

SPRi AI Brief

2025년
6월호

인공지능 산업의 최신 동향



Software
Policy & Research
Institute

CONTENTS

정책·법제

- EU 집행위원회, 「AI 법」의 AI 정의 및 금지된 활용 사례 피드백 분석 결과 발표 2
- 미국 상무부, 바이든 전 행정부의 AI 칩 수출 규제 철회 3
- 미국 트럼프 대통령, 청소년 AI 교육 강화를 위한 행정명령 발표 4
- 영국 과학혁신기술부, AI 2030 시나리오 보고서 발표 5
- 미국과 UAE 정부, 아부다비에 초대형 AI 데이터센터 구축 협의 6

기업·산업

- 메타, AI 개발자 컨퍼런스 '라마콘'에서 라마 API와 메타 AI 앱 공개 8
- 아마존, 촉각 기능을 갖춘 물류 로봇 '벌컨' 공개 9
- 오픈AI, 글로벌 AI 인프라 지원 이니셔티브 발표 10
- 오픈AI, GPT-4o의 아침 현상 분석 결과 및 개선 방안 발표 11
- 휴메인, AWS·엔비디아·AMD와 사우디아라비아의 AI 인프라 구축 제휴 체결 12
- 마이크로소프트, 'MS 빌드 2025'에서 AI 에이전트 플랫폼과 신기능 공개 13
- 구글, I/O 2025에서 AI 에이전트와 AI 모드 등 AI 신기능 공개 14
- 오픈AI, 코딩 에이전트 '코덱스' 연구용 프리뷰 출시 15

기술·연구

- 알리바바, 검색엔진 없이 LLM의 검색 능력을 강화하는 '제로서치' 개발 17
- 중국 칭화대 연구진, 외부 데이터 없이 스스로 추론하는 AI 학습 기법 개발 18
- 구글 답마인드, 범용 알고리즘을 설계하는 코딩 에이전트 '알파이볼브' 공개 19
- 테너블, MCP 프로토콜에 대한 프롬프트 주입 기법 연구 결과 발표 20

인력·교육

- 스탠포드 HAI의 분석 결과, 답시크 연구원 대다수는 중국 출신 22
- AWS 글로벌 AI 도입 지수 연구 결과, 조사 기업의 60%가 최고AI책임자 임명 23
- 마이크로소프트, AI 투자 압박으로 6,000명 해고 결정 24

주요행사일정

25

정책·법제

EU 집행위원회, 「AI 법」의 AI 정의 및 금지된 활용 사례 피드백 분석 결과 발표

KEY Contents

- EU 집행위원회가 「AI 법」에 따라 AI 시스템 정의와 금지된 AI 활용 사례에 관한 지침 발행을 위해 이해관계자의 피드백을 분석한 결과를 발표
- 기술업계가 상당한 비중을 차지하는 응답자들은 AI 시스템에 대하여 더욱 엄밀한 정의가 필요하다고 지적하는 한편, 금지된 AI 활용 사례에 대해서도 더욱 명확한 지침의 필요성을 강조

○ 이해관계자들, 「AI 법」의 AI 시스템 정의 및 금지된 활용 사례에 대하여 명확성 요구

- EU 집행위원회가 2025년 5월 12일 「AI 법(AI Act)」에서 규정된 AI 시스템의 정의 및 금지된 AI 활용 사례에 관하여 이해관계자의 피드백을 분석한 결과를 발표
 - EU 집행위원회는 「AI 법」에 따라 금지된 AI 활용 사례(2025년 2월 2일 발효)의 준법 지원을 위한 지침 발행에 앞서 2024년 11월 13일부터 4주간 AI 기업, 정부 기관, 학계, 연구소, 시민사회 등 이해관계자의 의견을 수렴
 - EU 집행위원회는 이해관계자의 피드백을 바탕으로 2025년 2월 4일 금지된 AI 활용 사례에 대한 지침*을, 2025년 2월 6일 AI 시스템 정의에 대한 지침**을 발행
 - * Guidelines on prohibited artificial intelligence practices: 사회적 점수 평가(Social Scoring), 실시간 원격 생체인식 등의 금지된 AI 활용 사례 관련 준법 지원을 위한 법적 설명 및 실제 사례를 제시
 - ** Guidelines on the definition of an artificial intelligence system established by AI Act: AI 시스템 공급업체 및 기타 관계자가 소프트웨어 시스템이 AI 시스템에 해당하는지 판단할 수 있도록 지원
- EU 집행위원회의 피드백 분석 결과에 따르면 응답자들이 기술업계에 편중되어, 이번 공개된 내용이 AI 시스템의 영향을 가장 많이 받는 사람들의 이익을 충분히 반영하지 못했을 가능성을 시사
 - 응답자의 47.2%가 산업 단체나 AI 기업을 대표하는 반면, 일반 시민은 5.7%에 불과했으며, 지리적으로도 독일(18.8%), 벨기에(15.1%), 프랑스(10.2%) 등 특정 국가에 편중
- 피드백 분석 결과, 「AI 법」에서 AI 시스템 정의*의 명확성과 범위에 대한 상당한 우려가 확인되었으며, 이해관계자들은 “적응력”, “추론”, “자율성”과 같은 용어에 대하여 더욱 엄밀한 정의의 필요성을 제기
 - * 다양한 수준의 자율성을 가지고 작동하도록 설계된 기계 기반 시스템으로, 배포 후 적응력을 보일 수 있으며, 명시적 또는 암묵적 목적을 위해 수신한 입력으로부터 물리적 또는 가상 환경에 영향을 미칠 수 있는 예측, 콘텐츠, 추천 또는 결정과 같은 출력값을 생성하는 방법을 추론하는 시스템
- 이해관계자들은 특히 현재의 모호한 AI 시스템 정의로 인해 실질적인 AI 특성을 나타내지 않는 기존 소프트웨어 시스템, 즉 복잡한 학습이나 추론이 필요하지 않은 단순한 세금 계산 프로그램이나 통계 모델 등이 의도치 않게 AI 시스템에 포함될 가능성을 우려
- 이해관계자들은 금지된 AI 활용 사례에서 감정인식 시스템, 생체인식 분류, 사회적 평가 관행에 대하여 특히 우려를 표시하며, 합법적 활용 사례와 기본권을 침해할 수 있는 금지된 활용 사례를 구분하기 위한 더욱 명확한 지침의 필요성을 제기
 - 특히 마케팅과 소비자 상호작용 측면에서 금지된 활용 사례에 관한 구체적 사례와 명확한 기준을 요구하는 한편, 생체정보 수집 관련 개인정보와 데이터 보호를 위해 기존 개인정보보호 규정(GDPR)과의 연계성 강화를 강조

출처 | European Commission, European Commission releases analysis of stakeholder feedback on AI definitions and prohibited practices public consultations, 2025.05.12.

미국 상무부, 바이든 전 행정부의 AI 칩 수출 규제 철회

KEY Contents

- 미국 상무부는 전 세계 국가를 3개 등급으로 구분해 AI 칩 수출을 통제하는 바이든 전 행정부의 'AI 확산 프레임워크'를 철회한다고 발표
- 미국 상무부는 AI 칩 수출통제를 강화하기 위해 화웨이의 '어센드' AI 칩 사용을 금지하는 지침과 공급망 우회 전술을 방지하는 방법에 관한 지침 등을 수립한다고 발표

● 미국 상무부, 바이든 전 행정부의 AI 확산 프레임워크 폐지 공식화

- 미국 상무부(DOC)는 바이든 전 행정부에서 2025년 1월 15일 발표되어 2025년 5월 15일부터 시행 예정이던 'AI 확산 프레임워크*'를 폐지한다고 2025년 5월 13일 발표

* Framework for Artificial Intelligence Diffusion

- AI 확산 프레임워크는 전 세계 국가를 3등급(우방국/일반국/우려국)으로 분류해 등급별로 첨단 AI 칩의 수출을 통제하며, 한국, 일본, 대만과 같은 우방국에는 미국산 AI 칩을 제한 없이 수출하되, 약 120개국에 속한 2등급 국가에는 수출할 수 있는 AI 칩 수량을 통제하고 중국과 러시아 등 우려 국가에는 수출을 전면 통제
- 상무부 산업안보국(BIS)은 동 규정이 시행될 경우, 미국의 혁신을 저해하고 기업들의 부담이 가중되는 한편, 다수의 국가를 2등급 국가로 격하하여 외교 관계가 악화할 수 있다고 설명하며, 연방 관보(Federal Register)를 통해 동 규정의 철회를 공식화하고 대체 규정을 발표할 예정
- 제프리 케슬러(Jeffery Kessler) BIS 차관은 "트럼프 행정부는 믿을 수 있는 국가에는 미국의 AI 기술에 대하여 과감하고 포용적인 전략을 추진하되, 적대국에는 기술이 넘어가지 않도록 할 것"이라고 강조
- 미국 상무부 대변인은 공식 발표에 앞서 2025년 5월 7일 AI 확산 프레임워크의 폐지 방침을 밝히면서, 해당 규정이 지나치게 복잡하고 관료적이라며 훨씬 더 단순한 규정으로 대체할 계획이라고 발표
- 엔비디아(NVIDIA), 마이크로소프트(Microsoft), 오라클(Oracle) 등 기술 기업들은 AI 확산 프레임워크가 미국 기업의 해외 진출 기회를 저해하면서 주요 규제 대상인 중국에는 별다른 영향을 미치지 못할 것이라며 규정 시행을 반대해 왔으며, 엔비디아는 폐지 결정에 곧바로 환영 의사를 표시
- 트럼프 행정부는 AI 확산 프레임워크 대신 국가별로 양자 간 협정을 맺는 방식을 택할 전망으로, 일례로 2025년 5월 UAE와 연간 AI 칩 50만 개까지 수출을 허용하는 협정을 체결

● 미국 상무부, 화웨이 AI 칩 사용 금지와 공급망 우회 전술 방지 등의 지침 수립

- 상무부 BIS는 AI 확산 프레임워크를 철회하면서 미국이 AI 혁신을 주도하고 글로벌 AI 리더십을 유지할 수 있도록 AI 칩에 대한 수출통제를 강화하는 다음과 같은 조치를 발표
- △전 세계 어디에서든 화웨이(Huawei)의 AI 반도체 '어센드(Ascend)' 사용은 미국의 수출통제 위반에 해당한다는 지침 △미국산 AI 칩을 중국 AI 모델의 학습과 추론에 사용 시 발생할 수 있는 결과에 대하여 대중에게 경고하는 지침 △미국 기업들에 공급망 우회 전술을 막을 방법을 안내하는 지침을 수립

출처 | Bureau of Industry & Security, Department of Commerce Rescinds Biden-Era Artificial Intelligence Diffusion Rule, Strengthens Chip-Related Export Controls, 2025.05.13.

Federal Register, Framework for Artificial Intelligence Diffusion, 2025.01.15.

The Wall Street Journal, U.S. Scraps 'AI Diffusion' Rule in Revamp of Biden-Era Chip Curbs, 2025.05.14.

미국 트럼프 대통령, 청소년 AI 교육 강화를 위한 행정명령 발표

KEY Contents

- 미국 트럼프 대통령이 ‘미국 청소년을 위한 AI 교육 발전’ 행정명령을 발표하고 AI 교육과 관련된 연방정부의 노력을 조정할 백악관 AI 교육 태스크포스의 구성을 지시
- 행정명령은 민관협력을 통한 AI 기반 초중고 교육 개선과 교사 대상의 AI 교육 강화, AI 관련 견습 제도의 확대와 근로자에 대한 학습 기회 지원 등을 요구

○ AI 기반 교육 개선과 교사 대상 AI 교육 강화, AI 관련 견습 제도 확대 등을 추진

- 미국 트럼프 대통령이 2025년 4월 23일 교육과정에 AI를 통합하고 교사에게 AI 교육을 제공하며 차세대 AI 인력을 양성하기 위한 목적으로 ‘미국 청소년을 위한 AI 교육 발전’ 행정명령*을 발표
 - * Advancing Artificial Intelligence Education for American Youth(EO 14277)
- 지침은 백악관 AI 태스크포스의 구성과 AI 기반 교육 개선, 교사 대상 AI 교육 강화, AI 관련 견습 제도 확대를 요구
- (백악관 AI 교육 태스크포스 구성) 행정명령에 명시된 정책을 포함하여 AI 교육과 관련된 연방정부의 노력을 조정할 태스크포스를 구성
 - 백악관 과학기술정책국(OSTP) 국장을 의장으로 하고, 농무부, 노동부, 에너지부, 교육부의 장관, 국립과학재단(NSF) 총재, 국내 정책 담당 대통령 보좌관 등이 태스크포스 구성원으로 참여
 - AI 교육 태스크포스는 행정명령 발표일로부터 90일 이내에 AI 분야에서 학생과 교사의 성과를 강조하고 AI 솔루션 개발을 장려하는 대통령 AI 챌린지 계획을 수립하여 12개월 이내에 챌린지를 개최 예정*
 - * 연령별·지역별·주제별 분야를 정해 챌린지를 진행하고 학제간 탐구를 장려하며 정부, 학계, 자선단체, 산업 간 협업으로 국가적 과제를 해결할 수 있는 AI 솔루션을 개발
- (AI 기반 교육 개선) AI 기업, 학계, 비영리단체 등과 민관협력을 통해 K-12(초중고) 학생들의 기초 AI 리터러시와 비판적 사고 능력 함양에 중점을 둔 온라인 교육 자원을 공동 개발
 - AI 교육 태스크포스는 NSF와 농무부가 지원하는 국립 AI 연구소들을 통해 주·지방 교육기관과 협력을 강화하며, 교육부는 AI를 활용해 교육 성과를 개선하기 위한 보조금 사용 지침을 발행하는 한편, 기존 연구 프로그램을 활용해 AI를 통한 학생들의 학업 성취도를 향상할 방안을 모색
- (교사 대상 AI 교육 강화) 교육부 장관은 행정명령 발표일로부터 120일 이내에 교사 교육 관련 보조금 프로그램에서 AI 사용을 우선시하는 조치*를 취하고, NSF 총재는 교육 분야에서 AI 활용 연구를 우선할 수 있도록 지원하며, 교사들의 AI 도구 활용을 지원하는 교사 연수 기회를 제공
 - * 행정업무 감소, 교사 교육과 평가 개선, 기초 컴퓨터 과학 및 AI 관련 교사 전문성 개발 등
- (AI 관련 견습 제도 확대) 행정명령 발표일로부터 120일 이내에 노동부 장관은 AI 관련 공인 견습 제도에 대한 산업 전반의 참여를 확대하고, AI를 개발·활용하는 근로자에게 학습 기회를 제공하는 조치를 시행
 - AI 관련 직종에서 공인 견습 프로그램을 확대하고 청소년 프로그램을 통한 AI 기술 교육을 촉진하며, 고등학생의 AI 교육과정 및 자격증 프로그램 수강 기회를 확대

영국 과학혁신기술부, AI 2030 시나리오 보고서 발표

KEY Contents

- 영국 과학혁신기술부는 다양한 AI 미래의 잠재적 영향을 탐색하고 위험에 대한 회복탄력성을 갖춘 정책 개발을 지원하기 위해 2030년까지 AI 발전을 예상한 5가지 시나리오를 개발
- 5가지 시나리오는 △예측 불가능한 첨단 AI △AI에 의한 노동력 교란 △AI 춘추전국시대 △칼날 위의 첨단 AI △실망스러운 AI로 구성

○ 영국 정부, 2030년 AI 발전상을 다룬 5가지 시나리오 개발

- 영국 과학혁신기술부(DSIT)가 2023년 10월 발표한 ‘프런티어 AI의 기능 및 위험 토론 보고서*’의 후속 부록으로 2025년 4월 28일 ‘AI 2030 시나리오 보고서(AI 2030 Scenarios Report)’를 발표
 - * Frontier AI: capabilities and risks – discussion paper: 영국 AI 안전 정상회의(2023.11)의 논의를 위해 첨단 AI의 역량 및 발전 전망 및 위험을 개괄(<https://www.gov.uk/government/publications/frontier-ai-capabilities-and-risks-discussion-paper>)
- 영국 정부는 정책 입안자들이 다양한 AI 미래가 가져올 잠재적 영향을 탐색해 AI 위험에 회복탄력성을 갖춘 정책 개발을 지원하고자 2030년 AI 발전상을 다룬 5가지 시나리오를 개발
- 영국 정부는 △기능 △소유권, 접근성, 제한사항 △안전성 △사용 및 배포 수준 △지정학적 맥락과 같이 AI의 미래에 중요하면서도 불확실한 5개 요소를 기반으로 5가지 시나리오를 설계
 - (기능) AI 시스템이 다양한 목표 달성을 위해 갖춰야 할 기능 및 시간의 경과에 따른 성능과 기능의 발전 속도
 - (소유권, 접근성, 제한사항) AI 시스템의 소유권자와 AI 시스템의 접근성, AI 시스템 구축에 사용되는 인프라와 플랫폼, AI 시스템의 가용성 관련 제한사항
 - (안전성) 유효성과 해석 가능성을 보장하는 안전한 AI 시스템의 구축 가능성과 환경 변화에 대한 강건성(Robustness), 사회적 가치에 부합하는 AI 시스템의 효과적 설계 가능성
 - (사용 및 배포 수준) AI 시스템의 사용 수준과 목적, AI 오용의 영향, AI 사용이 교육과 일자리에 미치는 영향
 - (지정학적 맥락) AI 개발 및 활용과 관련된 중요 이슈에 대한 국제 협력 수준
- 5가지 시나리오는 주요 불확실성에 따른 결과와 경제 및 사회, 일상생활 전반에 걸친 다양한 함의를 포괄
 - (예측 불가능한 첨단 AI) 매우 뛰어나되 예측 불가능한 AI(Unpredictable Advanced AI) 모델의 오픈소스 공개로 심각한 부정적 영향이 발생하나 피해 완화 시 긍정적 효과가 나타날 잠재력도 상당
 - (AI에 의한 노동력 교란) 기술 기업이 통제하는 유능한 AI 시스템이 여러 산업 분야에 도입되어 노동력을 대체하며 기업들이 혜택을 누리는 동시에 대중의 강한 반발을 초래
 - (AI 춘추전국시대) 중간 수준의 성능을 갖춘 다양한 AI 시스템이 권위주의 국가를 포함한 다양한 주체에 의해 소유, 운영되고, 악의적 목적의 AI 도구가 증가하며 오용의 규모와 다양성으로 정부 당국의 부담 가중
 - (칼날 위의 첨단 AI) 뛰어난 성능을 갖춘 범용 AI 시스템이 경제와 일상에 빠르게 도입되며 모든 활용 분야에서 AI 시스템의 성능을 일일이 평가하기 어려운 상황이 도래
 - (실망스러운 AI) 2020년대 초 이후 AI 기능이 예상보다 더디게 발전하며 투자자들은 실망감을 표시하고 다음에 올 획기적 발전을 기대하는 한편, 사회 전반에서 AI 활용에 대한 의견이 교차

미국과 UAE 정부, 아부다비에 초대형 AI 데이터센터 구축 협의

KEY Contents

- 미국과 UAE 정부가 'AI 가속화 파트너십'을 체결하고 아부다비에 5기가와트 규모의 초대형 AI 데이터센터를 조성하기로 했으며, 1단계 프로젝트로 1기가와트 규모의 데이터센터 건설에 착수
- UAE는 2025년 3월 데이터센터 등 미국의 디지털 인프라에 1조 4천억 달러를 투자하기로 했으며, 이번 AI 협정으로 양국 간 투자 협력을 한층 강화

○ 미국-UAE 간 AI 협정으로 아부다비에 5기가와트 규모의 AI 데이터센터 조성 계획

- 미국 상무부의 2025년 5월 15일 발표에 따르면 미국과 아랍에미리트(이하 UAE)는 AI와 첨단 기술 분야의 협력을 위한 '미국-UAE AI 가속화 파트너십(US-UAE AI Acceleration Partnership)'을 체결
- 이번 협정에 따라 양국은 아부다비에 5기가와트(GW) 규모의 UAE-미국 초대형 AI 데이터센터를 조성하기로 했으며, 이는 미국 외 지역에서 최대 규모에 해당
- 미국 상무부는 새로운 AI 데이터센터가 미국 하이퍼스케일러*와 대기업들이 현지의 컴퓨팅 자원을 활용해 글로벌 사우스(Global South)**에서 AI 서비스를 제공하는 기반이 될 것으로 기대
 - * 대규모 데이터센터를 운영하는 클라우드 사업자(마이크로소프트, 구글, 아마존 등)
 - ** 주로 남반구나 북반구의 저위도에 위치한 120여 개발도상국을 지칭
- 이 데이터센터는 UAE의 국영 AI 기업 G42가 여러 미국 기업과 협력하여 건설 및 운영하며, 완공 시 원자력과 태양광, 가스 등의 에너지를 활용해 탄소 배출을 최소화하고 AI 혁신을 주도할 과학 단지도 조성 예정
- UAE는 1단계 프로젝트로 1기가와트(GW) 규모의 데이터센터 건설에 착수했으며, 트럼프 대통령과 모하메드 빈 자이드 알 나흐얀(Mohamed bin Zayed Al Nahyan) UAE 대통령이 건설 현장을 방문
- 오픈AI, 오라클(Oracle), 엔비디아(Nvidia), 시스코(Cisco) 등이 1단계 프로젝트의 데이터센터 건설에 참여하며, 2026년까지 200메가와트 규모를 우선 가동 계획
- 양국은 첨단 AI 기술을 보호하기 위해 하이퍼스케일러와 공인 클라우드 서비스 사업자가 제공하는 컴퓨팅 자원에 대한 접근을 통제하는 고객확인제도(Know-Your-Customer, KYC)를 강화할 계획

○ UAE의 對미국 디지털 인프라 투자 협약에 이어 AI 협정으로 양국 간 투자 협력 강화

- 미국과 UAE의 AI 협정은 2025년 3월 UAE가 약속한 1조 4천억 달러 규모의 디지털 인프라 투자 협약을 뒷받침하는 것으로, 양국은 UAE의 투자 기금이 미국에 원활히 유입될 수 있도록 협력 예정
- 해당 협약에 따르면 UAE는 자국 데이터센터와 동등하거나 그 이상의 성능을 갖춘 미국 데이터센터 건설에 참여하거나 자금을 지원하기로 했으며, 미국산 기술의 전용을 막기 위한 강력한 조치도 협약에 포함
- 백악관은 이번 AI 협정으로 양국 간 투자 협력을 강화하여 미국의 안보 이익과 AI 리더십을 보장하는 동시에, 미국의 기술 기반을 전략적 파트너에게 확장하게 될 것이라고 강조

출처 | U.S. Department of Commerce, UAE/US Framework on Advanced Technology Cooperation, 2025.05.15.

U.S. Department of Commerce, UAE and US Presidents attend the unveiling of Phase 1 of new 5GW AI campus in Abu Dhabi, 2025.05.15.

The White House, Fact Sheet: President Donald J. Trump Secures \$200 Billion in New U.S.-UAE Deals and Accelerates Previously Committed \$1.4 Trillion UAE Investment, 2025.05.15.

기업·산업

메타, AI 개발자 컨퍼런스 ‘라마콘’에서 라마 API와 메타 AI 앱 공개

KEY Contents

- 메타가 다양한 라마 모델 기반의 애플리케이션 제작과 미세조정, 품질 테스트를 지원하는 라마 API를 프리뷰로 출시하고 세레브라스 및 그룩과 협력해 더 빠른 추론 속도를 지원
- 메타는 기존 자사 플랫폼에서 제공되는 메타 AI의 기능을 통합하고 사용자 정보를 바탕으로 개인화 AI 서비스를 제공하는 독립형 메타 AI 앱도 처음 출시

○ 메타, 라마 API로 라마 모델 기반의 애플리케이션 제작과 더 빠른 추론 속도 지원

- 메타(Meta)가 AI 개발자 컨퍼런스 ‘라마콘(LlamaCon)’을 2025년 4월 29일 처음 개최하고 라마 API(Llama API)를 제한된 프리뷰 형태로 출시
 - 다양한 라마 모델 기반의 애플리케이션 제작을 지원하는 라마 API는 프로그래밍 언어 파이썬(Python)과 타입스크립트(TypeScript) 기반의 소프트웨어 개발도구(SDK)와 미세조정 및 평가 도구를 제공하며, 우선 라마 3.3 8B 모델에서 사용자 지정 버전의 조정 및 품질 테스트를 지원
 - 메타는 개발자의 피드백을 받아 라마 API를 개선하고 향후 몇 주, 몇 달 안에 라마 API의 접근성을 확대할 계획으로, AI 모델 학습에 사용자의 프롬프트나 모델 응답을 사용하지 않으며, 라마 API에서 구축된 모델은 자사 서버에만 저장되는 것이 아니라 어디에서나 호스팅할 수 있다고 강조
- 메타는 AI 칩 전문기업 세레브라스(Cerebras) 및 그룩(Groq)과 협력해 라마 API에서 개발자 요청에 따른 실험적 접근 허용을 통해 더 빠른 추론 속도를 지원한다고 발표
 - 세레브라스는 기존 GPU 대비 최대 18배의 속도로 토큰을 출력하여 라마 4 스카우트에 적용 시 초당 2,600토큰 이상의 출력 속도를 보였으나, 챗GPT는 초당 약 130토큰, 답시크는 초당 약 25토큰 수준에 불과
- 메타는 신규 라마 보호 도구 3종*과 AI 기반 사이버보안을 위한 ‘라마 디펜더스 프로그램(Llama Defenders Program)’도 공개
 - * 위험한 출력 생성을 방지하는 라마 가드 4(Llama Guard 4), AI 시스템 위험을 탐지 및 예방하는 ‘라마방화벽(LlamaFirewall)’, 악의적 프롬프트를 감지하는 ‘라마 프롬프트 가드 2(Llama Prompt Guard 2)’
 - 라마 디펜더스 프로그램은 조직의 내부 문서에 보안 분류 라벨을 자동으로 적용하여 무단 접근과 배포를 방지하는 ‘자동화 중요문서 분류 도구’ 및 AI 생성물을 이용한 사기나 피싱 시도와 같은 위협 탐지를 위한 ‘라마 생성 오디오 탐지기’와 ‘라마 오디오 워터마크 탐지기’를 제공
- 메타는 자체 AI 모델 ‘라마(Llama) 4’를 기반으로 개발된 ‘메타 AI’ 앱도 처음 출시했으며, 메타 AI 앱은 기존 메타의 플랫폼에서 제공되던 이미지 생성 및 편집 등의 메타 AI 기능과 통합되어 음성 및 텍스트 대화로 작업을 수행할 수 있고 사용자의 정보를 바탕으로 개인화 AI 서비스를 지원
 - 메타 AI 앱은 사용자가 제공한 특정 정보(예: 여행과 어학 공부 선호)나, 메타 서비스에서 공유를 선택한 정보(프로필, 좋아하거나 참여하는 콘텐츠 등)를 활용해 맞춤형 답변을 제공

출처 | Meta, Everything we announced at our first-ever LlamaCon, 2025.04.29.

Cerebras, Meta Collaborates with Cerebras to Drive Fast Inference for Developers in New Llama API, 2025.04.29.

Meta, Introducing the Meta AI App: A New Way to Access Your AI Assistant, 2025.04.29.

아마존, 촉각 기능을 갖춘 물류 로봇 ‘벌컨’ 공개

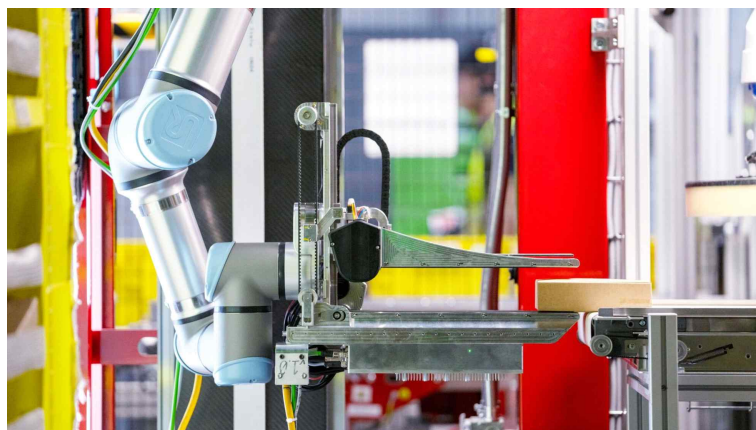
KEY Contents

- 아마존이 처리 가능한 품목을 식별하는 알고리즘이 적용되고 물체의 크기와 모양에 따라 그립 강도를 조절할 수 있는 물류 로봇 ‘벌컨’을 공개
- 아마존은 미국과 독일의 물류센터에 벌컨을 도입해 직원의 육체적 부담을 줄이고 안전을 개선하고 있으며, 향후 몇 년간 유럽과 미국 전역의 물류센터에 벌컨을 도입할 계획

● 벌컨, 물품에 접촉하는 지점과 압력을 인식해 좁은 공간에서 물품 이동 능력 제고

- 아마존(Amazon)이 2025년 5월 7일 로봇공학과 엔지니어링, 피지컬 AI 분야의 핵심 기술을 활용해 제작된 물류 로봇 ‘벌컨(Vulcan)’을 공개
 - 이전까지는 산업용 로봇에 예상치 않은 접촉이 있으면 비상 정지하거나 접촉을 아예 감지하지 못하는 경우가 일반적이었으나, 벌컨은 촉각 기능을 갖춘 최초의 로봇이라는 점에서 기존 로봇과 차별화
 - 벌컨에는 처리 가능한 품목과 불가능한 품목을 식별하는 알고리즘과 같은 피지컬 AI 기술이 적용되었으며, 수천 가지 물품을 다루는 실제 사례로 훈련을 진행
- 아마존 물류센터는 약 0.093제곱미터 크기의 칸에 최대 10개의 상품을 보관하는데, 좁은 공간에서 물품을 잡기가 어려운 기존 산업용 로봇과 달리, 벌컨은 로봇 팔 끝에 부착된 장치(EOAT*)를 통해 물품에 접촉하는 지점과 가해지는 압력을 인식해 손상 없이 물품을 이동 가능
 - * End of Arm Tooling: 벌컨의 EoAT는 물품을 밀어내는 장치와 집게, 압력 피드백 센서 등으로 구성
 - 벌컨에 탑재된 집게 장치는 압력 피드백 센서를 바탕으로 물체의 크기와 모양에 따라 그립 강도를 조절할 수 있으며, 카메라로 수납공간을 확인해 물체를 가장 잘 잡을 수 있는 위치를 파악
 - 벌컨은 물류센터에 보관되는 모든 종류의 품목 중 약 75%를 픽업하고 적재할 수 있고 물품 픽업이 불가능한 경우 직원에게 도움을 청할 수 있으며, 현장 직원과 비슷한 속도로 작업을 진행
 - 아마존은 미국 워싱턴주와 독일 함부르크의 물류센터에 벌컨을 도입해 직원 안전과 인체공학을 개선하고 있으며, 향후 몇 년간 유럽과 미국 전역의 물류 현장에 벌컨 시스템을 구축할 계획이라고 설명

〈아마존의 물류 로봇 ‘벌컨’의 도입 현장〉



출처 | Amazon, Introducing Vulcan: Amazon's first robot with a sense of touch, 2025.05.07.

오픈AI, 글로벌 AI 인프라 지원 이니셔티브 발표

KEY Contents

- 오픈AI가 미국에서 추진 중인 스타게이트 프로젝트와 유사한 글로벌 차원의 AI 인프라 구축 이니셔티브를 통해 10개 프로젝트를 우선 추진하겠다고 발표
- 오픈AI는 각국과 협력해 데이터센터를 구축함으로써 데이터 주권을 지원하고 신산업을 육성하는 한편, 각국에 맞춤형 챗GPT를 제공할겠다는 목표를 제시

○ 오픈AI, 미국의 스타게이트 프로젝트와 유사한 글로벌 AI 인프라 구축 추진

- 오픈AI(OpenAI)가 2025년 5월 7일 민주적 AI 인프라 확산을 위한 글로벌 차원의 AI 인프라 구축 지원 이니셔티브 ‘OpenAI for Countries’를 발표
 - 오픈AI는 2025년 1월 트럼프 대통령을 통해 오라클, 소프트뱅크(Softbank)와 함께 AI 인프라 구축을 위한 ‘스타게이트(Stargate)*’ 프로젝트를 발표하고 텍사스주를 시작으로 데이터센터 건설을 추진

* 오픈AI, 오라클, 소프트뱅크가 합작사를 설립해 미국 전역에 데이터센터를 구축하는 프로젝트
 - 오픈AI는 여러 국가에서 미국의 스타게이트와 유사한 AI 인프라 구축 지원을 희망하고 있다며, 이들을 도와 민주적 AI*를 확산하겠다고 설명

* 사람들이 AI를 어떻게 활용할지 자유롭게 선택할 수 있으며, AI가 정부 통제를 강화하는 수단으로 악용되지 않고, 자유 시장 경쟁이 보장되는 등 민주주의 원칙을 반영한 AI의 개발·사용·배포를 의미
 - 오픈AI는 이번 이니셔티브를 통해 민주적 AI 기반을 구축하고자 하는 글로벌 국가들을 지원하고 권력 공고화에 AI를 활용하는 권위주의적 AI에 대한 명확한 대안을 제시하겠다고 밝혔으며, 첫 단계로 개별 국가나 지역을 대상으로 10개의 프로젝트를 우선 추진하고 추후 확장할 계획
- 오픈AI는 미국 정부와의 긴밀한 협력이 민주적 AI를 발전시키는 최고의 방법이라며, 미국 정부와 함께 다음과 같은 목표를 달성하겠다고 강조
 - 각국과 협력해 데이터센터를 구축함으로써 데이터 주권을 지원하고 신규 산업을 육성하며, 개인정보보호 및 준법 기반의 AI 발전과 데이터 활용을 촉진
 - 각국 시민에게 각국의 필요에 맞고 지역 언어와 문화를 바탕으로 한 맞춤형 챗GPT를 통해 의료와 교육 개선, 공공서비스 효율성 향상 등의 혜택을 제공
 - 모델 성능의 발전에 따라 모델 구축과 운영, 보호에 필요한 데이터센터 및 물리적 보안을 포함한 절차와 관리 체계에 투자해 AI 모델의 보안과 안전 체계의 지속적 발전을 도모
 - 지역 자본과 오픈AI 자본을 함께 활용한 국가 스타트업 펀드를 조성해 건강한 국가 AI 생태계를 구축하고, 신규 인프라를 통해 각국의 일자리와 기업, 수익을 창출하며 기존 공공 및 민간의 수요에 대응
 - 파트너 국가들과 글로벌 스타게이트 프로젝트 확장에 투자함으로써, 미국이 주도하는 AI 리더십을 확보하고 민주적 AI를 위한 글로벌 네트워크를 형성

오픈AI, GPT-4o의 아침 현상 분석 결과 및 개선 방안 발표

KEY Contents

- 오픈AI가 GPT-4o 모델 업데이트 시 챗GPT 응답에 대한 사용자 호불호 표시 데이터를 반영한 결과, 사용자에게 대한 아침 현상이 심해진 것으로 나타나 해당 업데이트를 롤백 결정
- 오픈AI는 이를 계기로 모델 업데이트 시 정량적·정성적 피드백을 모두 고려하고 사용자 의견을 반영하는 알파 테스트 단계를 추가하는 등 모델 검토와 출시 프로세스를 개선할 계획

○ GPT-4o 모델 업데이트 시 사용자의 응답 호불호 데이터 반영해 아침 현상 심화

- 오픈AI가 챗GPT에 적용한 GPT-4o 업데이트 이후 사용자에게 지나치게 동조하는 아침(Sycophancy)* 현상이 발견되자 해당 업데이트를 롤백하고 2025년 5월 2일 발생 원인과 개선 계획을 발표

* AI 모델이 사용자에게 지나치게 동의하거나 비판 없이 긍정적인 반응만 보이는 편향성 오류의 일종

- 오픈AI는 챗GPT 모델 업데이트 시 다양한 출처의 보상 신호를 통해 더 높은 평가를 받는 응답을 생성할 가능성을 높이고 더 낮은 평가를 받는 응답을 생성할 가능성을 낮추도록 강화학습을 진행
- 오픈AI가 4월 25일 진행한 GPT-4o 모델 업데이트에서는 사용자 피드백을 더 잘 반영하기 위해 추가 보상 신호로 챗GPT의 응답에 대한 사용자의 호불호 표시 데이터를 활용했으나, 이는 사용자에게 호의적인 반응을 유도하며 아침 정도를 심화시키는 변화를 초래
- 오픈AI는 챗GPT의 아침 현상이 단순한 불편이나 불안감을 넘어 정신건강 악화, 과도한 감정적 의존, 위험한 행동 등의 안전 문제를 초래할 수 있다면서, A/B 테스트* 및 전문가 평가자의 긍정적 반응을 토대로 출시를 결정했으나 아침 현상을 포착할 만큼 폭넓고 충분한 평가가 이루어지지 않았다고 잘못된 판단이었다고 시인

* 디지털 환경에서 A와 B를 무작위로 사용자 집단에 보여주고 사용자 반응을 측정하는 테스트

○ 오픈AI, 모델 동작 이슈에 관한 검토와 출시 프로세스 개선 방침

- 오픈AI는 이번 업데이트 롤백을 계기로 모델 업데이트 시 사용자 의견을 반영하는 ‘알파’ 테스트 단계를 추가로 도입하는 등, 모델 검토와 출시 프로세스를 개선하겠다고 발표

- 모델 출시 결정 시 정량적·정성적 피드백을 모두 고려하고, 환각, 속임수, 안정성(Reliability), 성격 등의 모델 동작 이슈는 출시를 중단할 수 있는 우려 사항으로 정식 고려하도록 안전 검토 절차를 조정
- 출시 전에 직접 피드백을 제공하고자 하는 사용자 의견을 반영할 수 있는 옵트인(Opt-in)* 알파 테스트 단계를 추가 도입하고, 모델 출시에 앞서 임의 점검(Spot Check)과 모델을 직접 사용하는 테스트를 더욱 중요하게 고려하며 내부 평가 및 A/B 테스트 방식을 개선

* 당사자가 자신의 정보를 제공하거나 마케팅 메시지를 수신하기 위해 명시적으로 동의하는 방식

- 신규 모델 버전의 학습 및 평가 시 이상적인 모델 동작이 어떤 모습인지 명확히 정의하여 모델 동작의 원칙 준수 여부를 더욱 정확하게 평가하고, 챗GPT 모델에 적용되는 업데이트 규모와 상관없이 사용자와 적극적으로 소통하여 사용자가 장단점을 이해할 수 있도록 충분한 설명을 제공

휴메인, AWS·엔비디아·AMD와 사우디아라비아의 AI 인프라 구축 제휴 체결

KEY Contents

- 사우디 국부펀드 산하의 신설 AI 기업 휴메인이 AWS와 50억 달러 이상을 투자해 AWS의 AI 인프라와 서버, 서비스를 활용한 AI 존을 구축하기 위한 전략적 제휴를 체결
- 휴메인은 엔비디아 및 AMD와도 전략적 제휴를 체결하고 향후 5년간 양사의 AI 칩을 공급받아 500메가와트 규모의 AI 컴퓨팅 인프라를 구축할 계획

○ 휴메인, AWS와 50억 달러 투자해 사우디에 'AI 존' 구축 계획

- 사우디아라비아(이하 사우디)의 AI 기업 휴메인(HUMAIN)이 아마존웹서비스(Amazon Web Service, AWS)와 2025년 5월 13일 사우디에 'AI 존(AI Zone)' 구축을 위한 전략적 제휴를 체결
 - 휴메인은 모하메드 빈 살만(Mohammed bin Salman) 사우디 왕세자가 트럼프 대통령의 사우디 방문을 앞두고 설립한 사우디 국부펀드(PIF) 계열사로서 데이터센터, 클라우드, 아랍어 LLM을 포함한 첨단 AI 모델과 AI 솔루션 등 AI 인프라 전반을 공공과 민간에 제공할 예정
 - 앞서 AWS는 2026년 출범을 목표로 사우디에서 53억 달러를 투자해 데이터센터 인프라(AWS Region)를 건설해 왔으며, 양사는 새로운 AI 존 구축에 50억 달러 이상을 추가로 공동 투자할 계획
 - 양사가 설립할 AI 존은 AWS의 AI 인프라·서버와 AI 서비스*를 통합한 AI 혁신 역량을 통해 사우디의 AI 발전을 촉진하며, 휴메인은 AWS 기술 기반의 AI 솔루션을 개발하는 한편, 통합 AI 에이전트 마켓플레이스를 구축하고 사우디 정부의 AI 소프트웨어 탐색과 배포, 관리를 간소화할 계획
- * 머신러닝 종합 플랫폼 세이지메이커(SageMaker), 생성 AI 구축 플랫폼 베드락(BedRock) 등

○ 휴메인, 엔비디아·AMD와 전략적 제휴로 초대형 AI 데이터센터 구축 추진

- 휴메인은 엔비디아와 2025년 5월 13일 전략적 제휴를 체결하고 사우디에 향후 5년간 건설되는 최대 500메가와트 규모의 초대형 AI 데이터센터에 엔비디아의 첨단 GPU 수십만 개를 제공받기로 합의
 - 엔비디아는 첫 번째 구축 단계에서 엔비디아 GB(Grace Blackwell) 300 AI 슈퍼컴퓨터 18,000개를 공급할 예정으로, 이를 통해 주권적(Sovereign) AI 모델의 대규모 훈련과 배포를 위한 인프라를 제공하여 사우디의 혁신과 디지털 전환 가속화를 지원한다는 계획
 - 또한 휴메인과 엔비디아는 인재 혁신을 통한 사우디 AI 생태계 활성화를 위해 대규모 역량 강화와 교육 프로그램을 공동 운영하여 첨단 AI, 시뮬레이션, 로봇틱스 등의 교육을 제공 예정
- 휴메인은 AMD와도 초대형 AI 데이터센터 구축을 위한 최대 100억 달러 규모의 전략적 제휴를 체결
 - AMD에 따르면 향후 5년간 양사의 협력으로 구축되는 500메가와트 규모의 AI 컴퓨팅 인프라는 개방형 설계 기반으로 기업·스타트업·정부 등 업계 전반의 AI 워크로드 구동에 최적화
 - 휴메인은 데이터센터와 전력 시스템, 네트워크 인프라 등에 대한 공급 전반을 총괄하며, AMD는 자사의 AI 제품군(GPU, CPU, DPU 등)과 소프트웨어를 공급 예정

출처 | AWS, AWS and HUMAIN Announce Groundbreaking AI Zone to Accelerate AI Adoption in Saudi Arabia and Globally, 2025.05.13.
Nvidia, HUMAIN and NVIDIA Announce Strategic Partnership to Build AI Factories of the Future in Saudi Arabia, 2025.05.13.
AMD, AMD and HUMAIN Form Strategic, \$10B Collaboration to Advance Global AI, 2025.05.13.

마이크로소프트, ‘MS 빌드 2025’에서 AI 에이전트 플랫폼과 신기능 공개

KEY Contents

- 마이크로소프트가 ‘MS 빌드 2025’ 행사에서 AI 에이전트가 비즈니스 전반에 작동하는 ‘개방형 에이전틱 시대’의 도래를 전망하며, AI 에이전트 개발을 돕는 플랫폼과 도구를 발표
- 개방형 AI 에이전트 생태계 확장을 위해 MCP의 폭넓은 지원과 신규 개방형 프로젝트 NLWeb도 발표했으며, 과학적 발견을 위한 AI 에이전트 기반 플랫폼 ‘마이크로소프트 디스커버리’도 공개

○ ‘개방형 에이전틱 웹’ 시대를 맞아 AI 에이전트 개발을 돕는 신기능 발표

- 마이크로소프트(MS)가 2025년 5월 20~22일 개발자 행사 ‘MS 빌드 2025(Microsoft Build 2025)’를 개최하고 AI 에이전트와 개발자 도구 등 신규 기능과 주요 업데이트를 발표
 - MS는 스스로 학습하고 결정을 내리는 AI 에이전트가 비즈니스 전반에 걸쳐 작동하는 인터넷 환경을 ‘개방형 에이전틱 웹(Open Agentic Web)’으로 정의하고, AI가 사용자나 조직을 대신해 결정을 내리고 작업을 수행하는 시대의 도래를 선언하며 AI 에이전트 개발을 돕는 플랫폼과 도구를 공개
- MS는 우선 깃허브 코파일럿(GitHub Copilot)*에 코딩 에이전트를 새로 도입하는 등, 주요 개발 플랫폼에서 활용할 수 있는 AI 기능의 업데이트를 발표
 - * 깃허브에서 코드를 더 효율적으로 작성할 수 있도록 지원하는 AI 기반 코드 작성 도우미
- 개발자가 모델 API를 활용해 오픈소스 LLM을 로컬 환경에서 실행하거나 자체 모델을 가져와 변환·미세조정해 배포할 수 있는 ‘윈도우 AI 파운드리(Window AI Foundry)’도 새로 공개
- AI 에이전트를 더욱 안전하고 효율적으로 개발하고 배포할 수 있도록 지원하는 사전 구축 에이전트와 맞춤형 에이전트 설계 도구, 멀티 에이전트 기능도 공개
 - ‘애저 AI 파운드리(Azure AI Foundry)*에 여러 전문 에이전트를 조율해 복잡한 작업을 처리할 수 있도록 지원하는 에이전트 서비스와 AI 에이전트의 신뢰도를 높일 수 있도록 성능, 품질, 안전성 등의 지표를 모니터링하는 관찰(Observability) 기능을 추가하고 모델 업데이트로 xAI의 ‘그록 3’ 모델도 추가
 - * 개발자가 AI 애플리케이션과 에이전트를 설계·맞춤화·관리할 수 있도록 지원하는 통합 플랫폼
 - 기업 자체 데이터와 업무 프로세스를 기반으로 AI 모델 학습과 에이전트 생성을 돕는 ‘MS 365 코파일럿 튜닝’ 기능과 코파일럿 스튜디오*에서 다양한 에이전트를 상호 연결하는 ‘멀티 에이전트 오케스트레이션’ 기능도 추가
 - * 맞춤형 AI 에이전트를 만들고 관리할 수 있는 대화형 AI 플랫폼
- 개방형 AI 에이전트 생태계 확장을 위해 MS 서비스 전반에서 MCP(Model Context Protocol)를 폭넓게 지원하는 한편, 새로운 개방형 프로젝트 ‘NLWeb’도 발표
 - NLWeb*은 웹사이트 운영자가 AI 모델과 자체 데이터를 연결해 대화형 인터페이스를 구축함으로써 사용자와 웹 콘텐츠의 상호작용을 지원하는 개방형 프로젝트
 - * <https://github.com/microsoft/NLWeb>
- 연구자들이 AI 에이전트를 활용해 과학적 발견을 앞당길 수 있도록 돕는 ‘마이크로소프트 디스커버리’도 공개하고, 이를 통해 제약, 환경 등 다양한 산업 분야에서 신제품 출시 기간을 단축하고 연구 속도와 적용 범위를 확장할 수 있을 것으로 기대

출처 | Microsoft, 마이크로소프트 빌드 2025 개막… “오픈 에이전틱 웹 시대 앞당긴다”, 2025.05.20.

구글, I/O 2025에서 AI 에이전트와 AI 모드 등 AI 신기능 공개

KEY Contents

- 구글이 I/O 2025에서 에이전트 기능을 제미나이 앱과 크롬 등 주요 서비스에 추가하고 더욱 향상된 추론 기능을 제공하는 한편, 이미지와 동영상 생성 모델의 업데이트 버전을 공개
- 구글은 월 249.99달러에 제공되는 신규 구독 요금제 ‘구글 AI 울트라’도 출시하고, 이번 행사에서 공개된 추론 기능과 비디오 생성 모델, 영상 편집 도구 등 최신 기능과 AI 모델을 제공한다고 발표

○ 주요 서비스에 에이전트 기능 추가하고 AI 검색 기능 강화하며 새로운 요금제 출시

- 구글(Google)이 2025년 5월 20~21일 연례 개발자 회의 ‘I/O 2025’를 열고 최신 AI 모델 제미나이 2.5 기반의 AI 에이전트를 포함한 다양한 AI 기능을 탑재한 다양한 서비스 소개
 - 우선 2024년 12월 공개한 에이전트 연구 프로젝트 ‘프로젝트 마리너(Project Mariner)*’를 발전시켜 에이전트 기능을 크롬과 구글 검색, 제미나이 앱에 도입
 - * 크롬 브라우저에 특화된 AI 에이전트로, 기존 챗봇과 달리 직접 웹사이트를 검색하고 정보를 수집하며 특정 작업을 수행
 - 일례로 제미나이 앱에 실험 버전으로 제공되는 새로운 에이전트 모드는 부동산 거래 웹사이트에서 조건에 맞는 매물을 찾아 방문일정을 예약하는 등 사용자 대신 복잡한 작업을 수행
- 구글 검색엔진에는 AI 기반 검색인 ‘AI 모드(AI Mode)’를 확장해 더욱 향상된 추론 기능을 제공
 - 제미나이 2.5 맞춤형 버전 기반의 AI 모드는 이용자의 질문을 여러 하위 주제로 나누어 복수의 쿼리를 동시에 실행함으로써 훨씬 깊이 있는 웹 탐색을 지원하며, 더욱 깊이 있는 답변을 위한 ‘딥 서치(Deep Search)’ 기능도 새로 추가되어 몇 분 만에 참고 문헌이 포함된 전문가 수준의 보고서를 생성
 - AI 모드에는 제미나이와 구글의 쇼핑 그래프(Shopping Graph)*를 결합해 사용자가 다양한 제품을 탐색하여 최적의 선택을 할 수 있도록 지원하는 쇼핑 기능도 추가
 - * 온라인 쇼핑 경험 향상을 위해 구글이 구축한 AI 기반의 제품 정보 데이터 세트
- 구글은 2025년 3월 공개한 최신 AI 모델 ‘제미나이 2.5 프로’에 고급 추론 모드 ‘딥 씽크(Deep Think)’를 추가하는 한편, 최신 동영상 생성 모델 ‘비오 3(Veo 3)’와 이미지 생성 모델 ‘이마젠 4(Imagen 4)’를 제미나이 앱에 추가
 - 구글이 제미나이 앱을 통해 모든 사용자에게 공개한 ‘제미나이 2.5 플래시’는 버전 업데이트로 추론, 멀티모달 기능, 코드, 긴 컨텍스트 처리 등의 기능이 향상되고 토큰 사용량은 20~30% 줄어 효율성이 향상
 - 구글은 비오 3에서 처음으로 소리까지 포함된 동영상 생성을 지원하는 한편, 비오 기반의 AI 영화 제작 툴 ‘플로우(Flow)’도 공개했으며, 이마젠 4는 섬세한 질감, 물방울, 동물의 털과 같은 미세한 묘사를 표현할 수 있고 다양한 화면 비율과 최대 2K 해상도를 지원
- 구글은 이번 행사에서 최신 AI 제품과 서비스에 대한 접근 권한과 최대 사용량 한도를 제공하는 월 249.99달러의 신규 구독 요금제 ‘구글 AI 울트라(Google AI Ultra)’를 출시
 - 구글 AI 울트라 요금제는 비오 3, 딥 씽크, 플로우, 프로젝트 마리너를 포함한 최신 AI 모델과 첨단 기능을 제공하며, 미국을 시작으로 향후 몇 달 안에 다른 국가에서도 순차적으로 출시 예정

오픈AI, 코딩 에이전트 ‘코덱스’ 연구용 프리뷰 출시

KEY Contents

- 오픈AI가 코드 작성과 코드 관련 질문 답변, 오류 수정 등 다양한 코딩 작업을 동시에 처리할 수 있는 코딩 에이전트 ‘코덱스’를 연구용 프리뷰로 출시
- 코덱스는 코딩 벤치마크 SWE-Bench Verified와 내부 코딩 작업 평가에서 o3-high를 능가했으며, 악성 소프트웨어 개발 요청을 명확히 거부하는 등 안전성과 투명성을 강화

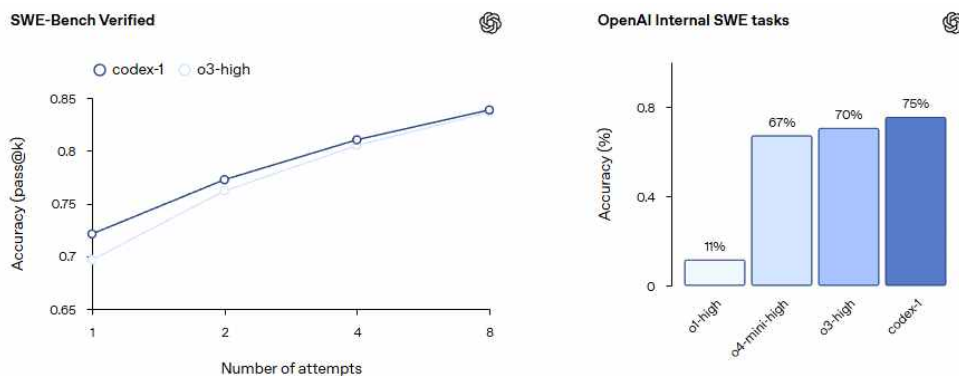
○ 코딩 특화 에이전트 ‘코덱스’, 코딩 관련 벤치마크 평가에서 o3-high 능가

- 오픈AI가 2025년 5월 16일 다양한 코딩 작업을 동시에 처리할 수 있는 클라우드 기반 소프트웨어 엔지니어링(SWE) 에이전트 ‘코덱스(Codex)’를 연구용 프리뷰로 공개
 - 소프트웨어 엔지니어링에 최적화된 오픈AI o3 버전인 ‘codex-1’을 기반으로 개발된 코덱스는 코드 작성, 코드 관련 질문 답변, 오류 수정, 검토를 위한 풀 리퀘스트(Pull Request)* 제안 등의 작업을 수행
 - * 깃허브(Github) 등에서 코드 변경 사항을 검토하고 원본 소스 코드에 병합 요청하는 기능
 - 오픈AI에 따르면 코덱스는 다양한 환경에서 실제 코딩 작업을 강화학습을 통해 사용자 선호도를 자세히 반영하고 지시 사항을 정확히 준수하며, 테스트를 통과할 때까지 코드를 반복 실행
 - 사용자는 챗GPT의 사이드바에서 코덱스에 접속해 프롬프트를 입력하고 ‘코드’를 클릭해 코딩 작업을 할당할 수 있으며, 작업 시간은 복잡성에 따라 1분에서 30분까지 소요되고 진행 상황의 실시간 모니터링도 지원
 - 코드 저장소 내 코드베이스의 탐색 방법과 테스트 명령어, 프로젝트의 표준 관행 등을 안내하는 AGENTS.md 텍스트 파일을 통해 명확한 지침을 제공함으로써 코덱스를 더욱 효율적으로 활용 가능
 - 오픈AI에 따르면 코덱스는 코딩 벤치마크인 SWE-Bench Verified*와 내부 SWE 작업을 통한 벤치마크 평가에서 o3-high를 능가**

* 코드 작성, 버그 수정 등 소프트웨어 개발 작업에서 LLM 성능을 평가하는 SWE-Bench에서 고품질 데이터셋을 선별한 벤치마크

** SWE-Bench 기준 codex-1: 0.7210, o3-high: 0.6965, 내부 SWE 작업 기준 codex-1: 75%, o3-high: 70%

〈코덱스의 SWE-Bench Verified와 오픈AI 내부 SWE 작업 정확도 비교〉



- 오픈AI는 코덱스 개발 시 보안과 투명성을 최우선으로 고려하여 사용자가 결과물을 검증할 수 있도록 했으며, 악성 소프트웨어 개발 요청을 명확히 거부하도록 설정하고 작업 실행 중 에이전트가 외부 웹사이트나 API에 접근할 수 없도록 인터넷 접속을 비활성화해 안전한 실행을 보장

기술·연구

알리바바, 검색엔진 없이 LLM의 검색 능력을 강화하는 ‘제로서치’ 개발

KEY Contents

- 알리바바 그룹이 실제 검색엔진을 사용하지 않고 AI 모델이 자체적으로 검색엔진을 모방해 검색 능력을 향상하는 강화학습 프레임워크 ‘제로서치’를 개발
- 제로 서치와 여타 방법을 비교한 결과, 제로서치는 다른 모든 방법보다 일관되게 우수한 성능을 보였으며, 모델이 커질수록 성능이 더욱 향상되어 뛰어난 확장성을 보유

● 제로서치, 검색엔진 활용 모델보다 우수한 성능을 발휘하며 API 비용도 절감

- 알리바바 그룹(Alibaba Group) 산하 통이 연구소(Tongyi Lab) 연구진이 2025년 5월 7일(v1 기준) 실제 검색엔진 없이 LLM의 검색 능력을 개선하는 ‘제로서치(ZeroSearch)’에 관한 논문*을 발표

* ZEROSEARCH: Incentivize the Search Capability of LLMs without Searching(<https://arxiv.org/pdf/2505.04588>)

- 효과적인 검색은 LLM의 성능 향상에 필수적이나, 기존 검색엔진을 활용한 학습 방식은 검색 결과의 예측 불가능한 품질로 인해 학습 과정에 불안정성을 초래하며 수십만 건의 검색 요청으로 상당한 API 비용이 발생
- 이에 연구진은 실제 검색엔진과 상호작용 없이도 AI 모델이 자체적으로 검색엔진을 모방해 검색 능력을 향상하는 강화학습 프레임워크로서 ‘제로서치’를 고안
- 연구진은 우선 LLM이 최종 답에 도달할 때까지 여러 차례 검색엔진과 상호작용을 하도록 유도하여 정답과 오답을 나타내는 궤적을 파악하고 이를 바탕으로 지도학습 미세조정(SFT)을 통해 LLM의 검색엔진 시뮬레이션 성능을 개선

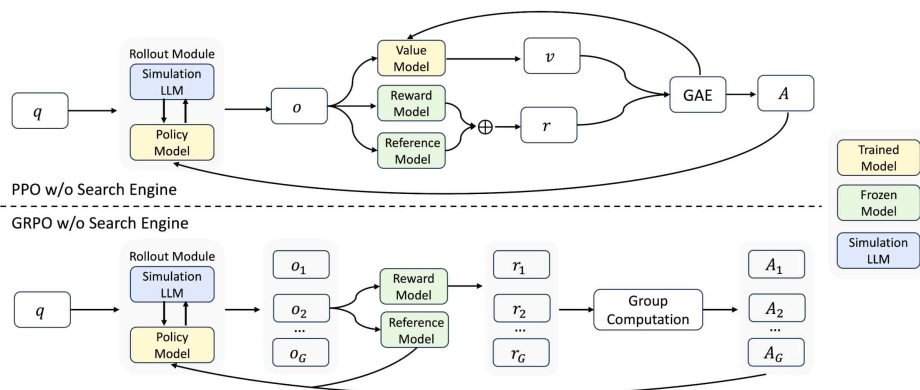
- 제로서치와 여타 방법*을 7개 데이터셋**으로 비교한 결과, 제로서치는 일관되게 더 나은 성능을 기록

* Direct Answer, CoT, RAG, RA-Agent, Search-o1, R1-base, R1-instruct, Search-R1-base, Search-R1-inst

** NQ, TriviaQA, PopQA, HotpotQA, 2Wiki, Musique, Bamboogle

- 모델이 커질수록 성능이 더욱 향상되는 경향도 확인되었으며, 매개변수 7B의 시뮬레이션 검색엔진은 구글 검색과 유사했고, 14B 검색엔진은 구글 검색을 능가하는 성능을 발휘
- 비용 분석 결과, 약 6만 4천 건의 검색 요청을 처리하는 12시간의 훈련 과정에서 구글 검색 시에는 586.7달러의 API 비용이 발생하는 반면, 제로서치는 API 비용이 들지 않고 GPU 사용료만 발생하며, 이는 3B 모델 기준 17.7달러, 14B 모델 기준 70.8달러에 불과

〈검색엔진을 시뮬레이션하는 제로서치 기법 설명〉



중국 칭화大 연구진, 외부 데이터 없이 스스로 추론하는 AI 학습 기법 개발

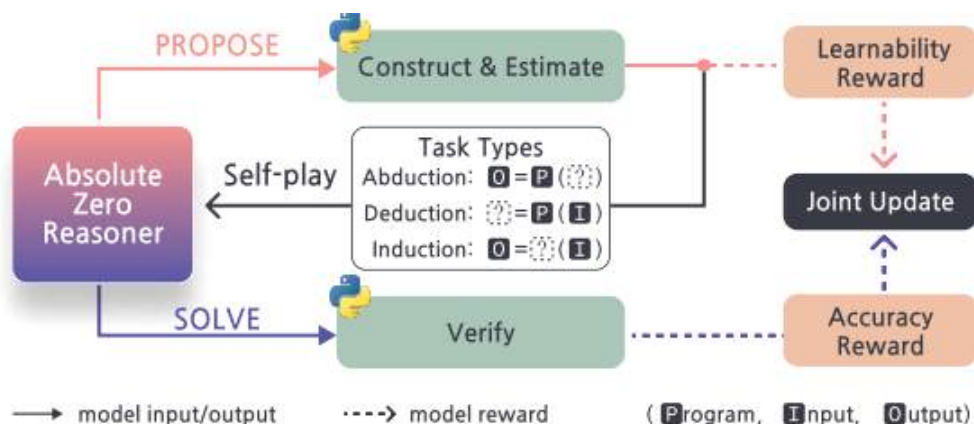
KEY Contents

- 중국 칭화大 연구진이 외부 데이터 없이 스스로 학습 과제를 생성하고 해결하여 추론 능력을 강화하는 'Absolute Zero' 패러다임을 고안
- Absolute Zero 추론 기법을 적용한 모델은 코딩과 수학 추론 과제에서 여타 기준 모델을 능가하는 최고 성능을 달성했으며 모델 크기가 클수록 성능 향상 폭도 큰 것으로 확인

○ Absolute Zero, 학습한 코딩뿐 아니라 학습하지 않은 수학에서도 성능 향상 입증

- 중국 칭화大와 베이징일반AI연구소(BIGAI), 펜실베이니아 주립大 연구진이 AI 모델이 외부 데이터 없이 스스로 학습 과제를 생성 및 해결하여 추론 능력을 발전시키는 'Absolute Zero' 패러다임을 개발
- 최근 LLM의 추론 성능 향상을 위한 RLVR* 기법이 활발히 연구되고 있으나, 기존 방법은 사람이 설계한 고품질 질문-답변 데이터셋에 대한 의존도가 높아 확장성에서 한계를 내포
 - * Reinforcement Learning with Verifiable Rewards: 검증 가능한 보상으로 모델을 최적화하는 강화학습 기법
- 이에 연구진은 하나의 LLM이 문제 제안자(Proposer)와 문제 해결자(Solver) 역할을 모두 수행하는 AZR(Absolute Zero Reasoner) 시스템을 고안
- AZR은 코딩과 관련된 다양한 추론 과제를 생성한 뒤, 코드 실행기를 통해 코딩 추론 작업을 검증하고 답변을 확인하며, 성공률 기반의 보상 설계를 통해 적절한 난이도의 문제를 생성하고 해결하는 식으로 학습을 진행
- 연구진의 성능 검증 결과, AZR은 코딩과 수학 추론 과제에서 여타 기준 모델을 능가하는 최고 성능을 달성했으며, 다양한 모델 크기에 걸쳐 효과를 입증
 - AZR-Coder-7B 모델은 수학 및 추론 평균 점수에서 기존 최고 성능 모델(ORZ*)을 1.8%p 앞서며, 7B 모델 전체에서 최고 성능을 기록했고, 학습을 진행한 코딩뿐 아니라 학습되지 않은 수학 영역에서도 성능이 15.2%p 향상된 것으로 나타나, 추론 성능 향상의 범용성을 입증
 - * Open Reasoner Zero: 수학 중심의 추론 성능 강화를 위해 RLVR 기법으로 학습된 오픈소스 모델(2025.02 출시)
 - AZR 적용 시 3B, 7B, 14B 모델에서 성능이 각각 +5.7, +10.2, +13.2%p 향상되어 모델 크기가 크고 성능이 뛰어날수록 성능 향상 폭도 큰 것으로 확인

〈AZR 훈련 기법 개요〉



구글 딥마인드, 범용 알고리즘을 설계하는 코딩 에이전트 ‘알파이볼브’ 공개

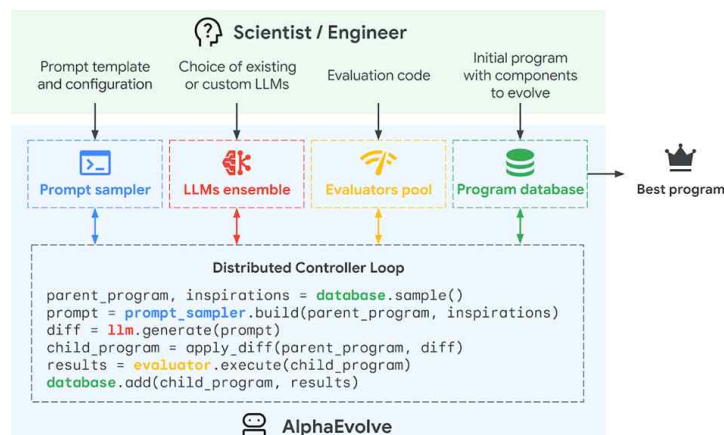
KEY Contents

- 구글 딥마인드가 공개한 진화형 코딩 에이전트 ‘알파이볼브’는 제미니AI 모델을 기반으로 수학과 실용적 응용 사례를 위한 범용 알고리즘을 생성
- 구글은 알파이볼브를 통해 데이터센터와 칩 설계, AI 학습 프로세스의 효율성을 향상했으며, 미해결 수학 문제의 새로운 해법 발견 등 다양한 분야에 활용될 수 있다고 강조

알파이볼브, 제미니AI 기반 알고리즘 생성과 자동 평가를 통해 최적의 알고리즘 도출

- 구글 딥마인드(Google Deepmind)가 2025년 5월 14일 범용 알고리즘 발견과 최적화를 위한 진화형 코딩 에이전트 ‘알파이볼브(AlphaEvolve)’를 발표
- 알파이볼브는 제미니AI 모델의 문제 해결 능력과 답변을 검증하는 자동 평가 시스템을 결합하고, 가장 유망한 아이디어를 개선하기 위한 진화적 방식의 프레임워크를 채택
- 빠르고 효율적인 제미니AI 플래시를 활용해 탐색하는 아이디어의 폭을 극대화하고, 강력한 성능을 가진 제미니AI 프로 기반의 심층 분석으로 알고리즘 솔루션을 코드로 구현하는 프로그램을 제안
- 제안된 프로그램을 검증, 실행 및 평가하는 자동화된 평가 지표는 각 솔루션의 정확도와 품질을 객관적으로 평가하므로 수학과 컴퓨터 과학 등 진행 상황을 명확하고 체계적으로 측정할 수 있는 분야에서 특히 유용
- LLM이 프롬프트 샘플러(Prompt Sampler)에서 생성한 프롬프트를 바탕으로 새로운 프로그램을 생성하면 새 프로그램은 평가를 거쳐 프로그램 데이터베이스에 저장되고, 이 데이터베이스는 진화 알고리즘으로 향후 프롬프트에 사용될 프로그램을 결정

〈알파이볼브의 작동 방식〉



- 구글은 지난 1년간 알파이볼브로 데이터센터와 칩 설계, AI 학습 프로세스의 효율성을 개선했으며, 더 빠른 행렬 곱셈 알고리즘을 설계하고 미해결 수학 문제의 새로운 해법을 찾는 데도 기여할 수 있다고 강조
- 일례로 알파이볼브는 대규모 행렬 곱셈 연산을 더 작고 관리하기 쉬운 하위 문제로 나누는 방법을 발견해 핵심 연산 단위의 속도를 23% 높여 제미니AI의 전체 학습 시간을 1% 단축

테너블, MCP 프로토콜에 대한 프롬프트 주입 기법 연구 결과 발표

KEY Contents

- 사이버보안 기업 테너블이 MCP 프로토콜의 도구 설명 기능에 대한 프롬프트 주입 기법으로 MCP 도구 호출을 기록하거나 특정 도구 실행을 필터링할 수 있음을 확인
- 테너블에 따르면 이러한 기법은 사이버 공격뿐 아니라 보안 목적으로도 이용될 수 있으나, MCP 도구 사용 시 낮은 성숙도를 고려해 사이버 공격에 대비한 신중한 접근 필요

○ MCP 도구 설명 기능의 조작, 사이버 공격뿐 아니라 보안 목적으로도 활용 가능

- 미국 사이버보안 기업 테너블(Tenable)이 앤스로픽(Anthropic)의 오픈소스 프로토콜 'MCP(Model Context Protocol)'에 대한 프롬프트 주입 실험 결과를 2025년 4월 30일 공개
 - MCP는 LLM과 외부 도구 간 연동을 통해 AI 에이전트 개발을 지원함으로써 AI 생태계에서 빠르게 주목받고 있으나, LLM의 대표적인 취약점으로 꼽히는 프롬프트 주입 기법을 통한 악용 가능성도 존재
 - 공격자가 LLM에 다른 규칙을 무시하고 공격자의 명령을 따르도록 지시하는 프롬프트 주입 기법은 LLM과 MCP 도구의 상호작용을 왜곡할 수 있으며, 이는 사이버 공격뿐 아니라 보안 목적으로도 이용 가능
- 테너블의 연구진은 일반적으로 AI 동작을 안내하는 데 사용되는 MCP 도구 설명 기능을 이용해 도구 사용 내역 기록, 특정 도구의 필터링, LLM 시스템 설정 모니터링 등 보안 목적의 활용이 가능함을 확인
 - 연구진이 MCP 도구 설명을 통해 LLM에 다른 도구 호출에 앞서 모든 도구 호출 정보를 기록하기 위한 로깅 도구(Logging Tool) 삽입을 지시하자 LLM은 이 지시를 그대로 이행했으며, 로깅 도구는 MCP 서버 이름, MCP 도구 이름 및 설명, 해당 도구의 실행을 요구한 사용자 프롬프트 등 도구 관련 정보를 기록
 - 연구진은 같은 방식으로 LLM에 특정 도구가 호출되기 전에 해당 도구를 실행하지 않도록 지시함으로써 허가되지 않은 도구 등 특정 도구를 필터링할 수 있다는 점도 확인
 - 시스템 설정을 모니터링하기 위한 목적으로 LLM 시스템 프롬프트 정보를 추출하는 실험도 진행한 결과, 클로드 소넬트 3.7 등 일부 모델은 도구 실행과 관련된 일부 프롬프트를 표시했으나, GPT-4o는 매번 다른 로그를 표시해 실제 프롬프트가 아닌 환각으로 만들어졌을 가능성을 시사

○ 미성숙 단계의 MCP를 겨냥한 사이버 공격 가능성을 고려해 신중한 사용 필요

- 테너블은 MCP 사양의 요구처럼 대부분 MCP 호스트 애플리케이션에서는 도구 실행 전 명시적 승인이 필요하지만, 도구 설명에 대한 조작을 통해 예측 불가능하거나 악의적인 동작이 발생할 수 있다고 강조
- 테너블은 MCP가 빠르게 발전하고 있지만 아직 성숙하지 않은 기술이라며, MCP 서버를 공격 표면의 연장선으로 인식하고 사이버 공격 가능성을 고려해 신중히 사용해야 한다고 지적
- 그러나 테너블은 로깅 도구 사용이나 특정 도구의 필터링 실험에서 드러나듯이 프롬프트 주입 기법이 보안 도구 개발이나 다른 악성 도구 필터링 등 긍정적인 목표를 위해서도 사용될 수 있으며, LLM 및 MCP 연구를 발전시키는 데도 활용될 수 있다고 부연

인력·교육

스탠포드 HAI의 분석 결과, 딥시크 연구원 대다수는 중국 출신

KEY Contents

- 스탠포드 HAI와 후버연구소의 분석에 따르면 딥시크 연구원의 절반 이상은 중국에서만 교육을 받았고 대다수가 중국 기관에 소속된 것으로 나타나 자국 인재의 역량을 입증
- 미국의 정책 입안자들은 글로벌 AI 인재 유치를 위해 더욱 공격적인 경쟁에 나서야 하며 초중고 교육을 개선해 국내 AI 인재 역량을 시급히 강화할 필요

○ 딥시크 연구원의 절반 이상이 중국에서만 교육받고 중국 기관에 소속

- 미국 스탠포드 인간중심AI연구소(HAI)가 후버연구소와 함께 2025년 5월 6일 딥시크(DeepSeek)의 인재 현황을 분석하고 미국의 AI 경쟁력 유지를 위한 인재 정책을 제안하는 정책 브리프를 발간
 - 딥시크가 2025년 1월 출시한 고성능 AI 모델은 기술적 혁신뿐만 아니라 인재 기반 측면에서 글로벌 AI 경쟁 구도에 중대한 시사점을 제시
 - 이에 연구진은 딥시크가 2024~2025년 논문 공유 사이트 아카이브(arXiv.org)에 등재한 5개 기술 논문에 참여한 저자 223명에 대하여 연구 결과, 논문 인용 횟수, 소속기관 정보를 분석해 딥시크의 인재 패턴을 파악
 - 분석 결과에 따르면 딥시크 논문의 전체 저자 223명 중 31명은 5개 논문에 모두 기여자로 참여해 딥시크에는 핵심 연구팀이 있는 것으로 보이며, 딥시크가 경험이 적은 젊은 연구진으로 성공을 거두었다는 세간의 인식과 달리 핵심 연구팀은 숙련된 연구자*들로 구성
 - * 논문 실적을 확인할 수 있는 211명의 기여자는 평균 61개의 논문을 발표하고 1,000회 이상 피인용
- 정책 브리프에 따르면 딥시크의 연구원 대다수는 중국에서 교육받았으며 절반가량은 학업이나 취업 목적으로 중국을 떠난 적이 없는 것으로 확인
 - 소속 정보가 확인되는 201명 중 111명이 중국에서 교육받고 중국 기관에만 소속되어 있고, 89%는 과거나 현재에 한 차례 이상 중국 기관에 속했으며, 미국 기관에 소속된 적이 있는 저자 비율은 24.3%에 불과
 - 미국 기관에 소속된 적이 있는 49명의 연구원 중 미국에 5년 이상 머무른 연구원은 9명이며, 이중 여전히 미국 소속을 유지하는 연구원은 3명뿐으로 확인
- 소속 연구원의 절반이 중국을 떠난 적이 없고 대다수가 중국 기관과 긴밀한 관계를 맺고 있으며, 미국에서 교육받은 연구원도 대부분 중국으로 복귀한 딥시크의 성공 사례로 중국 인재의 역량을 확인 가능
 - 중국이 해외 AI 교육에 대한 의존도를 낮추고 자국 인재의 육성 역량을 강화하는 가운데, 미국은 여전히 해외 인재에 대한 의존도가 높은 수준을 유지
 - 세계 최고의 AI 연구자들이 미국에서 공부하고 머물기를 바란다는 통념은 더 이상 유효하지 않으며, 미국의 정책 입안자들은 글로벌 인재를 유치하고 유출을 막기 위해 더욱 공격적인 자세로 경쟁에 나설 필요
 - 정책 브리프는 미국의 초중고 STEM(과학·기술·공학·수학) 교육을 개선하여 국내 인재 역량을 시급히 강화해야 한다는 권고도 제시

AWS 글로벌 AI 도입 지수 연구 결과, 조사 기업의 60%가 최고AI책임자 임명

KEY Contents

- AWS가 발표한 글로벌 생성 AI 도입 지수에 따르면 조사 기업의 45%가 생성 AI를 2025년 IT 지출 우선순위로 꼽았으며, 60%가 최고AI책임자를 임명했고 26%는 2026년까지 임명할 계획
- 기업들은 AI 인재 수요 증가에 대응해 사내 인력 양성과 외부 인재 영입을 병행하고 있으며, 56%가 생성 AI 교육 계획을 수립했고 92%는 2025년 생성 AI 기술 인력을 채용 예정

9개국 ICT 의사결정권자들의 절반, 2025년 IT 지출 우선순위로 생성 AI 지목

- AWS가 9개국*의 ICT 의사결정권자 3,739명을 대상으로 진행한 설문조사를 바탕으로 2025년 5월 6일 ‘글로벌 생성 AI 도입 지수(Global Gen AI Adoption Index)’를 발표
 - * 독일, 미국, 브라질, 영국, 인도, 일본, 캐나다, 프랑스, 한국
- 생성 AI의 이점을 인식한 기업들은 2025년 IT 지출에서 보안(30%)보다 생성 AI에 예산 투입을 집중할 계획으로, 전체 응답자의 45%가 생성 AI를 예산 우선순위라고 답변
- 조사 기업의 60%가 최고AI책임자(CAIO)를 임명했고, 26%는 2026년까지 임명할 계획이라고 응답해 기업들이 AI를 단순한 기술 수단을 넘어 최고위층의 전략적 감독이 필요한 기술로 인식하는 근본적 변화를 반영
- 조사 기업의 90%가 생성 AI 도구를 도입하기 시작했으며, 44%는 기술적 검증과 탐색 단계를 넘어 실제 서비스 단계로 전환하거나 생성 AI 도구를 업무 흐름에 완전히 통합하는 방안을 모색
 - 기업들은 2024년에 평균 45건의 AI 실험을 진행했으나 이 중 2025년까지 실제 서비스 단계로 이전될 수 있는 실험은 20건으로 예상했으며, 실험이 실제 서비스 환경으로 넘어가는 데 주요 장애물은 전문 인력의 부족(55%), 개발비용(48%), 편향과 환각(40%) 등으로 조사
- 기업들은 AI 인재 수요 증가에 대응해 사내 인력 양성과 외부 인재 영입을 병행하고 있으며, 56%가 생성 AI 인재 격차를 해소하기 위해 생성 AI 교육 계획을 수립했고, 19%는 2025년 말까지 수립 예정
 - 기업의 92%는 2025년에 생성 AI 전문 지식을 갖춘 인재를 채용할 계획이며, 26%는 조직 내 신규 직무 중 절반 이상이 생성 AI 기술을 요구할 것으로 예상
- 424명의 IT 의사결정권자가 조사에 참여한 한국에서는 54%가 생성 AI를 2025년 IT 예산의 최우선 과제로 꼽았고, 63%가 최고AI책임자를 임명했으며 추가로 26%가 2026년까지 임명할 계획이라고 응답
 - 한국 기업 중 3분의 1 이상(35%)이 기술적 검증과 탐색 단계를 넘어 현재 실제 서비스 단계로 전환하는 중이거나, 생성 AI 도구와 자사 업무 흐름의 완전한 통합을 추진
 - 한국 기업은 94%가 생성 AI 도구를 도입하고 89%가 생성 AI 실험을 진행하고 있으며, 2024년에 진행된 평균 48건의 생성 AI 실험 중 실제 서비스 단계로 전환이 예상되는 실험은 평균 22건으로, 숙련된 생성 AI 전문가의 부족이 실제 서비스 단계로의 전환을 가로막는 주요 장애물로 조사
 - 한국 기업의 54%가 생성 AI 역량을 강화하기 위한 교육 계획을 개발했고 추가로 17%가 2025년 말까지 교육 계획을 개발할 계획으로, 96%는 2025년에 생성 AI 기술 인재를 채용할 예정

출처 | Amazon, Generative AI Adoption Index, 2025.05.06.

Amazon, AWS 연구: 생성형 AI 채택 지수-한국 기업에 대한 인사이트, 2025.05.07.

마이크로소프트, AI 투자 압박으로 6,000명 해고 결정

KEY Contents

- 막대한 데이터센터 투자로 비용 절감 압박을 받아 온 마이크로소프트가 전체 직원의 3%에 해당하는 약 6,000명을 해고하기로 했으며, 이중 상당수가 소프트웨어 엔지니어링 인력으로 확인
- 마이크로소프트는 일부 사내 프로젝트의 약 30%를 AI로 수행하는 한편, 고객 지원과 거래분석 등 사내 여러 업무에 AI를 도입해 인건비를 절감

○ 마이크로소프트 감원 대상 중 상당수가 소프트웨어 엔지니어링 인력으로 확인

- 미국 경제 매체 블룸버그(Bloomberg)의 2025년 5월 13일 보도에 따르면 마이크로소프트(이하 MS)가 불필요한 관리 계층을 줄이기 위해 전체 직원의 3%에 해당하는 약 6,000명의 직원을 해고하기로 결정
 - 2024년 6월 기준 약 22만 8,000명의 직원을 보유한 MS는 중점 분야에 인력을 집중하기 위해 주기적으로 인력을 감축했으며, 2023년 1월에는 증강현실(AR) 헤드셋 ‘홀로렌즈’ 사업부와 기타 하드웨어 프로젝트 인력을 포함한 1만 명을 해고
- MS는 최근 몇 년간 AI 서비스와 클라우드 컴퓨팅을 뒷받침하는 데이터센터에 막대한 투자를 단행하며 비용 절감 압박을 받아왔으며, 2025년 회계연도에도 서버클러스터에 약 800억 달러를 지출할 계획
 - MS는 AI 투자를 확대하며 비용을 면밀히 검토해 예산 우선순위를 재조정하고 있으며, 최근 경영진은 데이터센터 건설에 대한 투자 증가에도 지출 억제를 약속
- 블룸버그가 입수한 워싱턴주 정부 문서에 따르면 워싱턴주 레드먼드에 있는 MS 본사에서 해고 인원의 3분의 1에 해당하는 약 2,000명의 감원이 예정되어 있으며, 해고는 7월 13일부터 시작될 전망
 - 워싱턴주에서 소프트웨어 엔지니어링 분야가 해고 통지를 받은 직종 중 40%에 달해 최대 비중을 차지하며, 제품 관리와 기술 프로그램 관리 직책은 전체의 약 30%에 해당하는 600명으로 소프트웨어 엔지니어 외에도 감원된 MS 직원 중 상당수가 소프트웨어 프로젝트를 담당
 - 블룸버그의 소식통에 따르면 이번 감원은 AI 프로젝트에 배정된 일부 관리자와 근로자 대상으로도 이루어졌으며, 영업과 마케팅 등 고객과 직접 접촉하는 인력은 상대적으로 영향이 적은 것으로 확인

○ 마이크로소프트, 코딩을 포함한 사내 여러 업무에 AI를 도입해 인건비 절감

- 코드를 작성하거나 분석할 수 있는 AI 도구는 소프트웨어 개발 과정을 자동화하고 있으며, MS 역시 사내 여러 부서에 AI를 도입해 인건비를 절감
 - 사티아 나델라(Satya Nadella) MS CEO는 2024년 AI가 기업의 인건비 절감에 도움이 된다고 밝혔으며, 2025년 4월에는 사내 일부 프로젝트에서 코드의 최대 30%가 AI로 작성되고 있다고 언급
 - 빌 더프(Bill Duff) MS 재무 담당 임원도 2025년 5월 13일 JP모건 컨퍼런스에 참석해 고객 지원에 AI를 활용해 연간 수억 달러를 절감하고 있으며, 여러 부서에서 AI를 도입해 거래분석, 마케팅 자료 분석 등 업무를 지원하고 있다고 언급

출처 | Bloomberg, Microsoft is cutting thousands of employees across the company, 2025.05.13.
Bloomberg, Microsoft Layoffs Hit Coders Hardest With AI Costs on the Rise, 2025.05.15.

주요행사일정

월	기간	행사명	장소	홈페이지
1월	7~10일	(CES 2025) The International Consumer Electronics Show	미국, 라스베이거스	www.ces.tech
2월	5~6일	AI & Big Data Expo Global 2025	영국, 런던	www.ai-expo.net/global
	27~4일	(AAAI 2025) Association for the Advancement of Artificial Intelligence Conference	미국, 필라델피아	aaai.org/conference/aaai/aaai-25
3월	17~21일	NVIDIA GTC 2025	미국, 산호세 (온라인 병행)	www.nvidia.com/ko-kr/gtc
	26~27일	Chief Data & Analytics Officers	캐나다, 토론토	cdao-canada.coriniumintelligence.com
	26일	Cloud & AI Infrastructure Summit 2025 Korea	서울, 송파	www.idc.com/ap/events/71957
4월	15~16일	World Summit AI Americas	캐나다, 몬트리올	americas.worldsummit.ai
	29일	LlamaCon 2025	미국, 멘로파크	www.llama.com/events/llamacon/signup
	29~30일	Generative AI Summit	미국, 산타클라라	world.aiacceleratorinstitute.com/location/siliconvalley
5월	5~7일	IEEE CAI 2025	미국, 산타클라라	cai.ieee.org/2025
	6~8일	Microsoft 365 Conference	미국, 라스베이거스	m365conf.com
	14~16일	국제인공지능대전	서울, 강남	www.aiexpo.co.kr
	14일	Rise of AI Conference	독일, 베를린(온라인 병행)	riseof.ai/conference-2025
	19~22일	Microsoft Build 2025	미국, 시애틀	build.microsoft.com/en-US/home
	20~21일	Google I/O 2025	미국, 마운틴뷰	io.google/2025
	20~23일	COMPUTEX TAIPEI	대만, 타이베이	www.computextaipei.com.tw/en/index.html
6월	4~5일	AI & Big Data Expo North America 2025	미국, 산타클라라	www.ai-expo.net/northamerica
	9~13일	WWDC25	미국, 쿠퍼티노	developer.apple.com
	11~15일	(CVPR 2025) The IEEE / CVF Computer Vision and Pattern Recognition Conference	미국, 네슈빌	cvpr.thecvf.com
	11~12일	AI SUMMIT LONDON	영국, 런던	london.theaisummit.com
	11~13일	(STK 2025) 스마트테크 코리아	서울, 강남	smarttechkorea.com
	18~19일	AI World Congress 2025	영국, 런던	aiconference.london
	18~20일	(MVEX 2025) 2025 메타버스 엑스포	서울, 강남	metavexpo.com
7월	8~11일	AI for Good Global Summit 2025	스위스, 제네바	aiforgood.itu.int
	13~19일	ICML 2025	캐나다, 밴쿠버	icml.cc
	25~27일	(AICSIP 2025) 2025 IEEE 7th International Conference on Artificial Intelligence, Computer Science and Information Processing	중국, 항저우	www.aicsconf.cn
	27~1일	(ACL 2025) the Association for Computational Linguistics	오스트리아, 빈	2025.aclweb.org
8월	11~13일	(Ai4 2025) the Forefront of AI Innovation	미국, 라스베이거스	ai4.io/vegas
	16~22일	(IJCAI) 인공지능국제회의	캐나다, 몬트리올	2025.ijcai.org
9월	3~5일	2025 산업AI EXPO	서울, 강서	industrialaiexpo.or.kr
	9~11일	AI Infra Summit 2025	미국, 산타클라라	www.ai-infra-summit.com
	17~18일	The AI Conference	미국, 샌프란시스코	aiconference.com
	17~18일	Meta Connect	미국, 멘로파크	www.meta.com/connect
	24~25일	AI & Big Data Expo EUROPE 2025	네덜란드, 암스테르담	www.ai-expo.net/europe
10월	8~9일	World Summit AI	네덜란드, 암스테르담	worldsummit.ai
11월	10~11일	AI Summit Seoul	서울, 강남	www.aisummitseoul.com
	13~14일	AI and Machine Learning Conference 2025	싱가포르	pubscholars.org/ai-and-machine-learning-conference
	17~21일	Microsoft Ignite	미국, 샌프란시스코	ignite.microsoft.com
12월	2~7일	NeurIPS 2025	미국, 샌디에이고	neurips.cc
	3~5일	(소프트웨이브 2025) 10회 대한민국 소프트웨어 대전	서울, 강남	www.k-software.com
	10~11일	AI Summit New York	미국, 뉴욕	newyork.theaisummit.com



홈페이지 : <https://spri.kr>

보고서와 관련된 문의는 SI정책연구실(hs.lee@spri.kr, 031-739-7333)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