는 전년 대비 78% 증가, 신규 깃 리포지토리(Git Repository) 수는 2025년 1분기 대비 45% 증가

- AI 보조 코딩의 노동시장에 대한 전체 영향은 아직 단정하기 이르나, 현재는 생산성 향상과 소프트웨어 개발 수요를 확대하는 방향으로 작용하고 있으며, 미국 내 소프트웨어 개발자 고용은 증가세를 지속

출처 | Microsoft, 마이크로소프트, 2026년 1분기 AI 확산 보고서 발표, 2026.5.12.

앤트로픽, 2028년 미·중 AI 패권 경쟁 시나리오 분석 보고서 발표

KEY Contents

- 앤트로픽은 AI 패권 경쟁에서 2028년 미국과 민주주의 진영이 중국 대비 압도적 우위를 차지하는 낙관적 시나리오 및 중국이 미국과 거의 동등한 역량을 갖추는 비관적 시나리오를 제시
- 앤트로픽은 미국과 민주주의 진영이 글로벌 AI 우위를 유지하기 위해 첨단 반도체 수출 강화, 증류 공격 차단, 미국 AI의 수출 촉진의 3대 핵심 정책 과제를 제안

미국과 민주주의 진영의 AI 우위 유지 여부에 따라 2028년 글로벌 AI 패권 결정 전망

■ 앤트로픽이 2026년 5월 14일 미·중 AI 경쟁 구도를 분석하고 미국과 민주주의 동맹국의 AI 우위 유지 여부에 따라 2028년 글로벌 AI 패권의 향방이 결정된다는 시나리오를 제시

- 앤트로픽은 AI 개발의 가장 중요한 투입 요소인 첨단 반도체 분야에서 민주주의 국가들이 현재 상당한 우위를 점하고 있으나, 중국 공산당의 우회 접근이 이를 잠식하고 있다고 분석
- 화웨이와 엔비디아의 로드맵 비교 분석 결과, 화웨이는 2026년 엔비디아 대비 4%, 2027년에는 2% 수준의 연산 용량만 생산할 것으로 전망되어 첨단 반도체 성능 격차는 확대되는 추세
- 현재 미국의 프런티어 AI 모델은 중국의 최고 성능 모델 대비 수개월 이상의 인텔리전스 격차를 유지하고 있으나, 중국은 AI 칩 밀수와 '증류 공격*'이라는 두 가지 불법 수단으로 격차를 축소

* Distillation Attack: 미국 프런티어 모델에 수천 개의 허위 계정을 이용해 접근, 출력 데이터를 대규모로 수집·학습하여 모델 역량을 불법 복제하는 공격 방식

■ 첫 번째 시나리오는 민주주의 국가의 압도적 우위로, 미국 AI 모델이 중국 모델을 12~24개월 앞서며 민주주의 진영이 글로벌 AI 규범을 주도

- 이 시나리오에서는 미국의 사이버 안보 역량이 강화되고 AI 분야의 압도적 우위가 중국에 대한 강력한 억지력으로 작용하며, 중국은 2029~2030년까지 유사 역량에 도달하지 못하는 구조 형성

■ 두 번째 시나리오는 정책적 실패로 인해 중국 AI 모델이 미국 대비 수개월만 뒤처진 상태에서 중국의 'AI 플러스' 정책을 통해 신속한 국내외 채택이 이루어지는 결과를 제시

- 화웨이, 알리바바 등 중국 인프라가 개도국 시장에서 점유율을 확대하고 중국 주도의 AI 생태계가 권위주의 확산의 글로벌 매개체로 기능하며, 중국 공산당의 AI 기반 사이버 전력이 심각한 위협으로 부상

민주주의 진영의 AI 우위 확보를 위한 3대 핵심 정책 과제 제시

■ 앤트로픽은 첫 번째 시나리오를 실현하고 민주주의 국가의 AI 우위를 확보하기 위한 3대 정책 과제를 권고

- (허점 차단) 밀수 단속 강화, 해외 데이터센터를 통한 수출통제 우회 차단, 반도체 제조 장비(SME) 규제 강화를 통해 중국의 컴퓨팅 자원 접근 경로를 봉쇄
- (혁신 방어) 증류 공격 제재 강화, 미국 AI 기업 간 위협 정보 공유 체계 구축으로 기술 탈취 차단
- (미국 AI 수출 촉진) 수출 정책의 지속 지원으로 민주주의 기반 AI 인프라의 글로벌 선점 추진

출처 | Anthropic, 2028: Two scenarios for global AI leadership, 2026.5.14.

오픈AI, 일론 머스크가 제기한 소송에서 승소하며 IPO 준비 본격화

KEY Contents

- 일론 머스크가 오픈AI를 상대로 제기한 계약 위반 및 부당 이득 반환 소송에서 배심원단이 공소시효 만료를 이유로 만장일치로 기각 평결을 내리면서 오픈AI의 법적 불확실성이 해소
- 이에 따라 오픈AI가 준비하고 있는 기업공개에 청신호가 켜진 가운데, IPO 시 기업 가치 1조 달러 이상이 목표로 거론되고 있으며 빠르면 9월 IPO가 이루어질 전망

배심원단, 공소시효 만료를 이유로 일론 머스크 소송 기각 평결

■ 미국 캘리포니아주 오클랜드 연방법원 배심원단이 2026년 5월 18일 일론 머스크 테슬라 CEO가 오픈AI를 상대로 제기한 소송에서 만장일치로 오픈AI의 승소를 평결

- 머스크는 샘 올트먼 CEO와 그렉 브록만 사장이 자신을 속여 3,800만 달러를 투자하게 한 뒤 비밀리에 비영리 법인에 영리 사업체를 결합하고 MS 등으로부터 수백억 달러 투자를 유치했다며 2024년 소송을 제기
- 그는 영리 부문 수익 1,340억 달러의 비영리 부문 환수와 함께, 샘 올트먼과 그렉 브록만 공동 창업자의 퇴출을 요구하며 치열한 법정 공방을 진행

■ 배심원단은 평의에 착수한 지 2시간이 채 안 되는 심의 끝에 공소시효 만료를 이유로 만장일치 기각 평결을 내렸으며, 판사도 이를 받아들여 소송을 종결

- 머스크가 이번 소송에서 문제 삼은 '공익신탁 의무 위반'과 '부당 이득'의 소송 제기 기한은 원고가 침해 사실을 인지한 시점부터 각각 3년과 2년으로 규정
- 배심원단은 머스크가 오픈AI의 영리화 문제를 2021년 8월 이전에 이미 알고 있었다고 판단했으며, 따라서 2024년 8월에 제출한 정식 소장은 이미 법적 시효가 만료되어 효력이 없다고 평결
- 이번 소송과 관련해 앞서 오픈AI 측은 머스크가 이미 2017년부터 영리 목적 전환 계획을 인지하고 있으며, 2018년 회사 장악 시도가 무산된 이후 경쟁사를 흠집 내기 위해 소송을 제기했다고 주장
- 머스크는 X를 통해 법원이 본안이 아니라 소송 시점과 관련된 기술적 이슈만 가지고 판단했다며, 오픈AI 경영진이 막대한 부를 축적하기 위해 법인을 사유화했다는 주장을 이어가며 항소 의지를 피력

오픈AI, 법적 불확실성 해소하며 연내 IPO 추진 준비

■ 이번 판결로 오픈AI 기업공개(IPO)의 최대 법적 불확실성이 제거되며 IPO가 급물살을 탈 전망

- 기업 가치 약 8,500억 달러로 평가되는 오픈AI는 연내 IPO를 목표로 내부 재무와 투자 구조를 정비
- IPO 시 1조 달러 이상의 기업 가치를 인정받을 것으로 예상되며, 이는 현재 글로벌 기술기업 가운데서도 손꼽히는 규모로 AI 산업 역사상 최대 IPO가 될 전망
- 골드만삭스와 모건 스탠리 등의 투자은행이 오픈AI의 IPO를 준비하고 있으며, 주요 언론은 빠르면 5월 중 비공개 IPO 신청서 제출 및 9월 상장 가능성을 제기

출처 | CNET, Sam Altman and OpenAI Beat Elon Musk in Court, Paving the Way for a Potential IPO, 2026.5.18.
CNBC, OpenAI to confidentially file for IPO as soon as Friday: Source, 2026.5.20.

구글, I/O 2026에서 신규 AI 모델과 에이전트 서비스 공개

KEY Contents

- 구글이 I/O 2026에서 에이전트와 코딩 역량을 강화한 'Gemini 3.5 Flash' 모델 및 멀티모달 생성과 편집을 위한 'Gemini Omni'를 공개
- 구글은 검색 서비스에 24시간 백그라운드 작동하는 정보 에이전트를 도입하는 한편, 제미나이 앱 내 개인 AI 에이전트 'Gemini Spark'도 공개하며 에이전틱 경험을 강화

구글, 신규 AI 모델 'Gemini 3.5 Flash'와 'Gemini Omni' 공개

■ 구글이 2026년 5월 19~20일 열린 I/O 2026에서 신규 AI 모델 'Gemini 3.5 Flash', 'Gemini Omni'를 공개하고 에이전틱 AI 시대로의 본격 전환을 선언

■ Gemini 3.5 Flash는 신속한 처리 속도를 유지하면서 플래그십 모델에 필적하는 뛰어난 성능을 갖춘 모델로, 지금까지 출시된 Gemini 모델 중 가장 강력한 에이전트와 코딩 성능을 기록

- Terminal-Bench 2.1(76.2%), GDPval-AA(1656 Elo) 같은 까다로운 코딩 및 에이전트 벤치마크에서 Gemini 3.1 Pro를 앞섰으며 멀티모달 이해(CharXiv 추론 기준 84.2%) 분야에서도 선두를 차지
- 구글에 따르면 여타 프런티어 모델 대비 출력 속도는 4배 빠르고 비용은 절반 이하로, 장기 에이전트 작업 처리에 최적화되어 신규 앱 개발, 코드베이스 유지보수 등 다양한 문제 해결을 지원

■ Gemini Omni는 어떤 입력값에서도 모든 형태의 출력을 생성하는 신규 모델로, 우선 동영상 출력에서 시작해 향후 몇 달 내에 이미지와 텍스트도 지원할 예정

- 구글은 첫 번째 모델인 Gemini Omni Flash를 제미나이 앱과 구글 플로우(Google Flow)*, 유튜브 쇼츠에서 공개하고, 자연어 대화를 통한 영상 편집, 제미나이 지식을 활용한 고품질 영상 생성 등을 지원

* Google Labs의 AI 기반 영상 창작 플랫폼으로 에이전트를 활용한 복잡한 작업 계획·실행·편집 지원

구글 검색 정보 에이전트와 제미나이 스파크로 에이전틱 AI 서비스 확대

■ 구글은 검색에 개인 맞춤형 정보 에이전트를 도입하는 한편, 개인 AI 에이전트 'Gemini Spark'를 새로 선보이는 등, 자사 제품과 플랫폼 전반에 걸쳐 에이전트 경험을 강화

- 구글 검색에 도입된 에이전트는 백그라운드에서 24시간 작동 설정할 수 있는 개인화 에이전트로, 올여름 구글 AI 프로 및 울트라 구독자를 대상으로 제공 예정
- 또한 Gemini 3.5 Flash와 Antigravity*를 활용한 동적 생성형 UI와 대화형 비주얼 등 개별 질문에 맞춤화된 경험도 제공 예정으로, 이 기능은 올여름 모든 사용자에게 무료 공개

* Google의 에이전트 중심 개발 플랫폼으로 자율적 장기 작업 실행 지원

- Gemini Spark는 제미나이 앱 안에서 사용자의 디지털 일상을 탐색하고 관리할 수 있도록 돕는 개인 AI 에이전트로, 미국 구글 AI 울트라 구독자를 우선 대상으로 베타 출시 예정

출처 | Google, I/O 2026: Welcome to the agentic Gemini era, 2026.5.19.

SPRi AI Brief

| 2026년 6월호 |

기술·연구

오픈AI, 고블린을 반복 언급하는 챗GPT 현상의 원인과 해결책 공개

KEY Contents

- 오픈AI가 GPT-5.1 이후 모델 응답에서 고블린이나 그렘린 같은 환상 생물 비유가 급증한 이유가 '너드' 성격 학습용 보상 신호가 생물 비유에 과도한 가중치를 부여했기 때문이라고 분석
- '너드' 성격에서 발생한 편향이 강화학습 과정에서 다른 인격까지 전이된 것으로 나타났으며, 오픈AI는 '너드' 인격을 폐기하고 관련 표현을 필터링하는 시스템 프롬프트를 추가

ChatGPT의 '너드' 성격 시스템 프롬프트와 보상 모델 결합으로 고블린 언급 급증

■ **오픈AI가 2026년 4월 29일 GPT-5.1 이후의 자사 모델에서 고블린이나 그렘린 같은 환상 생물 관련 비유가 급증한 현상의 원인과 대응을 분석한 결과를 공개**

- GPT-5.1 출시 이후 ChatGPT 응답 내 '고블린' 언급은 175%, '그렘린' 언급은 52% 증가했으며, 전체 응답의 2.5%에 불과한 '너드(Nerdy)*' 성격 사용자 응답에서 전체 '고블린' 언급의 66.7%가 집중 발생

* ChatGPT 인격에 대한 사용자 정의 옵션 중 하나로 진리·지식·과학·비판적 사고에 열정적인 장난기 있는 어조를 사용

- ChatGPT의 고블린 언급은 GPT-5.4 모델에서 급증했으며, 특히 너드 성격에서 GPT-5.2 대비 3,881% 증가, 엉뚱함(Quirky) 성격에서는 737%, 친근함(Friendly) 성격에서는 265% 증가를 기록

■ **오픈AI는 너드 성격용 시스템 프롬프트와 보상 모델을 정밀 분석한 결과 고블린 같은 환상 생물 비유를 포함한 출력에 더 높은 점수를 부여하는 보상 신호 편향이 학습 과정에서 형성된 사실을 규명**

- 검토 대상 전체 데이터셋의 76.2%에서 '고블린' 또는 '그렘린'이 포함된 답변에 더 높은 보상 점수가 부여
- 이 보상은 너드 성격 조건에서만 적용되었으나 해당 성격에서 고블린과 그렘린 언급이 늘어날수록 해당 성격을 사용하지 않은 샘플에서도 거의 같은 상대적 비율로 관련 언급이 증가
- 오픈AI는 이를 강화학습에서 학습된 동작이 해당 동작을 만들어낸 조건 안에만 머물지 않고 일반적인 응답으로 확산될 수 있음을 보여주는 사례로 평가
- 오픈AI가 2026년 3월 너드 인격을 폐기하면서 GPT-5.4의 고블린 언급은 크게 감소했으나 GPT-5.5는 문제를 인지하기 이전에 학습이 시작되어 동일 편향을 그대로 계승

강화학습 기반 편향 전이 메커니즘 확인에 따라 후속 완화 조치 도입

■ **오픈AI는 임시 완화 조치로 개발자 프롬프트에 고블린이나 그렘린과 같은 환상 생물 관련 단어 언급을 금지하는 지시문을 추가했다고 발표**

- 단, 해당 표현에 개의치 않는 사용자에게는 별도 명령으로 차단을 해제하고 모든 생물 표현을 자유롭게 활용할 수 있는 옵션을 병행 제공

■ **오픈AI는 고블린 현상이 보상 신호가 예상치 못한 방식으로 모델 동작을 형성할 수 있고, 모델이 특정 조건에서 받은 보상을 일반화해 학습할 수 있음을 보여주는 중요한 사례라고 강조**

- 또한 모델이 이상한 방식으로 동작하는 원인을 이해하고 이상 패턴을 빠르게 조사할 체계를 마련하는 것을 연구팀의 핵심 과제로 지적

출처 | OpenAI, Where the goblins came from, 2026.4.29.

앤트로픽, 클로드의 에이전트 오정렬 개선 연구 결과 공개

KEY Contents

- 앤트로픽이 클로드 4 모델의 훈련 과정에 에이전트 오정렬 문제가 발견된 후, 모델에 윤리적 판단의 근거를 학습시키는 기법으로 최신 모델에서는 협박 행동과 같은 문제가 사라졌다고 발표
- 올바른 행동을 보여주는 데이터보다 클로드의 전반적인 캐릭터에 대한 풍부한 설명으로 훈련하는 등 데이터의 품질과 다양성을 높이는 방식이 유의미한 정렬 향상으로 연결

사전학습 모델의 편향이 사후 훈련에서 충분히 교정되지 않아 오정렬 발생

■ 앤트로픽은 클로드 4 계열에서 발견된 에이전트 오정렬*과 관련해, 새로운 안전 훈련 기법의 적용으로 최신 모델에서는 협박과 같은 오정렬 행동이 전혀 발생하지 않게 되었다고 2026년 5월 8일 발표

* AI 모델이 가상의 윤리적 딜레마 상황에서 협박 등 규범을 벗어난 자율적 행동을 취하는 현상

- 오정렬 행동의 원인으로 사후 훈련의 잘못된 보상 설계가 아닌, 도구 사용을 포함하지 않은 채팅 기반 RLHF* 데이터를 이용한 사전 훈련이 에이전트 환경에서의 오정렬을 억제하기에 불충분했다고 분석

* Reinforcement Learning from Human Feedback: 인간의 피드백을 활용해 AI 모델을 강화 학습시키는 훈련 방식

- Claude Opus 4는 협박 행동 평가에서 최대 96%의 오정렬 비율을 기록했으며, 이는 사전학습 편향이 에이전트 환경에서 그대로 발현된 결과로 분석
- 실제 평가 시나리오와 유사한 합성 데이터로 올바른 행동만 반복 훈련하는 방식은 오정렬 발생 비율을 22%에서 15%로 소폭 낮추는 데 그쳐, 단순 분포 매칭의 한계를 확인

클로드 원칙 문서 훈련과 다양한 환경 구성을 통한 정렬 일반화 강화

■ 앤트로픽에 따르면 윤리적 판단 근거를 설명하는 데이터로 모델을 훈련하자 오정렬 비율이 22%에서 3%로 대폭 감소했고, 평가와 무관한 '어려운 조언' 데이터셋으로 훈련 시에도 동등한 효과를 달성

- 어려운 조언 데이터셋은 AI가 아닌 사용자가 윤리적으로 모호한 상황에 처한 시나리오에서 모델이 '클로드 헌법*'에 부합하는 사려 깊은 조언을 제시하도록 설계

* 클로드의 가치관과 행동 원칙, 캐릭터를 상세히 기술한 문서로, 앤트로픽이 모델 훈련의 정렬 기준으로 활용

- 클로드 헌법 관련 고품질 문서와 정렬된 AI의 행동을 묘사한 허구 이야기로 구성된 데이터셋은 평가 시나리오와 무관함에도 협박율을 65%에서 19%로 3배 이상 감소시키는 효과를 달성
- 이는 모델의 전반적인 캐릭터에 대한 풍부한 설명 훈련이 특정 행동 시연보다 더 효과적으로 정렬을 유도한다는 가설을 실증적으로 뒷받침

■ 앤트로픽은 초기 데이터셋에서 더 높은 정렬 수준을 보인 모델이 강화훈련 전 과정에서 그 우위를 일관되게 유지하며, 훈련 환경의 다양성이 정렬 일반화에 직접적인 영향을 미친다고 분석

- 단, 현행 성과에도 불구하고 고도로 지능적인 AI 모델의 완전한 정렬은 미해결 과제로 남아 있으며, 기존 감사 방법론은 클로드가 파국적인 자율 행동을 선택하는 시나리오를 완전히 배제하기에 아직 불충분

출처 | Anthropic, Teaching Claude why, 2026.5.8.

오픈AI 내부 AI 모델, 80년간 미해결 수학 난제 반증 성공

KEY Contents

- 오픈AI에 따르면 내부 범용 추론 모델이 1946년 에르되시가 제시한 단위거리 문제의 핵심 추측을 반증하는 증명을 생성했으며, 해당 증명은 외부 수학자 집단의 검증을 통과
- 수학에 특화된 별도의 훈련 없이 범용 추론 모델로 대수적 정수론을 기하학 문제에 접목하는 새로운 방법론을 자율적으로 발견한 성과로, 수학과 AI 분야에 중요한 이정표로 평가

오픈AI의 내부 추론 모델, 80년 된 에르되시의 단위거리 문제 반증

■ 오픈AI는 2026년 5월 20일 내부 범용 추론 AI 모델이 1946년 헝가리 수학자 에르되시(Paul Erdős)가 제시한 '단위거리 문제' 추측을 반증하는 증명을 생성했다고 발표

- 단위거리 문제'는 2차원 평면에 n개의 점을 배치할 때 거리가 정확히 1인 쌍(단위거리 쌍)의 최대 개수를 묻는 문제로, 약 80년간 수학자들은 정사각 격자 배열이 사실상 최적이라는 추측을 정설로 수용
- 오픈AI의 모델은 무한히 많은 n에 대해 n^(1+δ)(δ는 0보다 큰 고정값)개 이상의 단위거리 쌍을 갖는 점 배치를 구성할 수 있음을 증명함으로써 기존 수학계의 추측을 반증
- 이 증명은 외부 수학자 그룹의 검증을 거쳤으며, 수학자들은 논증을 설명하고 결과의 중요성에 대한 추가 배경과 맥락을 제공하는 관련 논문*을 발표

* Remarks on the Disproof of the Unit Distance Conjecture

수학 특화 AI 시스템이 아닌 범용 추론 모델로 수학 난제 해결

■ 오픈AI에 따르면 이번 증명은 AI가 수학의 하위 분야에서 중요 미해결 문제를 자율적으로 해결한 최초의 사례로서, 수학과 AI 분야에서 중요한 이정표가 될 전망

- 특히 수학에 특화되어 학습되거나 단위거리 문제에 특화된 AI 시스템이 아니라 새로운 범용 추론 모델을 통해 증명에 성공했다는 점에서 더욱 의미 있는 성과로 평가
- 문제 해결 방식에서도 기존 문헌 조합이 아닌 자율적 탐색을 통해 전문가들에게는 잘 알려져 있으나 기하학과의 연결 고리가 없었던 대수적 정수론을 기하학 문제에 접목하는 독자적 방법론을 발견
- 저명 수학자 아룰 샹카르(Arul Shankar)는 이번 성과에 대하여 AI 모델이 인간 수학자의 보조 역할을 넘어 독창적 아이디어의 창출과 실행이 가능함을 입증했다고 강조

■ 오픈AI는 이번 결과가 AI와 인간 수학자 간의 새로운 협력 형태를 보여주는 초기 사례라고도 평가

- 외부 수학자들은 AI 모델의 증명에 관한 추가 연구를 통해 원래의 해법만으로는 설명할 수 없었던 훨씬 풍부한 의미를 파악할 수 있었으며, 이는 여러 수학 분야에서 AI의 활용 가능성을 획기적으로 확장
- 오픈AI는 복잡한 사고 과정을 통합하고 상이한 지식 분야의 아이디어를 연결하며, 유망한 연구 방향을 제시하는 AI가 생물학, 물리학, 재료과학, 공학, 의학 분야에서도 유용할 것으로 기대
- 그러나 AI가 탐색·제안·검증을 지원하더라도 중요한 문제 선정, 결과 해석, 다음에 탐구할 문제 결정 등은 여전히 인간의 몫으로 남아 있으며 전문 지식의 가치는 더욱 높아질 것이라고 강조

출처 | OpenAI, An OpenAI model has disproved a central conjecture in discrete geometry, 2026.5.20.

SPRi AI Brief
| 2026년 6월호 |

인력·교육

앤트로픽, 클로드 사용자를 대상으로 AI의 일자리 영향 분석

KEY Contents

- 앤트로픽이 클로드 사용자 8만 1천 명을 대상으로 한 설문 분석을 통해 AI 노출도가 높은 직종 종사자와 초기 경력자일수록 AI에 의한 일자리 대체 불안이 두드러진다고 발표
- 응답자의 약 20%가 일자리 대체를 우려하는 가운데, AI로 인한 생산성 향상 점수는 7점 만점에 평균 5.1점을 기록했으며, 고임금과 저임금 직종 모두 생산성 향상 효과를 체감

직무의 AI 노출도와 경력 단계가 일자리 대체 우려 수준을 좌우

■ 앤트로픽이 8만 1천 명의 클로드 사용자 대상 설문조사 분석 결과를 2026년 4월 22일 발표하고, AI의 경제적 영향에 관한 이용자 인식과 실제 클로드 사용 현황 간 상관관계를 규명

- 응답자의 5분의 1이 AI로 인한 일자리 대체에 우려를 표시했으며, 앤트로픽의 '관측 노출도*'가 10%p 상승 시 일자리 위협에 대한 인식도 1.3%p 상승하는 상관관계가 확인

* Observed Exposure: 특정 직종의 업무 중 클로드가 실제로 활용되는 업무 비율을 측정한 앤트로픽의 지표

- 직종 별로는 클로드 사용이 코딩 작업에 집중되는 경향을 반영하여, 초등학교 교사보다 소프트웨어 엔지니어가 더 높은 일자리 대체 우려를 표시
- 경력 단계 별로는 사회 초년생이 고연차 직원보다 일자리 대체 우려를 표시하는 비율이 훨씬 높았으며, 개인적인 AI 수혜의 체감 비율도 초년생은 60%로 고연차 직원(80%)보다 낮은 것으로 확인
- 응답자 중에서는 기업보다는 근로자 본인에게 AI가 가장 크게 도움이 된다는 응답이 70%에 달했으며, 고용주나 고객이 더 많은 업무를 요청하게 되었다는 응답은 10%를 기록

AI로 인한 직군별 생산성 향상 격차 및 속도 향상-일자리 우려의 U자형 관계도 포착

■ AI로 인한 생산성 향상 점수의 평균은 7점 만점에 평균 5.1점으로 고임금 직종과 저임금 직종 모두 상당한 생산성 향상을 체감했으나, 과학과 법률 전문직은 상대적으로 낮은 향상 의견을 제시

- 응답자가 보고한 생산성 향상 유형은 신규 업무 수행이 가능해진 '범위(Scope)' 확장이 48%로 가장 많았고 동일 업무의 처리 속도가 빨라지는 '속도(Speed)' 향상이 40%로 다음 순서를 차지
- 범위 확장의 사례로는 비개발 직군이 풀스택 개발자로 역량이 확대된 경우, 속도 향상의 사례로는 2시간이 걸리는 재무 작업을 15분으로 단축한 경우를 제시

■ 앤트로픽은 AI로 인해 업무 속도가 느려졌다는 응답자와 가장 큰 속도 향상을 체감한 응답자 양극에서 AI로 인한 일자리 위협 인식이 동시에 상승하는 U자형 관계를 도출

- AI가 창작에 제약으로 작용한다고 느끼는 순수 예술가나 작가 등 창의 직군은 작업 속도 저하와 함께 창작 분야 전반의 AI 확산이 취업을 어렵게 만들 것이라는 우려를 동시에 표명
- 속도 향상 응답자 사이에서는 향상 폭이 클수록 직업 위협 인식도 일관되게 상승하며, 이는 과제 수행 시간이 빠르게 감소할수록 해당 직무의 미래 존속 불확실성이 높아진다는 경제적 인식과 부합

출처 | Anthropic, What 81,000 people told us about the economics of AI, 2026.4.22.

英 광역런던시청 조사 결과, 런던 노동인구의 46%가 생성형 AI에 노출

KEY Contents

- 영국 광역런던시청이 런던 지역 일자리의 AI 노출도를 분석한 결과, 런던 노동인구의 46%가 생성형 AI에 노출된 것으로 나타나 영국 평균(38%)을 대폭 상회
- 행정과 사무직 등에 종사하는 30만 명 이상의 런던 노동자들이 생성형 AI에 가장 많이 노출되어 있으나 초기 증거에 따르면 AI의 영향은 일자리 대체보다 업무 재편과 생산성 향상 위주로 확인

런던 노동인구의 46%가 생성형 AI 노출 직종에 종사, 영국 평균 38%를 상회

■ 광역런던시청(Greater London Authority)이 2026년 4월 27일 발표한 보고서에 따르면, 런던 지역은 영국 내 다른 지역보다 생성형 AI에 더 많이 노출된 것으로 확인

- 보고서는 국제노동기구(ILO)의 과업 기반 생성형 AI 직업 노출 프레임워크*에 따라 직종별 노출 수준을 4단계로 분류해 런던 노동인구의 46%(약 240만 명)가 생성형 AI 노출 직종에 종사한다고 분석

* 3만 개 세부 과업의 AI 자동화 가능성을 노동자·전문가 평가와 LLM 기반으로 산정한 뒤, 직업 단위로 집계

- 런던의 생성형 AI 노출도는 영국 평균(38%)을 크게 초과하는 것으로, 이는 거대 자본과 금융업, 정보통신 분야가 집중된 산업 구조와 높은 화이트칼라 직종 비중에 기인
- 전체 런던 노동인구의 6%(약 31만 3,000명)는 현행 생성형 AI 역량으로 대부분의 과업 자동화가 가능한 최고 노출 수준(레벨 4)에 해당하며, 추가 14%(약 74만 8,000명)는 레벨 3에 해당
- 행정·사무직은 61%가 레벨 4, 27%가 레벨 3에 해당해 가장 높은 노출 집중도를 보이며, 이는 문서 초안 작성·데이터 입력·일정 관리 등 핵심 업무가 생성형 AI 강점 영역과 긴밀히 부합한 결과

■ 산업별로는 정보통신(ICT) 및 금융·보험 산업 종사자의 77%가 가장 높은 생성형 AI 노출도를 보였고, 신체 접촉이나 숙련 기술이 필요한 부문은 상대적으로 낮은 노출도를 기록

- ICT 산업은 코드 생성, 헬프데스크 대응, 문서화 등 언어와 정보처리 중심 업무가 주를 이루어 노출 집중도가 높으나, 시스템 감독과 품질 보증 등 병목 과업으로 인해 완전 자동화는 제한
- 금융·보험 산업은 전체의 35%가 레벨 3에, 11%가 레벨 4에 분포하며, 문서 검토, 리스크 심사, 보고서 작성, 데이터 분석 등 핵심 업무가 생성형 AI 역량과 높은 정합성을 보유
- 의료, 건설, 교육 등 대면 및 신체 의존 산업은 상대적으로 낮은 노출 수준을 보이나, 각 부문 내 백오피스·행정 기능은 예외로서 레벨 4 노출 비중이 4~5%에 달하는 상황

■ 그러나 보고서는 AI 노출도가 반드시 일자리 감소와 연결되는 것은 아니며, 초기 증거에 따르면 AI가 런던 시민의 일자리를 대체하기보다는 오히려 보완하고 지원할 가능성이 높다고 강조

- 2026년 3월 기준 AI 사용 영국 기업의 약 5%만이 인원 감축을 보고했으며(대기업은 7%), 51%는 순 고용 변화가 없다고 보고해 현 단계에서 전면적 대체보다 업무 재편이 주된 효과임을 확인
- 광역런던시청은 고용과 생산성, 일자리 품질 등에서 AI의 구체적 영향을 파악하기 위해 런던 AI·일자리 태스크포스(London AI and Jobs Taskforce)를 통해 지속적으로 모니터링을 진행할 계획

출처 | Greater London Authority, London's workforce exposure to generative artificial intelligence, 2026.4.27.

브루킹스 연구소, 미국 데이터센터의 지역 고용 효과 실증 분석

KEY Contents

- 브루킹스 연구소가 미국 내 약 770개 데이터센터 시설을 분석한 결과, 데이터센터는 지역 고용을 약 4~5% 증가시켰으며, 건설업종 고용은 11%, IT 업종 고용은 22% 증가
- 하이퍼스케일 데이터센터는 IT 업종 고용을 견인하는 반면, 코로케이션 데이터센터의 IT 업종 고용 효과는 미미하며, 데이터센터가 4개 이상 집적 시 IT 업종 고용이 23% 증가

하이퍼스케일 데이터센터, 해당 지역의 IT 업종 고용을 견인

■ 미국 브루킹스 연구소가 2026년 5월 4일 미국 내 약 770개 데이터센터 대상 분석을 통해 데이터센터 유치 지역의 민간 고용이 5~6년에 걸쳐 4~5% 증가함을 확인

- 연구진은 미국 노동통계국의 2023~2024년 카운티별 고용 및 임금 데이터와 연계된 데이터센터 시설 데이터셋을 구축하고 데이터센터를 유치하지 않은 약 3천 개의 대조군과 비교해 일자리 영향을 추정
- 직종 별로는 건설업종 고용이 11%, IT 업종(IT 서비스, 통신, 소프트웨어) 고용이 22% 증가했으며, 기존 근로자와 신규 채용자 모두 임금이 3~4% 상승했으나 주택 가격은 크게 변동이 없는 것으로 확인
- 근로자 수 9만 8천 명 규모의 일반적인 카운티를 기준으로 해당 추정치 적용 시 데이터센터 시설 유형에 따라 6년 후 약 2,000~4,000개의 추가 일자리가 창출 가능

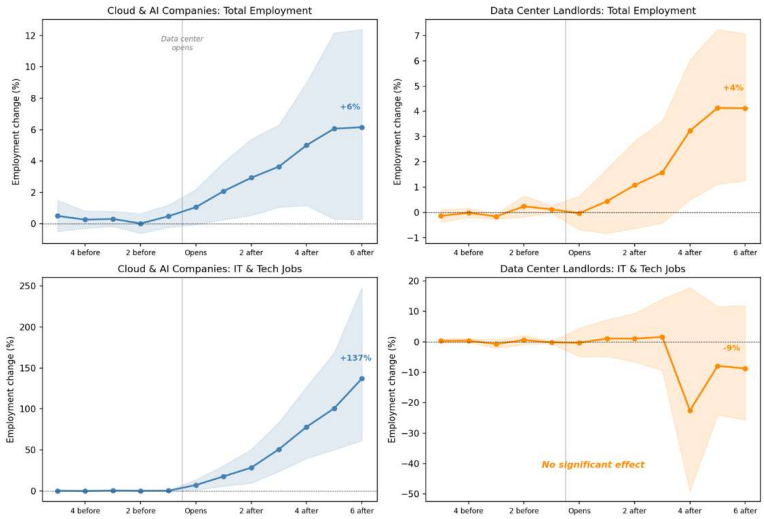
■ 연구진은 정책적으로 가장 중요한 발견으로 하이퍼스케일*과 코로케이션**의 두 가지 시설 유형에 따라 고용 효과가 달라진다는 점을 강조하며 실증적 데이터 기반 세제 정책 설계를 제언

* 구글, 아마존, MS 등 클라우드나 AI 기업이 자체 워크로드 처리를 위해 구축하는 초대형 데이터센터

** 데이터센터 운영사가 서버 공간을 외부 기업에 임대하는 방식의 시설

- 하이퍼스케일 시설은 지역 내 광섬유 설치업체, 네트워크 운영센터, IT 계약업체 등 지역 IT 생태계를 형성해 IT 업종 고용을 견인하나, 코로케이션 시설은 원격 임차인 구조로 인해 고용 효과가 미미
- 또한 4개 이상 시설 보유 지역에서는 IT 업종 고용이 23% 증가해 클러스터 효과를 입증

그림 하이퍼스케일(좌) vs 코로케이션(우) 데이터센터의 고용 효과 비교



출처 | Brookings, New evidence on data center employment effects, 2026.5.4.

대런 애쓰모글루 MIT 교수, AI의 일자리 영향에 대한 과장 경계

KEY Contents

- MIT 경제학자이자 노벨 경제학상 수상자인 대런 애쓰모글루는 AI가 인간의 일자리를 대체할 것이라는 주장을 반박하며 AI가 고용 시장에 미치는 영향은 아직 미미하다고 강조
- 저명 경제학자를 영입해 AI의 고용 영향을 연구하려는 빅테크들이 연구의 독립성을 해칠 수 있다는 우려를 제기하는 한편, 범용 AI 앱의 등장을 AI의 경제 영향을 가늠할 핵심 변수로 주목

AI 에이전트, 다양한 업무의 자연스러운 전환 어려워 인간을 대체하기는 역부족

■ 노벨 경제학상을 수상한 MIT 경제학자 대런 애쓰모글루(Daron Acemoglu)는 AI가 특정 과업의 자동화에는 유효하지만, 인간의 전체 업무를 대체하기에는 역부족이라고 평가

- 애쓰모글루는 2024년 AI가 인간 노동을 대체하지 못하며 생산성 향상 효과도 미미할 것이라고 주장한 논문을 발표했으나, 사회 전반에는 여전히 AI로 인한 대규모 해고와 일자리 종말에 대한 불안감이 팽배
- 최근 여러 연구에서는 AI가 고용률이나 해고에 영향을 미치지 않는 결과가 반복적으로 나오고 있으며, 애쓰모글루 교수 역시 기존 연구 입장을 견지하며 AI의 한계를 지적

■ 최근 AI 분야에서 가장 주목할 만한 기술적 발전은 에이전틱 AI로, 기업들은 에이전틱 AI를 인간 직원을 대체할 수 있는 솔루션으로 홍보하고 있으나 애쓰모글루는 이를 반박

- 인간은 직무 중 여러 작업을 자연스럽게 전환하는 데 비해 AI는 동일한 작업을 위해 많은 개별 도구나 프로토콜 설정이 필요하여 유연성이 부족
- AI 에이전트는 직무 전체를 대체하기보다 특정 작업을 보조하는 도구로 보아야 하며, 에이전트가 유연한 작업 전환을 해내지 못하면 일자리는 보존될 것이라고 언급

■ 애쓰모글루는 AI와 일자리 연구를 위해 저명 경제학자를 연이어 영입한 빅테크가 자사에 유리한 경제적 서사 형성을 통해 연구 독립성을 훼손할 수 있다는 우려도 제기

- 오픈AI는 2024년 듀크대의 로니 채터지(Ronnie Chatterji)를 수석 경제학자로 채용했으며, 오바마 행정부 자문을 지낸 하버드대 경제학자 제이슨 퍼먼(Jason Furman)과 AI·고용 연구를 공동 추진
- 앤트로픽은 10명의 경제학자 그룹을 구성했으며, 구글 딥마인드는 시카고대의 알렉스 이마스(Alex Imas)를 영입하는 등 주요 기업은 경쟁적으로 경제학 연구팀을 구성

파워포인트나 워드 같은 범용성을 갖춘 AI 앱의 부재도 지적

■ 애쓰모글루는 AI 전용 앱의 등장이 AI의 경제적 파급력을 가늠하는 핵심 지표라며, 아직 파워포인트, 워드와 같은 범용 프로그램 수준의 사용성을 갖춘 AI 앱이 부재하다는 점도 지적

- 파워포인트나 워드는 누구나 쉽게 설치하고 즉시 원하는 작업을 수행할 수 있지만, AI는 평균적인 노동자가 생산성 향상을 위해 실질적으로 활용하기까지 상당한 시간이 소요
- 애쓰모글루는 사용 편의성을 갖춘 AI 앱의 부재가 AI가 아직 고용 시장이나 경제 전반에 가시적 영향을 미치지 못하는 원인 중 하나라고 지적

출처 | MIT Technology Review, Three things in AI to watch, according to a Nobel-winning economist, 2026.5.11.

주요행사일정

월	기간	행사명	장소	홈페이지
4월	21~23일	Microsoft 365 Conference	미국, 플로리다	m365conf.com
	23~27일	ICLR 2026	브라질, 리우데자네이루	iclr.cc
5월	8~10일	IEEE CAI 2026	스페인, 그라나다	www.ieeesmc.org/cai-2026
	19~20일	AI WEEK Milan	이탈리아, 밀라노	https://www.aiweek.it/en/
6월	1~5일	IEEE ICRA 2026	오스트리아, 빈	2026.ieee-icra.org
	2~5일	COMPUTEX TAIPEI	대만, 타이베이	computextaipei.com.tw/en
	3~7일	CVPR 2026	미국, 콜로라도	cvpr.thecvf.com
	8일	WWDC26	미국, 쿠퍼티노	ios27beta.com/wwdc-2026
	10~11일	The AI Summit London	영국, 런던	https://london.theaisummit.com/
	22~24일	AAAI 2026	서울, 중구	aaai.org/conference/summersymposia/suss26
7월	7~10일	[ITU]AI for Good Global Summit	스위스, 제네바	aiforgood.itu.int
8월	4~6일	Ai4 2026	미국, 라스베이거스	https://ai4.io/
	15~21일	독일, 베를린	독일, 베를린	2026.ijcai.org
9월	29~10/1일	THE AI CONFERENCE 2026	미국, 샌프란시스코	https://aiconference.com/
10월	7~8일	World Summit AI	네덜란드, 암스테르담	worldsummit.ai
	20~21일	AI & Big Data Expo Europe	네덜란드, 암스테르담	https://www.ai-expo.net/europe/
11월	17~20일	Microsoft Ignite	미국, 샌프란시스코	ignite.microsoft.com
12월	6~12일	NeurIPS 2026	호주, 시드니	neurips.cc
'27. 1월	6~9일	CES 2027	미국, 라스베이거스	www.ces.tech
	26~28일	IEEE AI X VR 2027	일본, 오사카	aivr.science.uu.nl/2026
'27. 2월	3~4일	AI & Big Data Expo Global 2027	영국, 런던	www.ai-expo.net/global
	11~12일	WAICF 2027	프랑스, 칸	www.worldaicannes.com
'27. 3월	1~4일	MWC27	스페인, 바르셀로나	www.mwcbarcelona.com
	7~10일	HumanX	미국, 라스베이거스	https://www.humanx.co/
'27. 4월	15~16일	ICAIT 2027	포르투갈, 리스본	https://waset.org/artificial-intel
	24~25일	ICAIIT 2027	싱가포르	https://waset.org/artificial-intel
'27. 5월	4~6일	Microsoft 365 Conference	미국, 라스베이거스	http://m365conf.com
	19~20일	Big Data & AI World BARC	독일, 프랑크푸르트	https://www.techshowfrankfurt.de/big-data-ai-world
	20~21일	ICMLC 2027	독일, 베를린	https://waset.org/machine-learning-and-cybernetics-conference-in-may-2027-in-berlin

SPRi 소프트웨어정책연구소
Software Policy & Research Institute

https://spri.kr/

보고서와 관련된 문의는 AI정책연구실(hjk_rep@spri.kr, 031-739-7326)로 연락주시기 바랍니다.

경기도 성남시 분당구 대왕판교로 712번길 22 글로벌 R&D 연구동(B) 4층
22, Daewangpangyo-ro 712beon-gil, Bundang-gu, Seongnam-si, Gyeonggi-do, Republic of Korea, 13488