Research Report

연구보고서 RE-203

2025년 국내외 인공지능 산업 동향 연구

Artificial Intelligence Industry Trends in Korea and Overseas in 2025

이해수/유재흥/임정주/안성원

2026.08.

SPRi 소프트웨어정책연구소

이 보고서는 2025년도 과학기술정보통신부 정보통신진흥기금을 지원받아 수행한 연구결과로 보고서 내용은 연구자의 견해이며, 과학기술정보통신부의 공식입장과 다를 수 있습니다.

목 차

제1장 서론 ····· 1

제1절 연구 배경 및 목적 ····· 1

제2절 연구 내용 및 방법 ····· 5

제2장 인공지능 산업 현황 및 전망 ····· 7

제1절 AI 산업 현황 ····· 7

제2절 전 세계 AI 시장 전망 ····· 16

제3장 주요국 인공지능 정책 동향 ····· 18

제1절 미국 ····· 20

제2절 EU ····· 50

제3절 중국 ····· 68

제4절 일본 ····· 77

제5절 영국 ····· 92

제6절 캐나다 ····· 110

제7절 프랑스 ····· 124

제8절 독일 ····· 140

제9절 싱가포르 ····· 149

제10절 이스라엘 ····· 155

제11절 인도 ····· 161

제12절 UAE ····· 172

제13절 대만 ····· 183

제14절 호주 ····· 191

제15절 우리나라 ·····198

제4장 국제기구 인공지능 동향 ·····217

제1절 OECD ·····219

제2절 UN ·····231

제3절 기타 기구(G7, G20 등) ·····238

제5장 인공지능 관련 주요 기업 동향 ·····244

제1절 국외 기업 ·····246

1. 알파벳(구글) ·····246

2. 마이크로소프트 ·····267

3. 메타 플랫폼스 ·····278

4. 오픈AI ·····285

5. xAI ·····303

6. 엔비디아 ·····309

7. 기타 주요 기업(애플, 아마존, 알리바바 등) ·····316

8. AI 스타트업 ·····341

제2절 국내 기업 ·····368

1. 네이버 ·····368

2. LG AI 연구원 ·····375

3. SKT ·····384

4. NC AI ·····392

5. 업스테이지 ·····396

6. 기타 주요 기업(삼성, KT 등) ·····399

7. AI 스타트업 ·····407

제6장 인공지능 기술 및 연구 동향 ·····415
제1절 학술단체 ·····417
제2절 표준기관 ·····435
제3절 연구기관 ·····445

제7장 인공지능 교육 및 인력 동향 ·····463
제1절 교육 분야 ·····465
제2절 인력 및 고용 분야 ·····475

제8장 결론 ·····485

참고 문헌 ·····495

[첨부 1] 2025년 주요국의 AI 관련 분야별 정책 동향 요약 ·····503
[첨부 2] 2025년 주요국 AI 정책 전 세계 맵(World Map) ·····507
[첨부 3] 국내외 주요 기업 동향 ·····511
[첨부 4] 2025년 글로벌 AI 생태계 동향 ·····528
[첨부 5] AI 산업 동향 전문가 간담회 개최 결과 ·····530

표 목 차

<표 2-1> 전 세계 주목할 머신러닝 학습모델 현황 ·····8

<표 2-2> 2025 HAI AI 보고서의 주요 AI 트렌드(Top Takeaways) ·····8

<표 3-1> 미국의 주요 AI 정책 추진 현황 ·····20

<표 3-2> 미국 행정명령 14110호 6대 중점 분야 ·····23

<표 3-3> 미국의 AI 관련 규제와 진흥 정책 변화 ·····28

<표 3-4> 미국 AI 실행계획 3대 전략별 주요 실행 과제 ·····32

<표 3-5> EO 14319 ‘연방정부 조달에서 ‘Woke AI’ 방지 ‘ 행정명령 핵심 내용 ·····37

<표 3-6> EU의 주요 AI 정책 추진 현황 ·····50

<표 3-7> EU AI 혁신 패키지 이니셔티브 구조 ·····52

<표 3-8> AI 대륙 실행계획 5대 핵심 전략 ·····54

<표 3-9> EU AI 법 ‘금지된 AI 시스템’ ·····57

<표 3-10> AI 적용 전략 3대 추진 방향 ·····63

<표 3-11> 디지털 옴니버스(디지털 규제 간소화 패키지) 주요 내용 ·····65

<표 3-12> EU 데이터 연합 전략 주요 내용 ·····66

<표 3-13> 중국의 주요 AI 정책 추진 현황 ·····68

<표 3-14> 중국 AI 플러스 행동 심화 추진을 위한 6대 핵심 영역 ·····70

<표 3-15> 중국 컴퓨팅 자원 상호연결 실행계획의 주요 내용 ·····73

<표 3-16> 중국 행정 분야 AI 모델 도입을 위한 4개 중점 시나리오 ·····76

<표 3-17> 일본의 주요 AI 정책 추진 현황 ·····77

<표 3-18> ‘생성 AI의 조달·활용에 관한 가이드라인’의 주요 내용 ·····83

<표 3-19> 일본 AI 추진법의 개요 ·····85

<표 3-20> 생성 AI의 이용 상황 변화와 AI에 대한 민간투자액의 변화 ·····88

〈표 3-21〉 ‘AI 기본계획(초안)’의 4대 기본 방침과 강구해야 할 시책 ·····90
〈표 3-22〉 영국의 주요 AI 정책 추진 현황 ·····92
〈표 3-23〉 영국 국가 AI 전략(2021)의 부문별 정책 과제 ·····93
〈표 3-24〉 영국 AI 실행 계획 주요 내용 ·····94
〈표 3-25〉 영국 AI 규제에 대한 친혁신 프레임워크 핵심 요소 ·····95
〈표 3-26〉 영국 친혁신적 AI 규제 프레임워크의 5대 원칙 ·····96
〈표 3-27〉 영국 정부 조직을 대상으로 하는 생성형 AI 프레임워크의 10대 원칙 97
〈표 3-28〉 영국 경쟁시장청(CMA)의 AI 이행 원칙 ·····99
〈표 3-29〉 영국 AI 기회 실행계획과 정부 답변 ·····101
〈표 3-30〉 AIRR 이용 가이드 주요 내용 ·····103
〈표 3-31〉 캐나다의 주요 AI 정책 추진 현황 ·····110
〈표 3-32〉 연방 공공서비스 AI 전략 4대 목표 및 주요 내용 ·····115
〈표 3-33〉 캐나다 디지털규제포럼 기관별 합성미디어 쟁점 정리 ·····122
〈표 3-34〉 프랑스의 주요 AI 정책 추진 현황 ·····124
〈표 3-35〉 프랑스 「국가 AI 전략」 단계 및 중점 내용 ·····126
〈표 3-36〉 AI 혁신위원회 25개 권고 요약 ·····129
〈표 3-37〉 주요 권고와 「국가 AI 전략」 제3단계 연계 ·····130
〈표 3-38〉 파리 ‘AI 액션 서밋’ 주요 논의 ·····132
〈표 3-39〉 커런트 AI의 중점 영역 ·····134
〈표 3-40〉 프랑스 개인정보보호위원회 ‘2025-2028 전략 계획’ 4대 축 ·····135
〈표 3-41〉 프랑스 ‘Osez l’IA’ 전략 주요 내용 ·····136
〈표 3-42〉 프랑스 ‘AI와 보건의료 데이터’ 1차 전략 주요 내용 ·····138
〈표 3-43〉 독일의 주요 AI 정책 추진 현황 ·····140
〈표 3-44〉 독일 연방 정부 조직 변화 ·····143
〈표 3-45〉 AI 시스템의 필수 기술정보 클러스터 7가지 ·····145

〈표 3-46〉 독일 하이테크 어젠다 핵심 내용 ·····146
〈표 3-47〉 AI 시장 감독 및 혁신 촉진법 초안 내용 ·····147
〈표 3-48〉 싱가포르의 주요 AI 정책 추진 현황 ·····149
〈표 3-49〉 NAIS 1.0와 NAIS 2.0 비교 ·····151
〈표 3-50〉 이스라엘의 주요 AI 정책 추진 현황 ·····155
〈표 3-51〉 국가 AI 본부 관련 주요 승인 내용 ·····158
〈표 3-52〉 프로그램 지원 대상 분야 ·····160
〈표 3-53〉 인도의 주요 AI 정책 추진 현황 ·····161
〈표 3-54〉 ‘인도AI미션’ 7개 필라 및 주요 내용 ·····163
〈표 3-55〉 공무원을 위한 AI 역량 프레임워크 ·····167
〈표 3-56〉 ‘디지털 슈람세투’ 구현 로드맵 ·····168
〈표 3-57〉 AI 거버넌스 가이드라인 주요 내용 ·····169
〈표 3-58〉 UAE의 주요 AI 정책 추진 현황 ·····172
〈표 3-59〉 연방 공공서비스 AI 전략 4대 목표 ·····182
〈표 3-60〉 대만의 AI 정책 추진 현황 ·····183
〈표 3-61〉 대만 10대 AI 인프라 프로젝트의 구성 ·····187
〈표 3-62〉 대만 스마트 로봇산업 육성계획 ·····190
〈표 3-63〉 호주의 주요 AI 정책 추진 현황 ·····191
〈표 3-64〉 우리나라의 주요 AI 정책 추진 현황 ·····198
〈표 3-65〉 AI 일상화 주요 추진과제 ·····201
〈표 3-66〉 4대 AI 플래그십 프로젝트 ·····203
〈표 3-67〉 4대 분야 정책 추진 방향 ·····203
〈표 3-68〉 국가 AI 역량강화방안의 추진전략 및 핵심 과제 ·····204
〈표 3-69〉 국가 AI 역량강화방안의 후속조치 ·····205
〈표 3-70〉 생성형 AI 서비스 이용자 보호를 위한 기본 원칙 ·····206

<표 3-71> 생성형 AI 서비스 이용자 보호를 위한 실행 방식 ·····206
<표 3-72> K-휴머노이드 연합의 5대 과제 ·····207
<표 3-73> AI 바이오 확산전략의 주요 추진내용 ·····209
<표 3-74> 생성형 AI 개인정보 처리 안내서의 핵심 내용 ·····210
<표 3-75> 분과위원회의 역할 ·····212
<표 3-76> AI 액션플랜의 3대 정책축 및 12대 전략 분야 ·····213
<표 3-77> AI 인재양성 방안의 주요 정책과제 ·····214
<표 3-78> AI기본법 시행령 제정안의 주요내용 ·····215
<표 4-1> OECD의 연도별 AI 정책 추진 경과 ·····221
<표 4-2> OECD.AI를 통해 살펴본 한국의 AI 현황 예시(2025년 11월 기준) ·····223
<표 4-3> AI 역량 수준 개요(2024년 11월 기준) ·····225
<표 5-1> 알파벳(구글) 기업 현황 ·····246
<표 5-2> 마이크로소프트 기업 현황 ·····267
<표 5-3> 메타 플랫폼즈 기업 현황 ·····278
<표 5-4> 오픈AI 기업 현황 ·····285
<표 5-5> xAI 기업 현황 ·····303
<표 5-6> 엔비디아 기업 현황 ·····309
<표 5-7> 애플 기업 현황 ·····316
<표 5-8> 아마존 기업 현황 ·····320
<표 5-9> 아마존 노바 액트와 경쟁 AI 에이전트의 벤치마크 평가 비교 ·····321
<표 5-10> 알리바바 기업 현황 ·····326
<표 5-11> AI 스타트업 현황 ·····341
<표 5-12> 네이버 기업 현황 ·····368
<표 5-13> LG AI 연구원 기업 현황 ·····375
<표 5-14> SKT 기업 현황 ·····384

<표 5-15> NC AI 기업 현황 ·····392
<표 5-16> 업스테이지 기업 현황 ·····396
<표 5-17> 기업 일반현황 ·····399
<표 5-18> AI 스타트업 현황 ·····407
<표 6-1> IEEE SA의 AIS 관련 주요 표준(2025. 11월 기준) ·····423
<표 6-2> 신진연구자 II - AI 관련 연구 발표 주제 ·····432
<표 6-3> ISO/IEC JTC 1/SC 42에서 발행된 표준 목록 ·····436
<표 6-4> 한국정보통신기술협회의 인공지능 분야에서 발행된 최신 표준 목록 442
<표 6-5> KISDI의 인공지능 분야에서 발행된 주요 연구보고서 목록(2024~2025) ·····460
<표 6-6> SPRi의 인공지능 분야에서 발행된 주요 연구보고서 목록(2024~2025) ·····461
<표 6-7> NIA의 AI 관련 연구보고서 사례 ·····462
<표 7-1> 미국 청소년을 위한 AI 교육 자원 제공 이니셔티브 참여 기업 주요 내용 473
<표> 2025년도 주요국의 AI 관련 분야별 정책 동향 요약 ·····503
<표> 국내 AI 기업 동향 ·····511
<표> 해외 AI 기업 동향 ·····518
<표> 제1차 중소기업 AI 활용 간담회 개최 결과 주요 내용 ·····530
<표> 제2차 피지컬 AI 전문가 간담회 개최 결과 주요 내용 ·····532

그 림 목 차

[그림 2-1] 전 세계 AI 관련 법안 수 통과 현황: 2016-2024 ·····7
[그림 2-2] 글로벌 AI 지수 국가별 종합 순위 ·····10
[그림 2-3] 응답 조직의 AI 에이전트 활용 단계: 사업 기능별(복수 응답, n=1,933) ····12
[그림 2-4] 투자 유형별 글로벌 AI 기업 투자: 2013~2024 ·····13
[그림 2-5] 분야별 글로벌 민간 AI 투자 비교: 2023 vs 2024 ·····13
[그림 2-6] 글로벌 AI 투자 거래 현황 및 추이 ·····15
[그림 2-7] 생성형 AI 소프트웨어 시장 규모 전망: 2024~2030 ·····16
[그림 2-8] SaaS와 에이전트 AI의 전체 시장 규모(TAM) 전망 비교 ·····17
[그림 3-1] K-휴머노이드 연합의 추진체계 ·····208
[그림 3-2] AI 바이오 확산전략의 주요 추진내용 ·····209
[그림 3-3] 생성형 AI 개발·활용 단계별 고려사항 ·····211
[그림 3-4] AI 인재양성 방안의 주요 정책 방향 ·····214
[그림 4-1] 전 세계 AI 기초 모델 공급 추이 ·····228
[그림 4-2] 첨단 기술의 시장 규모 성장 비교 (2023년 vs. 2033년) (단위: 십억 달러) ····235
[그림 5-1] 젬마 3와 주요 AI 모델의 챗봇 아레나 점수 비교 ·····248
[그림 5-2] 알파이볼브의 작동 방식 ·····253
[그림 5-3] 구글의 메드젬마와 메드시그립 모델 개요 ·····256
[그림 5-4] 지니 3로 생성된 세계의 일관성 유지(예시) ·····257
[그림 5-5] 제미나이 2.5 플래시 이미지’의 이미지 합성 예시 ·····261
[그림 5-6] 과학적 문제 해결을 위한 코드 솔루션 도출 알고리즘 ·····265
[그림 5-7] 마이크로소프트의 ‘MAI-DxO’ 개요 ·····273
[그림 5-8] 생성 AI 도구 사용 시 인지 활동에 들이는 노력 수준 ·····277
[그림 5-9] V-JEPA 2의 2단계 학습 방식 ·····283
[그림 5-10] GPT-5-코덱스와 GPT-5의 SWE-bench 및 코드 재구성 작업 정확도 비교 ·····292

[그림 5-11] 코덱스의 SWE-Bench Verified와 오픈AI 내부 SWE 작업 정확도 비교 ·····297
[그림 5-12] 챗GPT 스터디 모드의 응답 예시 ·····301
[그림 5-13] 그록 코드 패스트 1' 과 경쟁 AI 모델의 성능 비교 ·····307
[그림 5-14] 그록 4.1과 주요 AI 모델의 EQ-Bench3 벤치마크 점수 비교 ·····308
[그림 5-15] 아마존의 물류 로봇 '벌컨' 의 도입 현장 ·····323
[그림 5-16] 큐원2.5-맥스와 경쟁 AI 모델의 벤치마크 점수 비교 ·····327
[그림 5-17] 과학적 문제 해결을 위한 코드 솔루션 도출 알고리즘 ·····328
[그림 5-18] Qwen3-Next의 사전훈련 효율성과 추론 속도 ·····329
[그림 5-19] '옴니휴먼-1' 로 생성된 영상 예시 ·····333
[그림 5-20] 'R-Zero' 프레임워크의 개요 ·····338
[그림 5-21] Agent Laboratory 개요 ·····339
[그림 5-22] 클로드 대화에서 업무 증강 vs 업무 자동화 비율 ·····343
[그림 5-23] 뉴론피디아에서 제공되는 귀속 그래프 탐색 인터페이스 ·····347
[그림 5-24] 클로드 오푸스 4와 주요 AI 모델 벤치마크 평가 점수 ·····349
[그림 5-25] 퍼플렉시티 딥 리서치와 주요 AI 모델과의 HLE 정확도 비교 ·····351
[그림 5-26] 다중 LLM AB-MCTS의 검색 알고리즘 개요 ·····355
[그림 5-27] 앱트로닉의 휴머노이드 로봇 '아폴로' ·····361
[그림 5-28] '키미 K2' 와 경쟁 AI 모델의 주요 벤치마크 평가 결과 비교 ·····364
[그림 5-29] 12개 벤치마크 기준 LLM 성능 종합 평가 비교 ·····366
[그림 5-30] 딥엘 에이전트의 화면 ·····367
[그림 5-31] 이미지 검색 AI 요약 서비스 ·····374
[그림 5-32] 엑사원 모델과 글로벌 모델 성능평가 비교 ·····377
[그림 5-33] 생성 AI 평가 기준 ·····379
[그림 5-34] 다양한 병렬화 기법에 따른 MT-NLG 학습 시간 및 GPU 사용률 변화 ··401
[그림 6-1] 지역별 AI 연구 출판물 인용(수) ·····417

[그림 6-2] 상위 100위권 AI 출판물(수) ·····418
[그림 6-3] 연구 분야별 AI 관련 출판물 발간 추이(2010-2022) ·····418
[그림 6-4] 주요 모델별 추론 비용 ·····419
[그림 6-5] 지역별 특허(수) 비율 ·····420
[그림 6-6] 대규모 학습 모델의 탄소 배출량(추정치) ·····421
[그림 6-7] AI+ 인증 절차 ·····440
[그림 6-8] 산업인공지능표준화포럼 ·····441
[그림 6-9] 소규모 고품질 학습 예제 기반 LIMO의 성능 향상 효과 ·····445
[그림 6-10] S1의 샘플 효율성 비교 ·····447
[그림 6-11] AZR 훈련 기법 개요 ·····448
[그림 6-12] 게이트로서의 메모리(MAG) 아키텍처 ·····449
[그림 6-13] ‘몰모액트’의 개요 ·····450
[그림 6-14] Agent Laboratory의 개요 ·····452
[그림 6-15] MoR(Mixture-of-Recursions) 개요 ·····453
[그림 6-16] MIRAS 아키텍쳐 개요 ·····454
[그림 6-17] 비전 트랜스포머 아키택쳐(예) ·····455
[그림 6-18] 인공지능안전연구소 조직도 ·····458
[그림 7-1] 교육 부문에서 생성 AI 사용 ·····465
[그림 7-2] 학생과 교사의 생성 AI 사용 ·····466
[그림 7-3] 생성 AI 교육 계획 ·····475
[그림 7-4] AI 노출 대비 매출 증가율 ·····476
[그림 7-5] AI 노출도가 가장 높은 산업과 가장 적은 산업 비교 ·····477
[그림 7-6] AI 기술 보유 근로자 수에 따른 임금 프리미엄 ·····478
[그림 7-7] AI에 노출된 직업에서 학위를 요구하는 비율 ·····479
[그림 7-8] 신규 대졸자 실업률과 전국 평균 실업률 비교 ·····481

[그림 7-9] 클로드 대화에서 업무 증강 vs 업무 자동화 비율 ·····484
[그림] 2025년 주요국 AI 정책(정책/법제 분야) 전 세계 맵(World Map) ·····507
[그림] 2025년 주요국 AI 정책(기업/산업 분야) 전 세계 맵(World Map) ·····508
[그림] 2025년 주요국 AI 정책(기술/연구 분야) 전 세계 맵(World Map) ·····509
[그림] 2025년 주요국 AI 정책(인력/교육 분야) 전 세계 맵(World Map) ·····510
[그림] 2025년 글로벌 AI 생태계 동향(국내 AI 기업) ·····528
[그림] 2025년 글로벌 AI 생태계 동향(해외 AI 기업) ·····529

요 약 문

1. 제 목 : 2025년 국내외 인공지능 산업 동향 연구

2. 연구 목적 및 필요성

ㅇ 생성형 AI와 초거대 AI 모델, 에이전트형 AI 및 피지컬 AI의 확산으로 인공지능 산업 생태계가 빠르게 재편

- AI 기술은 산업 혁신을 넘어 행정·노동·교육·안보 등 사회 전반의 구조적 변화를 촉발하고 있으며, 주요국은 이를 국가 경쟁력과 기술 패권의 핵심 요소로 인식하고 국가 AI 전략 고도화와 법·제도 정비를 본격화
- 특히, 2025년을 전후로 AI 관련 기본법 제정과 규제·거버넌스 체계가 ‘설계 단계’를 넘어 ‘시행 단계’로 진입하면서, 정책의 실질적 효과와 산업·사회적 영향에 대한 분석의 중요성이 증대
- 관련 하드웨어, AI 반도체, 연산 인프라에 대한 대규모 투자와 함께 전 산업에서 AI 융합과 활용이 가속화되고 있으며, 글로벌 빅테크를 중심으로 초거대 모델, 에이전트형 AI, 피지컬 AI 분야에서 주도권 경쟁이 심화
- 동시에 AI 안전성·신뢰성·책임성에 대한 우려가 확대되면서, 각국 정부와 국제기구, 표준화 기구를 중심으로 AI 윤리, 안전, 거버넌스 규범 정립과 국제 협력이 강화
- 이러한 환경 변화 속에서 국내 인공지능 산업의 경쟁력 강화, 활용 확산, 그리고 글로벌 수준의 신뢰성 확보를 위한 정책 마련을 위해 국내외 AI 산업 및 정책 동향에 대한 종합적 분석 필요성 증대

ㅇ 본 연구는 급변하는 국내외 인공지능 산업 생태계를 체계적으로 분석하여, 국내 산업 경쟁력 강화와 인공지능 활용 확산을 지원하기 위한 정책적 기초 자료를 마련하는 것을 주된 목표

- 주요국, 주요 기관 및 학술단체, 주요 기업의 동향을 조사하여 현재의 상황을 분석하고, AI 기술의 발전 및 산업 변화에 대한 전망을 통해 적시에 효과적인 AI 정책을 개발하고 정책적 의사결정을 지원하기 위한 기초 자료를 제공하며, 이를 통해 정책 과제를 발굴
- 이를 통해 향후 국가 AI 전략, 산업 정책, 법·제도 개선을 위한 주요 정책 과제를 도출하고, 중장기 정책 고도화에 기여
- 인공지능 산업과 관련된 폭넓은 조사를 바탕으로 국내외 AI 정책·법제·전략·권고안 등 핵심 자료를 체계적으로 확보하여, 정책 설계 및 발전 과정에서 활용 가능한 근거 자료로 제공

3. 연구의 구성 및 범위

- ㅇ AI 산업 현황 및 전 세계 AI 시장전망에 대한 개괄적 정리
- ㅇ 국내외 주요국·국제기구 정책 동향
  - 주요국에는 미국, EU, 중국, 일본, 영국, 캐나다, 프랑스, 독일, 싱가포르, 이스라엘, 인도, UAE, 대만, 호주 등 포함
  - 국제 기구/회의로 OECD, UN, G7정상회의, G20정상회의 동향 분석
- ㅇ 국내외 인공지능 관련 주요 기업 동향 분석
  - 해외 글로벌 빅테크(구글, 마이크로소프트, 메타, 오픈AI, xAI, 엔비디아 등) 기업 및 국내 주요 기업 (네이버, LG AI 연구원, SKT, NC AI, 업스테이지 등)을 포함한 주요 AI 스타트업 동향 포함
- ㅇ 국제 학술 단체 연구, 표준기관, 비영리 연구 기관 동향 분석
  - IEEE, ACM 및 최상위 AI 컨퍼런스 발표 및 ISO/EC 국제 표준화 동향
  - 국내 TTA 인공지능 표준화 동향 및 관련 인공지능 학회 주요 연구 동향
- ㅇ 인공지능 교육, 인력 및 고용 관련 동향 분석

4. 연구 내용 및 결과

- ㅇ 초거대 AI에서 에이전트형 AI·피지컬 AI로의 기술 패러다임 전환
  - 생성형 AI와 초거대 AI 모델은 단일 기능 제공을 넘어, 다단계 업무 수행이 가능한 에이전트형 AI로 진화
  - AI 에이전트는 의료, 교육, 금융, 행정 등에서 업무 프로세스 전반을 자동화·지능화하는 핵심 도구로 확산
  - 피지컬 AI는 로봇, 자율시스템, 제조·물류 자동화와 결합되며 실물 산업 구조를 재편하는 단계로 진입
  - 이에 따라 AI 산업은 ‘모델 개발 경쟁’에서 ‘에이전트·체화형 AI 기반 서비스·산업 경쟁’ 국면으로 전환
- ㅇ AI 산업 경쟁의 축 이동: 모델 경쟁에서 국가 AI 역량 경쟁으로 확대
  - 주요국은 초거대 모델 성능 경쟁을 넘어, 데이터·연산 인프라·AI 반도체·인재를 묶은 국가 AI 역량 확보에 주력
  - 글로벌 빅테크와 AI 스타트업을 중심으로 AI 인프라 및 기초모델 분야에 대규모 투자가 지속
  - 국내외 기업들은 자국 언어·산업 특성을 반영한 기초모델과 서비스 내재화를 통해 차별화 전략을 강화
  - AI 산업 경쟁은 기업 단위를 넘어 국가 차원의 전략 경쟁으로 확장되는 양상

ㅇ AI 거버넌스의 전환: 원칙・설계 단계에서 시행・집행 단계로 진입
- AI 기술 확산에 따라 안전성, 책임성, 설명가능성, 인권 보호에 대한 요구가 구체적 정책 과제로 전환
- 주요국은 AI 기본법 제정과 함께 영향평가, 사고 보고, 감독 체계 등 실질적 집행 수단을 마련
- 국제기구(OECD, UN 등)와 다자 협의체를 중심으로 AI 윤리・안전・신뢰성에 관한 국제 규범 형성이 가속
- AI 거버넌스는 기술 규범을 넘어 행정, 조달, 산업 정책 전반에 영향을 미치는 핵심 정책 영역으로 부상

ㅇ AI 산업 경쟁의 축 이동: 모델 경쟁에서 국가 AI 역량 경쟁으로 확대
- AI 기술은 연구・학계 전반에 확산되며 공학, 사회과학, 생명과학 등과의 융합 연구를 촉진
- 고용 구조 변화와 함께 재교육(Reskilling)과 인간-AI 협업 역량 강화가 핵심 정책 과제로 부상
- 주요국은 생애주기별 AI 교육과 인재 양성 전략을 통해 국가 차원의 AI 대응 역량을 강화
- AI는 산업 기술을 넘어 교육・노동・윤리・거버넌스를 포괄하는 사회 기반 기술로 자리

5. 정책적 활용 내용

ㅇ 본 연구의 분석 결과는 정부의 인공지능 정책 수립 및 고도화를 위한 기초 자료로 활용될 수 있음
- 생성형 AI, 에이전트형 AI, 피지컬 AI 확산에 따른 산업 구조 변화와 주요국 정책 동향 분석을 통해, 국가 인공지능 경쟁력 강화 및 활용 확산을 위한 정책 방향 설정에 참고 자료로 활용
- 주요국의 AI 전략, 법・제도 및 거버넌스 시행 동향 분석을 바탕으로, 국가 AI 전략 고도화 및 관련 법・제도 정비, 특히 인공지능 기본법의 시행・보완을 위한 현황 분석 자료로 활용

6. 기대효과

ㅇ 급변하는 국내외 인공지능 환경 변화에 대한 시의성 있는 정보 제공을 통해 정부의 정책 대응 역량 제고

ㅇ 인공지능 산업 관련 중장기 정책 연구 기획 및 정부의 시의성 있는 정책 수립에 지원

ㅇ 인공지능 산업과 관련된 정책, 시장, 기업 전략, 표준, 인력 및 연구개발 동향 등 다양한 정보를 연계・제공함으로써, 정책 담당자, 연구자, 산업계 등 다양한 이해관계자의 AI 정책 및 산업 이해도 제고와 전략적 대응 역량 강화에 기여

SUMMARY

1. Title: Artificial Intelligence Industry Trends in Korea and Overseas in 2025

2. Purpose and Necessity of the Research

○ The rapid diffusion of generative AI, large-scale foundation models, agent-based AI, and physical AI is fundamentally reshaping the artificial intelligence industry ecosystem.

- AI technologies are driving structural transformations across society, extending beyond industrial innovation to public administration, labor markets, education, and national security. Major countries increasingly recognize AI as a core pillar of national competitiveness and technological sovereignty, and are actively advancing national AI strategies alongside legal and institutional reforms.
- In particular, around 2025, AI-related framework laws and regulatory and governance systems are transitioning from the design phase to the implementation phase, heightening the importance of analyzing the practical effectiveness of policies and their industrial and societal impacts.
- Large-scale investments in AI-related hardware, AI semiconductors, and computing infrastructure are accelerating, while AI integration and utilization are expanding across all industries. At the same time, global big tech companies are intensifying competition for leadership in large-scale models, agent-based AI, and physical AI.
- Concurrently, growing concerns regarding AI safety, reliability, and accountability have prompted governments, international organizations, and standard-setting bodies to strengthen efforts to establish ethical principles, safety frameworks, and governance norms, as well as to enhance international cooperation.
- Against this backdrop, there is an increasing need for comprehensive analysis of domestic and international AI industry and policy trends in order to strengthen the competitiveness of Korea’s AI industry, promote the widespread adoption of AI, and ensure globally recognized AI trustworthiness.

○ Against this backdrop, there is an increasing need for comprehensive analysis of domestic and international AI industry and policy trends in order to strengthen the competitiveness of Korea’s AI industry, promote the widespread adoption of

AI, and ensure globally recognized AI trustworthiness.

- By examining trends across major countries, key institutions and academic organizations, and leading companies, the study assesses the current state of the AI landscape and forecasts developments in AI technologies and industrial transformation. This analysis aims to provide timely and effective evidence to support AI policy development and policy decision-making, while identifying key policy challenges.
- Through this process, the study seeks to derive major policy tasks for advancing national AI strategies, industrial policies, and legal and institutional frameworks, thereby contributing to the enhancement of mid- to long-term AI policy development.
- Based on a broad and systematic review of the AI industry, the study compiles and organizes key domestic and international AI-related policies, legal frameworks, strategies, and recommendations, providing evidence-based resources to support policy design and implementation.

3. Composition and Range

ㅇ Overview of the AI industry landscape and global AI market outlook

- A comprehensive review of current developments in the AI industry and projections for the global AI market, including key technological and industrial trends.

ㅇ Analysis of AI policy trends in major countries and international organizations

- Major countries covered include the United States, the European Union, China, Japan, the United Kingdom, Canada, France, Germany, Singapore, Israel, India, the United Arab Emirates, Taiwan, and Australia.
- Analysis of AI-related policy discussions and initiatives in international organizations and multilateral forums, including the OECD, the United Nations, the G7 Summit, and the G20 Summit.

ㅇ Analysis of major domestic and international AI company trends

- Examination of global big tech companies (e.g., Google, Microsoft, Meta, OpenAI, xAI, NVIDIA) as well as major Korean companies (e.g., Naver, LG AI Research, SK Telecom, NC AI, Upstage), including trends among leading AI startups.

- Focus on corporate strategies, technological developments, and ecosystem dynamics shaping the AI industry.

○ Analysis of research trends in international academic organizations, standard-setting bodies, and non-profit research institutions

- Review of research outputs from IEEE, ACM, and top-tier AI conferences, as well as international standardization trends led by ISO/IEC.
- Analysis of domestic AI standardization activities led by the Telecommunications Technology Association (TTA) and major research trends presented in Korean AI-related academic societies.

○ Analysis of trends in AI education, workforce development, and employment

- Examination of policies and initiatives related to AI education, talent development, and workforce transformation, including changes in employment structures driven by AI adoption.

4. Main Contents and Results

○ Technological paradigm shift from large-scale foundation models to agent-based AI and physical AI

- Generative AI and large-scale foundation models are evolving beyond single-function applications toward agent-based AI capable of performing multi-step tasks and autonomous workflows.
- AI agents are rapidly diffusing as core tools that automate and intelligently orchestrate end-to-end processes across healthcare, education, finance, and public administration.
- Physical AI is entering a new phase in which it is integrated with robotics, autonomous systems, and manufacturing and logistics automation, thereby reshaping the structure of the real economy.
- As a result, the AI industry is transitioning from competition centered on model development to competition focused on services and industries built on agent-based and embodied AI.

○ Shifting axis of AI industry competition: from model performance to national AI capabilities

- Major countries are moving beyond competition over model performance toward securing comprehensive national AI capabilities that integrate data resources,

computing infrastructure, AI semiconductors, and talent.

- Large-scale investments in AI infrastructure and foundation models continue to be concentrated among global big tech companies and leading AI startups.
- Companies in both advanced and emerging markets are strengthening differentiation strategies by internalizing foundation models and services tailored to national languages and industry-specific characteristics.
- Consequently, AI industry competition is expanding beyond firm-level rivalry to strategic competition at the national level.

○ Transformation of AI governance: from principles and design to implementation and enforcement

- As AI technologies diffuse across society, demands related to safety, accountability, explainability, and the protection of human rights are increasingly translated into concrete policy requirements.
- Major countries are establishing practical enforcement mechanisms—such as impact assessments, incident reporting obligations, and supervisory systems—alongside the enactment of AI framework laws.
- International organizations (including the OECD and the United Nations) and multilateral forums are accelerating the formation of global norms related to AI ethics, safety, and trustworthiness.
- AI governance is emerging as a core policy domain that extends beyond technical regulation to influence public administration, procurement systems, and industrial policy.

○ Societal transformation driven by AI diffusion: implications for research, education, and the workforce

- AI technologies are rapidly permeating research and academic communities, fostering interdisciplinary convergence across engineering, social sciences, and life sciences.
- Changes in employment structures have elevated reskilling and the strengthening of human–AI collaboration capabilities as central policy challenges.
- Major countries are enhancing national AI preparedness through lifecycle-based AI education and talent development strategies.
- AI is increasingly positioned as a foundational societal infrastructure that encompasses not only industrial technologies but also education, labor, ethics, and governance.

5. Policy use

ㅇ The analytical findings of this study can be utilized as foundational evidence for the formulation and advancement of government AI policies.

- By analyzing structural changes in the AI industry driven by the diffusion of generative AI, agent-based AI, and physical AI, as well as major policy trends in leading countries, the study provides reference materials for setting policy directions aimed at strengthening national AI competitiveness and promoting the widespread adoption of AI.
- Based on an analysis of AI strategies, legal frameworks, and governance implementation trends in major countries, the study offers empirical inputs to support the enhancement of national AI strategies and the refinement of related legal and institutional frameworks, particularly for the implementation and improvement of the Artificial Intelligence Basic Act.

6. Research Implication and Expected Effects

ㅇ Enhancing the government's policy response capacity through the provision of timely and relevant information on rapidly evolving domestic and global AI environments.

ㅇ Supporting mid- to long-term AI policy research planning and the development of evidence-based government policies related to the AI industry.

ㅇ Contributing to improved understanding and strategic response capabilities among diverse stakeholders—including policymakers, researchers, and industry participants—by integrating and disseminating comprehensive information on AI-related policies, markets, corporate strategies, standards, talent, and R&D trends.

CONTENTS

Chapter 1. Introduction ····· 1
1.1. Research Background and Purpose ····· 1
1.2. Research Content and Methods ····· 5

Chapter 2. Artificial Intelligence Industry Status and Outlook ····· 7
2.1. AI Industry Status ····· 7
2.2. Artificial Intelligence Industry Outlook ····· 16

Chapter 3. Trends in AI Policies of Major Countries ····· 18
3.1. United States ····· 20
3.2. European Union (EU) ····· 50
3.3. China ····· 68
3.4. Japan ····· 77
3.5. United Kingdom ····· 92
3.6. Canada ····· 110
3.7. France ····· 124
3.8. Germnay ····· 140
3.9. Singapore ····· 149
3.10. Israel ····· 155
3.11. India ····· 161
3.12. UAE ····· 172
3.13. Taiwan ····· 183
3.14. Australia ····· 191
3.15. Korea ····· 198

Chapter 4. Trends in AI of International Organizations ····· 217
4.1. OECD ····· 219
4.2. UNESCO ····· 231
4.3. Other Organizations(G7, G20, etc.) ····· 238

Chapter 5. Trends in Major Companies Related to AI ·····244
5.1. Foreign Companies(Google, MS, Meta, OpenAI, etc.) ·····246
5.2. Domestic Companies(Naver, LG AI Research, etc.) ·····368

Chapter 6. Trends in AI Technology and Research ·····415
6.1. Academic OrganizationsInstitutes ·····417
6.2. Standards Organizations ·····435
6.3. Research Institutes ·····445

Chapter 7. Trends in AI Education and Human Resources ·····463
7.1. Education ·····465
7.2. Human Resources and Employment ·····475

Chapter 8. Conclusion ·····485

References ·····495

Appendix 1. Summary of AI Policy Trends by Countries in 2025 ·····503
Appendix 2. Global Map of AI Policies in 2025 ·····507
Appendix 3. Trends in major domestic and foreign companies ·····511
Appendix 4. Global AI Ecosystem Trends in 2025 ·····528
Appendix 5. Results of Expert Meeting with AI Industry Trends ·····530

제1장 서론

제1절 연구 배경 및 목적

1. 연구 배경

생성 AI와 초거대 AI 모델이 등장한 이후 불과 몇 년 사이, 인공지능은 개별 서비스나 특정 산업의 혁신을 넘어 사회 전 영역의 구조를 바꾸는 범용 기술로 자리매김하였다. 오픈소스와 클라우드 인프라, 대규모 데이터, 고성능 연산 자원의 결합으로 AI 성능이 비약적으로 향상되면서, 2024년까지는 다양한 산업에서 시범·도입 단계가 빠르게 확산되었다면 2025년에는 실제 업무와 일상생활 깊숙이 내재화되는 단계로 기술적 진보가 양적 확산을 넘어 질적 전환 국면으로 접어들고 있다. 스탠퍼드 HAI의 「AI 인덱스 2025」는 연구개발, 기술 성과, 책임 있는 AI, 경제·의료·정책·거버넌스 등 9개 분야를 종합적으로 평가하며, 생성 AI를 포함한 최신 모델들의 성능이 다양한 벤치마크에서 전년 대비 큰 폭으로 향상되었음을 보여준다. 또한 2016년 이후 전 세계에서 '인공지능' 관련 법률이 꾸준히 증가하였는데 이는 AI가 더 이상 단순한 신기술이 아니라, 각국이 법·제도와 정책 체계를 통해 본격적으로 관리해야 하는 사회 기반 인프라로 인식되고 있음을 시사한다.

국제사회 역시 AI 거버넌스 체계 구축에 속도를 내고 있다. 2024년 3월 유엔총회는 '안전하고 신뢰할 수 있는 AI 시스템의 개발과 사용'을 촉구하는 결의안을 컨센서스로 채택하여, 인권과 국제 인권법을 준수하지 못하거나 인권에 과도한 위험을 야기하는 AI 시스템의 사용을 자제할 것을 각국에 권고하였다. 이어 2025년에는 글로벌 디지털 협약 이행 결의와 함께 AI 국제과학패널과 글로벌 AI 거버넌스 대화를 신설하여, AI의 기회와 위험에 대한 과학적 평가 제공, 허위정보·자율무기 등 고위험 영역에 대한 조기 경보, 다중 이해관계자 간 협력 플랫폼 구축 등을 추진하고 있다. 이러한 움직임은 AI 기술 발전 속도에 상응하는 국제 규범과 협력 메커니즘을 마련하려는 시도로, 향후 각국의 법·제도 설계에 중요한 기준점이 될 전망이다.

한편 AI의 경제적 잠재력과 위험은 동시에 확대되고 있다. 유엔무역개발회의(UNCTAD)는

2025년 기술·혁신 보고서에서 2033년까지 전 세계 AI 시장 규모가 약 4조 8천억 달러 수준에 달할 수 있다고 전망하면서도, 현재의 혜택이 소수 국가와 기업에 집중되고 있다고 지적한다. 또한 전 세계 일자리의 최대 40%가 AI의 영향을 받을 수 있으며, 많은 개발도상국이 아직 국가 AI 전략조차 수립하지 못해 AI 격차와 대표성 부족 문제가 심화될 가능성을 경고하였다. 이처럼 AI는 경제 성장과 효율성 제고의 동력이자, 국가 간·계층 간 불평등을 심화시킬 수 있는 양면적 기술로 부상하고 있다.

주요국은 이러한 기회와 위험을 동시에 관리하기 위해 기술혁신과 규범 정비를 병행하는 정책을 강화하고 있다. 미국은 국가 AI 이니셔티브법 제정 이후 AI 권리장전, 안전하고 신뢰할 수 있는 AI 행정명령 등을 통해 AI 연구개발 투자 확대와 함께 안전·신뢰성·권리 보호 원칙을 구체화하고 있으며, EU는 2024년 AI법을 최종 승인하여 위험 기반 규율을 도입하는 등 세계 최초의 포괄적 AI 규제 체계를 구축하였다. 영국, 캐나다, 일본, 중국 등도 국가 전략과 로드맵을 고도화하면서, 경쟁력 확보와 윤리·안전 확보를 동시에 달성하기 위한 정책 수단을 다층적으로 설계하고 있다.

우리나라도 이러한 흐름 속에서 AI를 경제·사회 전반의 혁신을 견인할 핵심 인프라로 인식하고, 정책과 제도 정비를 본격화하고 있다. 2019년 ‘인공지능 국가전략’을 시작으로, 2023년 ‘초거대 AI 경쟁력 강화 방안’, 2024년 ‘AI G3 도약을 위한 AI 디지털 혁신성장 전략’ 등을 통해 데이터·인프라·인재·연구개발 전 주기에 걸친 지원 방향을 제시해 왔다. 특히 2024년 말 국회를 통과하여 2025년 초 제정된 인공지능 기본법은 AI 진흥과 이용자 보호, 신뢰할 수 있는 AI 구현, 국제 협력 등을 포괄하는 첫 포괄적 기본법으로, 향후 시행령과 세부 지침을 통해 국내 AI 생태계의 법적 토대를 마련하는 전환점이 되고 있다. 또한 정부는 ‘국가 AI 역량 강화 방안’을 통해 AI 반도체 및 컴퓨팅 인프라, 차세대 기초모델, AI+X 확산 등 세 축을 중심으로 국가 차원의 역량을 높이는 전략을 발표하였으며, 교육부문에서도 ‘모두를 위한 AI 인재양성’ 계획을 통해 초·중등 교육과 고등·평생 교육 전반에서 AI 교육을 체계화하고 있다. 국내 표준기관과 연구기관은 국제 표준화 기구 논의에 적극 참여하면서, 신뢰성·안전성·윤리 원칙을 반영한 AI 표준과 인증 제도 정비를 추진하고 있어, 기술 경쟁력과 책임 있는 활용 간 균형을 도모하는 기반이 마련되고 있다.

산업계에서는 글로벌 빅테크와의 경쟁과 협력을 동시에 고려한 전략이 전개되고 있다. 국

내 주요 기업들은 한국어와 아시아 언어에 특화된 기초모델과 멀티모달 모델을 고도화하고, 이를 검색·커머스·콘텐츠·금융·클라우드 서비스 등 핵심 비즈니스에 내재화하는 '온-서비스 AI' 전략을 강화하고 있다. 예를 들어, 국내 대표 플랫폼 기업은 파운데이션 모델 '하이퍼클로바X'를 경량·고성능 구조로 업그레이드하고, 텍스트·이미지·영상까지 이해하는 멀티모달 기능과 효율적인 운영 구조를 통해 다양한 서비스에서 생성 AI 적용을 확대하고 있다. 검색 서비스에서는 사용자의 복합 질문을 요약·정리해 제시하는 AI 기반 검색 기능을 도입하고, 이를 향후 도메인 특화 에이전트 서비스로 확장하려는 시도도 활발하다. 이러한 움직임은 국내 기업이 언어·문화적 강점을 활용해 차별화된 AI 생태계를 구축하려는 전략적 대응으로 평가할 수 있다.

그러나 기술 확산 속도가 빨라질수록 오남용과 부작용에 대한 우려도 커지고 있다. 딥페이크 영상·음성을 활용한 사기, 개인정보 침해, 허위정보 확산, 알고리즘 편향에 따른 차별 등은 이미 구체적인 피해 사례로 나타나고 있으며, 관련 탐지·방지 기술과 제도적 대응의 중요성이 한층 부각되고 있다. 특히 생성 AI와 에이전트형 AI가 정보 생산과 의사결정 과정에 깊숙이 관여하면서, 설명가능성, 투명성, 책임성, 인권 보호를 어떻게 담보할 것인지에 대한 논의가 정책·산업·시민사회 전반에서 심화되고 있다.

이와 같은 국내외 환경 변화는 AI 산업 동향을 단편적인 기술 이슈나 개별 정책 수준이 아니라, 기술·시장·정책·기업·연구·인력·교육이 상호작용하는 거시적 생태계 차원에서 종합적으로 파악할 필요성을 더욱 높이고 있다.

2. 연구 목적

2025년은 AI 관련 법·제도와 국제 거버넌스가 본격적으로 구체화되기 시작한 시점으로, 2024년 한 해 동안 축적된 데이터와 정책 사례를 체계적으로 정리·분석하는 작업이 향후 정책 설계와 산업 전략 수립에 중요한 기준선(baseline)을 제공하게 될 것이다.

이에 따라 본 연구는 2025년에 발표된 국내외 AI 관련 자료를 중심으로 AI 산업과 정책·거버넌스의 구조적 변화를 조망하고자 한다. 구체적으로, 제2장은 AI 인덱스 2025와 주요 시장·투자 데이터, AI 에이전트 시장 전망 등을 바탕으로 전 세계 AI 발전 수준과 시장·

투자 구조 변화를 분석한다. 제3장과 제4장에서는 미국, EU, 중국, 일본, 영국 등 주요국과 OECD, UN, G7·G20 등 국제기구의 2025년 정책·거버넌스 동향을 살펴봄으로써, AI 진흥·규제 정책의 방향과 국제 규범 형성 과정에서의 쟁점을 정리한다. 제5장과 제6장에서는 국내외 주요 기업, 학술단체, 표준기관, 연구기관의 기술·서비스·연구 개발 흐름을 분석하고, 제7장에서는 교육·인력·고용 측면의 변화를 다룬 뒤, 마지막 제8장에서 종합 시사점과 향후 정책 과제를 제시한다.

본 연구의 궁극적인 목적은 다음과 같다. 첫째, 급변하는 글로벌 AI 산업·정책 환경을 체계적으로 정리함으로써 우리나라가 국제적 흐름 속에서 자신의 위치를 정확히 진단하고, 중장기 전략을 설계하는 데 필요한 기초 정보를 제공하는 것이다. 둘째, 인공지능 기본법과 후속 하위 법령, 국가 AI 전략, 산업·교육 정책 등 국내 정책 고도화 과정에서 참조할 수 있는 정량·정성 자료와 비교 분석을 제공하여, 혁신 촉진과 위험기관이 조화를 이루는 정책 설계에 기여하고자 한다. 셋째, 국내 기업과 연구기관이 글로벌 기술·시장 경쟁 구도를 이해하고, 기초모델·에이전트·AI 반도체·AI 인프라 등 전략 분야에서 차별화된 방향성을 모색할 수 있도록 시사점을 제시하는 것이다. 마지막으로, AI가 노동·교육·윤리·거버넌스 등 사회 전 영역에 미치는 영향을 균형 있게 조망함으로써, 공공·민간·시민사회가 함께 참여하는 책임 있는 AI 논의의 기반을 마련하고자 한다.

「2025년 국내외 인공지능 산업 동향 조사」는 최근의 데이터와 사례를 토대로 2025년의 기술·시장·정책 환경을 입체적으로 분석하고, 향후 우리나라 AI 산업이 글로벌 경쟁과 협력의 장에서 지속가능한 성장을 달성하기 위한 전략적 방향을 모색하는 데 목적이 있다. 본 보고서에서 제시하는 분석과 시사점이 정부의 정책 수립과 산업계·학계·연구계의 전략 수립에 폭넓게 활용되어, 우리나라가 혁신성과 책임성을 겸비한 AI 선도국가로 도약하는 데 기여하기를 기대한다.

제2절 연구 내용 및 방법

1. 연구 내용

본 연구에서는 주요국, 국제기구, 주요 기업, 인공지능 관련 기술 및 연구 동향을 조사·분석하고자 하였다. 주요국으로는 AI 기술 수준이 높고 국제적 영향력을 가진 국가들을 선정하였으며, 미국, EU, 중국, 일본, 영국 등을 포함한 주요국과 캐나다, 프랑스, 독일, 싱가포르, 이스라엘, 인도, UAE, 대만, 호주 등 기타 국가를 대상으로 하였다. 또한, 우리나라 동향도 조사 범위에 포함되었다. 국제기구로는 AI 관련 심도 있는 정책문서를 발간하는 OECD와 UN 등을 중심으로 조사했다. 국제기구로는 AI 관련 심도 있는 정책문서를 발간하는 OECD와 UN 등을 중심으로 조사했다. 기업 분야에서는 해외는 빅테크 기업인 알파벳(Google), 마이크로소프트, 메타플랫폼, 오픈 AI, xAI, 엔비디아와 기타 주요 기업(애플, 아마존, 알리바바 등) 그리고 주요 AI 스타트업들을 대상으로 했다. 국내 기업으로는 네이버, LG AI 연구원, SKT, NC AI, 업스테이지와 삼성, KT 등 주요 대기업 그리고 주요 AI 스타트업들을 조사하였다. 또한, 기술 및 연구 동향 조사를 위해 국내외 학술단체, 표준기관, 연구기관을 포함하였으며, 교육 및 인력 동향은 별도의 항목으로 구분하여 주요국과 국제기구, 교육기관 및 연구기관의 사례를 분석하였다. 이와 같은 조사 범위를 통해 국내외 AI 산업의 현황과 관련 요소를 종합적으로 파악하고자 했다.

2. 연구 방법

본 연구는 국내외 AI 산업 동향을 체계적으로 파악하기 위해 주요 국가와 국제기구의 공식 홈페이지에 공개된 보도자료 및 각종 공개 문서를 1차 자료로 활용하였다. 아울러, 미디어를 통해 보도된 관련 기사와 참고 자료를 보조적으로 수집하여 분석에 반영하였다. 기업 동향 조사의 경우, 기업의 공식 보도자료와 언론 보도, 운영 중인 공식 블로그 등을 주요 자료원으로 삼았으며, 학술단체 및 표준화 기구와 관련해서는 공식 홈페이지와 학술대회 웹사이트를 통해 공개된 문헌을 중심으로 조사하였다. 이와 함께 Gartner, IDC 등 신뢰성이 검증된 시장조사기관의 보고서와 Springer, IEEE Xplore 등 주요 학술 데이터베이스를 활용하였고, Google Scholar 등 AI 기반 검색 도구를 통해 최신 연구 자료를 추가적으로 탐색하였다. 조사 과정에서는 분야별·주제별 세부 키워드를 사전에 정의하여 일관되게 적용하였

으며, 필요에 따라 AI 전문가와 정책 분석가 등 관련 분야 전문가의 검토를 거쳐 조사 결과의 정확성과 신뢰성을 제고하였다. 연구의 시간적 범위는 2025년에 발표된 자료를 중심으로 설정하되, 동향의 배경 이해와 정책 흐름 분석에 필수적인 일부 선행 자료도 함께 포함하였다. 이를 통해 2025년 AI 산업의 최신 동향과 관련 정책의 전개 맥락을 종합적으로 제시하고자 하였다.

제2장 인공지능 산업 현황 및 전망

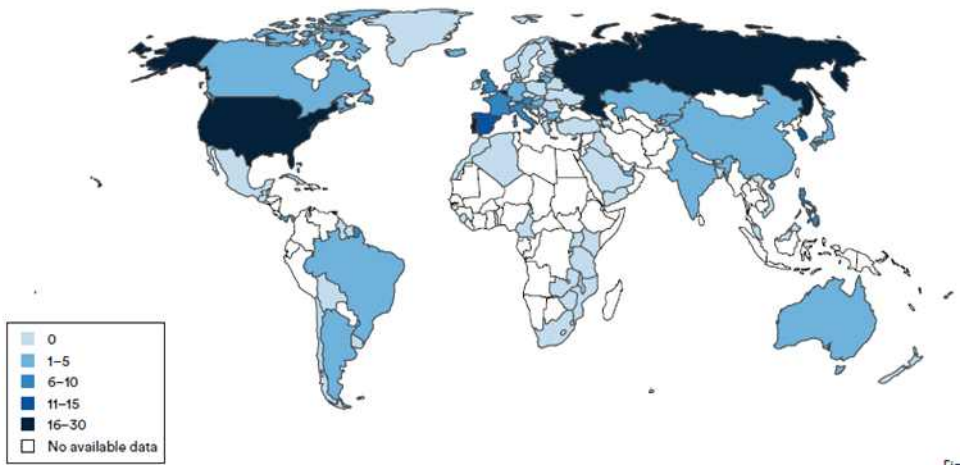
제1절 AI 산업 현황

1. 전 세계 AI 현황 및 수준

2025년 4월, 미국 스탠퍼드 대학교 HAI(Human-Centered AI)는 AI 분야의 국제적 수준을 비교 분석한 ‘AI 인덱스 보고서(Artificial Intelligence Index Report) 2025’[1]를 발표하였다. AI 인덱스 보고서는 연구개발, 기술적 성과, 책임 있는 AI, 경제, 과학 및 의료, 정책과 거버넌스, 다양성, 대중 의견 측면(총 9개 분야)의 AI 발전 현황을 평가하며, 통계정보와 추세를 제시했다.

2025년 HAI 보고서는 전 세계 국가의 AI 관련 입법 현황을 분석했다. 2016년부터 2024년까지 114개국에서 ’인공지능’이라는 용어가 포함된 입법을 분석한 결과, 39개국이 최소 한 건 이상의 인공지능 관련 법률을 제정했으며, 총 204건의 인공지능 관련 법률이 통과된 것으로 나타났다. 통과된 AI 관련 법률의 총 수는 2023년 30건에서 2024년 40건으로 증가하여, 2024년은 2022년 다음으로 두 번째로 높은 수치를 기록했다. 2016년 이후 통과된 AI 관련 법률의 수는 단 1건에서 40건으로 크게 증가했다.

[그림 2-1] 전 세계 AI 관련 법안 수 통과 현황: 2016-2024

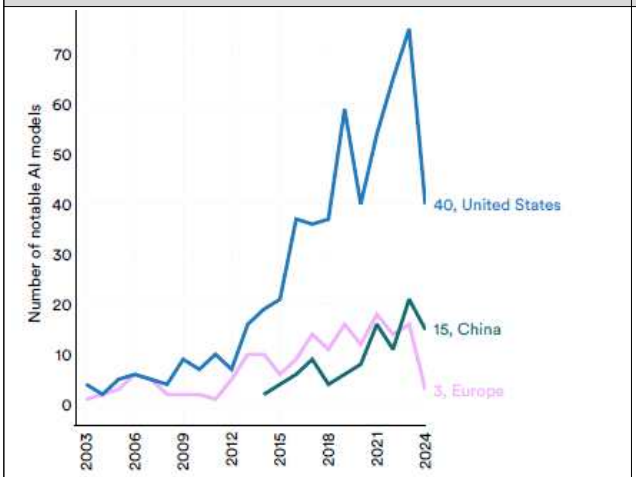
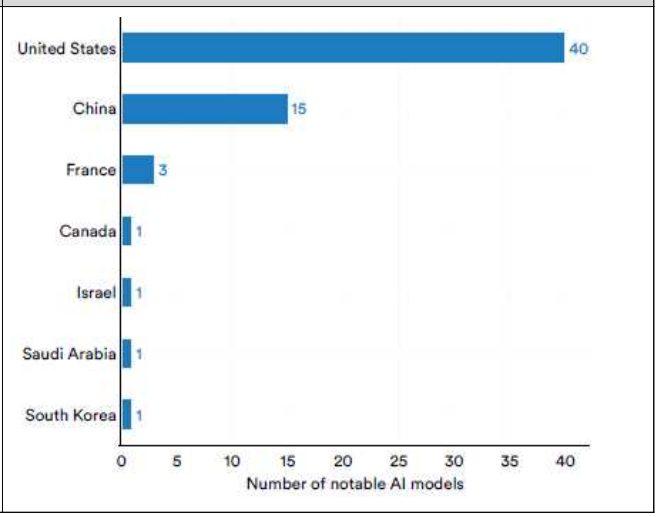


출처: Stanford University, Artificial Intelligence Index Report 2025.

1) Stanford HAI, The 2025 AI Index Report, 2025.4

또한, HAI는 Epoch AI의 자료를 인용하여 2003~2024년 전 세계 국가의 유력(notable) AI 모델의 개수를 정리했다. 아래 그림과 같이 2024년 당해년도 기준, 미국은 40개로 최대이며, 다음으로 중국, 프랑스가 뒤를 잇고 있으며, 우리나라도 캐나다, 이스라엘, 사우디아라비아와 함께 1개의 유력 AI 모델이 집계되었다. 2024년은 최초 생성형 AI 개발 경쟁이 절정에 달했던 데에 반해 개발 모델 숫자는 둔화된 것으로 나타났다.

<표 2-1> 전 세계 주목할 머신러닝 학습모델 현황

2003~2024년 지역별 유력 AI 모델 개수	2024년 국가별 유력 AI 모델 개수
	

출처: Stanford University, Artificial Intelligence Index Report 2025

HAI의 2025년 AI 지수 보고서에 최근 전 세계의 AI 발전 양상을 다음과 같이 12개의 주요 트렌드로 제시하였다.

<표 2-2> 2025 HAI AI 보고서의 주요 AI 트렌드(Top Takeaways)

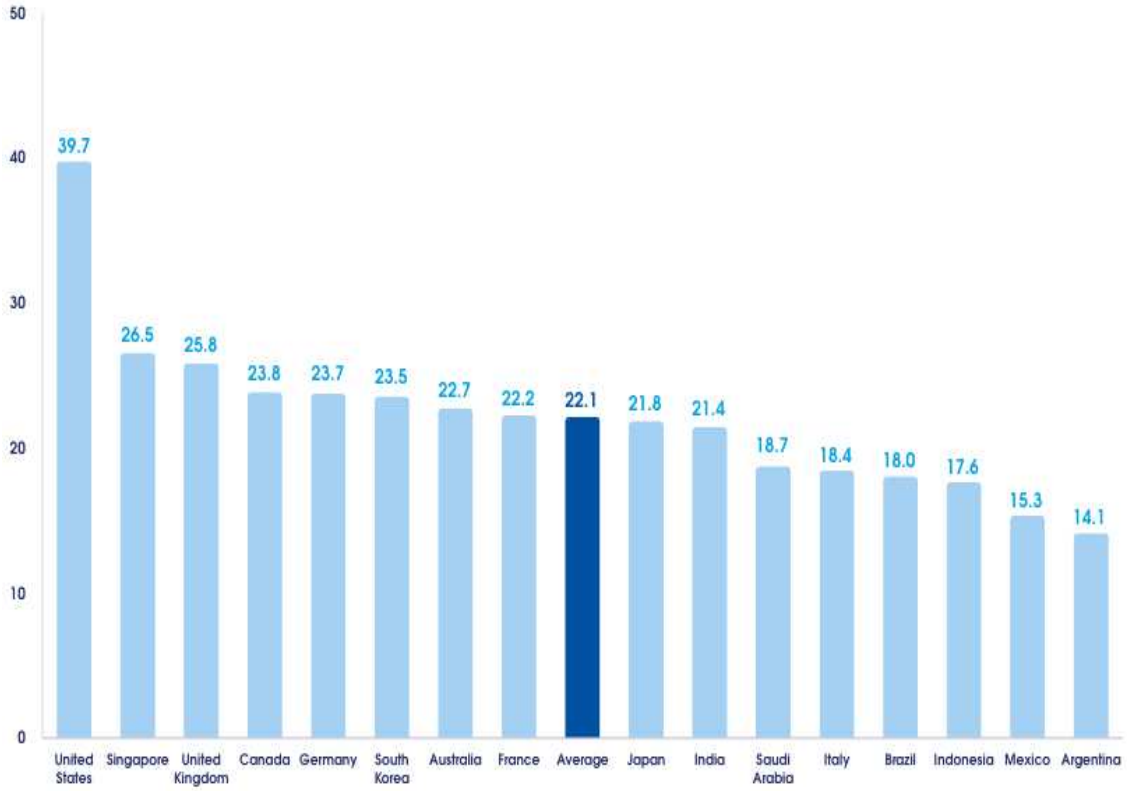
구분	주요 내용	
1	**AI 성능 향상 및 벤치마크 점수의 지속적 증가**	• 전년 평가 대비 MMMU, GPQA, SWE-bench에서 각각 18.8%, 48.9%, 67.3% 포인트의 점수가 급격히 상승 • AI 시스템은 고품질 비디오 생성에서 큰 발전을 이루었고, 일부 프로그래밍 작업에서는 언어 모델 에이전트가 인간보다 우수한 성능을 보임
2	**AI의 일상생활 내재화 가속**	• 2023년 FDA는 2015년의 6개에 비해 증가한 223개의 AI 지원 의료기기를 승인 • 미국 Waymo는 매주 150,000건 이상의 자율 주행 서비스를 제공하고, 중국 Baidu의 Apollo Go 로보택시도 여러 도시에서 운영
3	**기록적인 투자 및 생산성 향상**	• 2024년 미국 민간 AI 투자는 $1,091억으로, 중국이나 영국의 투자액과 큰 격차 발생

구분	주요 내용	
		• 생성형 AI는 전 세계적으로 $339억의 민간 투자를 유치하여, 2024년 조직의 78%가 AI를 사용(전년 55%)
4	**미국 주도 속 중국의 성능 격차 축소**	• 미국은 2024년에 40개의 주목할 만한 AI 모델을 개발하며 (중국 15개, 유럽 3개) 개발을 주도하고 있는 가운데, 중국은 성능 격차를 좁히고 있는 중 • 중국은 MMLU 및 HumanEval과 같은 주요 벤치마크에서 2023년 두 자릿수였던 성능 차이가 2024년에는 거의 동등한 수준으로 축소되었으며, AI 간행물 및 특허 분야에서 계속 선두 유지
5	**책임 있는 AI(RAI) 거버넌스의 불균형한 발전**	• AI 관련 사고는 급격히 증가하고 있는 한편, 주요 산업 모델 개발자들 사이에서 표준화된 RAI 평가는 드문 상황 • 2024년 OECD, EU, UN 등 기관들이 투명성과 신뢰성에 중점을 둔 프레임워크를 발표하며 AI 거버넌스에 대한 글로벌 협력을 강화하는 등 글로벌 정부 차원의 긴급 대응책 마련 중
6	**AI 낙관론 증가와 지역적 격차 공존**	• 중국(83%), 인도네시아(80%) 등에서는 AI가 유익하다는 낙관론이 강한 한편, 캐나다(40%), 미국(39%) 등 일부 지역에서는 낙관론이 훨씬 낮는 상황 • 그럼에도 불구하고, 2022년 이후 독일(+10%), 프랑스(+10%), 미국(+4%) 등 이전에 회의적이었던 여러 국가에서 낙관론이 확산되는 중
7	**AI의 효율성, 가격, 접근성 향상**	• GPT-3.5 수준 성능 시스템의 추론 비용은 2022년 11월부터 2024년 10월 사이에 280배 이상 감소 • 개방형 가중치(Open-weight) 모델은 폐쇄형 모델과의 성능 차이를 1년 만에 8%에서 1.7%로 줄이면서, 고급 AI에 대한 장벽을 빠르게 낮추고 있음
8	**규제와 투자를 통한 정부의 개입 강화**	• 2024년 미국 연방 기관들은 2023년의 두 배 이상인 59개의 AI 관련 규제를 도입했으며, 전 세계적으로 75개국에서 입법 분야에서의 AI 관련 언급이 2023년 이후 21.3% 증가 • 캐나다는 $24억, 중국은 $475억 반도체 펀드, 사우디아라비아는 $1,000억 이니셔티브를 약속하는 등 정부 차원의 대규모 AI 투자 지속
9	**AI 교육 확대와 접근성 및 준비 상태 격차 지속**	• 현재 국가의 3분의 2가 K-12 CS 교육을 제공하거나 계획 중(2019년의 두 배)이나 아프리카에서는 전기와 같은 기본 인프라 격차로 접근성이 제한적 • 미국 K-12 CS 교사의 81%는 AI가 기초 CS 교육에 포함되어야 한다고 보지만, 절반 미만이 실제로 가르칠 준비가 되었다고 느낌
10	**산업계의 선두와 최전선 경쟁 심화**	• 2024년 유력 AI 모델의 거의 90%가 산업계에서 나왔으며 (2023년 60%). 모델 규모도 훈련 컴퓨팅, 데이터셋, 전력 사용량 모두 빠르게 증가 • 하지만 성능 격차는 줄어들어, 상위 모델과 10위 모델 간의 Elo 기술 점수 차이는 1년 만에 11.9%에서 5.4%로 감소했고, 최상위 두 모델 간의 차이는 0.7%에 불과
11	**과학적 영향에 대한 최고 영예 획득**	• 딥러닝과 단백질 분야 응용으로 이어진 AI 연구에 대해 2개의 노벨상을 수여 • 튜링상은 AI 기술의 획기적 향상에 기여한 강화학습에 수여
12	**복잡한 추론 능력의 한계**	• AI 모델은 국제 수학 올림피아드 문제와 같은 작업에서는 능력을 보이지만, PlanBench와 같은 복잡한 추론 벤치마크에서는 여전히 어려움을 겪고 있는 중 • 1로 해결하지 못하여, 정밀도가 중요한 고위험 설정에서 효과가 제한적

출처: Stanford University, Artificial Intelligence Index Report 2025

미국의 클라우드 서비스 제공사인 Salesforce가 2025년 7월 발표한 글로벌 AI 지수(The Global AI Index)[2]에 따르면, 조사 대상 16개 국가 중 우리나라는 미국, 싱가포르, 영국, 캐나다, 독일에 이어 6위에 기록됐다. 대부분의 국가들은 지수 평균에 근접한 점수를 기록하며 전반적인 준비 태세가 상향 평준화되는 경향을 보였다. 해당 지수는 정책, 확산, 혁신, 투자, 인적 자본의 다섯 가지 주요 차원에서 AI 준비도를 평가하여 종합 순위를 결정했다.

[그림 2-2] 글로벌 AI 지수 국가별 종합 순위



출처 : Salesforce, Global AI Readiness Index, 2025.7

주요 차원별 글로벌 지수 평가 결과는 다음과 같이 요약할 수 있다.

○ 규제 프레임워크(1차원): 대부분의 국가에서 가장 높은 준비도를 보인 영역으로, 거의 모든 국가가 국가 AI 전략과 데이터 보호 체계를 갖추고 있으며, 싱가포르, 영국, 호주가 특히 실용적인 정책 실행에서 두각

2) Salesforce, Global AI Readiness Index, 2025.7

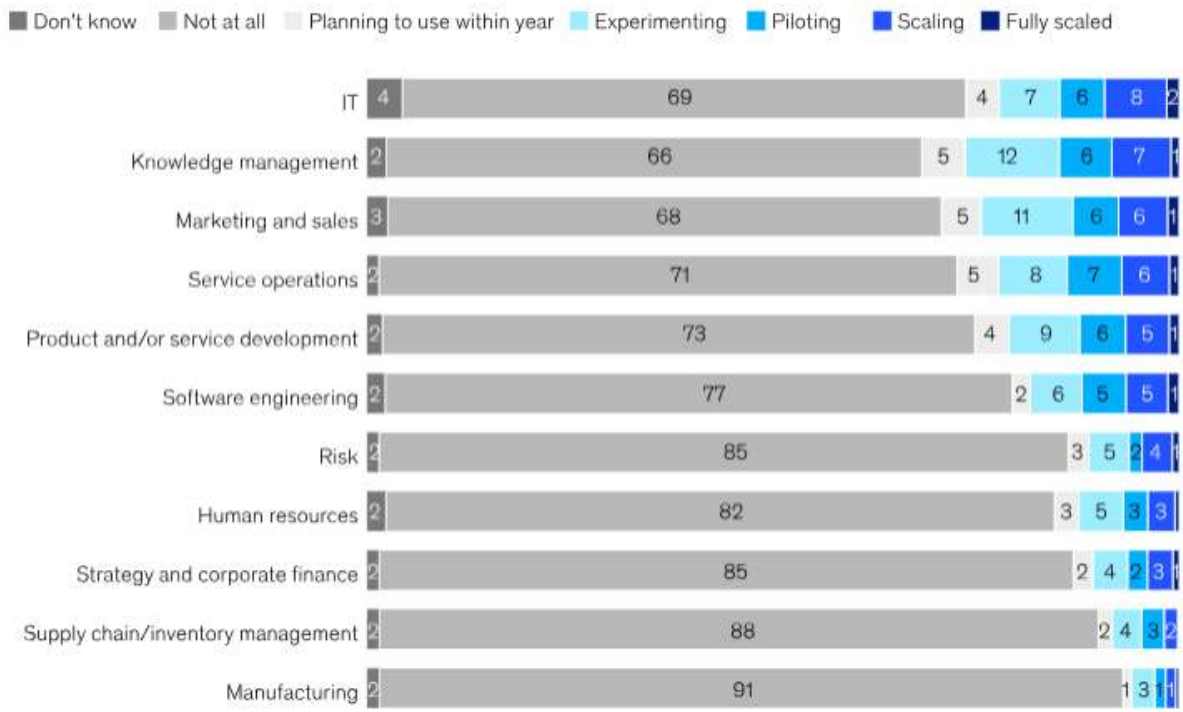
- 확산 및 도입(2차원): 많은 국가들이 시범 사업을 넘어 공공 서비스와 비즈니스 프로세스에 AI를 통합하며 강한 성장세
- 혁신 (3차원) 및 투자 (4차원): 국가 간 격차가 가장 큰 영역으로, 많은 국가에서 병목 현상을 겪고 있으며, 미국이 연구, 상용화, 자본 동원 능력에서 압도적 우위 유지
- 인적 자본(5차원): AI 준비도의 결정적인 동인으로 작용하며, 강력한 직업 교육 시스템과 기술 향상 전략을 갖춘 국가들이 우수한 평가 획득

맥킨지앤컴퍼니는 2025년 11월 전 세계 기업들의 AI 도입 및 활용 현황과 관련한 조사 결과[3]를 발표했다. 이 조사에 따르면, 전 세계적으로로 기업 AI 도입은 보편화되었으나 '확장'은 초기 단계에 머무르고 있는 것으로 분석되었다. 구체적인 내용을 살펴 보면, 응답 기업의 약 88%가 적어도 하나의 비즈니스 기능에 AI를 정기적으로 사용하고 있다고 답해 전년 78% 대비 10%p가 증가해, AI 사용은 기업 환경에서 일반적인 현상으로 자리잡았다. 그러나 전체 기업 수준에서 AI를 본격적으로 확장(Scaling)하고 있는 기업은 23%에 불과하다. 대다수(약 3분의 2)는 여전히 실험이나 파일럿 단계에 머물러 있어, 파일럿에서 실제 대규모 적용으로 넘어가는 데 어려움을 겪고 있다. 기업 규모별로도 차이가 발생하고 있는데, 매출 50억 달러 이상의 대기업은 절반 가까이가 확장 단계에 진입한 반면, 소규모 기업은 그 비율이 현저히 낮은 것으로 나타났다.

특히 이 조사에서는 'AI 에이전트(AI Agent)'의 부상에 주목했는데, 응답자의 62%가 조직 내에서 AI 에이전트를 실험 중인 것으로 확인됐다. 약 23%의 기업이 적어도 한 곳 이상의 부서에서 에이전트 시스템을 확장하고 있다고 답했지만, 특정 기능(부서) 내에서 완전히 확장된(fully scaled) 비율은 3% 미만으로 아직 초기 단계인 것으로 나타났다. AI 에이전트는 'IT(서비스 데스크 관리)'와 '지식 관리(심층 연구)' 분야에서 가장 활발하게 도입되고 있는 추세로 밝혀졌다.

3) McKinsey&Company, The state of AI in 2025: Agents, innovation, and transformation, 2025.11.5

[그림 2-3] 응답 조직의 AI 에이전트 활용 단계: 사업 기능별(복수 응답, n=1,933)



출처 : 맥킨지앤컴퍼니(2025.11)

이외 산업별로 볼 때, 기술(Tech) 분야가 여전히 앞서가고 있지만, 미디어·통신 및 보험 산업도 AI 도입률이 빠르게 증가하여 기술 기업 수준에 근접했다. 고용 측면에서는 AI 도입이 내년도 인력 규모에 미칠 영향에 대해서는 의견이 엇갈렸으며, 혁신 기여와 관련해서는 응답자의 64%는 AI가 조직의 혁신을 가능하게 하고 있다고 평가했다.

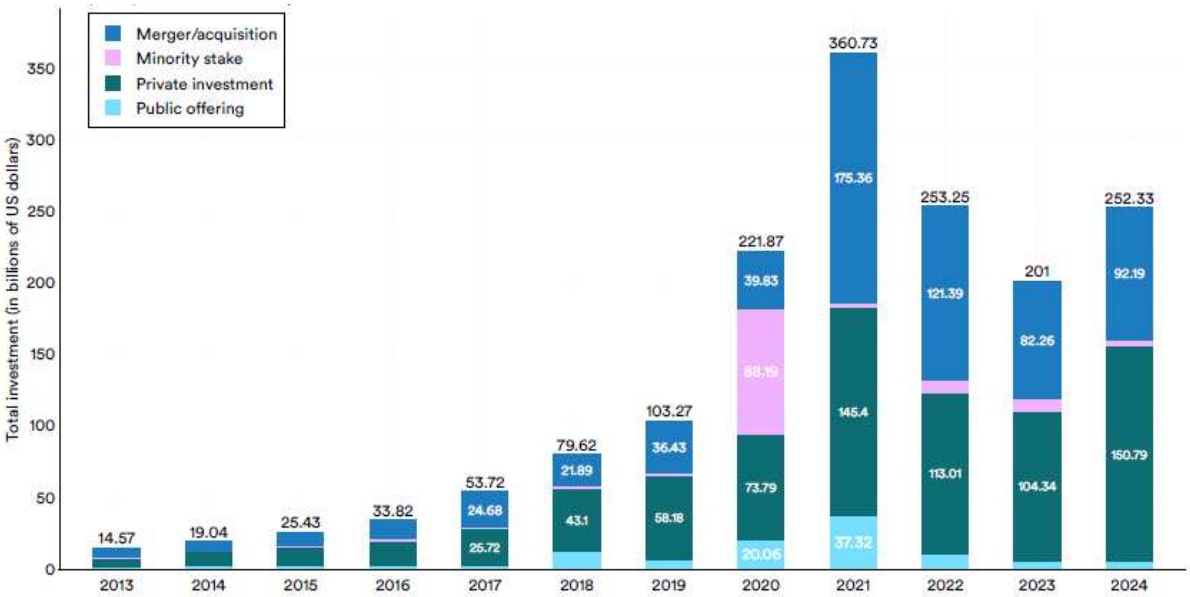
2025년은 인공지능이 화면 속 텍스트와 이미지를 넘어, 물리적 세계와 직접 상호작용하는 '피지컬 AI(Physical AI)'의 원년으로 기록되고 있다. 과거의 로봇이 특정 작업만 수행하는 '수동적 기계'였다면, 현재의 피지컬 AI는 VLA(Vision-Language-Action) 모델을 통해 스스로 사고하고 행동하게 됐다. 특히 NVIDIA는 11시간 만에 780,000개의 합성 궤적을 생성하는 기술을 선보이는 등 현실 데이터 수집의 한계를 극복하기 위해 합성 데이터(synthetic data) 생성 기술의 진보를 보였다. 모건 스탠리는 향후 10년 간 휴머노이드 로봇 시장이 60조 달러로 성장할 것으로 전망[4]한 가운데, 피지컬 AI는 새로운 국면을 맞이하고 있다.

4) Morgan Stanley, Mapping the Humanoid Robot Value Chain, 2025.2.6

2. 전 세계 AI 투자 현황

HAI AI 인덱스 보고서에 따르면, 2024년 총 투자액은 2,523억 달러로 2023년 대비 25.5% 증가했다. 가장 큰 상승세를 보인 분야는 민간 투자로 전년 대비 44.5% 증가했으며, 인수합병(M&A)은 12.1% 증가했다. 지난 10년간 인공지능(AI) 관련 투자는 약 13배 가까이 증가했다.

[그림 2-4] 투자 유형별 글로벌 AI 기업 투자: 2013~2024

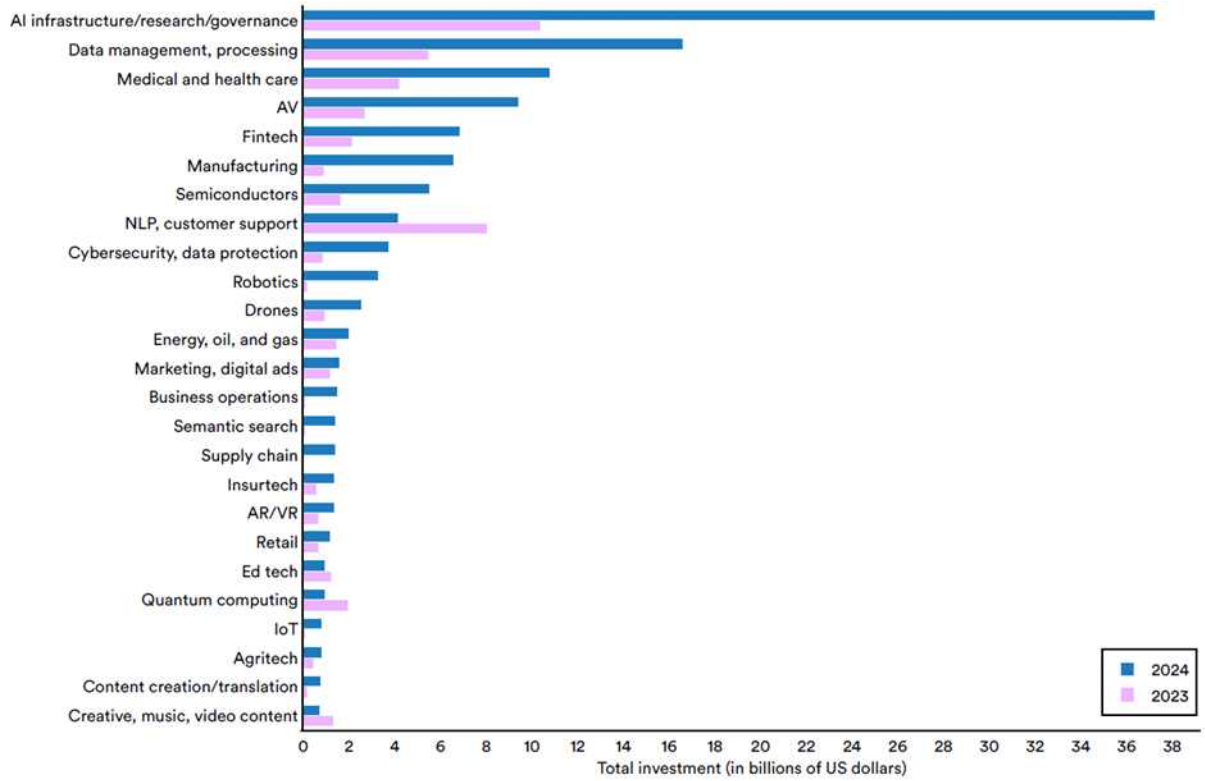


출처: Stanford University, Artificial Intelligence Index Report 2025

HAI의 민간 부문의 AI 투자 분야별 현황 분석에 따르면, 2024년 가장 많은 투자를 유치한 분야는 AI 인프라/연구/거버넌스(373억 달러), 데이터 관리 및 처리(166억 달러), 의료 및 헬스케어(110억 달러)가 차지했다. AI 인프라/연구/거버넌스 분야의 두드러진 투자 비중은 OpenAI, Anthropic, xAI 등 AI 애플리케이션 개발에 특화된 기업들에 대한 대규모 투자에 기인하고 있다. 특히 이 분야는 전년 대비 투자가 대폭 늘어나며 2024년 AI 투자의 중요한 트렌드를 형성했다. 이외 산업 분야별로는 의료 및 헬스케어 이외에 자율주행, 핀테크, 제조 등의 분야에서 전년 대비 고르게 높은 투자 신장세를 보였다.

출처: Stanford University, Artificial Intelligence Index Report 2024

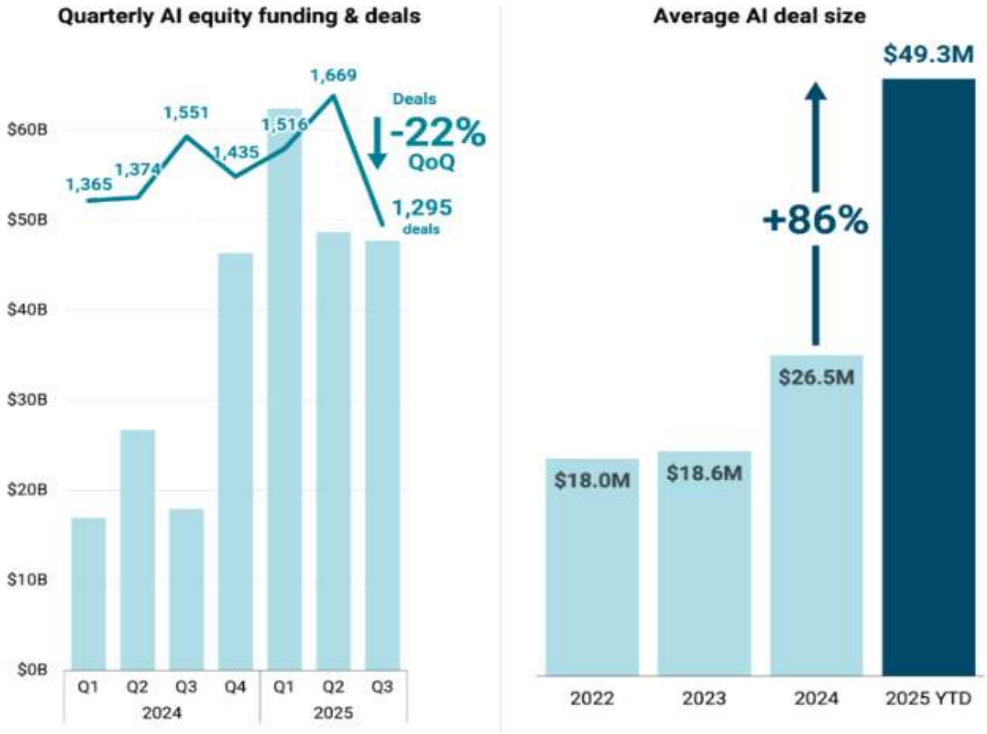
[그림 2-5] 분야별 글로벌 민간 AI 투자 비교: 2023 vs 2024



출처: Stanford University, Artificial Intelligence Index Report 2025

글로벌 시장조사기관인 CB인사이트는 2025년 3분기 전 세계 AI 투자 규모를 업데이트 공개한 가운데, 전 세계 AI 투자 규모가 전 분기 대비 22% 감소했으나, 규모 면에서는 4분기 연속 450억 달러 이상을 유지하고 있는 것으로 집계했다. 2025년 3분기까지의 평균 거래 규모는 4,930만 달러로, 2024년 대비 86% 증가했다. 지난 4분기 동안 메가 라운드(1억 달러 이상 거래)가 전체 자금의 75% 이상을 차지했으며, 2021년~2024년 3분기까지 메가 라운드의 전체 자금 평균 비중은 53%를 유지했다. 이는 투자자들은 대규모 인프라 비용과 파운데이션 모델 개발의 경쟁적 역학에 따라, AI 분야에서 유망한 기업으로 인식되는 소수 대상에 더 큰 규모의 자본을 집중하고 있다는 사실을 시사한다.실제 2025년 3분기에는 10억 달러 이상의 투자 라운드가 6건이나 발생했다. 상위 3개 거래는 LLM 개발사인 Anthropic(130억 달러, 시리즈 F), OpenAI(83억 달러, 사모펀드), Mistral AI(15억 달러, 시리즈 C) 등 프론티어 모델 개발에 집중됐다. OpenAI는 2025년 7월 연간 매출 120억 달러를 달성했으나, CB인사이트의 발표에 따르면 올해 약 80억 달러의 현금 소모를 예상하고 있다.

[그림 2-6] 글로벌 AI 투자 거래 현황 및 추이



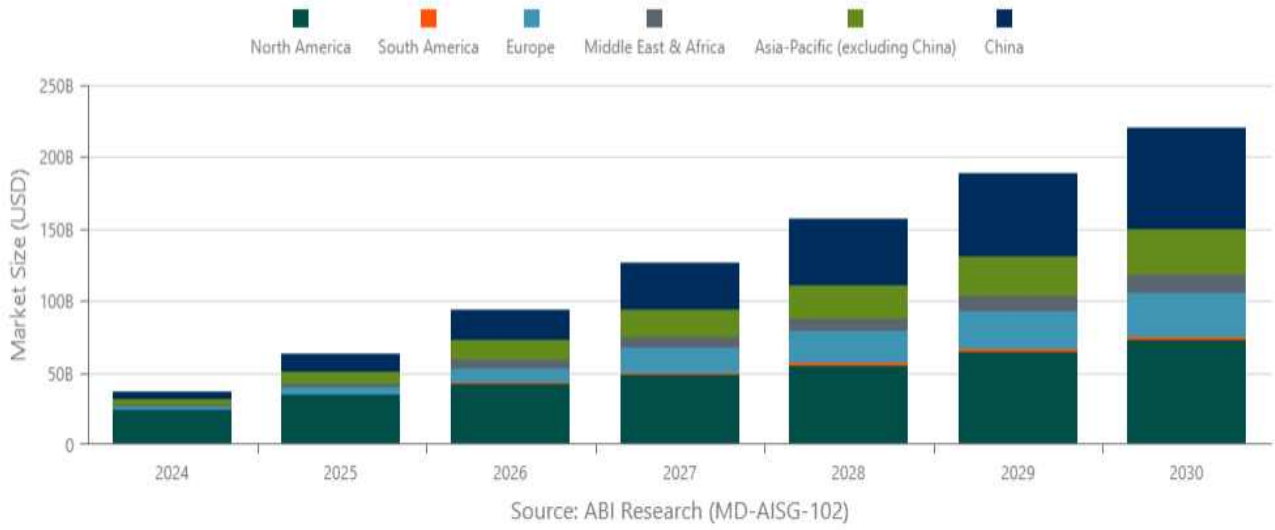
출처 : CB INSIGHTS, State of AI Q3’ 25 Report (2025.11.17. 방문)

CB인사이트가 트랙킹하는 1,500개 이상의 기술 시장 중에서는 투자 수 면에서 2025년 3분기 동안 산업용 휴머노이드 로봇 개발사(15건)와 코딩 AI 에이전트 및 코파일럿(13건)이 여전히 상위권을 유지한 가운데, LLM 개발사(10건)들도 2025년 2분기 대비 순위가 다시 상승했다. 최근 부상 중인 주목할 만한 시장은 생성형 엔진 최적화(GEO)로, ChatGPT 및 Perplexity와 같은 AI 검색 플랫폼에서 브랜드의 가시성을 최적화하는 도구를 개발하는 분야다. GEO 분야는 최근 LLM 인터페이스 상에서 쇼핑 기능을 제공하는 트렌드를 반영하고 있는데, OpenAI 역시 2025년 9월 ChatGPT 내 플랫폼 쇼핑 기능을 출시한 바 있다. 이는 향후 AI 플랫폼을 새로운 상거래 채널로 자리매김시키는 데 중요한 역할을 하게 될 것이다.

제2절 전 세계 AI 시장 전망

글로벌 시장조사 기관인 ABI Research는 2025년 2/4분기 시장 데이터 공개를 통해 AI 소프트웨어 시장 규모가 2024년 기준 1,220억 달러로 집계한 것으로 발표[5]했다. 이 시장은 연평균 복합 성장률(CAGR) 25%로 성장하며, 2030년 4,670억 달러에 도달할 전망이다. 글로벌 AI 소프트웨어 시장의 주요 성장 동력은 기업 전반과 다양한 사용 사례에 걸친 생성형 AI 도입에 크게 기인한다. 2025년에는 전체 AI 소프트웨어 투자의 54%가 북미 기반 기업에서 나올 전망이다. 아시아태평양 지역은 2025년 AI 소프트웨어 매출의 33%를 차지하지만, 중국이 미국과의 AI 경쟁에 본격적으로 가세함에 따라 2030년까지 해당 지역이 시장의 47%까지 늘어날 것으로 예상했다. ABI Research의 예측에 따르면 2030년까지 아시아태평양 지역 AI 소프트웨어 총 매출(1,495억 달러)의 3분의 2를 중국이 단독으로 차지할 전망이다. ABI 리서치는 미중간 AI 주도권 경쟁으로 인해 10년 말까지 북미의 인공지능 소프트웨어 매출 점유율이 33%로 감소할 것으로 내다보고 있다.

[그림 2-7] 생성형 AI 소프트웨어 시장 규모 전망: 2024~2030



출처 : ABI Research(2025, Q2)

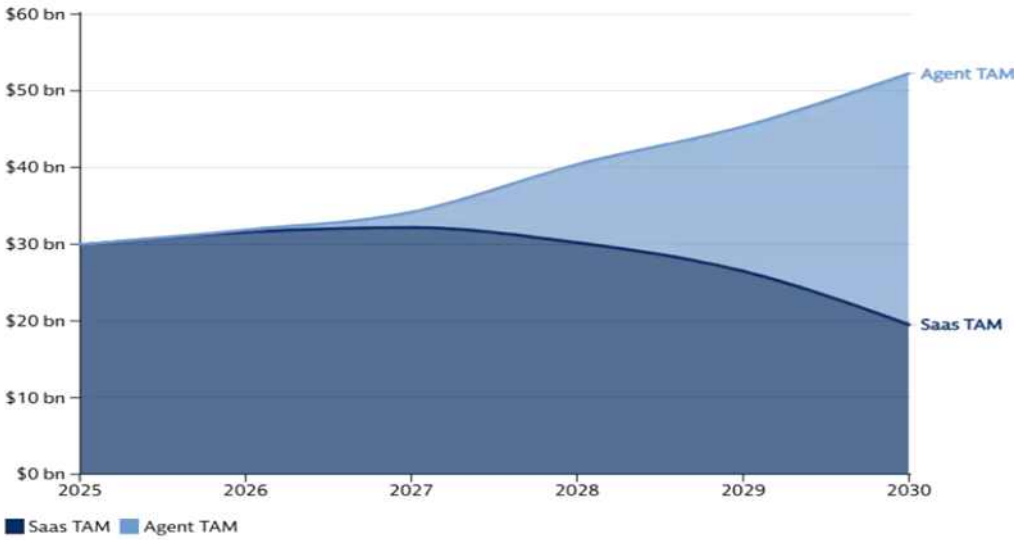
생성형 AI의 경우, 시장 예측 기간 동안 34.5%의 CAGR로 가장 빠르게 성장하는 분야로, 파운데이션 모델, 최적화 소프트웨어 및 모델 배포 도구 분야의 확산이 기대된다. 기업 단위의 AI 전략 성숙도가 높아지며 기존 AI(AI 센싱, 예측 AI, 자연어 처리(NLP))는 AI 소프트웨어 시장

5) ABI Research, Artificial Intelligence (AI) Software Market Size: 2024 to 2030, 2Q 2025

의 총 매출을 지속적으로 주도하게 될 것이다. 또한, 대규모 언어 모델(LLM)은 기존 챗봇이나 유사한 형태에서 다단계 작업을 수행할 수 있는 에이전트형 AI로 전환되고 있다. 이는 운영적 특성을 지닌 서비스로 전환되면서 지출 구조를 변화시키고 다양한 서비스 활용을 유도하게 될 것으로 보인다. 예측 기간 동안 생성형 AI의 산업 및 사용 사례 전반에 걸친 광범위한 적용성을 고려할 때, 생성형 AI와의 매출 격차는 지속적으로 축소될 전망이다.

한편, 글로벌 투자 기관인 Goldman Sachs는 글로벌 소프트웨어 시장 관점에서 AI 에이전트의 확산이 큰 시장 전환점을 마련할 것으로 관측했다[6]. 골드만삭스는 AI 에이전트의 도입으로 인해 2030년까지 글로벌 소프트웨어 시장 규모가 기존 전망치 대비 최소 20% 이상 추가 확대될 것으로 전망했다. 가장 큰 수혜가 예상되는 분야로는 고객 서비스 소프트웨어 분야로, 기존 SaaS 제품에 AI 에이전트가 결합되면서 시장이 20%에서 최대 45%까지 추가 성장할 것으로 예상했다. 시장 점유율면에서는 2030년이 되면 전체 소프트웨어 산업(SaaS + 에이전트)의 60% 이상을 '지능형 AI'가 차지할 것으로 내다봤다.

[그림 2-8] SaaS와 에이전트 AI의 전체 시장 규모(TAM) 전망 비교



출처 : Gartner(2024.10) 자료를 토대로 Goldman Sachs 분석(2025.7)

한편, 시장 조사기관 Precedence Research에 따르면, 전 세계 피지털 AI 시장은 2024년 기준 41.2억 달러로 2034년까지 연평균 31.3%로 성장하며 611.9억 달러의 규모를 형성할 것으로 전망된다. 로봇공학, 체화 지능(embodied intelligence), 센서 기반 AI 시스템의 급속한 발전이 의료, 제조, 물류 분야에서 전례 없는 자동화를 주도하며 피지털 AI 솔루션에 대한 수요를 촉진시킬 것으로 예상된다.[7]

6) Goldman Sachs, AI Agents to Boost Productivity and Size of Software Market, 2025.7.3
7) Precedence Research, Physical AI Market Size to Expand at 31.26% CAGR, to Capture USD 61.19 Billion 2034, 2025.11.25

제3장 주요국 인공지능 정책 동향

2025년 전 세계 주요국 정부가 추진한 AI 정책을 살펴보면 다음과 같은 주요 트렌드를 확인할 수 있다.

● ‘강한 규제’보다 규제를 정리·완화하면서 적극적인 산업 적용 도모

EU는 2024년의 AI 혁신 패키지·AI 팩토리를 산업 성과로 연결하는 실행 플랜인 Ai 적용 전략(2025년 10월)을 발표하면서 AI법과 인프라를 토대로 정책 초점을 ‘위험 규율’에서 ‘활용·확산’으로 옮기는 작업을 단행했다. 또한 디지털 옴니버스 규제 제안(2025년 11월)을 통해 규제 단순화·정비를 실시하고 “규제의 틀 안에서 최대한 AI를 활용한다”는 2단계로 진입했다. 일본은 ‘AI 추진법’ 제정에서 “AI 연구개발·활용을 적극적으로 추진해 산업 경쟁력 강화와 사회문제 해결”을 목적으로 밝혔으며, 미국은 트럼프 정부의 행정명령 제14179호 ‘미국의 AI 리더십의 장벽 제거’(2025년 1월)에서 기존 정부의 “과도한 규제와 이데올로기적 제약을 완화해 AI 혁신을 가속한다”는 친혁신·탈규제 기조를 천명했다. 이밖에 캐나다는 강한 AI 규제법으로 설계됐던 AIDA(Artificial Intelligence and Data Act) 입법을 2025년 1월 결국 폐기했다.

● 포괄적 AI 기본법·AI 규제체계 마련과 ‘시행 단계’ 진입

EU는 2025년부터 AI법과 데이터법 시행에 들어가면서, 위험기반 규제·데이터 접근·클라우드/데이터 이동성까지 포괄하는 디지털 규제 패키지를 가동하였다. 한국은 2025년 1월 「인공지능 기본법」을 제정하고 2026년 시행을 목표로 시행령 제정 작업을 추진하였고, AI 정의·책임·영향평가·사고 보고 등 포괄적 규율체계를 구축 중이다. 일본 역시 「AI 관련 기술의 연구개발 및 활용 추진에 관한 법률」(AI 추진법)을 통과시키고, AI 활용 촉진과 위험기관 원칙을 동시에 담은 첫 종합적인 기본법을 마련하였다.

● 데이터·반도체·연산 자원을 묶은 ‘국가 AI 역량’ 경쟁 심화

주요국 AI 전략의 공통 축이 ▲GPU·HPC 인프라 ▲초거대 AI 학습용 데이터 ▲AI 반도체·클라우드 생태계 확충으로 이동하였다. 한국의 「국가 AI 역량 강화 방안」(2025년 2월)은 초거대 AI 인프라, 데이터, 인력을 묶어 국가 역량으로 관리하는 대표 사례라고 할 수 있다. 영국의 AI 기획 실행계획도 안전연구, 연산자원, 데이터 접근, 인력 양성을 묶은 패키지 정책으로 설계되어 있고, 향후 2~3년간 단계적 이행 로드맵을 제시하고 있다.

● 책임 있는 AI 거버넌스와 규범 경쟁의 동시 심화

스탠퍼드 HAI와 보고서(2025)는 2019년 이후 책임 있는 AI(RAI) 원칙·규정은 빠르게 늘었지만, 책임성 벤치마크·평가 기준이 통일되어 있지 않아 규제 집행과 국제 비교에 어려움이 있다고 지적하였다. E는 AI법 집행을 위해 AI 사무국·감독기관을 설립하고, 일본과 영국은 데이터보호기관·시장경쟁당국(CMA) 등에 AI 감독 역할을 부여하는 등 기존 규제기관의 권한 재조정을 진행했다. 미국은 2025년 7월 행정명령(EO 14319)을 통해 연방조달용 LLM에 대해 '진실 추구'·'이념적 중립성' 원칙을 명문화하는 등 AI 내용·편향까지 규율하는 새로운 형태의 규범을 공개했다.

● AI 정책의 교육·노동·지역균형까지 아우르는 '전사회 전략화'

2025년을 기점으로 AI가 산업·정책을 넘어 교육·노동·사회 구조 전체를 재편하는 핵심 인프라로 부상했으며, 국가 간·계층 간 'AI 역량 격차'가 크게 벌어지고 있다. 한국의 「모두를 위한 AI 인재 양성 방안」(2025년 11월)은 생애주기별 AI 기본교육 강화, 지역 AI 인재 양성체계, 우수 인재 조기양성, 산업·학문 전반의 융합인재 양성을 포괄하는 대형 패키지로, AI 정책이 교육·지역정책과 결합되는 흐름을 보여준다.

● 미래 AI 대응으로 에이전트형 AI와 피지컬 AI에 대한 정책 부상

2025년 주요국 인공지능 정책에서 생성형 AI를 넘어 에이전트형 AI(Agentic AI)와 피지컬 AI가 새로운 전략 축으로 부상함을 엿볼 수 있다. 먼저 에이전트형 AI는 여러 시스템·데이터·규칙을 엮어 공공·기업 업무의 전 과정을 수행하는 워크플로형 도구로 인식되면서, 미국·OECD 국가를 중심으로 공공조달·샌드박스·거버넌스 기준이 '모델'이 아니라 '에이전트'를 전제로 재설계되는 추세다. 동시에 중국과 한국 등은 체화형 AI(Embodied AI)·휴머노이드·AI 로봇을 차세대 전략산업으로 격상해 제조·물류·모빌리티·서비스 산업의 핵심 인프라로 육성하고, 일본·EU·미국도 기존 로봇·스마트제조·자율시스템 R&D를 "AI+로봇" 관점에서 재정렬하고 있다. 이로써 2025년 AI 정책은 초거대 모델 중심 1단계를 넘어, 에이전트형 AI로 행정·화이트컬러 업무를, 피지컬 AI로 실물경제와 산업현장을 재구조화하려는 2단계 전환기라는 성격을 띠고 있다고 정리할 수 있다.

제1절 미국

1. 그간의 인공지능 정책 추진 경과

<표 3-1> 미국의 주요 AI 정책 추진 현황

시기	발표기관	정책
2015년	백악관(OSTP)	'미국을 위한 혁신 전략' 개정 - AI·로봇 등 전략기술 제시
2016년	백악관(OSTP)	국가 AI R&D 전략계획 발표 - 연방 AI R&D 투자 방향 제시
2019년 2월	대통령(백악관)	행정명령 제13859호 'AI에서의 미국 리더십 유지' - 국가 AI 이니셔티브 출발점으로 AI 연구·데이터·인재·국제협력 방향 규정
2019년 6월	백악관(OSTP)	국가 AI R&D 전략계획: 2019 업데이트 - 기존 전략 업데이트
2020년 1월	미 연방의회	국가 AI 이니셔티브법 제정 - AI 조정기구 및 연구·교육 프로그램 법제화
2021년 5월	연방정부	AI.gov 포털 개설
2023년 10월	대통령(백악관)	행정명령 제14110호 '안전하고 신뢰할 수 있는 AI를 위한 행정명령'
2023년 12월	국가과학재단	EducateAI 이니셔티브 발표
2025년 1월	대통령(백악관)	바이든 정부 행정명령 제14141호(AI 인프라 리더십 강화) 발표
2025년 1월	대통령(백악관)	트럼프 정부 행정명령 제14179호(미국 AI 리더십 장벽 제거) 발표
2025년 7월	대통령(백악관)	미국 AI 실행계획과 3가지 행정명령(EO 14318; 데이터센터 인프라 허가 가속), EO 14319; 연방정부 내 Woke AI 방지), EO 14320; 미국 AI 기술 스택 수출 촉진) 발표
2025년 11월	대통령(백악관)	트럼프 정부 행정명령 제14363호(제네시스 미션, Genesis Mission) - AI를 활용하여 혁신적인 과학적 발견을 가속화하고, 금세기의 가장 중대한 문제들을 해결하기 위한 국가적 차원의 노력 착수

1) 초기 AI 리더십 확보 및 진흥 정책(2019~2022)

미국은 2019년부터 2022년까지 AI 선도국 지위 확보와 산업 진흥에 중점을 둔 정책들을 추진했다. 트럼프 행정부 시기인 2019년 2월, 인공지능에서의 주도권 유지를 위한 대통령 행정명령(Executive Order 13859)이 발령되어 연방 차원에서 '미국 AI 이니셔티브(American AI Initiative)'를 출범시켰다. 이 행정명령은 연방 기관들에게 AI 연구개발(R&D)을 우선시하고, 연구 데이터와 컴퓨팅 자원의 개방을 독려하며, AI 기술 표준 수립과 국제 협력을 촉진하는 등의 방향을 제시하여 미국의 AI 기술 리더십 유지 기반을 마련했다. 같은 해 미국 백악관 과학기술정책실(OSTP)은 2016년 수립된 '국가 AI 연구개발 전략 계획(The National AI R&D Strategic Plan)'을 업데이트하여 2019년판 AI R&D 전략계획을 발표, AI 연구 투자 우선순위와 협력체계를 재정비했다.

2020년 말에는 미 의회가 「2020 국가 AI 이니셔티브 법」(National AI Initiative Act of 2020, NAIIA)[8]을 제정함으로써 미국 AI 정책의 법적 토대를 구축했다. 이 법은 2021년 1월 1일 발효되어 범정부적인 국가 AI 이니셔티브를 공식화한 것으로, AI 연구개발의 지속적 미국 리더십 보장, 신뢰할 수 있는 AI 시스템 개발과 사회 각 분야로의 AI 통합, AI 인재 양성 등을 국가 차원의 목표로 선언했다. NAIIA에 따라 백악관 OSTP 산하에 국가 AI 이니셔티브실(NAIIO)이 신설되어 각 부처의 AI 정책을 조율하고, AI 정책 데이터베이스 역할을 수행하게 되었다. 또한 동 법에 의해 국가 AI 자문위원회(National AI Advisory Committee)가 구성되어 민간 전문가들이 정부에 AI 정책 자문을 제공하기 시작했다. 연방 기관간 조정위원회도 구성되어 모든 연방 부처의 AI 연구·개발·시범·교육 활동을 강화하고 조율하는 프레임워크가 마련되었다. 이로써 미국은 혁신, 신뢰성, 인력양성, 인프라, 활용, 국제협력의 6대 전략 분야를 아우르는 종합적인 AI 추진체계를 법률로 갖추게 되었다.

이 시기 예산 투자와 인프라 확충도 활발했다. 2022년 한 해에만 미국 연방정부는 AI 분야에 약 17억 달러(한화 약 2조 5천억 원)를 투입하여 연구개발을 지원했다. 국방 분야에서도 인공지능 역량 강화를 위해 국방부 산하 합동인공지능센터(JAIC)를 출범(2018년)시킨 후, 2022년에는 이를 국방부 최고디지털·AI책임자실(CDAO)로 격상하여 AI 국방 적용을 본격화했다. 2022년 8월에는 반도체 공급망 강화와 과학기술 R&D 투자에 총 2,800억 달러를 배정한 「반도체 및 과학법」(CHIPS and Science Act)이 제정되어, 차세대 반도체와 AI 등 첨단기술 분야에 대한 대규모 투자도 시작되었다. 동 법에 포함된 과학기술 예산을 통해 국립과학재단(NSF)은 AI 연구 장학금·펠로우십 확대와 지역사회 기술인력 재교육(reskilling) 프로그램 등을 시행하고, AI 인재 양성을 뒷받침했다.

AI 연구 인프라와 데이터 개방 측면에서도 기반을 다졌다. 2021년 5월 백악관은 각 부처의 AI 정책과 자원 정보를 모아놓은 포털 사이트 'AI.gov'[9]를 개설하여 국민에게 정부의 AI 노력을 투명하게 공개하기 시작했다. 2022년 1월에는 국가 AI 이니셔티브실이 'AI 연구자 포털'을 구축하여 연구자들이 연방정부의 데이터셋, 컴퓨팅 인프라 프로그램, 연구비 지원 정보를 한 곳에서 활용할 수 있도록 했다. NSF는 2020년부터 국가 AI 연구소(National AI Research Institutes) 프로그램을 통해 대학·기업 컨소시엄 형태의 AI 연구센터들을 전국에 지정·지원하여, 기초연구와 인력양성 거점을 확대했다. 이와 함께 트럼프 행정부는 산업계의 자율적 혁신을 강조하며 연방규제기관에 신기술에 대한 과도한 규제를 지양하도록 권고했다. 이후, 2022년 10월에는 과학기술정책국(OSTP)이 AI 및 기타 자동화 시스템의 설계·개발·도입 과정에서 발

8) 미국 연방 하원 과학·우주·기술위원회, H.R. 6216 — National Artificial Intelligence Initiative Act of 2020, 2020.12.28. https://science.house.gov/2020/12/hr-6216-national-artificial-intelligence-initiative-act-2020

9) 미국 연방정부 AI 공식 포털 (2025.11.17. 방문) https://www.ai.gov

생할 수 있는 잠재적 부작용을 최소화하고, 시민의 권리를 보호하며, 민주적 가치를 반영할 수 있도록 「AI 권리장전」 청사진(Blueprint for an AI Bill of Rights)」[10]을 발표했다. 이 지침은 △안전하고 효과적인 시스템 △알고리즘 차별 방지 △데이터 프라이버시 △고지 및 설명 △인간 개입 및 대체 수단 보장 등의 핵심 5대 원칙을 제시하며 AI의 발전과 혁신을 안정적으로 추구하기 위한 기본 방침을 마련했다. 2019~22년 미국의 AI 정책은 '선 혁신 후 규율' 및 자율 규제 기치 하에 AI 편향성 완화와 윤리 연구에도 투자하였지만 전반적으로는 산업 발전과 기술 리더십 확보에 초점이 맞춰졌다.

2) 초기 AI 리더십 확보 및 진흥 정책(2019~2022)

2023~2024년 바이든 행정부는 AI 정책 방향을 "안전하고 신뢰할 수 있는 AI"에 방점을 찍으며, 기존의 진흥 일변도 접근에 책임성과 규범을 접목하는 노력을 기울였다. 이는 앞서 2022년 10월의 비전 문서인 「AI 권리장전 청사진」의 연장선에서 이뤄진 조치로 이러한 원칙들은 이후 연방 및 주정부의 입법과 행정명령에 기준점을 제공하며 AI 거버넌스 담론을 주도했다.

2023년 1월, 상무부 산하 국립표준기술연구소(NIST)는 'AI 위험관리 프레임워크'(AI Risk Management Framework) 1.0 버전[11]을 발표하여, AI 시스템의 위험요소 식별·측정·관리 방법에 대한 가이드라인을 제공했다. 이 프레임워크는 자발적 표준 형태로 민간과 정부에 배포되어, AI 개발·운영 시 안전성, 투명성, 공정성 확보를 위한 실천 지침 역할을 했다. 또한 바이든 대통령은 2023년 2월 행정명령 14091호 「소외계층 지원을 통한 인종 평등 증진(Further Advancing Racial Equity and Support for Underserved Communities through the Federal Government)」을 통해 연방기관들이 AI를 사용할 때 사회적 약자에 대한 편향을 시정하고 형평성을 증진하도록 지시하는 등, AI 기술의 윤리적 활용을 강조했다.

바이든 행정부 AI 정책의 정점은 2023년 10월 30일 발표된 대통령 행정명령 14110호 「안전하고 신뢰할 수 있는 AI의 개발 및 사용(Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence」[12]이다. 이 행정명령은 AI의 잠재력을 극대화함과 동시에 위험을 선도적으로 관리하기 위한 100여 개에 달하는 조치를 담은 포괄적 명령으로 평가된다. 행정명령 14110호는 연방정부에 다음과 같은 6대 중점 분야의 조치를 지시했다.

---

10) The White House, Blueprint for an AI Bill of Rights: Making Automated Systems Work for the American People, 2022.10

11) NIST, Artificial Intelligence Risk Management Framework (AI RMF 1.0), 2023.1. https://nvlpubs.nist.gov/nistpubs/ai/nist.ai.100-1.pdf

12) The Whitehouse, Fact Sheet: Biden-Harris Administration Announces Key AI Actions Following President Biden's Landmark Executive Order. 2024.1.19 https://www.whitehouse.gov/briefing-room/statements-releases/2024/01/29/fact-sheet-biden-harris-administration-announces-key-ai-actions-following-president-bidens-landmark-executive-order

<표 3-2> 미국 행정명령 14110호 6대 중점 분야

구분	주요 내용	
1	AI 안전성과 보안 강화	• 가장 강력한 AI 시스템을 개발하는 기업들을 대상으로 국방물자생산법을 활용하여 모델 안전성 테스트 결과 및 대규모 학습용 컴퓨팅 자원 현황을 연방 정부(상무부)에 보고 의무화 • 중요한 인프라에 대한 AI 위험 평가, 악의적 AI 사용에 대한 대책 마련 등을 각 부처에 90일 이내 수행 지시
2	개인정보 보호	• AI 개발 시 개인정보 오남용을 막기 위해 프라이버시 강화 기술 개발을 지원하고, 연방기관들이 개인 데이터 보호 조치를 강화 • 안전하지 않은 AI 사용으로부터 소비자의 프라이버시를 지키기 위한 권고안을 마련
3	형평성과 시민권 보호	• AI 시스템이 차별이나 편견을 확대하지 않도록 법무부, 국가과학기술위원회 등에게 기술 표준 개발과 평가 지시 • 고용·신용·주거 등 영역에서 AI로 인한 차별 감시를 강화하도록 EEOC(고용평등위)와 CFPB(소비자금융보호국) 등에 지침을 부여
4	소비자와 노동자 보호	• FTC(연방무역위원회)와 노동부 등에 AI를 활용한 사기, 조작으로부터 소비자를 보호하고, 근로자 감시나 자동화로 인한 부당한 처우를 방지하기 위한 대책을 마련 명령 • AI가 일자리와 노동조건에 미치는 영향을 연구하도록 지시
5	혁신과 경쟁 촉진	• 과도한 규제가 혁신을 저해하지 않도록 규제혁신 촉진팀을 구성 • AI 스타트업 지원 예산을 확충하며, 오픈소스 AI 개발을 지원하는 등 책임있는 혁신을 장려
6	미국의 글로벌 리더십 강화	• 국제 표준과 규범을 주도하기 위해 동맹국들과의 협력을 확대하고, AI 분야 해외 투자 규제 및 첨단 부품 수출통제를 통해 적대국의 AI 역량 강화를 견제

출처 : The White House, Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence, 2023.10

이처럼 2023년 말의 행정명령을 기점으로 미국은 연방 행정부 내부 운영부터 민간 기술 개발, 국제협력에 이르는 전방위적 AI 거버넌스 체계를 구축하기 시작했다. 백악관은 2024년 1월 브루스 리드(Bruce Reed) 부비서실장 주재로 백악관 AI 위원회를 출범시켜 각 부처 장관급 인사들이 모인 가운데 행정명령 이행상황을 점검했다. 회의 결과 행정명령이 지시한 90일 이내 과제들은 대부분 완료되었고, 장기 과제들도 이행 계획이 보고되어 AI로 인한 위험으로부터 국민을 보호하면서 혁신을 촉진하는 데 상당한 진전이 있었음을 발표했다.

동 기간 입법·예산 측면에서도 신뢰성 강화 기조가 반영되었다. 2023년 말부터 2024년까지 미 의회에서는 AI 투명성과 안전을 다루는 다수의 법안들이 발의되었고, 양당 모두 AI 대응 필요성에 공감하면서 초당적 AI위원회 구성 논의가 이루어졌다. 예산 면에서는 2024년도 연방 예산에 AI 윤리 연구 및 규제기관 역량 강화를 위한 항목이 증액되었다. 예를 들어 NIST의 신뢰할 수 있는 AI 연구 예산이 확대되고, FTC 등 규제기관에 AI 감시 인력 충원을 위한 예산이 배정되었다.

인재 양성 및 교육 분야에서도 중요한 정책들이 시행되었다. 바이든 행정부는 AI 시대에 맞춘 인력 정책으로서 연방정부 AI 인재 확보와 국민 AI 역량 제고를 병행했다. 행정명령 14110호에는 "대규모 AI 인재 확보(talent surge)" 계획이 포함되어, 인사관리처(OPM)에 각 부처가 AI 인재를 유치할 수 있는 특별채용 인센티브와 유연한 보수 지급 지침을 마련하도록 했다. 이에 따라 연방인사관리처(OPM)는 2024년 4월 AI 및 데이터 분야 인재 채용을 위한 역량 기반 채용 가이드라인(Skills-Based Hiring Guidance and Competency Model for Artificial Intelligence Work)[13]을 발표하여, 학위 대신 기술 역량으로 인재를 평가하는 새로운 기준(일반역량 43개, 기술역량 14개)을 제시했다. 또한 2024년 2월에는 우수 AI 인재에게 기본급의 최대 25%까지 보너스를 지급할 수 있도록 급여 규정을 변경하고, 2025년까지 연방 IT직군을 전면 스킬 기반 채용제로 전환하는 시범 프로그램도 시작했다. 한편 NSF는 미래 AI 인력 양성을 위해 2023년 12월 'EducateAI' 이니셔티브를 출범[14]하여 초·중·고교부터 고등교육, 직업훈련에 이르기까지 AI 교육 커리큘럼 개발과 교사훈련, 장학금 프로그램 등을 지원하기 시작했다. 이러한 노력은 미래 인재 풀 확보와 AI 활용 역량 제고를 통해 장기적으로 신뢰할 수 있는 AI 생태계를 구축하려는 전략적 투자로 평가된다.

기술 인프라 및 산업 지원 정책 역시 신뢰성 프레임워크 하에 추진되었다. 2024년 3월, 백악관 관리예산국(OMB)은 메모랜덤 M-24-10 「연방기관 AI 활용을 위한 거버넌스·혁신·위험관리 강화(Advancing Governance, Innovation, and Risk Management for Agency Use of Artificial Intelligence)」[15]를 발령하여, 처음으로 연방 기관들의 AI 활용에 법적 구속력 있는 책임 지침을 도입했다. 이 정책은 모든 연방기관이 AI 시스템을 도입할 때 위험관리 프로세스를 수립하고, 투명성·공정성·보안성 등의 요건을 충족하도록 의무화한 것이다. 이어 2024년 9월 OMB는 메모랜덤 M-24-18 「정부의 책임 있는

13) OPM, Skills-Based Hiring Guidance and Competency Model for Artificial Intelligence Work, 2024. https://www.opm.gov/chcoc/transmittals/2024/Skills-Based%20Hiring%20Guidance%20and%20Competency%20Model%20for%20Artificial%20Intelligence%20Work.pdf

14) NSF, Advancing education for the future AI workforce (EducateAI), 2023.12.1

15) OMB, Advancing Governance, Innovation, and Risk Management for Agency Use of Artificial Intelligence, 2024.3.28. https://www.whitehouse.gov/wp-content/uploads/2024/03/M-24-10-Advancing-Governance-Innovation-and-Risk-Management-for-Agency-Use-of-Artificial-Intelligence.pdf

AI 조달 증진(Advancing the Responsible Acquisition of Artificial Intelligence in Government)」[16]을 추가로 발표하여, 연방정부가 솔루션을 도입할 때 공급업체에게 AI 위험정보 공개와 편향 방지 조치를 요구하고, 신뢰성 검증을 거치도록 규정을 마련했다. 한편 NSF와 에너지부는 국가 AI 연구 자원(NAIRR) 구축을 추진하여, AI 연구자들이 고성능 컴퓨팅 인프라와 대형 데이터셋을 쉽게 활용할 수 있는 클라우드 플랫폼 시범사업을 2024년에 시작했다. 중소기업청(SBA)도 2024년 12월 애틀랜타에서 최초의 AI 중소기업 회의인 'SBA Atlanta Summit 2024'을 개최하여 '중소기업을 위한 AI 리소스 허브(AI for Small Businesses Resource Hub)'를 출범시켜, 중소기업들이 최신 AI 솔루션과 정부 지원 프로그램 정보를 접할 수 있게 했다. 또한 NIST는 2024년 7월 ' 생성형 AI 프로파일(Generative AI Profile) '[17]을 발간하여 생성형 AI의 특유 위험요소(예: 환각, 저작권 문제 등)를 식별하고 관리하기 위한 세부 지침을 업계에 제공했다. 이처럼 바이든 행정부는 민간 자율과 혁신을 저해하지 않는 범위 내에서 AI의 안전성과 책임성을 담보하려는 정책적 균형을 추구했다는 평가를 받는다.

3) 정권교체기 과도기 정책 동향(2024년 말~2025년 1월 중순)

2024년 11월 대통령 선거 결과에 따라 2025년 1월 정권 교체가 예고되면서, 2024년 말부터 2025년 1월 취임식까지 과도기적 AI 정책 동향에도 변화 조짐이 나타났다. 바이든 행정부는 임기 말까지 앞서 수립한 AI 거버넌스 계획들을 최대한 제도화하고자 노력했으며, 한편으로 차기 행정부는 새로운 우선순위에 맞춰 정책 전환을 준비했다.

우선 바이든 행정부는 2024년 10월 역사상 최초의 AI 관련 국가안보각서(NSM)를 발령[18]하여, 인공지능이 국방 및 외교에 미치는 영향과 대응원칙을 정리했다. 이 각서는 2023년 10월 행정명령 14110호의 지시에 따라 작성된 것으로, 안전하고 신뢰할 수 있는 AI의 국방 활용과 AI로 인한 신안보 위협 대응에 대한 정부 차원의 접근방식을 규정했다. 이를 통해 군사용 AI 개발 가이드라인, AI 대응 정보작전 전략, 동맹국과의 국방 AI 협력 등의 원칙이 마련되어, 차기 행정부에서도 국가안보 분야 AI 정책의 연속성을 일부 유지할 수 있게 했다.

기술 안보 측면에서, 바이든 행정부는 중국 등 경쟁국에 대한 첨단 반도체 및 AI 기술 통제 정책을

16) OMB, Advancing Trustworthy Artificial Intelligence through Acquisition, 2024.10. https://www.whitehouse.gov/wp-content/uploads/2024/10/M-24-18-AI-Acquisition-Memorandum.pdf
17) NIST, Artificial Intelligence Risk Management Framework: Generative Artificial Intelligence Profile, 2024.7.26
18) The White House, Memorandum on Advancing the United States' Leadership in Artificial Intelligence; Harnessing Artificial Intelligence to Fulfill National Security Objectives; and Fostering the Safety, Security, and Trustworthiness of Artificial Intelligence, 2024.10.24

임기 말까지 강화했다. 앞서 2022년 10월 미 상무부 산업안보국(BIS)은 대중국 첨단 AI 반도체 및 장비 수출 제한 규정을 발표한 데 이어, 2023년 10월에는 규제대상 칩의 성능 기준을 한층 낮추고 중국 기업의 해외 자회사에도 통제가 적용되도록 하는 추가 조치를 내놓은 바 있다. 또한 2023년 8월 바이든 대통령은 행정명령을 통해 미국인의 중국 첨단기술 분야 투자 제한을 도입, 중국의 AI칩 생산기업에 대한 미국 자본 및 전문성 유출을 막고자 했다. 이러한 대중 기술견제 조치들은 2024년 말까지 지속되었는데, 특히 2024년 12월 상무부는 고대역폭메모리(HBM) 칩과 AI 칩 제조용 핵심 장비의 대중 수출을 금지하는 추가 규제를 발표하며 AI 관련 핵심기술의 해외 유출 차단을 한층 강화했다. 이로써 바이든 정부는 퇴임 직전까지 미국의 AI 기술우위 수성과 국가안보 보호를 위한 방어막을 확고히 하려 했다.

한편, 정권 인수인계 과정에서 차기 행정부의 AI 정책 기조 변화가 서서히 감지되었다. 2024년 말 트럼프 당선인 측은 기존 행정부의 AI 규제 위주의 접근을 재검토하겠다는 의지를 밝히며, 과도한 규제가 미국 AI 경쟁력을 약화시켜선 안 된다는 입장을 피력했다. 이에 따라 인수위 단계부터 친기업·친혁신적 AI 정책 전환을 준비하는 움직임이 포착되었다. 실제로 차기 행정부는 취임과 동시에 발령할 새로운 AI 행정명령 초안을 2025년 1월 초에 마련한 것으로 알려졌다. 또한 바이든 행정부 말기에 예정되었던 일부 규제 조치는 보류되거나 변경 조짐을 보였다. 예를 들어, 바이든 정부는 퇴임 직전인 2025년 1월 "AI 확산 규칙(AI Diffusion Rule)"이라는 새로운 대외 수출통제 규칙을 예고하였는데, 이는 전 세계 국가들을 세 등급으로 구분해 AI 반도체 등이 경쟁국 및 적대국으로 유출되는 것을 방지하려던 조치였다. 그러나 2025년 5월 상무부 BIS의 발표에 따르면, 이 규칙은 결국 시행되지 않고 철회되었으며 차기 행정부에서 새로운 통제 방안을 마련하기로 했다. 이는 정권 교체로 인해 대중 기술통제 정책 역시 방향 수정이 이루어졌음을 시사한다.

과도기에는 또한 연방 차원의 AI 입법 움직임이 일시적으로 속도가 둔화되었다. 2024년 말 의회에서는 AI 관련 포괄법 제정 논의가 있었으나 정권 교체에 따른 우선순위 변화로 최종 입법까지 이어지지 못했다. 대신 2025년 새 의회에서 친기업 성향의 다수당을 중심으로 AI 혁신 장려법안들이 새로이 제출될 조짐을 보였다. 예를 들어 AI 혁신 촉진법, 불필요한 AI 규제완화 결의안 등이 거론되며, 행정부의 정책 전환과 보조를 맞추려는 움직임이 있었다.

전반적으로 2024년 말에서 2025년 1월 중순까지는 기존 정부의 AI 정책 마무리와 새로운 정부의 정책 준비가 교차하는 시기로서, 향후 방향성에 불확실성이 존재했다. 그러나 곧 출범할 트럼프 행정부가 이전 정부의 안전·윤리 위주의 기조를 대폭 수정하고 혁신 가속화에 나설 것이라는 점은 이미 정책 기

조 발언과 인수위 활동을 통해 예견되었다. 이는 곧바로 현실화되어 2025년 1월 취임 직후 새로운 행정명령 발동으로 이어지게 된다.

4) 트럼프 행정부의 친혁신·탈규제 중심 정책 전환(2025년 이후)

2025년 1월 출범한 트럼프 행정부는 AI 정책에서 혁신 촉진과 규제 완화를 최우선 가치로 내세우며, 전임 정부의 '안전·신뢰' 프레임워크를 급격히 전환하였다. 취임 직후인 2025년 1월 23일 트럼프 대통령은 행정명령 14179호 「미국의 AI 리더십의 장벽 제거」(Removing Barriers to American Leadership in Artificial Intelligence)[19]에 서명했다. 이 문서는 트럼프 2기 행정부 AI 정책 패러다임의 방향타로서, 바이든 행정부가 강조한 안전성과 규제 중심 접근에서 벗어나 자유로운 혁신과 글로벌 경쟁력 강화에 중점을 둔 새 방향을 천명하였다. 행정명령 14179호는 "이데올로기적 편향이나 특정 사회적 의제로부터 자유로운 AI 개발"의 필요성을 강조하면서, 이전 행정부 기간 마련된 AI 정책과 지침 중 미국 AI 혁신에 장애가 되는 요소들을 전면 철폐하겠다는 강력한 의지를 담았다. 특히 2023년 10월의 바이든 행정명령 14110호를 비롯한 AI 관련 정책·규정들을 포괄적으로 재검토하여 새로운 혁신 중심 노선과 불일치하거나 걸림돌이 되는 경우 즉시 중단·개정 또는 폐기하도록 지시했다. 즉, 안전성 강화 목적의 규제가 미국 AI 패권 유지에 방해된다면 예외 없이 완화하겠다는 것이다.

이를 통해 트럼프 행정부는 "인류 번영, 경제 경쟁력, 국가안보 증진을 위해 미국의 글로벌 AI 지배력을 지속·강화"하는 것을 AI 정책의 최우선 목표로 천명했다. 이러한 비전 하에 대통령은 과학기술 보좌관(OSTP), 국가안보보좌관, AI·암호화폐 특별고문 등을 주축으로 180일 이내에 구체적 실행계획을 수립하도록 명령하여, 새 정부 출범 반년 이내에 종합 전략을 마련토록 했다. 이후 2025년 7월, 미국은 사상 처음으로 다른 주요국들처럼 포괄적인 국가 AI 전략을 담은 '미국 AI 실행계획(America's AI Action Plan)'[20]을 공식 발표하게 된다. 2025년 11월에는 제네시스 미션(Genesis Mission)이라는 명칭의 행정명령을 통해 미국의 기술적 우위와 글로벌 전략적 리더십을 공고히 하기 위해 에너지부(DoE)가 중심이 되어 '미국 과학 및 안보 플랫폼(American Science and Security Platform)' 운영을 지시했다. 이 플랫폼은 연방 정부의 방대한 과학 데이터셋, 고성능 컴퓨팅 자원, AI 모델 및 실험 도구를 통합하여 과학 연구의 패러다임을 전환시키는 역할을 한다. 이 미션은 첨단 제조, 생명공학, 핵에너지, 양자 과학 등 국가적으로 중요한 과학기술 도전과제들을 해결하는 데 초점을 맞추며, 공공-민간 파트너십, 인재

19) The White House, Removing Barriers to American Leadership in Artificial Intelligence, 2025.1.23. https://www.whitehouse.gov/presidential-actions/2025/01/removing-barriers-to-american-leadership-in-artificial-intelligence

20) The White House, America's AI Action Plan, 2025.7 https://www.whitehouse.gov/wp-content/uploads/2025/07/Americas-AI-Action-Plan.pdf

양성, 국제 협력을 통해 그 목표를 달성할 계획이다.

이러한 일련의 행정명령과 정책들을 통해, 트럼프 행정부는 공격적인 친혁신 전략으로 방향을 선회했다. 규제기관 및 거버넌스 기구의 재편도 이어졌다. 2025년 6월 상무부는 기존에 운영되던 'AI 안전연구소(U.S. AI Safety Institute)'를 'AI 표준혁신센터(Center for AI Standards and Innovation, CAISI)'로 개편하였다[21]. 이는 AI 정책 우선순위를 안전보다 표준과 혁신으로 두겠다는 의지의 반영으로, AI 관련 표준화와 기술 확산을 주도하는 조직으로 재정비한 것이다. 비록 명칭은 바뀌었지만, 동 센터는 2025년 1월 워크숍을 개최하고 그 결과를 발표하는 등 여전히 활발히 운영되고 있는 것으로 전해진다. 아울러 백악관 OSTP도 AI 안전성 위원회 대신 AI 혁신 자문그룹을 구성하여 민간 전문가들로부터 규제 완화와 기술 촉진에 대한 조언을 구하는 한편, AI 윤리 관련 위원회의 활동은 축소 조정되었다.

국제 협력 측면에서, 미국은 동맹국들과의 AI 동맹 강화를 모색하고 있다. 트럼프 행정부는 2025년 7월 실행계획에서 밝힌 대로 첨단 AI 기술 동맹 구상을 추진하고, 영국, 일본 등과 AI 연구협정을 체결하여 대중 견제 연대를 구축하고 있다. 반면, OECD 및 G7 등 기존의 AI 거버넌스 논의에서는 과도한 글로벌 규제 논의에 제동을 거는 입장을 취하며, 국익과 기술패권을 최우선에 둔 실용적 협력을 강조하고 있다.

결과적으로 2025년 이후 미국의 AI 정책은 이전 정부의 '안전하고 책임있는 AI' 기조에서 '혁신 가속과 자유시장 중심' 기조로 급선회하였다. 연성 지침이었던 「AI 권리장전」 등의 접근은 트럼프 행정부의 시각에서는 과도한 규제 분위기를 조성할 우려가 있는 담론으로 평가되었고, 대신 "규제 장벽 제거와 미국 기술우위 공고화"가 정책 전면에 등장했다. 이러한 변화는 AI 정책에 있어서 민주당 정부의 신중한 책임론 vs 공화당 정부의 적극적 성장론이라는 극명한 대조를 보여준다. 향후 미국은 연방 차원의 AI 종합전략(Action Plan)을 통해 정부 역량을 총동원하여 AI 패권 경쟁에 나서는 한편, 민간의 혁신을 최대로 활성화하는 정책환경을 조성할 것으로 전망된다. 그러나 동시에 윤리적·사회적 측면의 고려 약화에 따른 부작용 우려도 제기되고 있어, 향후 정책 추진 과정에서 어떻게 혁신과 책임의 균형을 잡을지가 지속적인 도전 과제가 될 것이다.

21) Voices for Innovation, Commerce Department Announces Center for AI Standards and Innovation, 2025.6.8

<표 3-3> 미국의 AI 관련 규제와 진흥 정책 변화

구분	연도	정책
진흥정책	2015년	‘미국을 위한 혁신 전략(A New Strategy for American Innovation)’
진흥정책	2016년	‘AI R&D 전략’(The National AI R&D Strategic Plan)
진흥정책	2019년	‘AI에서의 주도권 유지를 위한 행정명령(Executive Order on Maintaining American Leadership in Artificial Intelligence)’
진흥정책	2019년	‘국가 AI R&D 전략 계획(National Artificial Intelligence Research and Development Strategic Plan: 2019 Update)’
진흥정책	2020년	‘2020 국가 AI 이니셔티브 법(National AI Initiative Act of 2020 (DIVISION E, SEC. 5001)’제정
진흥정책	2021년	‘AI.gov’ 설치
진흥정책	2022년	‘국방수권법(NDAA, National Defense AuthorizationAct)’통과
진흥정책	2022년	‘AI 연구자포털(AI Researchers Portal)’ 구축
규제정책	2023년	인공지능의 안전하고 보안성 있는 개발과 활용을 위한 행정명령 14110(Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence) 발표
진흥정책	2024년	EducateAI 이니셔티브 개시
규제정책	2024년	대통령 행정명령에 관한 주요 AI 조치 발표
규제정책	2024년	AI 혁신을 지원하는 주요 리소스에 관한 액세스 민주화를 통해 공유 연구 인프라 비전 실현을 위한 국가 인공지능 연구 리소스(NAIRR: National Artificial Intelligence Research Resource) 파일럿에 착수
규제정책	2024년	백악관 AI 행정명령 이후 후속 조치 현황
진흥정책	2025년	미국의 AI 리더십의 장벽 제거 행정명령 14179(Removing Barriers to American Leadership in Artificial Intelligence) 발표
진흥정책	2025년	미국 AI 실행계획(America’s AI Action Plan) 발표
진흥정책	2025년	미국의 AI 인프라 리더십 강화를 위한 행정명령 14141(Advancing United States Leadership in Artificial Intelligence Infrastructure) 발표
진흥정책	2025년	데이터센터 인프라 연방 허가절차 가속화를 위한 행정명령 14318(Accelerating Federal Permitting of Data Center Infrastructure) 발표
규제정책	2025년	연방정부 내 Woke AI 방지를 위한 행정명령 14319(Preventing Woke AI in the Federal Government) 발표
규제정책	2025년	미국 AI 기술 스택 수출 촉진을 위한 행정명령 14320(Promoting the Export of the American AI Technology Stack) 발표

구분	연도	정책
규제정책	2025년	「Take It Down Act」 제정
진흥정책	2025년	미국 AI 실행계획(America's AI Action Plan) 발표
진흥정책	2025년	제네시스 미션(Genesis Mission) 발표

2. 2025년 인공지능 정책 동향[22)]

1) 트럼프 행정부의 AI 정책 방향과 비전: EO 14179[23)]

(1) 배경 및 목적

바이든 행정부는 2023년 10월 AI 안전성 확보를 위한 행정명령(EO 14110)을 통해 연방 정부 차원의 통합적 규제 프레임워크를 수립하고 각 부처에 AI 책임자를 임명하는 등 안전과 신뢰를 중시하는 접근을 취했다. 그러나 2025년 1월 트럼프 대통령 취임과 동시에 해당 행정명령 이행은 중단되었고, 이전 정부의 규제 중심 기조에서 혁신과 경쟁력 강화로 전환하기 위한 새로운 정책이 추진되었다. 이러한 맥락에서 2025년 1월 23일 트럼프 2기 행정부는 첫 AI 관련 행정명령으로 「미국의 AI 리더십의 장벽 제거(Removing Barriers to American Leadership in Artificial Intelligence)」(EO 14179)를 발령하였다.

EO 14179는 글로벌 AI 패권 유지를 위해 기존의 과도한 규제와 이데올로기적 제약을 완화하고, AI 혁신을 가속화하는 것이 핵심 목표이다. 바이든 행정부가 강조했던 AI 안전성·규제 중심 접근에서 벗어나, 혁신과 경쟁력 강화를 중심에 둔 새로운 방향성을 제시하였다. 특히 "인간 번영, 경제 경쟁력, 국가안보 증진을 위해 미국의 글로벌 AI 지배력을 지속하고 강화"한다는 비전을 표명하며, AI 개발에서 특정 이념이나 사회적 의제에 치우치지 않은 중립성을 강조하였다. 이를 통해 AI 분야에서 미국의 선도적 지위를 공고히 하겠다는 의지를 천명한 것이다.

22) 2025년 이후 발표된 주요 정책을 중심으로 구성·기술하였으나, 과거에 발표된 정책 중 2025년 이후까지 연속성을 띠며 최근 정책의 특성을 중요하게 반영하고 있는 정책 항목도 일부 포함
23) Federal Register, Removing Barriers to American Leadership in Artificial Intelligence, 2025.1.31

(2) 주요 내용 및 실행방안

EO 14179는 행정부 차원의 AI 정책 패러다임 전환을 위한 구체 조치들을 담고 있다. 우선, 과학기술보좌관(OSTP), 국가안보보좌관(NSC), AI·암호화폐 특별고문을 주축으로 180일 이내에 미국 AI 실행계획(America's AI Action Plan)을 마련하도록 지시하여, 행정부의 AI 비전을 실행에 옮길 세부 계획 수립을 명령했다. 또한 바이든 행정부 시기의 "신뢰할 수 있는 AI 개발 및 사용" 관련 정책·지침·규정을 전면 재검토하고, 새로운 기조와 충돌하거나 혁신에 장애가 되는 기존 AI 규제는 중단·수정·폐지하도록 했다. 즉각 철폐가 어려운 경우에는 모든 가용한 면제 조치를 제공하여 규제 부담을 완화하라고 규정함으로써, 과감한 규제 완화를 통해 AI 혁신을 가속화하겠다는 구상을 구체화했다.

(3) 관계 부처 및 기관

백악관 과학기술정책실(OSTP)과 국가안보회의(NSC)가 중심이 되어 본 행정명령의 이행을 주도한다. 특히 과학기술보좌관과 국가안보보좌관이 특별고문과 함께 실행계획 수립을 총괄하며, 각 연방 부처들은 소관 분야에서 기존 AI 관련 규제와 지침을 재검토하고 필요한 완화 조치를 취하도록 지시받았다. 이 과정에서 관리예산국(OMB)은 정부-wide 지침 정비에 관여하고, 법무부 등은 행정명령에 따른 법적 검토를 수행한다. 나아가, OSTP 산하 국가 AI 이니셔티브실(NAIIO) 등 기존 AI 거버넌스 기구들도 새로운 정책 기조에 맞춰 조율 역할을 담당하게 된다.

2) 미국 AI 실행계획(America's AI Action Plan)[24]

(1) 배경 및 목적

미국은 그동안 행정명령을 수단으로 부처별 AI 정책을 추진해왔으나, 주요 경쟁국들이 2010년대 후반부터 국가 AI 종합전략을 갖춘 것에 비해 일원화된 전략이 부재하다는 지적이 있었다. 2021년 「국가 AI 이니셔티브법」 제정과 AI.gov 설치 등으로 정책 기반을 다졌으나, 여전히 각개 정책들의 나열에 머물렀다. 이러한 가운데 마침내 2025년 7월, 트럼프 행정부는 사상 첫 포괄적 AI 종합계획인 「미국 AI 실행계획」을 수립·발표하여 다른 주요국과 어깨를 나란히 하게 되었다. 이 계획은 앞서 EO 14179가 요구한 바에 따라 작성된 것으로, 행정부의 AI 비전을 구현하기 위한 연방 차원의 단기 실행목표들을 망라하고 있다.

---

24) The White House, White House Unveils America's AI Action Plan, 2025.7.23

‘미국 AI 실행계획’은 글로벌 AI 리더십 확보라는 국가 비전을 실현하기 위해 두 갈래 전략을 동시에 추구한다. 하나는 AI 혁신 역량의 극대화이고, 다른 하나는 경쟁국과의 기술 격차 확대 및 적대국의 AI 남용으로부터 국가 방어이다. 즉, AI 기술을 경제 성장과 국가안보, 산업경쟁력의 핵심 수단으로 인식하여 과감한 혁신 투자로 미국의 기술우위를 공고히 하는 한편, 중국 등 적성국의 AI 악용에 대비하여 안보 리스크를 차단하는 것이 목표이다. 이를 통해 미국 주도 기술 스택을 전 세계 표준으로 만들고, 동맹 결속을 강화함으로써 미국의 장기적 우위를 확보하려는 전략적 목적을 갖고 있다.

⑵ 주요 내용 및 실행방안

실행계획은 △AI 혁신 가속화, △미국 AI 인프라 구축, △글로벌 AI 외교 및 안보 주도의 세 개 전략축(three pillars) 아래 30개 실행과제를 제시한다. 실행계획에 담긴 30대 과제들은 위 세 축을 중심으로 구체적인 정책제언으로 연결되었다. 각 과제는 담당 주체와 이행 권고사항까지 포함하고 있다. 예컨대, ‘글로벌 AI 안보 강화’ 분야에서는 첨단 AI 컴퓨팅 수출통제 집행 강화, 반도체 수출통제 강화, 동맹국과의 수출통제 공조, 프런티어 AI 모델의 안보위험 평가체계 구축 등의 실행방안이 권고되고 있다.

⑶ 관계 부처 및 기관

‘미국 AI 실행계획’은 백악관 차원에서 수립된 국가 전략인 만큼, OSTP와 NSC가 정책 조정의 중심축을 이룬다. 계획 수립을 주도한 AI 특별고문과 과학기술·국가안보보좌관 외에도, 실행계획 이행에는 거의 모든 주요 부처가 관여한다. 예를 들면, 상무부는 AI 기술 수출 프로그램 및 표준개발을, 에너지부는 AI 인프라 전력망 및 초고성능 컴퓨팅을, 국방부는 군사 분야 AI 활용을 각각 책임지는 식으로 역할이 분담된다. 또한 NSF는 AI 연구인력 양성과 기초연구 지원을, OMB는 예산 편성과 연방기관 대상 지침을 통해 AI 활용·조달을 뒷받침한다. 이밖에 국가 AI 이니셔티브 거버넌스 기구들(NAIIO, AI 특별위원회 등)이 30여 개 기관을 아우르는 수평적 협의체로서 정책 추진 과정의 중복·단절을 줄이고 조율하는 역할을 한다. 한편, 의회도 실행계획의 입법적 뒷받침을 위해 관련 법안 검토와 예산 승인에 참여하게 된다.

<표 3-4> 미국 AI 실행계획 3대 전략별 주요 실행 과제

전략 축	실행 과제
전략축 1: AI 혁신 가속화	① 관료주의적이고 과도한 AI 규제 완화 ② AI 시스템 편향 제거 및 표현의 자유 보호 ③ 오픈소스 및 오픈웨이트 생태계 장려 ④ 산업 전반에 AI 도입 문화 조성 ⑤ 노동자의 AI 전환기 적응 지원 ⑥ 차세대 기술 투자로 제조업 선도 ⑦ AI 기반의 과학 혁신 촉진 ⑧ 세계 최대·최고의 AI 데이터셋 구축 ⑨ 프런티어 AI 분야 집중 투자 ⑩ AI 신뢰성 강화 기술 개발 투자 ⑪ AI 시스템 성능 측정 및 평가 체계 구축 ⑫ 정부 AI 도입 가속화 ⑬ 글로벌 군사 우위를 위한 국방 AI 촉진 ① 보안 위험으로부터 AI 혁신 보호 ⑭ AI 생성 미디어 활용에 따른 위협 대응
전략축 2 : 미국 AI 인프라 구축	① 관료주의적이고 과도한 AI 규제 완화 ② AI 시스템 편향 제거 및 표현의 자유 보호 ③ 오픈소스 및 오픈웨이트 생태계 장려 ④ 산업 전반에 AI 도입 문화 조성 ⑤ 노동자의 AI 전환기 적응 지원 ⑥ 차세대 기술 투자로 제조업 선도 ⑦ AI 기반의 과학 혁신 촉진 ⑧ 세계 최대·최고의 AI 데이터셋 구축
전략축 3: 글로벌 AI 외교 및 안보 주도	① 관료주의적이고 과도한 AI 규제 완화 ② AI 시스템 편향 제거 및 표현의 자유 보호 ③ 오픈소스 및 오픈웨이트 생태계 장려 ④ 산업 전반에 AI 도입 문화 조성 ⑤ 노동자의 AI 전환기 적응 지원 ⑥ 차세대 기술 투자로 제조업 선도 ⑦ AI 기반의 과학 혁신 촉진

출처 : 미국 AI 실행계획(2025.7)

3) AI 인프라 확충 정책: EO 14141[25]와 EO 14318[26]

(1) 배경 및 목적

프런티어 AI 시대를 맞아 대규모 데이터센터와 초고속 연산 인프라가 국가 AI 경쟁력의 기반으로 부상했다. 2023년 초거대 AI 모델들이 등장하면서 미국 내에서도 연산 자원 부족에 대한 우려와 투자가 확대되었다. 이에 바이든 행정부 말기에는 AI 인프라 강화를 위한 정부 차원의 조치가 논의되어, 2025년 1월 첨단 AI 인프라 구축을 위한 행정명령을 발령하기에 이르렀다. 반면, 2025년 초 정권 교체 후 트럼프 행정부는 기존 계획을 신속하고 민간 주도로 재편하고자 하였으며, 특히 환경규제 등 인허가 장벽 제거를 통해 AI 인프라 구축에 박차를 가하는 정책을 시행했다.

AI 인프라 확충 정책의 일차적 목표는 세계 최고 수준의 AI 계산자원과 산업기반을 확보하는 데 있다. 이를 통해 AI 연구개발과 서비스 구현에 필요한 토대를 강화하고, 나아가 국가안보 및 경제적 우위를 뒷받침하고자 한다. 2025년 1월 바이든 대통령이 서명한 EO 14141 「미국의 AI 인프라 리더십 강화」는 연방정부 주도로 AI 데이터센터를 확충함으로써 AI 연산능력 부족을 해소하고, 청정에너지와 결합된 지속가능 인프라를 구축하려는 목적이었다. 이후 7월 트럼프 대통령의 EO 14318 「데이터센터 인프라에 대한 연방 허가 절차 가속화」는 전 행정부의 계획을 대체하여 규제 완화와 신속 집행을 통해 민간과 협력한 초대형 AI 인프라 건설을 가속화하는 데 목적이 있다.

(2) 주요 내용 및 실행방안

EO 14141(미국의 AI 인프라 리더십 강화, 2025.1)은 연방 국유지를 활용한 대규모 AI 데이터센터 건설이 핵심이다. 구체적으로 국방부(DOD)와 에너지부(DOE)에 2025년 2월까지 각 최소 3개의 연방 부지를 선정하여 민간기업에 임대하고, 2027년 말까지 해당 부지에 프런티어 AI 데이터센터와 청정에너지 시설을 구축하도록 지시했다. 부지 선정 시 지역사회 건강과 복지에 미치는 영향을 고려하고, 고전압 송전망 접근성 및 청정에너지 자원 근접성을 우선 고려하도록 규정하여, 신규 데이터센터가 지역 전력 인프라와 조화롭게 들어서도록 했다. 아울러 모든 연방 허가기관에 인허가 절차 신속 처리를 지시하여 2025년 말까지 모든 건설허가를 발급하도록 하였고, 에너지부에는 2025년 6월까지 송전 인프라 업그레이드와 첨단 송전 기술 적용 등 전력망 확충 방안을 마련하도록 요구했다. 이러한 EO 14141은 연방

25) The White House, Advancing United States Leadership in Artificial Intelligence Infrastructure, 2025.1.14
26) The White House, Fact Sheet: President Donald J. Trump Accelerates Federal Permitting of Data Center Infrastructure, 2025.1.23

주도의 장기 인프라 투자 로드맵을 제시한 것이나, 실행 6개월 만인 2025년 7월 트럼프 행정부에 의해 철회되고 새로운 정책으로 대체되었다.

EO 14318(데이터센터 인프라 연방 허가절차 가속화, 2025.7)은 AI 인프라 구축방식의 전면 재편을 골자로 한다. 첫째로, 대규모 AI 데이터센터와 그에 필요한 반도체·네트워킹 장비 제조시설, 에너지 인프라 등을 모두 '적격 프로젝트(Qualifying Project)'로 지정하여 특별 지원 대상으로 분류했다. 이로써 AI 인프라 관련 핵심 설비에 대해 정부 차원의 우선 지원이 가능해졌다. 둘째, 환경 규제 완화 조치가 대거 도입되었다. 모든 연방기관에 「국가환경정책법(NEPA)」 상의 범주별 제외(categorical exclusion) 조항을 최대한 활용하거나 신설할 것을 지시했고, 정부 재정 지원이 프로젝트 비용의 절반 미만인 경우 환경영향평가 대상에서 제외하도록 규정했다. 또한 「멸종위기종법」, 「청정수질법」 등 환경 관련 허가요건을 완화하고 신속 처리하도록 지침을 제시하여, AI 인프라 건설에 따른 환경적 심사 절차를 크게 단축시켰다. 셋째, 신속 심사(FAST-41) 대상 지정과 브라운필드(brownfield) 활용 권장이다. 적격 프로젝트는 연방정부의 우선심사 대상으로 지정하여 절차를 앞당기고, 한편으로 이미 오염되어 활용도가 낮은 부지들을 데이터센터 용도로 전환할 수 있도록 가이드라인 개발을 지시했다. 넷째, 연방 자원 총동원을 통한 인프라 금융지원이다. 내무부와 에너지부에는 적합한 연방 보유 부지를 신속히 식별하여 민간 기업에 임대 승인하도록 하고, 상무부는 관련 부처들과 협의하여 대출·보조금·세제혜택·선구매 계약 등 포괄적인 금융지원 패키지를 마련하도록 규정했다. 또한 EO 14318은 바이든 행정부의 EO 14141을 즉시 철회함으로써, 이전 계획상의 절차를 모두 무효화하고 트럼프 행정부의 방침으로 일원화했다. 즉, 민관협력을 극대화하고 규제장벽을 제거함으로써 AI 인프라를 조속히 확충하려는 실행방안이었다.

(3) 관계 부처 및 기관

AI 인프라 정책은 범부처적 협력이 요구된다. EO 14141 하에서는 국방부와 에너지부가 부지 임대와 건설의 주관 부처였다. 또한 환경영향평가 및 인허가에는 환경보호청(EPA), 교통부, 내무부 등의 협조가 필요했다. EO 14318에서는 내무부와 에너지부가 연방 토지 활용 및 에너지 인프라 지원의 핵심 역할을 맡고, 상무부가 재정지원을 총괄 조율하는 주체로 명시되었다. 특히 상무부 내에는 해당 EO 이행을 위한 경제외교행동그룹(EDAG)까지 동원하여 금융 지원을 조율하도록 했다. 허가 절차 측면에서는 EPA와 「국립환경정책법(The National Environmental Policy Act, NEPA)」 소관기관들이 주요 플레이어이며, 빠른 심사를

위해 백악관 OMB 산하 연방 인프라 허가 개선 조정관실(FPIC) 등의 기존 인프라 프로젝트 허가 촉진 기구도 활용된다. 이밖에 재무부는 세제 혜택 실행, 국방부는 군사용 HPC 시설과 연계, 에너지부는 전력망 업그레이드 실행 등 각 부처가 세부 역할을 분담한다.

4) 정부조달과 'Woke AI' 방지 정책: EO 14319[27]

(1) 배경 및 목적

AI 시스템의 편향성 문제가 사회적 쟁점으로 떠오르면서, 그동안 미국 정부는 AI에서 인종·성별 등 불공정한 편견 제거를 주요 과제로 다루어왔다. 바이든 행정부도 AI 윤리 원칙에 공정성과 포용을 강조하고 연방기관에 편향 완화 조치를 요구해왔다. 한편 보수 진영에서는 거대언어모델(LLM) 등이 자유보수적 관점에 불리한 '편향'을 지닌다는 비판, 이른바 'Woke AI' 논쟁이 대두되었다. 'Woke'란 다양성·형평성·포용(DEI) 등 진보적 가치를 추구하는 경향을 비판적으로 지칭하는 표현으로, 일각에서는 ChatGPT 등 최신 AI가 정치적 올바름에 치우친 응답을 내놓는다고 지적했다. 트럼프 행정부는 이러한 우려를 정책에 반영하여, 정부가 활용하는 AI의 이념적 중립성을 보장하는 조치를 추진하였다.

본 정책의 목적은 연방정부가 조달·사용하는 AI 시스템에서 어떠한 이데올로기적 편향도 배제하는 것이다. 다시 말해 'Woke'(각종 진보 이념에 경도된) AI 모델이 정부 의사결정이나 서비스에 활용되지 않도록 하고, 객관적 진실에 기반한 중립적 AI만을 도입하겠다는 의도다. 이를 통해 정부 AI 활용에 대한 국민 신뢰를 높이고, 정치적 논쟁으로부터 자유로운 기술 활용 환경을 조성하려는 것이다. 또한 이러한 조치는 앞선 EO 14179의 이념 중립 AI 지향과 궤를 같이하며, 궁극적으로 AI에 대한 보수진영의 불신을 해소하고 AI 기술의 공정한 활용을 보장하려는 목적이 있다.

(2) 주요 내용 및 실행방안

2025년 7월 발령된 EO 14319 '연방정부 조달에서 'Woke AI' 방지' 행정명령은 연방기관이 획득하는 AI 특히 대규모 언어모델(LLM)에 적용될 일련의 지침을 담고 있다. 핵심 내용으로는 △이데올로기적 중립성 원칙 준수, △DEI 등 특정 개념 구현 금지, △OMB 지침 및 공급업체 준수 확인, △조달 계약 조건 강화 등을 골자로 하고 있다.

---

27) The White House, Preventing Woke AI in the Federal Government, 2025.7

<표 3-5> EO 14319 ‘연방정부 조달에서 ‘Woke AI’ 방지‘ 행정명령 핵심 내용

구분	주요 내용
이데올로기적 중립성 원칙 준수	• 모든 연방기관은 신규로 조달하거나 현재 사용 중인 LLM에 대해 ‘진실 추구(truth-seeking)’와 ‘이념적 중립(ideological neutrality)’ 두 가지 원칙 준수 • 이는 AI 모델이 어떤 정치적·사회적 편향 없이 사실에 기반한 응답을 제공하고, 특정 이념을 옹호하거나 반영하지 않아야 함을 의미 • 즉, 역사적 사실이나 통계 제공 시 객관성을 유지하고, 논쟁적 주제에 대해 한쪽 입장만을 대변하거나 금지하는 편향 필터가 없어야 함
DEI 등 특정 개념 구현 금지	• 다양성, 형평, 포용(DEI) 등의 개념을 AI 모델 내부에 프로그램으로 구현하는 것을 금지 • 모델이 답변 생성 시 인종·성별 균형 고려나 소수자 우대 같은 알고리즘적 조정을 하지 않도록 하라는 취지 • 즉, 모델 훈련이나 응답 규칙에 DEI 가치나 “정치적 올바름”을 강제로 반영하지 말라는 것
OMB 지침 및 공급업체 준수 확인	• 관리예산국(OMB)은 연방기관의 AI 조달 시 위 원칙을 구현하기 위한 상세 지침을 마련 • 이 지침에는 AI 제품을 공급하는 업체가 모델의 시스템 프롬프트, 기술 사양, 평가 결과 등을 정부에 공개하여 해당 AI가 진실성·중립성 원칙을 충족하는지 검증받는 절차가 포함 • 이를 통해 단순한 선언에 그치지 않고, 실증적 자료에 근거한 검증 메커니즘을 갖추려는 것
조달 계약 조건 강화	• 각 연방기관은 앞으로 체결하는 AI 제품/서비스 조달 계약에 진실성·중립성 원칙 준수 의무를 명시하고, 위반 시 공급업체가 계약 종료에 따른 비용을 부담하도록 규정 • 기존에 맺은 계약들도 가능한 범위 내에서 동일 조건을 반영하도록 수정 • 이를 통해 공급업체들이 원칙 위반 시 경제적 불이익을 명확히 인식하도록 하여, 계약 단계부터 모델 편향성을 제거하는 효과를 노린 것
예외 및 기타 사항	• 국가안보 시스템에 대해서는 이러한 요구사항을 예외적으로 적용 면제할 수 있도록 허용 • 첩보·사이버 등의 분야에서는 특수 목적상 특정 편향(예: 적대 세력 탐지에 편향된 모델 등)이 필요할 수 있음을 감안 • 그 외, 각 기관은 OMB 지침이 발표된 후 자체적인 준수 절차(compliance procedure)를 수립하여 내부 통제를 갖추도록 요구

출처 : EO 14319 ‘연방정부 조달에서 ‘Woke AI’ 방지‘ 행정명령(2025.7)

(3) 관계 부처 및 기관

본 정책의 주무 부처는 관리예산국(OMB)이다. OMB는 행정명령에 따라 연방 차원의 AI 조달 가이드라인(M-25-22 등)을 작성·배포하여 각 기관이 따를 세부 기준을 정하게 된다. 실제 조달 업무는 각 부처별 수요에 따라 진행되므로 각 연방기관의 최고정보책임자(CIO)와 조달 담당 부서가 현장에서 정책을 구현한다. 기관들은 AI 제품 구매 시 OMB 지침에 따른 계약조항 삽입, 업체 제출자료 검토, 위반 시 제재조치 등을 수행할 것이다. 공급업체 측면에서는, 연방정부에 AI 솔루션을 공급하는 기업들이 (MS, OpenAI, IBM 등) 이번 지침의 직접적인 규율 대상이 된다. 이들 기업은 모델 훈련 및 튜닝 시 정부 기준을 충족하도록 컨설팅을 받아야 하고, 필요 시 모델 아키텍처나 콘텐츠 필터링을 조정해야 할 수 있다. 또한 백악관 OSTP와 행정관리감사원(GAO) 등은 정책 이행 상황을 모니터링하고 효과성을 평가하는 역할을 맡을 수 있다. 전반적으로, 본 행정명령의 이행은 정부 조달 프로세스 내에서 이루어지므로 연방조달규정(FAR) 개정이나 공무원 대상 교육도 수반될 것이다.

5) AI 수출 정책: EO 14320[28] 및 수출통제 전략

(1) 배경 및 목적

AI 기술의 지정학적 중요성이 커지면서, 미국은 자국의 핵심기술이 전략적 경쟁국으로 유출되는 것을 통제하는 한편, 동맹국에는 기술 공급을 확대하는 이중 전략을 구사해왔다. 특히 2022년 10월 상무부 산업안보국(BIS)이 발표한 대중국 첨단 반도체 수출통제 규정은 미국산 AI 칩과 장비의 중국 수출을 사실상 금지하며 큰 파장을 일으켰다. 이어 2023년 8월 바이든 대통령은 미국인의 중국 첨단기술 분야 투자를 제한하는 행정명령을 발동했고, 2024년 12월에는 고대역폭 메모리(HBM)와 첨단 장비에 대한 추가 수출제한 규정을 내놓는 등 대중국 기술압박 수위를 높여갔다. 이러한 맥락에서 트럼프 행정부는 2025년 출범 후 AI 기술패권 전략을 한층 강화하였다. 즉, 미국 AI 기술 스택의 수출 촉진과 첨단 AI 컴퓨팅 자원의 대적국 유입 차단을 동시에 추진하는 능동적 수출전략을 수립한 것이다.

트럼프 행정부의 AI 수출 정책은 크게 두 갈래 목적을 지닌다. 첫째, 미국 AI 기술의 세계 확산을 통해 동맹을 강화하고 글로벌 기술질서에서 주도권을 유지하는 것이다. 이는 미국산 AI 솔루션이 국제 표준이 되도록 지원함으로써 우방국들의 기술 의존을 미국으로 묶어두고, 중국 등 경쟁국의 영향력을 약화시키려는 목적이다. 둘째, 첨단 AI 기술이 적대국의 손에

28) The White House, Promoting the Export of the American AI Technology Stack, 2025.7

들어가는 것을 막아 미국의 안보 리스크를 차단하는 것이다. 즉 수출통제를 통해 중국 등의 AI 역량 강화를 억제함과 동시에, 미국 산업의 핵심기술 우위를 지속하려는 전략적 목표가 있다. 즉, 기술 전파는 더 넓게, 통제는 더 촘촘하게 함으로써 미국 중심의 AI 공급망을 구축하는 것이 본 정책의 지향점이다.

(2) 주요 내용 및 실행방안

EO 14320(미국 AI 기술 스택의 수출 촉진, 2025.7)은 미국의 AI 관련 풀스택(full-stack) 기술 패키지를 우방에 적극 수출하기 위한 프로그램 신설을 명령했다. 여기서 풀스택 패키지란 AI 시스템 구현에 필요한 하드웨어, 소프트웨어, 모델, 응용, 데이터, 표준 등 전반을 포함하는 통합 기술 묶음이다. 상무부 장관 주도로 경제외교행동그룹(EDAG)이 재정지원을 조율하며, 선정된 우선 대상국에 대해 연방 정부의 금융지원 수단(차관, 수출신용, 보조금 등)을 총동원하여 패키지 수출을 뒷받침하도록 했다. 이를 통해 동맹국들이 중국산이 아닌 미국산 AI 생태계를 채택하도록 유인하고, 미국 기술·표준의 글로벌 채택을 촉진하려는 것이다. 예컨대, 개발도상국이나 안보 파트너국이 AI 인프라를 구축할 때 미국 기업이 원스톱 솔루션을 제공하고, 정부가 금융·외교적으로 지원하여 중국 영향력을 차단하는 시나리오다. EO 14320은 또한 미국 내 관련 기업에게 수출 기회를 확대하여 민간의 해외시장 개척을 독려함으로써, 산업경쟁력과 외교를 연계하려는 산업정책적 성격도 지닌다.

한편, 수출 촉진과 동시에 첨단 AI 기술에 대한 수출통제도 한층 강화되었다. 바이든 정부 시기의 대중 수출통제 조치들에 이어, 트럼프 행정부 BIS는 기존 규정의 영향을 재평가하고 미비점을 보완할 방침을 밝혔다. 예를 들어, 2024년 말까지의 통제조치에 따르면 일본·네덜란드 등 일부 우방도 특정 첨단 장비에서는 미국과 동일 제한을 받게 되어 불만이 있었는데, 트럼프 행정부는 이를 개선하여 우방에 대한 불필요한 규제 부담을 낮추고, 중국 등에 대한 통제는 더 효과적으로 할 새로운 규칙을 모색하고 있다. 또한 실행계획에서 밝힌 바와 같이, 첨단 AI 반도체와 컴퓨팅 자원의 수출통제 집행을 강화하기 위해 위치 검증 기술 활용, 공급망 모니터링 등의 수단을 검토하고, 반도체 제조 하위 시스템까지 통제대상을 확대하는 방안을 정책권고로 제시했다. 이는 중국 등 경쟁국이 우회로로 첨단 기술을 확보하지 못하도록 통제의 그물망을 세분화하는 조치다. 아울러, 외국산직접제품규칙(FDPR)의 적극 활용과 보복관세 등 우방국의 대중 수출통제 동참을 유도할 수단도 거론되었다. 실제로 동맹국들과는 수출통제에 관한 협의체계를 가동하고, 방위산업 공급망까지 연계하여 전

략적 디커플링(decoupling)을 추진할 계획이다. 한편, 이러한 강경한 통제 기조는 경제적 손실과 기술 분업 구조 훼손 우려도 수반하므로, 트럼프 행정부는 공세와 설득을 병행하는 모습이다. 즉, 한편으로 동맹에 기술 지원과 시장접근 기회를 제공하면서(앞서 EO 14320), 다른 한편으로 통제 불이행 시 불이익을 주는 체계를 만들어 친미 기술블록 결성을 꾀하고 있다.

또한, 내부적으로는 수출통제의 지속 가능성을 높이기 위해 관련 거버넌스를 정비하고 있다. 상무부 BIS를 중심으로 국방・국무부 등과 협력하여 차세대 수출통제 규정을 마련하고, 기업들의 부담을 줄이면서도 효과적인 통제가 되도록 규제 효율화를 추진한다. 또한 국제공조 측면에서는 Wassenaar Arrangement(바세나르 체제) 등 다자 수출통제 regimes에서 AI 관련 품목 통제를 강화하도록 주도하고, 신흥기술에 대한 다자 통제룰 수립에도 미국이 앞장서고 있다. 이와 함께 신뢰할 수 있는 AI 글로벌 기준을 미국이 선도 개발하여, 추후 수출통제의 국제 표준으로 활용하려는 움직임도 있다. 예를 들면, NIST가 개발한 AI 위험관리 프레임워크(AI RMF) 등을 각국에 확산시켜 책임있는 AI 활용 규범을 퍼뜨리는 한편, 이를 근거로 민감 AI 기술에 대한 국제공동 통제를 설득하는 전략이다.

(3) 관계 부처 및 기관

AI 수출통제 및 촉진 정책은 범정부적 역량을 동원한다. 상무부 산업안보국(Bureau of Industry and Security, BIS)은 수출통제 규정 제정・집행의 핵심기관으로, 반도체 장비 수출허가 심사, 기업들과의 소통 창구 역할을 담당한다. 국무부는 동맹국과의 외교협상 및 다자협의체 활동을 통해 우방의 공조를 이끌어낸다. 예를 들어 일본, 유럽 등과 수출통제 협의를 진행하고, 중국을 견제할 기술동맹(예: Chip4, TTC 등) 강화를 위해 노력하고 있다. 국방부와 에너지부는 전략 기술 목록 정의와 안보평가를 지원하고, 국가안보회의(NSC)는 전체 전략 조율 및 정보 공유를 총괄한다. EO 14320에서 언급된 경제외교행동그룹(EDAG)은 국무・상무・재무 등 관계자들로 구성된 태스크포스로, 수출 촉진 패키지의 대상 국가 선정 및 재원 배분을 조율한다. 재무부도 대외투자 규제와 보복관세 등의 경제제재 수단을 통해 참여하며, 국가정보국(DNI)이나 CIA는 기술 유출 경로와 우회 수법 식별에 정보 자원을 제공한다. 또한 대통령 직속 경제안보조정관 등을 통해 민관 협력을 도모하고, 주요 첨단 기업들과 정기적으로 소통하여 산업계 애로를 청취한다. 수출 촉진 측면에서는 미국 수출입은행(EXIM), 국제개발금융공사(DFC) 등이 동원되어 금융지원을 맡는다.

6) AI 교육 및 인력 정책: AI 교육 행정명령, EducateAI 이니셔티브[29)]

(1) 배경 및 목적

AI 초강대국으로서의 지위를 유지하려면 국내 AI 인재 풀(pool)의 확대가 필수적이다. 그러나 미국은 한편으로 AI 인력 수요 급증을 겪으면서, 다른 한편으로 공공·민간 모두 전문 인재 부족 현상이 심화되어 왔다. 특히 최첨단 AI 연구개발을 이끌 박사급 딥테크 인재의 공급이 수요를 따라가지 못해, 인재 쟁탈전이 벌어지고 해외 인력 의존도가 높다는 지적이 있었다. 이에 따라 연방 정부는 AI 인력 양성을 국가 전략의 중요한 축으로 삼고 다각도의 정책을 펼쳤다. 2020년 「국가 AI 이니셔티브법」에서도 AI 인재 양성이 핵심 과제로 포함되었고, 2022년 CHIPS 법과 함께 NSF 예산이 크게 늘어나 STEM 교육·훈련 프로그램이 확충되었다. 바이든 행정부는 2023년 10월 행정명령(EO 14110)을 통해 OPM(인사관리처)에 연방 정부 AI 인력 '대규모 충원(talent surge)'을 지시하기도 했다. 이러한 노력의 연장선에서, 2024년 말부터 EducateAI 이니셔티브 등의 신규 프로그램이 가동되면서 AI 인재 기반 확대에 박차를 가하게 되었다.

AI 교육 및 인력 정책의 목적은 전 국민의 AI 역량 저변을 확대하고, 나아가 세계 최고 수준의 AI 전문인력을 양성하는 데 있다. 구체적으로는 K-12 초중등 교육에서 AI 소양 함양을 지원하고, 대학·대학원 수준에서 AI 연구자를 육성하며, 산업계와 공공부문에 숙련된 AI 인력 공급을 원활히 하는 것이 목표다. 이를 통해 AI 시대의 노동력 재편에 대응하고 일자리 충격을 완화하는 한편, AI 기술혁신을 주도할 인재군을 확보하여 미국의 기술경쟁력을 지속하고자 한다. 또한 공무원 등 정부 인력의 AI 활용 능력을 높여 행정효율을 증진하고, AI 정책을 이해하는 디지털 문해력(literacy)을 전사회적으로 제고하려는 목적도 있다. 한 마디로, AI 시대에 사람에 대한 투자를 강화하여 기술과 인재의 선순환을 이루려는 것이다.

(2) 주요 내용 및 실행방안

미국의 AI 인재 전략은 교육 파이프라인 구축, 재직자 훈련, 정부 인력 강화 등의 세 가지 측면에서 추진되고 있는 것으로 관찰된다.

29) NSF, Advancing education for the future AI workforce (EducateAI), 2023.12.1

① AI 교육 이니셔티브(EducateAI)

2024년 NSF와 교육부가 공동 착수한 ‘EducateAI’ 이니셔티브는 AI 시대에 걸맞은 교육 혁신 및 인재 양성을 위한 대표 프로그램이다. 이 프로그램은 유치원(K)-12부터 대학원, 직업훈련에 이르기까지 전 주기에 걸친 AI 교육을 지원한다. 핵심은 교육자와 학습자를 위한 AI 자원 접근성 향상인데, 구체적으로 교사, 학생, 지역대학 등에 AI 학습 리소스(콘텐츠, 데이터셋, 코드, 컴퓨팅 자원)를 직접 제공하여 디지털 격차를 줄이고 AI 학습환경을 개선한다. 이를 위해 NSF는 교육기관 대상 보조금 프로그램을 운영하고, 민간 기술기업들과 파트너십을 맺어 클라우드 컴퓨팅이나 AI 소프트웨어를 학교에 무상 또는 저가 제공한다. 또한 AI 커리큘럼 개발을 지원하여 초중등 단계부터 데이터 과학, 알고리즘 사고, 기초 ML 실습 등을 접할 수 있도록 장려한다. EducateAI는 연간 공모를 통해 혁신적 AI 교육 프로젝트에 자금을 배정하며, 그 우수 사례를 전국에 확산시키는 역할도 한다. NSF가 2024년 말 해당 이니셔티브에 자금을 최초 배정하였으며, 2025년 현재까지 수백 개 학교와 교육기관이 참여하고 있다. 궁극적으로 이 프로그램은 AI 문해력을 사회 전반으로 넓히고 다음 세대 AI 인재 풀을 키우는 토대가 되고 있다.

② 산업 및 지역 인력 재교육(reskilling)

최첨단 AI 분야의 석박사 인재 양성과 별개로, 산업현장의 기존 인력에 대한 재교육도 중점 추진되고 있다. NSF는 CHIPS법을 통해 대폭 증액된 STEM 교육 예산을 활용하여 지역사회 기술인력 재교육 프로그램을 운영 중이다. 예를 들어 제조업 노동자나 중소기업 종사자들이 AI 및 자동화 기술을 습득하도록 커뮤니티 칼리지 기반 단기훈련을 제공하고, AI 활용 직무능력 인증 프로그램 등을 마련했다. 또한 산·학·관 협력으로 구성된 Manufacturing USA[30] 등 18개 제조업 혁신 연구소 네트워크를 통해, 각 지역 산업 수요에 맞는 맞춤형 AI 기술훈련을 지원하고 있다.

이러한 민간 인력 훈련 프로그램은 고숙련 기술인력의 저변 확대와 AI로 인한 일자리 전환 완충 효과를 노린 것이다. 한편, 고급 R&D 인력 양성 면에서 NSF는 장학금·펠로우십 확대를 통해 대학원생과 박사후 연구원들이 AI 분야에 진입하도록 장려하고 있다. 이를 통해 AI 핵심인재 파이프라인을 강화함과 동시에 다양성 확대(저소득층·소수인종 참여 촉진)에도 신경쓰고 있다.

30) Manufacturing USA (2025.11.17. 방문) https://www.manufacturingusa.com

③ 연방정부 인재 확보 및 교육

연방 차원에서는 공공부문 AI 인력 역량 강화에 힘쓰고 있다. 우선 2022년 10월 제정된 「AI 연수법(AI Training Act)」은 연방 조달 담당자를 포함한 공무원들을 대상으로 체계적인 AI 교육 프로그램을 제공하도록 했다. 이에 따라 OMB와 조달청이 주관하는 온라인·대면 교육이 시행되어, 2024년 한 해에만 200여 개 기관, 1만4천명 이상의 공무원이 AI 연수에 참여하였다. 교육 내용은 AI의 과학적 기초, 위험과 완화방안, 윤리 및 국가안보 이슈 등을 망라하여 공직자가 AI의 기회와 위험을 균형있게 이해하도록 설계되었다.

또한 바이든 행정부 말기에 OPM이 추진한 AI 인재 대거채용(talent surge) 계획에 따라 각 부처는 AI 관련 직위를 신설·확대하고, 트럼프 행정부도 이를 계승하여 특별고용 인센티브와 기술인재 특별채용 경로를 운영하고 있다. 아울러, OMB는 2025년 4월 연방기관의 AI 활용을 가속화하기 위한 지침(M-25-21)[31]을 발령하여 공공부문의 AI 인재 활용과 기관 간 인재교류(talent exchange) 프로그램 수립을 독려했다. 이러한 조치들로 연방 정부 내부의 AI 역량이 높아지고 있으며, 각 기관에서 최고 AI 책임자(Chief AI Officer) 임명이 확대되어 거버넌스 체계도 강화되고 있다.

(3) 관계 부처 및 기관

국립과학재단(NSF)은 AI 인재양성 정책의 핵심 수행기관이다. EducateAI와 각종 교육 보조금 프로그램을 총괄하며, CHIPS법에 따른 예산으로 교육·훈련 프로젝트를 지원한다. 교육부(DoEd)도 K-12 AI 교육과 교사훈련 방안을 마련하고 있으며, EducateAI 등에서 NSF와 공조한다. 인사관리처(OPM)는 연방 공무원 대상 AI 교육과 채용 정책을 담당하고, OMB는 정부 전체의 인재 전략을 예산과 규정 측면에서 뒷받침한다. 상무부 산하 NIST도 간접적으로 기여하는데, 예를 들어 AI 관련 인증 프로그램이나 평생학습 표준을 개발해 기업과 교육기관이 활용하도록 지원한다. 지역 현장에서는 주 정부 교육위원회 및 노동청 등이 연방 지원금을 받아 지역 특화 인력 양성 사업을 전개하고, 민간에서는 대형 IT 기업들이 자사 AI 온라인 과정을 무료로 공개하거나, AI 챌린지 대회 등을 개최해 인재 발굴에 참여하고 있다. 또한 국방부 산하 DARPA는 AI 연구장학생 프로그램으로 군사분야 AI 인재를 육성하고, 국토안보부는 사이버-AI 인력 이니셔티브 등 특수 분야 인재 프로그램을 운영하는 등 각 부처별로 특화된 노력이 진행 중이다.

---

31) OMB, Accelerating Federal Use of AI through Innovation, Governance, and Public Trust, 2025.4.3. https://www.whitehouse.gov/wp-content/uploads/2025/02/M-25-21-Accelerating-Federal-Use-of-AI-through-Innovation-Governance-and-Public-Trust.pdf

7) AI 안전·표준 거버넌스

(1) 배경 및 목적

AI 기술이 고도화됨에 따라 안전하고 신뢰할 수 있는 AI에 대한 요구가 세계적으로 대두되었다. 2023년 11월 영국이 주최한 AI 안전 서밋(AI Safety Summit)[32]을 계기로, 각국은 AI 위험을 완화하기 위한 국제 공조방안을 모색하기 시작했다. 미국은 전통적으로 민간 주도 자율규범을 중시해왔으나, 2023년 바이든 행정부는 대형 AI 기업들로부터 자발적 안전 확보 서약을 받아내고 NIST를 통해 'AI 위험관리 프레임워크(AI RMF)'를 발표하는 등 일정 수준의 개입을 시도했다. 또한 바이든 행정명령(EO 14110)은 특정 AI 개발기업에 대해 「국방물자생산법(DPA)」을 근거로 모델 안전성 테스트 결과 제출 의무까지 부과하며 강제 조치를 도입했다. 이러한 움직임 속에 상무부 산하 NIST에는 AI 안전연구소(AI Safety Institute)가 설립되어 산·학·관 협력 허브로 기능하기 시작했다. 2024년 2월에는 OpenAI, 구글 등 290여 개 조직이 참여하는 'AI 안전연구소 컨소시엄(AISIC)'이 발족되어, AI 능력 평가, 레드팀 운영, 합성콘텐츠 위험 완화 가이드라인 개발 등 업계 전반의 안전성 제고 노력을 주도했다.

AI 안전·표준 거버넌스의 목적은 AI 기술 발전과 리스크 관리의 균형을 찾는 데 있다. 혁신을 저해하지 않으면서도 대형 AI 모델의 오남용, 사고 위험 등을 선제적으로 다루기 위한 민관 협력 체계를 구축하는 것이다. 바이든 행정부 시기에는 여기에 규범 정립을 더해, AI 윤리 원칙과 신뢰성 확보를 위한 가이드라인 제정을 목표로 삼았다. 트럼프 행정부로 넘어와서는 안전보다는 표준과 혁신에 초점을 맞추도록 조정되었지만, 그 기본 취지는 AI 기술에 대한 사회적 신뢰 확보라고 할 수 있다. 즉, AI로 인한 잠재 위험을 최소화하고 국제표준 경쟁에서 미국이 주도권을 쥐기 위한 거버넌스 체계를 마련하는 것이 핵심 목표다. 이를 위해 민간의 창의를 활용하면서도 필요한 최소한의 공통 규칙(rule)을 정립하여 AI 생태계의 책임성을 확보하려는 것이다.

(2) 주요 내용 및 실행방안

① AI 안전연구소 컨소시엄(AISIC)[33]

2024년 출범한 AISIC(AI Safety Institute Consortium)는 산·학·관이 참여하는 다자 협의체로서, 신뢰성 있는 AI 개발·배포를 위한 모범사례를 공유하고 기술적 협력을 도모한다. 2025년 11월

32) GOV.UK, AI Safety Summit 2023 (2025.11.17. 방문)
https://www.gov.uk/government/topical-events/ai-safety-summit-2023
33) NIST, AI Safety Institute Consortium (2025.11.17. 방문)
https://www.nist.gov/artificial-intelligence/artificial-intelligence-safety-institute-consortium-aisic

기준으로 280여 개 조직(기업, 대학, 시민단체 등)이 참여하였으며, 주요 작업으로 AI 모델의 성능과 안전성 평가, 위험관리 프레임워크 적용, 레드팀(red-team) 테스트, 합성 미디어 탐지 기술 개발 등이 진행되고 있다. 이 컨소시엄은 NIST 산하 AI 안전연구소가 주도하여 운영되며, 정기 워크숍과 작업반을 통해 기술 표준과 지침을 논의한다. 이러한 협의체를 통해 업계는 정부 규제에 선제 대응할 자율규범을 만들고, 정부는 민간 전문성을 활용해 효과적인 정책을 설계하는 상호보완적 관계를 구축했다.

② CAISI 설립(표준혁신센터로 개편)[34]

2025년 1월 트럼프 행정부 출범 직후, 상무부는 NIST 산하 AI 안전연구소를 AI 표준혁신센터(Center for AI Standards and Innovation, CAISI)로 재편했다. 이는 AI 거버넌스의 정책 우선순위를 안전에서 표준과 혁신으로 전환한 조치였다. CAISI는 기존 안전연구소의 인적·물적 자산을 그대로 이어받아, AI 기술 표준 개발과 혁신 지원에 무게를 두고 운영되고 있다. 예를 들어, 2025년 1월 개최된 CAISI 워크숍에서는 대규모 언어모델 평가방법 표준화와 AI 시스템 투명성 지표 개발 등이 논의되었고, 그 결과를 분석·공개하는 등 활발히 활동 중이다. CAISI는 연방 예산으로 설립 이후 초기 2년 간 최대 2억 7,500만 달러의 운영자금이 필요하다고 예측되고 있으며[35], NIST 내 정보기술실 산하 조직으로서 상무부 장관의 관할 하에 있다. 향후 CAISI는 미국의 AI 표준화 전략을 총괄하며, 국제표준 제정 과정에서 미국 입장 대변과 기술 선도를 꾀할 예정이다.

③ NIST의 AI 리스크 관리 프레임워크(AI RMF)[36]

AI 안전 거버넌스의 한 축으로 표준화된 가이드라인 제공이 강조된다. 2023년 1월 NIST가 발표한 'AI RMF 1.0'은 이러한 노력의 결실로, AI 시스템 개발·운영시 고려해야 할 위험관리 체계를 제시하고 있다. 이 프레임워크는 프라이버시, 투명성, 안전, 공정성 등의 원칙에 따라 위험 식별, 완화, 모니터링 절차를 가이드하며, 기업과 기관들이 자발적으로 활용하도록 권고된다. 바이든 행정부는 연방기관과 연방지원 연구과제 등에 AI RMF 적용을 장려했고, 트럼프 행정부도 이를 민간 주도의 규범으로 존중하면서 강제규제 대신 권고표준으로서 기능하도록 하고 있다. AI RMF는 미국이 주도하여 개발했다는 점에서 국제적인 영향력이 기대되며, 현재 ISO/IEC 등에서도 표준 채택을 논의 중이다. 이는 미국이 법적 규제 없이도 표준을 통해 글로벌 AI 거버넌스에 영향력을 행사하는 통로로 활용되고 있다.

34) NIST, CAISI (2025.11.17. 방문) https://www.nist.gov/caisi
35) FAS, A National Center for Advanced AI Reliability and Security, 2025.6.11. (2025.11.17. 방문) https://fas.org/publication/a-national-center-for-advanced-ai-reliability-and-security/
36) NIST, AI Risk Management Framework (2025.11.17. 방문) https://www.nist.gov/itl/ai-risk-management-framework

④ 국제협력 및 다층 거버넌스

미국의 AI 안전·표준 거버넌스는 국내 자율규범 강화와 더불어 국제 협력을 통해 시너지를 내고 있다. 미·EU 간 무역기술협의회(TTC)에서 AI 투명성 원칙을 공동선언하거나, G7 히로시마 AI 프로세스에 참여하여 생성형 AI 지침 마련에 기여하는 등이 그 예다. 다만 트럼프 행정부는 다자주의 선언에 소극적일 수 있다는 전망도 있는데, 그럼에도 간접적 방법으로 동맹 규합을 도모하는 중이다. 예컨대 영국의 AI 안전 서밋에는 미국도 지속 참가하며 글로벌 AI 안전 협력을 이끌고 있고, Five Eyes 등의 정보동맹과 AI 위협 정보 공유 체계를 구축하고 있다. 국내적으로는 연방 차원의 가이드라인과 더불어 주(州) 정부 및 의회와의 조율도 과제다. 현재 일부 주는 자체적으로 AI 윤리위원회 등을 두고 안면인식 규제 등 안전조치를 시행 중인데, 연방정부는 최소공통의 기준을 제시함으로써 주별 과도한 규제 편차를 줄이고자 한다.

(3) 관계 부처 및 기관

상무부는 NIST를 통해 AI 안전·표준 거버넌스의 주관 부처 역할을 한다. NIST-CAISI가 표준과 기술적 가이드 개발을 담당하며, OSTP는 국가과학기술위원회(NSTC) 산하 AI 특별위원회 등을 통해 정부 내 전략 조율을 맡는다. 또한 백악관 국가 AI 이니셔티브 사무국(NAIIO)이 민간 및 주정부와의 연계를 지원하고, 의회가 설치한 NAI 자문위원회(NAIAC)가 거버넌스에 대한 권고를 제시한다. 국토안보부(DHS)와 법무부(DOJ)는 AI 안전 이슈 중 사이버보안, 알고리즘 차별 등과 관련하여 집행 역할을 가진다(예: FTC의 AI 오용 조사 권한 등).

앞서 언급된 AISIC 컨소시엄에는 주요 빅테크 기업(FAANG 등)과 스타트업, 시민단체, 학계 전문가들이 폭넓게 참여하여 민간 주도의 거버넌스를 구현하고 있다. 이들은 실제 위험사례 정보를 공유하고, 다양한 이해관계자의 의견을 수렴하여 권고안을 도출하는 데 기여한다. 국제적으로는 OECD AI 원칙그룹, ISO/IEC JTC1 SC42(인공지능 분과), UN AI 윤리위원회 등에도 미국 측 대표가 파견되어 글로벌 거버넌스에 목소리를 내고 있다.

8) 입법 및 사회적 위험 대응: 「Take It Down Act」 등

(1) 배경 및 목적

AI 기술의 발전은 새로운 사회적 위험을 초래하고 있으며, 이에 대응하기 위한 입법 움직임이 활발해지고 있다. 특히 딥페이크(deepfake)로 대표되는 허위·조작 콘텐츠는 개인의 사생활 침해와 선거 개입,

명예훼손 등의 위험을 낳아 심각한 사회 문제로 대두되었다. 2010년대 후반부터 몇몇 주(州)에서 딥페이크 포르노 금지법을 제정했지만, 연방 차원의 포괄법은 부재했다. 2023년 말 잇따른 유명인 합성 음란물 사건, 2024년 선거를 앞둔 딥페이크 우려 등이 겹치면서 초당적 대응 필요성에 공감대가 형성되었다. 이러한 배경에서 2025년 5월, 드디어 연방 의회는 비(非)동의 누드 이미지 및 딥페이크 확산 방지를 위한 「Take It Down Act」을 만장일치에 가까운 지지로 통과시켰다[37]. 한편, AI의 사회·경제적 영향 확산으로 산업 육성 vs 위험 통제 간 균형을 둘러싼 입법 논쟁도 진행되었다. 트럼프 행정부는 친기업 성향으로 연방 차원의 과도한 규제 유예를 모색했으나, 주 정부들은 개별적 법안을 쏟아내면서 규제 분권화가 나타났다. 2025년 7월에는 트럼프 행정부의 여러 핵심 정책을 담은 거대 조정법안인 「One Big Beautiful Bill Act」(일명 '하나의 크고 아름다운 법')[38]이 제정되었는데, 여기에 AI 및 첨단기술 분야에 대한 안보적 통제 조항이 포함되며 이슈가 됐다.

「Take It Down Act」의 목적은 AI 악용으로 인한 개인 피해를 방지하고 가해자에 대한 처벌근거를 마련하는 것이다. 특히 동의 없이 유포된 사적 이미지로 인한 성적 착취 피해를 막고, AI 합성포르노 같은 신종 범죄에 대응하는 데 초점을 맞춘다. 이를 통해 기술발전으로 인한 사회적 피해자의 보호와 디지털 플랫폼 책임을 강화하고자 한다. 「One Big Beautiful Bill Act(OBBBA)」의 관련 조항 목적은 두 가지로 요약된다. 하나는 국내 첨단 제조업 지원을 통한 AI 등 전략산업 육성, 다른 하나는 중국 등 적성국의 미국 기술 생태계 침투 차단이다. 이를 위해 세제 혜택과 예산지원을 묶어 자국 산업 보호와 안보 통제를 동시 달성하려는 의도를 가진다. 또한 OBBBA 입법 논의 과정에서 제기되었던 주(州) 차원의 AI 규제 유예 조항은 연방정부가 AI 규제의 통일성을 확보하려는 목적에서 나왔지만, 결국 제외되었다. 이는 연방 차원에서 과도한 규제 경쟁을 막고 혁신을 장려하려는 의도가 있었음을 시사한다. 요약하자면, 최근 입법은 AI로 인한 새로운 위험을 신속히 제재하는 한편 AI 산업 생태계를 건강하게 육성하려는 투트랙 목적을 추구하고 있다.

(2) 주요 내용 및 실행방안

① 「Take It Down Act」(2025년 5월 제정)[39]

이 법은 공식명칭이 다소 복잡하지만(웹사이트와 네트워크상 기술적 딥페이크 차단을 통한 확인된 착취 행위 대응법[40]), 핵심은 비동의 음란물에 대한 형사처벌 및 플랫폼 책임 규정이다. 제1부는 형법 개

37) NACo, TAKE IT DOWN Act signed into law, 2025.5.20
38) U.S. House of Representatives, H.R. 1, 2025.5.21. https://www.congress.gov/bill/119th-congress/house-bill/1/text
39) U.S. Senate, S. 146, 2025.5.19. https://www.congress.gov/bill/119th-congress/senate-bill/146
40) Tools to Address Known Exploitation by Immobilizing Technological Deefakes on Websites and Networks Act

정을 통해 비동의 사적 영상물(누드 이미지 등)의 고의적 유포를 연방법상 범죄로 규정하였다. 성인 피해자의 경우 유포행위 시 최대 징역 2년 또는 벌금, 미성년 피해자일 경우 최대 징역 3년까지 부과할 수 있다. 여기에는 AI나 기계학습으로 생성된 디지털 위조물(digital forgery), 즉 딥페이크 영상도 명시적으로 포함되어 처벌 대상이 되었다. 또한 이러한 영상물로 피해자를 협박하는 행위도 별도로 범죄화하여, 성인의 경우 최대 18개월, 미성년의 경우 최대 30개월의 징역형을 규정했다. 제2부는 온라인 플랫폼의 책임 의무를 다룬다. 특정 플랫폼(사이트, 소셜미디어 등)은 피해자가 신고한 비동의 영상물을 48시간 이내 삭제해야 하는 법적 의무를 지닌다. 신고는 피해자 서명 진술서, 해당 콘텐츠의 식별정보(URL 등), 비동의 게시에 대한 선의의 확신 진술, 연락처 등을 포함해야 하며 절차상의 허들을 두어 악의적 신고를 방지한다. 플랫폼이 선의로 규정에 따라 조치한 경우에는 민형사상 책임을 면제하는 세이프하버(safe harbor) 조항도 포함되었다. 본 법의 집행 기관으로는 연방거래위원회(FTC)가 지정되어, 플랫폼이 삭제 의무를 이행하지 않을 시 불공정 거래행위로 제재할 수 있게 되었다. 동 법은 가해자에 대한 형벌과 함께 플랫폼에 대한 신속삭제 의무를 결합함으로써 딥페이크 확산에 법적 억지력을 부여하고 있다.

② 「One Big Beautiful Bill Act」(OBBBA, 2025년 7월 제정)[41] 중 AI 관련 조항

OBBBA는 세제·예산 등을 포괄한 방대한 법안으로, 그 중 AI 및 첨단기술 안보와 관련된 부분이 눈에 띈다. 먼저 이 법은 '금지 대상 외국 법인(Prohibited Foreign Entity; PFE)' 개념을 도입하여, 특정 적대국(사실상 중국)과 연계된 기업을 목록화하고 있다. 그리고 PFE와 사업적으로 밀접한 관련이 있는 기업에 대해서는 미국 연방정부가 제공하는 세액공제 대상에서 제외하도록 규정하였다. 예컨대 AI 반도체 등 첨단 제조 분야에서 정부 지원을 받는 기업이, PFE로 지정된 중국 업체와 합작투자나 기술제휴를 할 경우 세제 혜택을 박탈당한다는 의미다. 이에 따라 연방 지원을 원하는 기업은 자사가 PFE와 무관함을 입증해야 하며, 거짓이나 위반이 적발되면 지급받은 지원을 전액 환수당하고 제재를 받을 수 있다.

해당 조치는 세법(IRS 코드) 개정을 통해 이루어졌으며, 주로 첨단 제조 투자 세액공제(IRC §45X) 등에 적용되도록 설계되었다. 한편 OBBBA 입법 초기에는 "주 정부의 AI 규제를 10년간 유예한다"는 파격적 조항이 포함되어 논란이 되었으나, 의회 심의 과정에서 삭제되었다. 결국 최종 법안에는 주 정부 관련 규정은 빠지고, 대신 연방 차원에서 안보 위험이 있는 외국 기술 영향 차단 조치 위주로 담긴 것이다. OBBBA는 세법 개정 외에도, 연방 예산을 통한 AI 등 첨단산업 지원 근거를 마련했다. 예를 들어 특정 PFE와 관련 없는 기업에 대해서는 추가 세액공제나 보조금 우대를 부여할 수 있는 권한이 포함되었다. 종합적으로 보면, OBBBA는 AI를 비롯한 첨단 제조업 육성과 국가안보 고려를 한데 묶은 포

41) U.S. House of Representatives, H.R. 1, 2025.5.21.
https://www.congress.gov/bill/119th-congress/house-bill/1/text

괄적 법률 패키지로서 의의가 있다. 이 법의 통과로 미국은 연방 재정을 활용해 국내 첨단 산업 생태계의 자율성과 안전성을 확보하는 새로운 접근을 시도한 것이다.

③ 주(州) 입법 동향[42]

연방 입법과 별개로, 각 주에서도 AI 관련 법률 제정이 봇물을 이루었다. 2025년 상반기 기준 47개 주가 총 260건 이상의 AI 관련 법안을 상정했고, 이 중 22건이 최종 법률로 제정되었다. 공화·민주 양당이 제출한 법안의 성격은 상이했다. 민주당 주도 법안(전체 66.5%)은 주로 AI 개발자와 기업에 대한 엄격한 의무를 부과하는 포괄적 규제안이 많았고, 공화당 법안(32.3%)은 특정 유해 AI 사용 금지나 혁신 장려에 초점을 맞추는 경향이 있었다. 그럼에도 선거 딥페이크 금지, 아동 성착취물 단속 등 몇몇 분야에서는 초당적 합의로 법이 통과되기도 했다. 예를 들면, 텍사스 등 일부 주는 선거 10일 이내 후보를 합성한 영상 배포를 불법화하는 법을 공화·민주가 함께 통과시켰다. 반면 캘리포니아, 뉴욕 등 민주당 우세 주들은 광범위한 AI 거버넌스 체계를 마련하려는 움직임을 보였고, 공화당 우세 주들은 자율주행차나 의료AI 등 특정 분야 규제완화 법안을 추진하는 등 차이가 컸다. 이러한 주별 입법은 미국 내 AI 규제의 파편화 가능성을 높이고 있으며, 연방 차원에서 우려하는 지점이다. 실제로 OBBBA 논의 중 포함되었다 탈락한 주 규제 모라토리엄 조항은 연방-주 간 규제 충돌의 잠재성을 내포하고 있음을 시사한다.

(3) 관계 부처 및 기관

「Take It Down Act」 시행과 관련해서는 연방수사국(FBI)과 사이버범죄 수사기관이 딥페이크 유포 범죄 단속의 전면에 나선다. 형사처벌 조항 집행은 법무부 산하 연방검찰이 담당하며, 특히 미성년자 대상 범죄의 경우 아동 착취 특임 검사들이 투입될 것이다. 플랫폼 측면에서는 연방거래위원회(FTC)가 신고·삭제 의무 준수 여부를 감시·집행한다. FTC는 위반 플랫폼에 대해 막대한 과징금 부과나 서비스 정지 명령 등을 내릴 수 있다. 또한 법무부 민권국은 딥페이크가 혐오표현이나 인권침해에 쓰이는 경우 별도 조사에 착수할 수 있다. OBBBA의 PFE 규정 시행은 재무부 산하 국세청(IRS)이 세제집행을 통해 담당하며, PFE 지정은 상무부 산업안보국(BIS)이나 국무부의 적성국 리스트를 활용하게 된다. 또한 법 이행을 위해 국토안보부 산하 국가안보 산업기반국(NSIB)이 기업들의 공급망 관계 조사를 지원할 수 있다. 주(州) 입법의 집행은 각 주 법무장관(Attorney General)과 규제기관의 소관이지만, 연방 법률과 충돌할 경우 연방법이 우선하기 때문에 연방법무부(DOJ)가 필요시 개입하여 정리할 수 있다.

42) Brookings, Commentary- How different states are approaching AI, 2025.8

제2절 EU

1. 그간의 인공지능 정책 추진 경과

<표 3-6> EU의 주요 AI 정책 추진 현황

시기	발표기관	정책
2018년 4월	EU 집행위원회	유럽을 위한 인공지능 통신문 발표 - 인간 중심·신뢰 가능한 AI 비전과 초기 투자·규범 방향 제시
2020년 2월	EU 집행위원회	AI 백서 발표 - 위험 기반의 규제 접근방식 공식화 및 고·중·저위험 체계 구상
2021년 4월	EU 집행위원회	AI 법(AI Act) 초안 제안 - 고위험 AI에 대한 요건과 금지 관행을 규정한 세계 최초의 포괄적 AI 규제안
2021년 4월	EU 집행위원회	AI 조정 계획 2021 리뷰 - 회원국 공동투자·전략조정 계획 개정, 유럽 AI 생태계 강화
2024년 1월	EU 집행위원회	AI 혁신 패키지 발표 - AI 팩토리, AI 사무국, 샌드박스 등 혁신 지원 이니셔티브
2024년 6월	유럽의회·EU이사회	AI 법 채택
2025년 1월	EU 집행위원회	경쟁력 나침반 정책 로드맵 발표 - AI 기술의 전략적 중요성 언급
2025년 2월	EU 집행위원회	금지된 AI 관행에 관한 지침(Guidelines on prohibited AI practices) 발표 및 AI 법 제5조 '금지 시스템' 조항 시행
2025년 2월	EU 집행위원회	InvestAI 투자촉진 이니셔티브 발표
2025년 2월	EU 집행위원회	인공지능책임지침안(AILD) 철회
2025년 4월	EU 집행위원회	AI 대륙 실행계획 발표 - AI법 서비스 데스크, 규제 간소화, 인력·인프라 투자를 포함한 실행 로드맵
2025년 7월	EU 집행위원회	GPAI 실천강령 채택
2025년 10월	EU 집행위원회	AI 팩토리 부지 추가 선정(총 19개) 및 AI 팩토리 안테나 부지 선정(총 13개)
2025년 10월	EU 집행위원회	AI 적용 전략 - 10대 핵심 산업과 공공 분야에 대한 AI 적용 전략
2025년 11월	EU 집행위원회	디지털 옴니버스 AI 규제안 제안
2025년 11월	EU 집행위원회	데이터 연합 전략 발표

EU의 인공지능 정책은 2018년 4월 '유럽을 위한 AI(Artificial Intelligence for Europe)' 통신문 발표 이후 본격화되었다.[43] EU의 초기 AI 전략의 중점은 '인간 중심 AI'와 '신뢰할 수 있는 AI'라는 EU

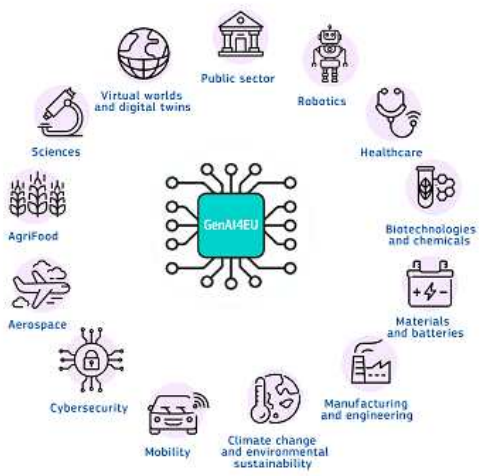
고유의 핵심 가치 정립에 있었다. 기술의 우수성 확보와 더불어 윤리·법적 프레임워크 구축을 목표로 삼은 것이다. EU는 2020년 2월 발간한 ‘AI 백서(White Paper on AI)’에서 이러한 비전을 구체화하고 AI 시스템을 위험도에 따라 차등적으로 접근하는 ‘위험 기반 접근법’을 공식적으로 제시했다.[44] 이는 향후 EU가 추진할 모든 AI 규제 논의의 핵심 토대가 되었다.

‘AI 백서’에서 제시된 위험 기반 접근법은 2021년 4월 「AI 법(AI Act)」 초안 제안으로 이어졌다. 「AI 법」은 세계 최초로 AI 기술 전반을 포괄하는 법률적 규제 체계를 시도했다는 점에서 세계적인 주목을 받았다. 「AI 법」은 EU 집행위원회, EU 이사회, 유럽의회 간의 복잡한 3자 협상을 포함해 수년간의 심도 깊은 입법 과정을 거쳐 2024년 3월 유럽의회 본회의에서 최종 채택되었고, 같은 해 8월 공식 발효되어 단계적인 시행 일정에 돌입했다. 「AI 법」은 AI 시스템이 사회와 개인에게 초래할 수 있는 위험 수준에 따라 명확한 규제 의무를 차등적으로 부과하는 특징을 가지고 있다. 이를 통해 EU는 AI 기술 발전의 전제 조건으로 시장의 안전과 시민의 기본권 보호를 최우선에 두는 정책 방향을 확립했다.[45]

그러나 「AI 법」의 엄격한 규제 요건이 유럽 기업들, 특히 스타트업과 중소기업의 혁신을 저해할 수 있다는 지적이 산업계 내부에서 지속적으로 제기되면서 EU는 산업 경쟁력 저하 가능성에 대응해야 했다. 이에 EU 집행위원회는 「AI 법」 최종 승인과 거의 동시에 2024년 1월 「AI 혁신 패키지 초안(AI innovation package to support Artificial Intelligence startups and SMEs)」을 발표했다.[46] 「AI 혁신 패키지」는 AI 규제와 산업 진흥의 균형을 맞추기 위한 EU의 주요 실행 조치 중 하나로, 같은 해 7월 「AI 혁신 패키지(AI innovation package)」 업데이트 발표에 따라 AI 기업을 재정 및 기술적으로 지원하는 데 정책의 초점을 맞췄다.[47] 「AI 혁신 패키지」의 핵심 요소는 AI 모델 훈련에 필수적인 슈퍼컴퓨팅 자원에 대한 접근성을 높이는 ‘AI 팩토리(AI Factories)’ 이니셔티브였다. 또한 「AI 법」의 일관된 집행을 감독하고 혁신 생태계를 지원할 AI 사무국(AI Office) 설립 계획도 공식화했다.

43) European Commission, Communication Artificial Intelligence for Europe, 2018.4.25.
44) European Commission, White Paper on Artificial Intelligence, 2020.2.19.
45) Official Journal of the European Union, REGULATION (EU) 2024/1689 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL, 2024.6.13.
46) European Commission, Commission launches AI innovation package to support Artificial Intelligence startups and SMEs, 2024.1.24.
47) European Commission, AI Innovation Package, 2024.7.25.

<표 3-7> EU AI 혁신 패키지 이니셔티브 구조

<table>
<tr><th colspan="2">AI 스타트업 및 혁신 커뮤니케이션</th></tr>
<tr><td colspan="2">① ‘호라이즌 유럽’과 ‘디지털 유럽 프로그램’을 통해 2027년까지 약 40억 유로 지원
② EU 내 AI 인재 풀 강화
③ AI 스타트업과 스케일업에 대한 공공 및 민간 투자 장려
④ ‘공동 유럽 데이터 공간’ 개발 및 배포 가속화</td></tr>
<tr><th>AI 사무국</th><th>AI 팩토리</th></tr>
<tr><td>① 신뢰할 수 있는 AI의 개발 및 활용 지원
② AI 위험으로부터 보호할 수 있도록
유럽 AI 거버넌스 시스템의 기반</td><td>EuroHPC를 활용한 신뢰할 수 있는
최첨단 생성 AI 모델 개발</td></tr>
<tr><th colspan="2">GenAI4EU</th></tr>
<tr><td>① 유럽 14개 산업 생태계 및 공공 부문에서
AI 신규 활용 사례 발굴
② 신흥 애플리케이션 개발에 기여</td><td> 
</td></tr>
<tr><th>ALT-EDIC</th><th>CitiVERSE-EDIC</th></tr>
<tr><td>① AI 학습용 언어 데이터 부족 문제 해결
② 유럽 공통 언어 기술 인프라 구축
③ 유럽의 LLM 개발 지원</td><td>① 최첨단 AI 적용 커뮤니티를 위한
로컬 디지털 트윈 개발 및 개선
② 도시 생태 프로세스 시뮬레이션 및 최적화 지원</td></tr>
</table>

출처 : European Commission(2024.7.), SPRi 정리

2025년에 들어서면서 이러한 산업 진흥 및 실질적 적용을 위한 움직임은 더욱 가속화되었다. 2025년 4월 EU 집행위원회는 「AI 대륙 실행계획(AI Continent Action Plan)」을 발표하며 AI 분야의 기술 주권 확보 의지를 분명히 했다. 「AI 대륙 실행계획」은 기존 'AI 팩토리'를 유럽 전역으로 신속하게 확충하고, 나아가 차세대 대규모 모델 개발을 위한 'AI 기가팩토리(AI Gigafactory)' 구축 구상까지 제시했다.

이러한 인프라 및 생태계 조성 노력은 2025년 10월 「AI 적용 전략(Apply AI Strategy)」 발표로 이어지며 정책의 방향성을 더욱 명확히 했다. 「AI 적용 전략」은 AI 기술이 연구 수준에 머무르지 않고, 유럽이 전통적으로 강점을 가진 ▲제조 ▲의료 ▲모빌리티 등 핵심 산업 분야 및 공공 부문에 실질적으로 도입 및 확산되는 것을 목표로 한다.

EU의 AI 정책은 「AI 법」을 통해 '신뢰할 수 있는 AI'의 제도적 기반을 글로벌 표준으로 확고히 다지는 단계에서 출발했지만, 현실 문제와의 마찰로 인해 기존에 구축한 규제 환경을 다소 수정하더라도 합의할 수 있는 규제 수준 내에서 산업 혁신을 촉진하는 방향으로 변화하고 있다. 나아가 「AI 대륙 실행계획」과 「AI 적용 전략」을 발표하며 AI 기술의 경제적 활용과 산업 적용을 극대화하는 방향으로 정책적 무게 중심을 이동시키고 있는 것으로 보인다.

2. 2025년 인공지능 정책 동향

1) AI 대륙 실행계획[48]

(1) AI 대륙 실행계획 발표 배경

EU 집행위원회는 2025년 4월 「AI 대륙 실행계획(AI Continent Action Plan)」을 공식 발표했다. 「AI 대륙 실행계획」 수립 배경에는 유럽의 현 상황에 대한 인식이 자리하고 있는 것으로 보인다. EU는 우수한 기초연구 기반과 강력한 전통 산업이라는 강점을 보유하고 있음에도 불구하고 AI 기술의 핵심 동력에서 분명한 약점을 드러냈다. 특히 AI 모델 개발에 필수적인 고성능 컴퓨팅 인프라가 회원국별로 파편화되어 있었고, 고품질 데이터에 대한 접근성이 부족했다. 2024년까지 유럽 내 AI 정책은 세계 최초의 포괄적 AI 규제 체계인 「AI 법」과 같이 규범의 정립에 초점을 맞추고 있었다. 그러나 생성 AI의 실용화 등 AI 산업이 발전함에 따라 산업 경쟁력 확보와 기술 주권 강화라는 실용적인 목표로 초점이 본격 전환되기 시작했다.

48) European Commission, AI Continent Action Plan COM(2025), 2025.4.9.

2025년 1월, EU 집행위원회는 향후 5년의 핵심적인 정책 로드맵을 담은 '경쟁력 나침반(Competitiveness Compass)'[49]에서 EU가 직면한 가장 시급한 과제 중 하나로 '혁신 격차 해소'를 지목하며, AI 기술의 전략적 중요성을 강조했다. 이러한 EU 집행위원회의 경쟁력 강화 의지는 2025년 2월 '파리 AI 액션 서밋(Paris AI Action Summit)'에서 구체화되었다. 우르줄라 폰 데어 라이엔(Ursula von der Leyen) EU 집행위원장은 이 자리에서 "유럽을 선도적인 'AI 대륙'으로 만들겠다"는 정책 비전을 공식적으로 선언했는데 이는 미국, 중국의 빅테크 기업들이 주도하는 글로벌 AI 패권 경쟁 속에서 유럽만의 독자적인 리더십을 확보하겠다는 의지의 표명이었다.

이 가운데 발표된 「AI 대륙 실행계획」에서 EU 집행위원회는 개발된 AI 기술이 실제 산업 현장, 특히 유럽 경제의 근간을 이루는 중소기업(SME)으로 확산하는 속도가 현저히 느리다는 문제점을 지적하며, 최우선 목표로 이러한 구조적 약점을 극복하고 EU의 기술 주권을 확보하여 글로벌 AI 리더십을 공고히 할 것을 제시했다. 또한 27개 회원국에 분산된 자원과 개별적으로 추진되던 정책들을 'AI 대륙'이라는 단일한 비전 아래 통합하고 조정하겠다고 밝혔다. AI 분야의 우수한 연구성과가 연구 및 실험에 머무르지 않고 실제 산업 현장의 부가가치 창출로 신속하게 연결되는 선순환 생태계를 구축하는 것 역시 핵심 목표의 하나로 설정했다. 특히 「AI 법」을 통해 확립한 '신뢰할 수 있는 AI'라는 가치를 단순한 규제적 장벽이 아닌, 차별화된 산업 경쟁력으로 전환시키려는 전략적 목표를 명확히 했다.

(2) AI 대륙 실행계획 주요 내용

「AI 대륙 실행계획」은 5대 핵심 전략인 ▲컴퓨팅 인프라 구축 ▲데이터 접근성 확대 ▲알고리즘 개발 및 도입 촉진 ▲AI 인재 육성 ▲규제 간소화를 중심으로 구성된다. 이 계획의 가장 큰 특징은 포괄성과 연계성으로, 완전히 새로운 정책을 제시하기보다 기존의 다양한 이니셔티브를 하나의 상위 전략으로 통합하여 정책적 시너지를 극대화하는 데 중점을 둔다.

<표 3-8> AI 대륙 실행계획 5대 핵심 전략

구분	내용
컴퓨팅 인프라 구축	• 'AI 팩토리'의 신속한 확산 ○ 총 13개의 'AI 팩토리' 네트워크 구축 - ▲스타트업 ▲중소기업 ▲연구자 AI 모델 훈련 및 테스트 - 개방형 고성능 컴퓨팅(HPC) 자원을 제공하는 허브 역할 • 'AI 기가팩토리' 구축 ○ 최대 5개의 'AI 기가팩토리' 건설 - 'EU 클라우드 및 AI 개발법' 추진

49) European Commission, Competitiveness compass, 2025.1.29.

구분	내용
데이터 접근성 확대	• ‘데이터 연합(Data Union)’ 전략 수립 ○ 데이터 상호운용성 표준을 마련 ○ 실질적인 EU 단일 데이터 시장 조성 - ‘AI 팩토리’ 내부에 ‘데이터 랩(Data Lab)’ 설치 - 공동 유럽 데이터 공간 연계
알고리즘 개발 및 도입 촉진	• 유럽디지털혁신허브(EDIH) ‘AI 경험 센터’로 기능적 전환 - 중소기업들의 AI 솔루션 테스트 - 자금 조달 및 기술 도입에 대한 전문적 자문
AI 인재 육성	• ‘AI 스킬 아카데미(AI Skills Academy)’ 출범 - AI 관련 고급 기술 교육 및 재교육 프로그램 제공
규제 간소화	• AI 사무국 내 ‘AI법 서비스 데스크’ 설치 - 중소기업 및 스타트업에 맞춤형 정보와 지침 제공 - '신뢰할 수 있는 AI' 생태계 연착륙 지원

출처: European Commission(2025.4.), SPRI 정리

먼저, ‘컴퓨팅 인프라 구축’은 AI 개발의 물리적 기반이 되는 컴퓨팅 인프라의 대대적인 확충을 위한 계획으로 그 핵심은 2024년 1월 「AI 혁신 패키지」에서 처음 제시된 ‘AI 팩토리(AI Factories)’의 신속한 확산이다. EU는 2024년 12월 7개 컨소시엄을 선정한 데 이어 2025년 3월과 10월에 추가 선정을 진행해 현재 총 19개의 ‘AI 팩토리’ 네트워크 구축을 본격화하고 있다. 이 시설들은 유럽 전역의 스타트업, 중소기업, 연구자들이 AI 모델을 훈련하고 테스트할 수 있도록 개방형 고성능 컴퓨팅(HPC) 자원을 제공하는 허브 역할을 수행한다.

또한 한 걸음 더 나아가, 차세대 프론티어 AI 모델 개발을 위한 ‘AI 기가팩토리(AI Gigafactory)’ 구축이라는 장기 목표를 포함한다. 최대 5개의 ‘AI 기가팩토리’를 건설하여 유럽에 전례 없는 규모의 연산 능력을 확보한다는 구상이다. 이러한 하드웨어 인프라 구축을 제도적으로 뒷받침하기 위해 「EU 클라우드 및 AI 개발법(EU Cloud and AI Development Act)」 입법도 함께 추진된다. 이 법안은 데이터센터 설립에 필요한 행정절차를 간소화하여 민간 투자를 유치하고, 향후 EU의 전체 데이터센터 용량을 크게 확대하는 것을 목표로 한다.

다음으로 ‘데이터 접근성 확대’는 AI 기술의 핵심 연료인 데이터 접근성을 근본적으로 강화하는 것이다. EU 집행위원회는 AI 개발과 혁신을 저해하는 주요 요인으로 회원국 간에 파편화된 데이터 환경을 지목하고, 「AI 대륙 실행계획」에서 이를 극복하기 위해 2025년 하반기에 새로운 「데이터 연합 전략

(Data Union Strategy)」을 수립하겠다고 밝혔으며, 실제로 2025년 11월 디지털 정책 패키지 제안에 포함해 이 전략을 발표했다.

'데이터 접근성 확대'의 목표는 데이터 상호운용성 표준을 마련하고 고품질 데이터의 가용성을 높여, 실질적인 EU 단일 데이터 시장을 조성하는 것이다. 이를 실행하기 위한 핵심 도구로 'AI 팩토리' 내부에 '데이터랩(Data Lab)'이 설치된다. '데이터랩'은 각 'AI 팩토리' 간의 데이터 공유를 촉진하고, '공동 유럽 데이터 공간(CEDS: Common European Data Spaces)'와 연계해 AI 개발자들에게 필요한 데이터를 안전하고 효율적으로 제공하는 기술적 창구 역할을 수행하게 된다.

'알고리즘 개발 및 도입 촉진'은 개발된 알고리즘이 실험실에 머무르지 않고 실제 산업 현장에 신속하게 도입되도록 촉진하는 것이다. EU 집행위원회는 특히 ▲첨단 제조 ▲항공우주 ▲제약 ▲에너지 등 유럽이 전통적인 강점을 보유한 핵심 산업 부문에 AI 솔루션 도입을 가속화하고자 한다. 이를 위한 실행 수단으로 「AI 대륙 실행계획」은 기존 유럽디지털혁신허브(EDIH) 네트워크를 'AI 경험 센터(Experience Centres for AI)'로 기능적으로 전환할 것을 제안했다. 이 센터는 중소기업들이 복잡한 AI 솔루션을 직접 테스트하고, 자금 조달 및 기술 도입에 대한 전문적인 자문을 받는 원스톱 지원센터로 활용될 예정이다.

'AI 인재 육성'은 AI 생태계의 지속가능성을 좌우할 AI 역량 및 인재를 육성하는 것이다. EU 집행위원회는 AI 분야 전문 인력에 대한 수요가 전 세계적으로 급증하는 상황에 대응하여, EU 외부로부터의 핵심 인재 유치와 내부 인력 양성을 병행한다. AI 인재 육성을 위한 핵심적인 실행 방안은 'AI 스킬 아카데미(AI Skills Academy)'의 공식 출범이다. 이 아카데미는 AI 관련 고급 기술 교육과 재교육 프로그램을 제공하는 원스톱 플랫폼 역할을 한다. EU 전역의 기술 격차를 해소하고, 다양한 산업 분야에서 요구하는 맞춤형 AI 인력 수요에 체계적으로 대응하는 것을 목표로 한다.

마지막으로 '규제 간소화'는 AI 혁신의 걸림돌로 작용할 수 있는 규제 부담을 완화하고 준수 과정을 간소화하는 것이다. 「AI 법」이 2024년 8월 발효되고 2025년부터 단계적 시행에 들어감에 따라, 기업들의 법 준수를 지원하는 것이 시급한 과제로 부상했다. 이에 EU는 「AI 대륙 실행계획」에서 AI 사무국 내에 'AI법 서비스데스크(AI Act Service Desk)'를 설치할 것을 명시했다. 'AI법 서비스 데스크'는 기업, 특히 규제 대응 여력이 부족한 중소기업과 스타트업이 복잡한 규제 요건을 명확히 이해하고 준수할 수 있도록 맞춤형 정보와 지침을 제공하는 중앙 창구 역할을 수행한다. 이를 통해 규제 불확실성을 해소하고 '신뢰할 수 있는 AI' 생태계의 연착륙을 지원한다.

2) AI 법 및 주요 입법 동향

(1) AI 법 추진 경과

2024년 EU는 「AI 법」의 모든 공식적인 입법 절차를 완료했다. 「AI 법」은 2024년 3월 유럽의회와 같은 해 5월 EU 이사회의 최종 승인을 거쳐 2024년 8월 1일부로 공식 발효되었다. 「AI 법」의 발효는 사전 설계된 '단계적 적용(staged application)' 일정의 법적 시작을 의미한다.

법적 구속력을 동반한 첫 조치로 2025년 2월 「AI 법」 제5조 '금지된 AI 시스템(Prohibited AI systems)'의 EU 내 ▲시장 출시 ▲서비스 제공 ▲사용이 금지되었다. 이는 EU의 핵심 가치인 기본권 보호 원칙이 가장 강력하게 반영된 법률의 발효로 AI 기술의 잠재적 위험성에 대해 EU가 설정한 'AI 레드라인'이 작동하기 시작했음을 의미한다.

<표 3-9> EU AI 법 '금지된 AI 시스템'

구분	금지된 AI 관행	주요 내용
인간 행동 조작	인지적 조작 시스템	• 잠재의식적 기법 혹은 기만적 방식으로 개인 행동을 왜곡하는 시스템 • 시스템으로 개인이 정보에 입각한 결정을 내릴 능력을 손상해 자신이나 타인에게 중대한 해를 끼치도록 유도하는 경우
	취약점 악용 시스템	• ▲연령 ▲신체적 장애 ▲정신적 장애 등 특정 집단의 취약점을 AI가 인지하고 이를 악용하는 시스템 • 해당 집단의 행동을 중대한 해를 끼치는 방향으로 왜곡시킬 가능성이 높은 경우
사회적 평가	범용 사회적 점수 평가	• 공공 기관이 개인의 ▲사회적 행동 ▲개인적 특성 등을 기반으로 신뢰도를 평가(social scoring)하는 시스템 • 평가 점수가 개인에게 불리하거나 무관한 맥락(예: 공공 혜택 거부)에서 불이익한 대우로 이어지는 경우
생체 인식	실시간 원격 생체 인식	• 법 집행 목적으로 공공장소에서 '실시간' 원격 생체 인식 시스템을 사용하는 행위 • *(예외) ▲테러 위협 방지 ▲중대 범죄 피해자 및 용의자 수색 등 엄격히 규정되고 사법 당국의 허가를 받은 경우는 제외*
	민감성 특성 추론	• 개인의 생체 인식 데이터를 사용하여 ▲인종 ▲정치적 견해 ▲종교적 신념 ▲성적 지향 ▲노동조합 가입 여부 등 민감한 개인 특성을 추론하거나 분류하는 시스템
	무차별적 이미지 스크래핑	• 인터넷이나 CCTV 영상 등에서 얼굴 이미지를 목표 없이 무차별적으로 수집하여 얼굴 인식 데이터베이스를 생성하거나 확장하는 행위

감정 및 행동 예측	직장 및 교육기관 내 감정 인식	• 직장이나 교육 기관 환경에서 AI 시스템을 사용하여 근로자나 학생의 감정을 추론하는 행위 • *(예외) 의료 또는 안전상의 명확한 이유가 있는 경우는 제외*
	프로파일링 기반 예측 치안	• 개인의 ▲프로파일링 ▲성격 특성 ▲과거 범죄 기록만을 기반으로 특정 개인의 범죄 발생 위험도를 평가하는 예측 치안 시스템
처벌 대상 및 내용		
처벌 대상	‘금지된 AI 시스템’의 ▲시장 출시 ▲서비스 제공 ▲업무 적용하는 기업 또는 기관	
처벌 내용	최대 3,500만 유로 또는 위반 기업의 직전 회계연도 기준 전 세계 연간 총매출액의 7% 중 더 높은 금액에 해당하는 과징금	

출처: EU AI Act(2024.2.), SPRI 정리

EU 집행위원회는 중대 조항의 시행착오를 줄이고 시장 혼란을 방지하기 위해 「금지된 AI 관행에 관한 지침(Guidelines on prohibited artificial intelligence practices)」 초안[50]을 발표했다. 이 지침은 그 자체로 법적 구속력은 없으나 법 조항들을 실제 사례에서 어떻게 해석하고 적용해야 하는지에 대한 기준을 제시한다.

이어 2025년 7월 EU 집행위원회는 범용 AI(GPAI) 실천 강령 최종본[51]을 발표하고, 특히 시스템적 위험을 가진 GPAI 모델에 대한 새로운 규정을 적용했다. 해당 모델의 제공자는 모델의 역량, 한계, 테스트 결과 등을 포함한 상세한 기술 문서를 작성하고 유지해야 하며, 저작권법 준수를 위해 훈련 데이터에 사용된 콘텐츠의 요약본을 공개해야 하는 의무를 지게 되었다.

그러나 「AI 법」을 통한 규제가 본격화되기 시작하자 EU의 관련 산업계는 규제의 복잡성과 불확실성을 이유로 우려를 표출했고, 2025년 7월 주요 기업들은 EU 집행위원회에 GPAI 규정 및 고위험 AI 규정의 시행을 2년간 연기해달라는 공개서한을 발송했다. 이처럼 거센 반발 속에서 EU는 2025년 11월 발표한 디지털 규제 간소화를 위한 「디지털 옴니버스(Digital Omnibus)」 패키지에서 AI 관련 고위험 분야 규제 적용 시점을 2026년 8월에서 최대 16개월 연기한다고 밝혔다.

50) European Commission, Commission publishes the Guidelines on prohibited artificial intelligence (AI) practices, as defined by the AI Act, 2025.2.4.
51) European Commission, The General-Purpose AI Code of Practice, 2025.7.10.

(2) 데이터법 적용[52)]

2025년 9월 EU는 「데이터법(Data Act)」을 모든 회원국에서 전면 적용한다고 발표했다. 「AI 법」이 AI 기술의 신뢰성과 안전한 사용을 규율하는 프레임워크라면, 「데이터법」은 AI 개발과 데이터 경제의 핵심 자원인 데이터의 공정한 접근과 자유로운 유통을 보장하는 법률로, 두 법안은 유럽의 AI 생태계 구축을 위한 상호보완적인 기반으로 작동한다.

「데이터법」의 전면 적용은 2024년 1월 11일 법안이 공식 발효된 시점부터 이미 예고되었으며, 데이터 보유 제품 제조사들이 새로운 데이터 공유 의무를 이행하고 관련 시스템을 준비할 수 있도록 기업에 법 발효 이후 약 20개월간의 유예 기간을 부여했다.

이 법이 마련된 배경에는 사물인터넷의 확산이 있다. 스마트 가전, 웨어러블 기기, 커넥티드 카, 산업용 로봇 등 인터넷으로 연결된 기기에서 막대한 양의 데이터가 생성됨에도 불구하고 데이터의 접근 권한은 대부분 기기 제조사가 독점해왔다. EU는 이러한 데이터 독점 구조가 공정한 시장 경쟁을 저해하고, 데이터 기반의 혁신적인 서비스 등장을 가로막는다고 판단했다.

「데이터법」의 가장 핵심적인 변화는 연결된 제품에서 생성된 데이터에 대한 통제권을 사실상 사용자에게 부여한다는 점이다. 스마트 기기 소유자는 자신이 사용하며 생성한 데이터에 직접 접근할 권리를 보장받는다. 더 나아가, 사용자는 데이터 보유자(주로 제조사)에게 해당 데이터를 제3자에게 무상으로, 그리고 기계가 읽을 수 있는 형식으로 제공하도록 요청할 수 있다.

「데이터법」은 데이터의 자유로운 흐름을 막기 위해 제조사가 설정할 수 있는 불공정한 계약 조항의 체결을 금지한다. 특히 '수리권(Right to Repair)'과 같은 애프터마켓 시장의 경쟁을 촉진하여, 사용자는 고비용의 제조사 공식 AS 대신 자신이 선택한 독립 수리업체에 기기 상태 데이터를 공유하고 더 저렴하고 효율적인 유지보수 서비스를 받을 수 있게 된다. 이는 애프터마켓의 시장 참여자가 공장의 스마트 설비나 기계에서 발생하는 운영 데이터를 분석하여 공정 최적화, 예방 정비 등 새로운 산업 서비스를 개발할 수 있음을 뜻한다.

---

52) European Commission, Regulation (EU) 2023/2854 of the European Parliament and of the Council of 13 December 2023 on harmonised rules on fair access to and use of data and amending Regulation (EU) 2017/2394 and Directive (EU) 2020/1828 (Data Act), 2023.12.13.

또한 「데이터법」은 클라우드 서비스 시장의 종속성 문제를 해결하기 위한 조치도 포함한다. 사용자가 특정 클라우드 제공업체의 서비스에서 다른 제공업체로 데이터를 이전할 때 발생하는 기술적 장벽과 수수료를 크게 낮춰, 서비스 간 전환의 자유를 보장한다.

(3) 인공지능책임지침안 철회[53]

2025년 EU 집행위원회는 2025년 2월 12일 발표한 '2025년 업무계획'에서 EU의 산업 경쟁력 강화를 위한 '규제 간소화'를 핵심 정책 기조로 제시하며 「인공지능책임지침안(AILD: AI Liability Directive)」[54]의 관련 업무를 공식 철회했다. 「인공지능책임지침안」은 2022년 9월 EC가 AI 시스템으로 인해 피해를 본 시민이 손해배상을 청구할 때 발생하는 입증 책임을 완화하려는 목적으로 제안된 법안이었다. AI 기술 고유의 복잡성과 '블랙박스' 특성은 피해자가 기존의 회원국별 민사 책임 법규에 따라 가해자의 과실이나 시스템의 결함을 입증하는 것을 매우 어렵게 만든다. 「인공지능책임지침안」은 이러한 법적 공백을 메우기 위해 고안되었었다.

EU 집행위원회는 「인공지능책임지침안」에 대해 유럽의회와 EU 이사회 등 입법 기관 간 정치적 합의 도달 가능성이 낮다는 점을 들어 이 지침안을 철회한다고 설명하고, 다른 대안적인 접근 방식을 검토하겠다고 밝혔다. 이러한 결정은 이미 「AI 법」이 도입된 상황에서 추가적인 규제가 혁신을 저해할 수 있다는 산업계의 우려와 '파리 AI 정상회의' 등에서 제기된 외부의 비판을 의식한 조치로 해석된다.

또한 2024년에 최종 개정된 「제품책임지침」이 AI를 포함한 소프트웨어를 '제품'의 범위에 명확히 포함시킨 것도 철회 결정의 중요 배경이 되었다. AI가 탑재된 '제품'이 결함으로 인해 손해를 유발한 경우, 피해자가 개정된 「제품책임지침」을 통해 구제받을 수 있는 경로가 이미 확보되었기 때문이다.

향후 EU에서 AI로 인한 피해 구제는 두 가지 경로로 나뉘어 처리될 전망이다. 결함이 있는 AI '제품'으로 인한 손해는 개정된 「제품책임지침」에 따라 배상을 청구할 수 있다. 제품 결함 외의 AI 시스템 운영이나 서비스로 인한 비계약적 손해는 27개 회원국 각각의 민사 책임법에 의존해야 한다. 이 법적 공백은 향후 「AI 법」의 시장 적용 경과를 지켜보며 EU 차원에서 다시 논의될 가능성이 높아 보인다.

---

53) iapp, European Commission withdraws AI Liability Directive from consideration, 2025.2.12.

54) European Commission, Proposal for a Directive on adapting non contractual civil liability rules to artificial intelligence, 2022.9.28.

3) InvestAI 투자촉진 이니셔티브[55)]

「InvestAI 투자촉진 이니셔티브」는 2025년 2월 개최된 '파리 AI 액션 서밋'에서 처음 공개되었다. 「InvestAI 투자촉진 이니셔티브」의 등장 배경에는 생성 AI 개발 경쟁에 필요한 막대한 자본과 인프라에 대한 위기감이 자리하고 있다. EU는 「AI 법」을 통해 신뢰할 수 있는 AI의 표준을 제시했지만, 초대형 AI 모델 훈련에 필요한 컴퓨팅 자원과 투자는 미국과 중국의 빅테크 기업들을 따라가지 못하고 있었다. 이에 EU 집행위원회는 유럽의 기술 주권과 산업 리더십을 확보하기 위한 대규모 자본 동원 계획으로 「InvestAI 투자촉진 이니셔티브」를 발표했다.

「InvestAI 투자촉진 이니셔티브」의 핵심 목적은 총 2,000억 유로 규모의 공공 및 민간 투자를 유치하는 것이다. 이는 EU가 직접 예산을 투입하는 방식이 아니라, 공공 재원을 활용해 민간 부문의 투자를 촉진하는 구조를 가진다. EU는 '디지털 유럽 프로그램(Digital Europe Programme)', '호라이즌 유럽(Horizon Europe)', 'InvestEU' 등 기존 EU 기금에서 약 500억 유로의 예산을 확보한다. 이 자금은 민간 투자자들의 위험 부담을 완화하는 일종의 보증 역할을 수행한다. 이를 통해 나머지 민간 자본과 회원국별 재원을 결합해 총 2,000억 유로라는 목표를 달성한다는 구상이다.

동원된 자금의 최우선 투자처는 2025년 4월 「AI 대륙 실행계획」에서 강조된 'AI 기가팩토리(AI Gigafactory)'의 구축이다. 'AI 기가팩토리'는 약 10만 개 이상의 최신 AI 칩을 집적한 초대형 컴퓨팅 거점으로, EU 내 스타트업, 중소기업, 연구기관 등 대기업이 아닌 조직 및 연구자들도 외부 기술에 의존하지 않고 독자적인 프론티어 AI 모델을 개발할 수 있는 핵심 인프라이다. EU는 InvestAI 펀드를 통해 EU 전역에서 4곳의 차세대 'AI 기가팩토리'에 자금을 지원할 예정이다. 이렇게 구축된 새로운 인프라는 매우 복잡하고 규모가 큰 AI 모델을 훈련하는 데에 활용될 전망이다.

4) AI 팩토리 확충[56)]

'AI 팩토리'는 2024년 1월 「AI 혁신 패키지」에서 제시된 개념으로, AI 모델 개발과 생성 AI 훈련에 필수적인 HPC 자원에 대한 접근성을 해결하기 위해 고안되었다. EU 스타트업과 중소기업이 미국과 중국의 빅테크 기업에 비해 컴퓨팅 자원 확보에 어려움을 겪는 컴퓨팅 격차가 유럽 AI 생태계의 성장을 저해하는 주된 요인으로 지목되었기 때문이다. 이러한 배경에서 유럽 고성능컴퓨팅공동사업단(EuroHPC

55) European Commission, EU launches InvestAI initiative to mobilise €200 billion, 2025.2.11.
56) EuroHPC JU, The EuroHPC JU Selects Additional AI Factories to Strengthen Europe's AI Leadership, 2025.3.12.

JointUnder(JU))은 2024년 12월 7개 회원국(그리스, 독일, 룩셈부르크, 스웨덴, 스페인, 이탈리아, 핀란드)에서 1차 AI 팩토리 컨소시엄 선정을 완료했다.

'AI 팩토리' 확충의 목적은 유럽 내 모든 AI 혁신 주체가 지리적 위치에 구애받지 않고 AI 개발에 필요한 핵심 자원에 접근할 수 있도록 보장하는 데 있다. 또한, 「AI 법」이 규정한 '신뢰할 수 있는 AI'를 개발하고 검증할 수 있는 공공 인프라를 제공함으로써 규제 준수와 기술혁신을 동시에 지원하려는 목표를 가진다. 본격적인 확충 작업은 2025년 3월 12일 EuroHPC JU가 6개의 'AI 팩토리'를 추가 선정하며 시작되었다. 2차 선정 대상은 오스트리아, 불가리아, 프랑스, 독일, 폴란드, 슬로베니아로 특히 프랑스와 독일의 'AI 팩토리'는 엑사스케일(Exascale) 슈퍼컴퓨터를 기반으로 구축된다. 독일은 유럽 최초의 엑사스케일 슈퍼컴퓨터 'JUPITER', 프랑스는 2026년 배치될 'Alice Recoque'를 활용한다. 이어서 2025년 10월 10일에는 2026년 'AI 팩토리'를 구축할 추가 부지로 체코, 리투아니아, 네덜란드, 폴란드, 루마니아, 스페인 등 6곳을 선정했다. 이들은 앞서 선정된 13개 'AI 팩토리' 부지와 함께 AI 허브 네트워크를 형성할 전망이다. 특히 스페인의 'AI 팩토리'에는 혁신적인 AI 모델 및 애플리케이션 개발과 테스트를 위한 최첨단 인프라 역할을 하는 실험 플랫폼이 포함될 예정이다.[57]

'AI 팩토리'는 AI 모델 개발에 필요한 자원을 통합 제공하는 '원스톱 허브'가 되는 것을 목표로 한다. 이를 위해 ▲AI 최적화된 슈퍼컴퓨팅 시설 접근권 제공 ▲「AI 대륙 실행계획」에서 언급된 '데이터랩' 연계로 고품질 학습용 데이터 지원 ▲관련 기술 전문 지식과 인재 양성 프로그램을 제공하며, 'AI 팩토리' 네트워크는 EU 집행위원회가 추진하고 있는 「InvestAI 투자촉진 이니셔티브」와 「AI 적용 전략」을 뒷받침하는 실질적이고 중추적인 물리적 기반으로 작동할 전망이다.

이와 함께 2025년 10월 13일 EuroHPC JU는 벨기에, 키프로스, 헝가리 등 유럽 전역에 13개[58]의 'AI 팩토리 안테나(AI Factory Antennas)'를 선정했다.[59] 각국의 안테나는 자국 연구자·산업·공공부문에 AI 학습·테스트·검증 인프라를 제공하고, 인접국의 'AI 팩토리'와 협력하는 분산형 AI 거버넌스 모델로 운영될 예정이다. 예로 핀란드의 'AI 팩토리'는 벨기에·아이슬란드·라트비아·스위스의 안테나와 직접 연결되게 되는데, 이는 기존 'AI 팩토리' 거점들을 보완하여 AI 최적화 슈퍼컴퓨터(HPC) 접근성을 확대하고 국가별 AI 혁신 생태계의 균형 발전을 도모하기 위한 조치이다. EU는 이를 통해 총 19개

57) EuroHPC JU, The EuroHPC JU Selects Six Additional AI Factories to Expand Europe's AI Capabilities, 2025.10.10.
58) AI 팩토리 안테나는 벨기에, 키프로스, 헝가리, 아이슬란드, 아일랜드, 라트비아, 몰타, 몰도바, 북마케도니아, 세르비아, 슬로바키아, 스위스, 영국
59) EuroHPC JU, EuroHPC JU Selects AI Factory Antennas to Broaden the AI Factories Initiative, 2025.10.13.

의 AI 팩토리와 13개 안테나 네트워크를 확보하고 2026년 중반까지 통합 운영을 달성할 계획이다.

5) AI 적용 전략[60]

2025년 10월, EU 집행위원회는 「AI 적용 전략(Apply AI Strategy)」을 발표했다. 「AI 적용 전략」은 2025년 상반기 EU가 제시한 정책들을 실질적인 산업 성과로 연결하기 위한 실행 계획으로 「AI 법」 등 규제 프레임워크와 'AI 팩토리'라는 제도적 인프라를 바탕으로 AI 정책적 초점을 '활용'과 '확산'으로 전환시키고자 하는 방향성을 가지고 있다.

「AI 적용 전략」의 추진 배경에는 유럽 내 AI 도입률이 저조하다는 위기감이 자리하고 있다. 특히 EU 내 중소기업의 AI 활용이 부진하여, AI 기술의 잠재력이 실제 생산성 향상이나 사회적 혜택으로 이어지지 못하고 있다는 분석이 제기되었다. 「AI 적용 전략」은 EU를 'AI 대륙'으로 실현하기 위한 후속 조치이자, 「AI 대륙 실행계획」의 5대 전략 중 '알고리즘 개발 및 도입' 부분을 구체화하는 정책으로 볼 수 있다.

「AI 적용 전략」은 유럽의 기술 주권과 전략적 부문의 산업 경쟁력 강화를 목표로 한다. 이를 위해 AI를 문제해결의 우선적 도구로 채택하는 'AI 우선 정책(AI first policy)'을 장려한다. 또한 AI 기술의 혜택이 일부 빅테크 기업에 집중되지 않고 유럽 경제 전반, 특히 중소기업을 중심으로 광범위하게 확산되는 것을 목표로 설정했다. 보다 구체적으로 「AI 적용 전략」은 세 가지 핵심 추진축을 바탕으로 추진된다.

<표 3-10> AI 적용 전략 3대 추진 방향

구분	내용
부문별 플래그십 도입	• AI 도입이 시급하거나 유럽 핵심 산업 분야를 선정해 맞춤형 지원 제공 – 10개 핵심산업(▲의료·제약 ▲로봇공학 ▲제조·엔지니어링·건설 ▲국방·안보·우주 ▲모빌리티·운송·자동차 ▲전자통신 ▲에너지 ▲기후·환경 ▲농식품 ▲문화·창작·미디어) 및 공공부문에 맞춤형 지원
범분야적 기술 주권 강화	• 유럽 디지털 혁신 허브(EDIH)를 'AI 경험 센터'로 기능적 전환 • 'AI 스킬 아카데미' 설립
새로운 거버넌스 체계 확립	• 'AI 얼라이언스'를 AI 정책 논의 단일 창구 'AI 적용 얼라이언스'로 전환 • 'AI 관측소' 신설

출처 : Apply AI Strategy(2025.10.), SPRi 정리

60) European Commission, Apply AI Strategy, 2025.10.23.

첫째는 ‘부문별 플래그십’의 도입이다. 이는 AI 도입이 시급하거나 유럽이 강점을 가진 핵심 산업 분야를 선정하여 맞춤형 지원을 제공하는 방식이다. 이를 위해 ‘호라이즌 유럽’, ‘디지털 유럽 프로그램’, ‘EU4Health’ 등 기존 EU 프로그램 기금에서 약 10억 유로의 예산을 동원한다. 의료, 로봇공학, 국방 등 10개의 핵심 산업과 공공 부문에 지원이 집중된다.

둘째는 ‘범분야적 과제 해결을 통한 기술 주권 강화’이다. 이는 특정 산업에 국한되지 않는 공통의 문제를 해결하기 위한 조치로, 중소기업의 AI 도입 지원을 위해 기존의 유럽디지털혁신허브(EDIH)를 ‘AI 경험 센터(Experience Centres for AI)’로 기능적으로 전환하는 계획이 포함된다. ‘AI 경험 센터’는 AI 솔루션 공급자와 수요자인 기업을 연결하는 원스톱 지원 창구 역할을 수행한다. 또한, 인력 양성을 위해 ‘AI 스킬 아카데미(AI Skills Academy)’를 출범시킨다. ‘AI 스킬 아카데미’는 산업 부문별로 특화된 실무 AI 문해력 교육을 제공하고, 산업 전문 지식과 AI 기술을 갖춘 하이브리드 인력을 양성하는 AI 석사과정 등을 신설한다.

셋째는 ‘새로운 거버넌스 체계 확립’이다. 전략의 효과적인 이행을 위해 기존의 거버넌스 구조를 개편하는 것이 골자로, 2018년 출범했던 ‘AI 얼라이언스(AI Alliance)’를 이해관계자와 정책 입안자들이 부문별 AI 정책을 논의하는 단일 창구인 ‘AI 적용 얼라이언스(Apply AI Alliance)’로 전환한다. 이와 함께 ‘AI 관측소(AI Observatory)’를 신설하여 ▲AI 기술 동향 ▲산업별 영향 ▲노동시장 변화 등을 모니터링하고, 향후 AI 투자 목표 설정을 지원할 예정이다. 「AI 법」에 따라 설립된 ‘AI 위원회(AI Board)’는 회원국 간의 주요 토론 포럼으로서 그 역할을 유지한다.

6) 디지털 옴니버스 AI 규제안[61]

2025년 11월 19일 EU 집행위원회는 디지털 부문의 규제 간소화를 위한 정책 패키지인 「디지털 옴니버스 규제안(Digital Omnibus Regulation Proposal)」[62]을 공개했다. 이는 데이터, 사이버보안, AI 규제와 관련한 여러 디지털 법령을 한 번에 개정·폐지하여 규제 중복을 해소하고 시행을 용이하게 하는 포괄적인 입법 제안 패키지로, 여러 복잡하고 중복된 디지털 규제를 간소화하는 것이 핵심이다. 특히 이 입법 패키지에는 ‘AI 규제 유연화’를 위해 「AI 법」에 따른 고위험 AI 규제 적용 시점을 최대 16개월 연기하는 조치를 포함하는 별도의 「디지털 옴니버스 AI 규제안(Digital Omnibus on AI Regulation Proposal)」이 포함되었다.

---

61) European Commission, Digital Omnibus on AI Regulation Proposal, 2025.11.19.
62) European Commission, Digital Omnibus Regulation Proposal, 2025.11.19.

<표 3-11> 디지털 옴니버스(디지털 규제 간소화 패키지) 주요 내용

구분	주요 내용
데이터 규제 통합 및 정비	• 기존 데이터 관련 규정인 비개인 데이터 자유이동 규정(Regulation (EU) 2018/1807), 데이터거버넌스법(Regulation (EU) 2022/868), 공공 데이터 지침(Directive (EU) 2019/1024)에 분산된 데이터 공유 규칙을 데이터법으로 일원화
프라이버시 및 쿠키 규정 개혁	• 웹사이트 쿠키 규제(E-Privacy 지침)를 전면 개편해 GDPR 틀로 통합하고, 이용자 친화적인 간소화 도모
사이버보안 원스톱 신고체계 도입	• 사이버보안 사고나 데이터 침해 발생 시 여러 법령에 따라 여러 기관에 각각 보고해야 했던 사고 보고 절차를 단일 창구(EU 사이버보안청)로 일원화
플랫폼 규제 합리화	• 온라인 플랫폼-기업(P2B)간 공정성 규정(Regulation (EU) 2019/1150) 폐지
AI 규제 유연화	• 「AI 법」의 주요 조항 발효 시점을 기술지원 수단의 준비 상황과 연계하고 중소기업 지원 확대
AI 거버넌스 효율화	• EU AI 사무국의 역할을 강화해 감독 권한을 EU 차원으로 중앙화함으로써 AI 규제의 감독·집행 체계를 합리화하고 이행 효율 개선

출처 : Digital Omnibus Regulation Proposal(2025.11.) SPRi 정리

「디지털 옴니버스 AI 규제안」에 따르면, 고용·법 집행 등 민감한 분야에서 활용되는 고위험 AI 규정 시행 일정은 기존 2026년 8월에서 2027년 12월까지 최대 16개월 연기된다. 이는 고위험 AI 요건의 적용 시기를 관련 표준 등 기업 지원 도구가 마련된 시점과 연동하여 기업들이 준비 부족으로 과도한 규제 충격을 받지 않도록 조정하기 위한 조치이다. EU 집행위원회는 고위험 AI의 분류 유형에 따라 유예 기간을 구분해 일부 유형에는 더 긴 준비 기간을 부여할 방침이다.

또한 「AI 법」에서 기존에는 중소기업에만 적용되었던 일부 의무(기술문서 작성 의무 등) 완화 조치는 종업원 750명 미만의 중견기업까지 확대 적용된다. 이를 통해 중소기업과 마찬가지로 AI 규제 준수에 큰 부담을 안고 있는 중견기업의 어려움이 완화될 수 있을 전망이다. 이에 더해 모든 AI 공급자·사용자에게 획일적으로 부과되었던 '직원 대상 AI 교육 의무'는 실효성이 낮고 특히 중소기업에 부담이 된다는 지적에 따라 각 회원국 정부와 EU 집행위원회 주도로 전환된다. 다만, 고위험 AI 시스템을 운영하거나 사용하는 인력에 대한 훈련 의무를 계속 유지될 전망이다.

EU의 디지털 규제 간소화 방안은 현재 입법 제안된 상태로 최종 통과 및 확정이 되기까지는 회원국 간 추가 논의와 유럽의회 승인 등과 같은 절차가 남아있다. 이러한 EU의 제안은 미국과 EU의 기술기업, 미국 트럼프 행정부 등에서 추진한 규제 개혁 요구를 반영한 것이며, EU가 글로벌 기술 경쟁에서 뒤처지지 않도록 하기 위한 목적도 담겨 있다. 그러나 시민단체와 개인정보보호 활동가들은 이번 제안을 '디지털 권리 후퇴'로 비판하며, 구글이나 메타 등 빅테크 기업에 유리하게 작용할 수 있다고 경고했다.

7) 데이터 연합 전략[63]

2025년 11월 19일 EU 집행위원회는 디지털 옴니버스 정책 패키지와 함께 완결된 형태의 「데이터 연합 전략(Data Union Strategy)」을 공개했다. 이는 2020년 2월 발표했던 「유럽 데이터 전략(A European strategy for data)」의 후속 전략으로, EU 집행위원회는 「데이터 연합 전략」을 마련하기 위해 2025년 5월부터 공개 의견 수렴을 진행했고, 11월에 EU의 데이터 단일 시장 완성과 AI 학습용 데이터를 확대하기 위한 실행 계획 등을 담은 최종 전략을 발표했다.

「데이터 연합 전략」은 EU의 데이터 정책의 초점을 '규칙 제정'에서 '활용 성과'로 전환하고, AI 시대를 위한 데이터 인프라 구축을 가속하기 위한 전략이다. 이를 위한 우선 작업 영역으로 ▲데이터 접근성 향상 ▲데이터 규제 효율화 ▲데이터 주권 확보 및 국제 협력을 강조한다.

<표 3-12> EU 데이터 연합 전략 주요 내용

구분	주요 내용
데이터 접근성 향상	• 의료, 모빌리티 등 핵심 분야별로 공공·민간 주체들이 신뢰 기반으로 데이터 공유를 할 수 있게 구축된 데이터 생태계인 '공동 유럽 데이터 공간(CEDS)' 확대 • 공공·민간 데이터를 안전하게 결합해 활용할 수 있는 신뢰할 수 있는 실험 공간인 '데이터랩' 구축 추진 • AI 학습용 데이터 부족을 해결하기 위해 합성 데이터 활용 촉진
데이터 규제 효율화	• 데이터 관련 권리가 현장에서 지나친 복잡성과 부담으로 작용하지 않도록 불필요한 규정을 정비하고 절차 간소화 • 데이터 접근 및 사용 관련 모범 계약 조건, 클라우드 컴퓨팅 계약에 대한 표준 조항 등 새로운 가이드라인 마련 • 기업의 행정 업무에 대한 부담 완화를 위한 데이터법 헬프데스크(Data Act legal helpdesk) 도입

63) European Commission, Data Union Strategy - Unlocking data for AI, 2025.11.19.

구분	주요 내용
데이터 주권 확보와 국제 협력	• 신뢰할 수 있는 국제 파트너에게 데이터를 개방하되, EU의 가치와 이익에 부합하는 공정하고 안전한 조건이어야 한다는 단서 제시 • 데이터 교류에 대한 EU의 접근방식을 국제적으로 홍보하고, 글로벌 데이터 거버넌스에서 EU의 영향력 강화

출처 : Data Union Strategy(2025.11.) SPRi 정리

핵심적인 내용은 민간 및 공공데이터를 한 곳에 모아 중소기업, 연구자, 기업 등에 고품질의 부문별 데이터를 제공하는 '데이터랩'을 구축하는 것이다. '데이터랩'은 '공동 유럽 데이터 공간(CEDS)'과 AI 생태계를 연결해주는 허브이자, AI 개발·연구에 필요한 기술 인프라와 도구, 데이터, 자문 및 교육을 제공해주는 서비스 센터로서의 기능을 수행할 전망이다. EU는 우선 'AI 팩토리' 내에 '데이터랩'을 시범 설치해 운영하면서 이를 토대로 서비스 및 비즈니스 모델을 조정한 후 다른 지역으로 확대해갈 계획이다.

제3절 중국

1. 그간의 인공지능 정책 추진 경과

<표 3-13> 중국의 주요 AI 정책 추진 현황

시기	발표기관	정책
2017년 7월	국무원	차세대 AI 발전 규획
2017년 12월	공업정보화부	차세대 AI 산업 발전 3개년 행동계획
2018년 4월	교육부	고등교육기관 AI 혁신 행동계획
2019년 8월	과학기술부	국가 차세대 AI혁신발전시험구 건설 업무 지침
2020년 7월	국가표준화관리위원회 외	국가 차세대 AI 표준체계 구축 지침
2021년 12월	국가인터넷정보판공실 외	인터넷정보서비스 알고리즘 추천 관리규정(2022년 3월 시행)
2022년 7월	과학기술부	AI 활용 고도화를 통해 경제의 질적 발전을 촉진하기 위한 활용 사례 혁신 가속화에 관한 지도의견
2022년 11월	국가인터넷정보판공실 외	인터넷정보서비스 심층합성 관리규정(2023년 1월 시행)
2023년 7월	국가인터넷정보판공실 외	생성 AI 서비스 관리 임시 시행방법(2023년 8월 시행)
2023년 10월	국가인터넷정보판공실	글로벌 AI 거버넌스 이니셔티브
2024년 3월	국무원	정부업무보고에서 ‘AI 플러스’ 개념 제시
2025년 3월	국무원	정부업무보고에서 ‘AI 플러스’ 지속 추진 방침 제시
	국가인터넷정보판공실 외	AI 생성·합성 콘텐츠 표시 방법
2025년 5월	공업정보화부	컴퓨팅 자원 상호연결 행동계획
2025년 7월	외교부	AI 글로벌 거버넌스 실행계획
2025년 8월	국무원	AI 플러스 행동 심화 추진에 관한 의견
	공업정보화부 외	AI 과학기술 윤리 관리서비스 방법(시범 시행)’ 초안
2025년 9월	국가인터넷정보판공실	AI 안전 거버넌스 프레임워크 2.0
	국가에너지국 외	AI+ 에너지 고품질 발전 추진에 관한 실시의견
	교통운수부 외	AI+교통운수에 관한 실시의견
2025년 10월	국가발전개혁위원회 외	행정 분야 대규모 AI 모델의 도입과 응용에 관한 지침
2025년 12월	민용항공국	AI+민항 고품질 발전 추진에 관한 실시의견

중국 국무원이 2017년 7월 발표한 ‘차세대 AI 발전규획’은 AI 분야의 체계적 발전 방향을 제시한 중국 최초이자 현재까지 유일한 종합 AI 국가전략으로, 2030년까지 AI에서 세계 1위 도약을 선언하고 중장기 발전계획을 제시하였다. 발전규획은 1단계(~2020년)까지 기술과 응용에서 세계를 선도, 2단계(~2025년)에서 AI 이론 발전, 3단계(~2030년)까지 AI 이론과 기술, 응용 전 분야에서 세계 선도를 목표로 제시하였다.

중국은 ‘차세대 AI 발전규획’을 핵심 축으로 하여 부처 차원에서 AI 산업 발전계획과 실행계획을 수립함으로써 AI 발전을 위한 정책 기반과 제도적 보장 체계를 구축하였다. 일례로 공업정보화부는 차세대 AI 발전 규획의 목표 달성을 위해 2017년 12월 ‘차세대 AI 산업 발전 3개년 행동계획(2018~2020)’을 수립하였으며, 교육부는 2018년 ‘고등교육기관 AI 혁신 행동계획’을 발표하였다. 또한 중국 과학기술부는 2022년 7월 AI 육성을 위해 ‘AI 활용 고도화를 통해 경제의 질적 발전을 촉진하기 위한 활용 사례 혁신 가속화에 관한 지도의견’을 발표하였다.

중국 정부는 AI 산업 규제를 위한 법률도 선제적으로 제정하였다. 중국의 인터넷 규제당국인 국가인터넷정보판공실이 마련해 2022년 3월부터 시행 중인 ‘인터넷정보서비스 알고리즘 추천 관리규정’은 알고리즘을 이용해 정보를 추천 또는 정렬, 제공하는 플랫폼을 대상으로 부정적 영향 방지와 이용자 권리 보장, 미성년자 보호 등의 의무를 부과하였다. 2023년 1월부터 시행된 ‘인터넷 정보서비스 심층합성 관리규정’은 심층합성 서비스 제공업체에 심층합성 서비스 사용자의 신원 인증과 콘텐츠 관리, 안전관리 책임 등을 부과하였다.

중국 정부는 2022년 11월 오픈AI의 챗GPT 출시 이후 생성 AI 규제에 중점을 두고 있으며, 2023년 7월 국가인터넷정보판공실이 발표해 8월 15일부터 시행한 ‘생성 AI 서비스 관리 임시 시행 방법’은 생성 AI 서비스 제공자의 데이터 처리 활동과 사용자 책임에 관한 요구사항을 규정하였다. 또한 국가인터넷정보판공실은 2025년 3월에는 AI로 생성된 합성 콘텐츠에 대한 워터마크 표시 방법을 상세히 정한 규정을 발표하고 9월 1일부터 시행에 돌입하였다. 2025년 9월에는 2024년에 발표된 1.0 버전을 업그레이드한 위험 기반 AI 거버넌스 기술 프레임워크로 ‘AI 안전 거버넌스 프레임워크 2.0’을 발표하기도 했다.

한편, 중국 정부는 2023년 10월 AI 관련 논의에서 주도권을 확보하고자 ‘글로벌 AI 거버넌스 이니셔티브’를 발표했다. 동 이니셔티브는 AI 제품과 서비스 제공 시 타국의 주권을 존중하고 법률을 준수하며, 타국의 주권을 위협하는 데 AI를 사용해서는 안 된다는 점, 그리고 타국의 AI 발전을 악의적으로 방해하기 위한 개발 장벽 조성과 글로벌 AI 공급망을 가로막는 기술 독점과 강압 조치에 반대한다는 점을 명시

하였다. 리창 중국 총리는 2025년 7월 ‘AI 글로벌 거버넌스 실행계획’을 발표하며 AI 글로벌 거버넌스를 강화하기 위한 세계 AI 협력기구의 창설을 공식 제안하기도 했다.

중국 정부는 2024년 들어 처음으로 정부업무보고에서 ‘AI 플러스’ 개념을 새롭게 포함하며, AI와 실물경제의 융합에 초점을 두고 있다. 2024년 정부업무보고는 당해 정부 과제 중 하나로 디지털 기술과 실물경제의 심층적 통합으로 디지털 경제의 혁신적 발전을 촉진할 정책의 수립을 선언하였으며, 특히 빅데이터와 AI 등의 연구개발과 응용을 심화하고 AI 플러스 행동을 실시하며 국제 경쟁력을 갖춘 디지털 산업 클러스터 구축을 추진하겠다는 계획을 밝혔다. 중국 정부는 2025년 8월 정책 로드맵으로 ‘AI 플러스 행동 심화 추진에 관한 의견’을 발표하며 2035년까지 지능형 경제와 사회의 새로운 발전 단계에 전면 진입하겠다는 목표를 제시하였다.

2. 2025년 인공지능 정책 동향

1) 국무원

(1) AI 플러스 행동 심화 추진에 관한 의견

중국 정부는 2024년 정부업무보고에서 ‘AI 플러스’ 개념을 처음 제시한 이래 2025년 8월 26일 중앙정부 최초의 AI 플러스 정책 로드맵인 ‘AI 플러스 행동 심화 추진에 관한 의견’을 공개했다. 의견은 2027년까지 △과학기술 △산업 발전 △소비 품질 향상 △민생 △거버넌스 역량 △국제협력의 6대 핵심 영역에서 AI의 광범위하고 심도 있는 융합을 실현하고 차세대 지능형 단말과 지능형 에이전트 등의 보급률 70% 돌파를 목표로 제시했다. 2030년까지는 차세대 지능형 단말과 에이전트 등의 응용 보급률을 90% 이상으로 높이고 지능형 경제를 중국 경제발전의 핵심 동력으로 만들며, 2035년까지 지능형 경제와 사회의 새로운 발전 단계에 전면 진입한다는 계획이다.

<표 3-14> 중국 AI 플러스 행동 심화 추진을 위한 6대 핵심 영역

핵심 영역	세부 내용
AI+과학기술	• AI 기반 연구 패러다임 탐색, 과학 AI 모델과 데이터셋으로 과학적 발견 가속화 • 기술 R&D 모델 혁신과 효율성 제고, 철학과 사회과학 연구 혁신
AI+산업 발전	• AI 내재화 기술을 위한 새로운 비즈니스 모델 육성 • 산업 전 과정(설계, 생산, 서비스 운영)에 지능화 적용, 농업과 서비스업의 지능화

AI+소비 품질 향상	• 체험 소비, 개인 맞춤형 소비 등 새로운 서비스 소비 시나리오 확대 • 지능형 커넥티드카, AI 기반 휴대폰과 컴퓨터, 스마트 로봇 등 차세대 지능형 단말을 적극 개발해 지능형 상호작용 환경을 구축
AI+민생	• 인간과 기계의 협업을 촉진하는 새로운 조직 구조와 비즈니스 모델 모색 • 교육 전반에 AI를 통합해 더욱 효과적인 학습 방법을 장려 • 의료·보건, 문화 산업, 육아와 노인 돌봄 등에 AI를 활용해 삶의 질 개선
AI+거버넌스 역량	• 도시와 농촌 인프라 지능화, 행정 서비스 전반에 AI를 접목 • 공공안전과 재난 대응, 국가안보에 AI 활용을 강화 • 환경 보호와 생태 관리에서 AI를 기반으로 정밀한 거버넌스 모델 개발
AI+국제협력	• AI 역량 강화를 위한 개방 생태계를 조성 • UN 주도의 글로벌 AI 거버넌스 체계를 공동 구축

출처 : AI 플러스 행동 심화 추진에 관한 의견(2025.8)

‘의견’은 AI 플러스 정책 추진의 성공을 뒷받침하기 위해 8대 기본 지원 역량의 강화도 제안하였으며, 기본 지원 역량은 AI 모델과 데이터, 지능형 컴퓨팅, 응용 발전 환경, 오픈소스 생태계, 인재, 정책·법제, 안전 등을 핵심 요소로 한다. 구체적으로 AI 모델의 기초 능력을 향상하고 데이터 공급의 혁신을 강화하며, 지능형 컴퓨팅의 통합을 강화하기로 하였다. 또한 응용 발전 환경을 최적화하고, 오픈소스 생태계의 번영을 촉진하며 인재 역량과 정책·법제 보장을 강화하고 안전 역량도 강화하겠다는 방침이다.

한편, 주요 정부 부처는 ‘의견’에 따라 각 관할 분야에서 AI 플러스를 이행하기 위한 정책을 발표하고 있다. 국가에너지국은 2025년 9월 ‘AI+에너지 고품질 발전에 관한 실시의견’을 발표하고, 2027년까지 에너지와 AI를 융합한 혁신 체계를 구축하며 5개 이상의 에너지 특화 AI 모델을 전력망, 석탄, 석유와 가스 등 핵심 분야에 적용하기로 하였다. 교통운수부는 유관 부처와 공동으로 2025년 9월 ‘AI+교통운수 실시의견’을 발표하고 핵심기술 공급 확대, 혁신 시나리오 가속화, 핵심 요소 보장 강화, 산업 발전 생태계 최적화라는 4개 영역의 과제를 이행하여 2027년까지 주요 교통 시나리오에 AI를 널리 적용하겠다는 목표를 수립하였다. 중국 민용항공국은 2025년 12월 민항 전 분야에서 AI의 심층 융합을 추구하는 ‘AI+ 민항 고품질 발전 추진에 관한 실시의견’을 발표하고 AI+안전, AI+운영, AI+여객, AI+물류, AI+관리감독, AI+계획·건설의 6대 시나리오를 중심으로 혁신을 중점 추진하겠다는 계획을 제시했다.

⑵ AI 글로벌 거버넌스 행동계획

중국 정부는 2025년 7월 상하이에서 AI 글로벌 거버넌스 고위급 회의를 개최하고, 글로벌 AI 개발과 거버넌스에서 세계 각국의 참여와 협력을 촉구하는 ‘AI 글로벌 거버넌스 행동계획’을 발표했다. 행동계획은 AI 기회의 공동 포착, AI 혁신 발전의 촉진, AI를 통한 산업 발전, 디지털 인프라 개발 가속화, 혁신 생태계 구축, 고품질 데이터 기반 AI 발전 촉진, 에너지와 환경 문제 해결, 공공 부문의 AI 도입 주도, AI 보안 거버넌스, AI 역량 강화를 위한 국제협력, 표준과 규범 합의, 포용적 거버넌스 모델 구축 등을 제안했다.

리창 중국 총리는 행사 개막식의 기조연설을 통해 AI 글로벌 거버넌스를 강화하기 위한 세계 AI 협력 기구의 창설을 공식 제안하기도 했다. 리창 총리는 AI가 인류의 이익을 위한 공공재로서 더 많은 국가와 집단이 혜택을 누릴 수 있어야 한다며, 개발도상국을 중심으로 전 세계에 AI 개발 경험과 기술 제품을 공유하겠다고 강조했다. 또한 AI 기술의 돌파구를 마련하기 위한 협력에 중점을 두고 기초 과학기술 R&D 협력과 기술과 인재 교류를 강화하겠다는 방침을 밝혔다.

2) 공업정보화부

⑴ 컴퓨팅 자원 상호연결 행동계획‘

중국 공업정보화부는 2025년 5월 21일 공공 컴퓨팅 성능 향상과 자원 활용 효율성 개선, 컴퓨팅 자원의 고품질 개발을 촉진하기 위한 ‘컴퓨팅 자원 상호연결 행동계획’을 발표하였다. 계획은 2026년까지 컴퓨팅 자원 상호연결과 상호 운용을 위한 표준 확립, 국가·지역·산업 단위 컴퓨팅 자원 상호연결 플랫폼 구축, 2028년까지 전국 공공 컴퓨팅 자원의 상호연결 표준화 실현을 목표로 설정했다. 이를 위한 6대 중점 과제로 △컴퓨팅 자원 상호연결 인프라 구축 △컴퓨팅 자원 상호연결 최적화 △컴퓨팅 자원의 상호 이용 촉진 △컴퓨팅 자원 업무의 상호연결 △컴퓨팅 자원 상호연결을 위한 응용 시나리오 구축 △컴퓨팅 자원 네트워크와 데이터 보안 강화를 제시하였다.

<표 3-15> 중국 컴퓨팅 자원 상호연결 실행계획의 주요 내용

중점 방향	세부 내용
컴퓨팅 자원 상호연결 인프라 구축	• 첨단 컴퓨팅 산업의 핵심기술 연구 • 컴퓨팅 자원 상호연결 표준과 사양 개발
컴퓨팅 자원 상호연결 최적화	• 서버 기업이 주도해 노드 내 상호연결을 가속화 • 네트워크 간 상호연결 강화 및 국가 인터넷 백본과 신규 네트워크 구축 추진
컴퓨팅 자원의 상호 이용 촉진	• 통합 컴퓨팅 자원 식별 시스템을 구축해 컴퓨팅 자원을 할당 • 컴퓨팅 자원 인터페이스의 상호운용성 개선 • 국가 차원의 컴퓨팅 자원 상호연결 플랫폼 구축
컴퓨팅 자원 비즈니스 상호연결	• 응용 수요에 따른 컴퓨팅 자원 공급 및 데이터와 스토리지 상호운용성 개선 • 클라우드 네이티브 및 AI 기반 네트워크 개선으로 네트워크 통합 역량 향상
컴퓨팅 자원 상호연결을 위한 응용 시나리오 구축	• 컴퓨팅 시설과 자원, 애플리케이션 상호연결을 촉진하는 시범 사업을 진행 • 지능형 제조, 의료 등 다양한 분야에서 컴퓨팅 자원 상호연결 응용을 확대
컴퓨팅 자원 네트워크 및 데이터 보안 강화	• 컴퓨팅 자원 네트워크의 사이버보안 강화 • 컴퓨팅 자원 상호연결 과정에서 데이터 보안 강화

출처 : 컴퓨팅 자원 상호연결 실행계획(2025.5)

(2) AI 과학기술 윤리 관리서비스 방법

중국 공업정보화부는 2025년 8월 22일 국가발전개혁위원회, 과학기술부, 국가인터넷정보판공실 등과 공동으로 AI 과학기술 활동의 윤리적 거버넌스를 규범화하기 위한 ‘AI 과학기술 윤리 관리서비스 방법(시범 시행)’ 초안을 발표했다. 초안은 AI 과학기술 활동에 종사하는 고등교육기관과 과학 연구기관, 의료기관, 기업 등을 대상으로 하며, 생명 건강, 인간 존엄, 환경, 공공질서 등에 윤리적 위험을 초래할 수 있는 활동에 대한 윤리 심사 방침을 제시하였다.

AI 활동을 수행하는 각 기관은 AI 기술, 응용, 윤리, 법률 전문가로 구성된 과학기술 윤리 위원회를 설치하여야 하며, AI 활동 책임자가 해당 위원회에 활동 계획, 윤리 위험 평가와 예방 계획, 준수 서약서 등을 제출해야 한다. 위원회는 공정성, 통제 가능성과 신뢰 가능성, 투명성과 해석 가능성, 책임성 등에 초점을 두고 위험을 평가하여, 활동 승인 또는 수정 후 재심, 불허 결정을 내린다. 승인된 AI 과학기술 활동에 대해서는 추적 심사를 진행하며, 일반적으로 추적 심사 간격은 12개월을 초과하지 않는다.

3) 국가인터넷정보판공실

(1) AI 생성·합성 콘텐츠 표시 방법 제정

중국 국가인터넷정보판공실은 공업정보화부, 공안부, 국가라디오텔레비전총국과 함께 2025년 3월 7일 ‘AI 생성·합성 콘텐츠 표시 방법’을 제정하고 2025년 9월 1일부터 시행하였다. AI의 건전한 발전 촉진과 AI 생성·합성 콘텐츠의 식별 방법 표준화를 위한 규정으로, AI 생성물에 대하여 출처 표시 의무가 있는 인터넷정보 서비스 제공자를 적용 대상으로 한다. AI 생성·합성 콘텐츠는 AI 기술로 생성·합성된 텍스트, 이미지, 오디오, 동영상, 가상 장면 등의 콘텐츠를 의미한다.

출처 표시 유형은 텍스트·음성·이미지 등으로 사용자가 명확히 인식할 수 있는 ‘명시적 표시’와 파일 메타데이터에 콘텐츠 속성 정보, 서비스 제공자 명칭이나 코드, 콘텐츠 번호 등의 정보를 삽입하는 ‘묵시적 표시’로 구분된다. 서비스 제공자는 AI 생성·합성 콘텐츠의 메타데이터에 묵시적 표시를 추가하고, 이용 약관에 AI 생성·합성 콘텐츠의 표시 방식과 관리 정책을 명확히 기재하고 이용자에게 고지해야 한다.

온라인 콘텐츠 전송 서비스 제공자는 AI 생성·합성 콘텐츠 게재 시 파일 메타데이터에 묵시적 표기가 포함되었거나, 사용자가 AI 생성·합성 콘텐츠임을 밝힌 경우, 해당 콘텐츠가 AI로 생성 또는 합성되었음을 명확히 알려야 한다. 사용자는 AI 생성·합성 콘텐츠 생성 및 게재 시 이를 자발적으로 밝히고 적절한 표시 기능을 사용해야 하며, 표시의 삭제·변조·위조·은폐 행위, 이를 위한 도구나 서비스 제공 행위는 금지된다.

(2) AI 안전 거버넌스 프레임워크 2.0

중국 국가인터넷정보판공실은 2025년 9월 국가 사이버보안 선전 주간에 열린 포럼에서 ‘AI 안전 거버넌스 프레임워크 2.0’ 버전을 발표했다. 2024년 9월 1.0 버전을 발표한 이후 AI 기술의 급속한 발전을 반영해 위험 분류와 예방 및 관리 조치를 개선한 것으로, AI 혁신 촉진과 위험 억제에서 균형을 달성하기 위하여 기술·관리, 개발·안전, 국내외 노력을 통합하는 포괄적 거버넌스 시스템을 구축하는 것을 목적으로 한다. 프레임워크는 거버넌스의 5대 원칙으로 △포용적 혁신과 신중한 안전관리 △위험 기반 거버넌스 △기술적 조치와 관리적 조치의 결합 △국제협력을 위한 글로벌 정합성 중시 △신뢰할 수 있는 AI 실현으로

통제력 상실 방지를 제시하였다.

프레임워크는 AI 안전 위험을 △내재적 위험 △응용 위험 △파생 위험으로 분류하고 그에 따른 기술적 대책과 종합 거버넌스 조치를 제시하였다. 내재적 위험은 모델, 알고리즘, 데이터 등 AI 핵심기술 자체에서 비롯되는 위험을 뜻하며, 모델과 알고리즘 위험(설명 부족, 편향, 환각), 적대적 공격, 데이터 위험을 포괄한다. 응용 위험은 AI가 실제 시스템에 통합 및 배포되는 과정에서 발생하는 위험으로서 네트워크와 시스템 위험, 딥페이크나 허위 정보 등의 콘텐츠 위험, 인프라 장애나 대량 살상무기 지식 확산 등 현실적 위험, 정보 편식이나 여론조작과 같은 인지 위험을 포괄한다. 파생 위험은 AI의 광범위한 도입으로 인해 발생하는 사회적 위험으로 일자리 변화나 자원 소비 확대, 윤리 위험 등을 포함한다.

프레임워크는 내재적 위험 대책으로 AI의 설명 가능성과 투명성 개선, 인간의 감독체계 도입, 데이터 보호 규정 준수, 합법적 훈련 데이터 사용을 장려하였으며, 응용 위험 대책으로는 네트워크 시스템 접근제어 강화, 보안 취약점을 줄이기 위한 안전 규정 수립과 이행, 공급망 보안 강화, AI 모델 안전 가드레일 구축 등을 권고하였다. 또한 파생 위험 대책으로 친환경 AI 모델 및 컴퓨팅 기술, 기술 표준 개발을 모색하고 생명과 건강에 영향을 미치는 핵심 AI 시스템에 대한 비상 통제 대책 마련을 제시했다.

기술적 대책 마련과 함께 정부, 사회단체, 사용자 등 다양한 이해관계자가 참여하는 종합 AI 보안 거버넌스 체계 수립이 필요하다고 강조했다. AI 안전에 관한 법률 제정과 개선을 추진하고 AI 기술 관련 윤리 기준을 확립하며, AI 연구개발과 응용의 전체 수명주기에 걸쳐 보안 역량을 강화하도록 요구하였다. 또한 오픈소스 생태계와 공급망의 보안을 강화하고 AI 애플리케이션의 보안 위험을 체계적으로 평가해 위험에 비례한 안전 조치 마련을 제안했다.

(3) 행정 분야 대규모 AI 모델의 도입과 응용에 관한 지침

중국 국가인터넷정보판공실은 2025년 10월 10일 국가발전개혁위원회와 함께 ‘행정 분야 대규모 AI 모델의 도입과 응용에 관한 지침’을 발표했다. 이번 지침은 행정 분야에서 AI 모델의 활용을 체계적으로 규범화하여 정부의 디지털 전환을 가속화하고 행정 효율성과 공공서비스 품질을 향상하며 안전하고 질서 있는 AI 기술 도입을 촉진하고자 마련되었다. 지

침은 시나리오 기반의 AI 모델 적용을 강조하며 행정 분양서 공통적이고 빈번하게 발생하는 수요를 중심으로 아래와 같은 4개 시나리오를 제시하였다.

<표 3-16> 중국 행정 분야 AI 모델 도입을 위한 4개 중점 시나리오

중점 방향	세부 내용
행정 서비스	• (지능형 응답) 정부 데이터 기반의 온라인 행정 상담 서비스 제공 • (업무 처리 보조) 원스톱 행정 지원 서비스 제공 • (원스톱 정책 서비스) 시민과 기업 대상의 맞춤형 정책 매칭 서비스 구현
사회 관리	• (스마트 모니터링/순찰) 영상과 이미지, 센서 데이터 분석으로 실시간 모니터링 • (법 집행/감독 보조) 사건 보고서 작성 또는 유사 사건 검색 지원 • (시장 위험 예측) 시장 데이터 모니터링과 분석으로 잠재적 시장 위험 예측
사무 지원	• (초안 작성 보조) LLM 생성 능력을 활용한 문서 초안 작성 지원 • (자료 검색) 지식 그래프 구축과 검색 기술을 활용해 검색 효율성과 정확성 향상 • (지능형 작업 할당) 문서 업무 자동 식별과 분류, 우선순위 설정 등
정책 결정 지원	• (재해 조기경보) 위성, 센서 등 데이터 분석으로 자연재해 조기 예측 및 경보 발령 • (비상 대응) 강화학습 등의 기술을 활용해 비상 상황의 대응 속도와 효율성 개선 • (정책 평가) AI 대규모 모델을 활용해 시장 피드백 반영해 정책 최적화 • (지능형 심사 보조) 프로젝트 심사 효율성과 정확성 향상

출처 : 행정 분야 대규모 AI 모델의 도입과 응용 지침(2025.10)

지침은 또한 AI 모델의 표준화된 배포를 강조하고, 질의응답과 문서 작성 등의 일반적인 활용 사례에는 기존 상용 모델을 우선 활용하고, 법 집행과 같은 전문 영역에는 맞춤형 수직 모델 개발을 고려하는 등 다양한 행정 업무에 적합한 합리적인 배포 방식과 통합 관리를 구현할 것을 요구하였다.

제4절 일본

1. 그간의 인공지능 정책 추진 경과

<표 3-17> 일본의 주요 AI 정책 추진 현황

시기	발표기관	정책
2017년 3월	내각부	AI 산업화 로드맵
2019년 6월	내각부	AI 전략 2019
2021년 6월	내각부	AI 전략 2021
2022년 4월	내각부	AI 전략 2022
2023년	내각부	통합 이노베이션 전략 2023
2024년 4월	총무성·경제산업성 등	AI 사업자 가이드라인(1.0판)
2024년 7월	방위성	방위성 AI 활용 추진 기본방침
2025년 2월	내각부	AI 제도검토회의 중간보고서
2025년 5월	국회·내각부	AI 기본법 제정
2025년 9월	내각부	AI 기본계획 수립

일본 정부는 4차 산업혁명의 핵심 영역으로 일컬어지는 인공지능(AI) 분야의 주도권 확보를 위해 각국 간 경쟁이 치열해짐에 따라 AI 정책을 적극 수립, 추진하고 있다.

일본은 AI를 자국 성장 전략의 핵심으로 설정한 아베 정부가 2018년 9월 설치한 ‘통합혁신전략추진회의’에서 AI 종합 전략의 수립을 위해 ‘AI 전략 패키지’ 검토에 착수했으며, 2019년 6월 국가 AI 종합 전략인 ‘AI 전략 2019’를 발표했다. 2021년 6월에는 지금까지의 대응 진척 상황 및 향후 대응 방향을 정리해 ‘AI 전략 2021’로 개정, AI 대응을 지속·추진하였다.

2022년에 수립된 ‘AI 전략 2022’에서는 ‘AI의 사회 구현 추진’에 중점을 두고 2021년도 전략에 추가된 ‘임박한 위기 대응’인 코로나19 팬데믹, 대규모 자연재해 등과 같은 보다 명백해지는 많은 리스크 요인 등을 반영해 구체적인 목표 등을 검토, 대폭 개정하였고, 2022년 양대 전략 목표로 ▲사회 구현 내실화를 위한 새로운 목표 설정・추진 ▲팬데믹, 대규모 재해 등의 임박한 위기 대응책 구체화를 제시하였다.

2023년에는 최근 ‘챗GPT(ChatGPT)’ 등 AI 기술의 이용이 급속히 확산되면서 기업・행정 업무에서의 작업 효율화가 기대되는 한편, 저작권・개인정보 침해 등 부정적 우려도 지적되며, 관련 규제 방식에 대한 재검토가 필요한 상황이었다. 이에 2023년 5월 생성 AI를 포함한 AI에 대한 국가 전략을 수립하고 정책 방향성을 제시하는 AI 정책을 총괄하는 자문기구인 ‘AI 전략회의’를 신설했다.[64] 또한 현재의 생성 AI를 중심으로 한 AI 논점을 정리한 ‘AI에 관한 잠정적 논점 정리’[65]를 발표하고, 일본이 중점 추진해야 할 구체적인 AI 전략 방향성으로 ▲리스크 대응 ▲AI 이용 촉진 ▲AI 개발력 강화 등을 제시하고, AI 관련 경쟁력 강화를 위해서 해당 분야를 중점 지원해 나갈 방침이다.

2023년 5월 개최된 G7 히로시마 정상회의(5.19~21)에서는 생성 AI에 관한 국제적인 규칙 검토를 위한 ‘히로시마 AI 프로세스’를 출범시켜, AI의 개발・공표 전 리스크의 파악・경감, AI의 성능에 관한 정보 공유 등 다양한 국제 지침을 명문화했다.[66]

2024년에는 2월에 경제산업성 산하 IT 정책 실시기관인 정보처리추진기구(IPA) 내에 AI의 안전성을 검토하는 기관인 ‘AI 안전연구소(AISI)’를 설립하고, 내각부나 경제산업성 등과 협력해 AI의 안전성 평가 방법 검토 등을 실시하고 있다.

또한 ‘AI에 관한 잠정적 논점 정리’(2023.5)의 AI 전략 방향에 따라 가이드라인 등을 정비하기 위해 경제산업성・총무성의 합동 검토회, 의견 수렴 등을 거쳐, 기존 가이드라인들

64) 일본은 지금까지 ‘AI 전략 2022’, ‘인간 중심의 AI 사회 원칙’ 등 정부의 AI 기본 전략·이념을 제시해 왔지만, 최근의 생성 AI 등 기술 변화와 국제적인 논의에 대응하기 위해 AI 정책을 총괄하는 자문기구인 ‘AI 전략회의’를 신설해, 생성 AI ‘챗GPT’를 포함한 AI에 대한 국가 전략을 수립하고 정책 방향성을 제시하는 새로운 사령탑 역할 수행

65) 内閣府, AI に関する暫定的な論点整理, 2023.5.26
https://www8.cao.go.jp/cstp/ai/ronten_honbun.pdf

66) 経団連, AI政策・制度検討の動向, 2024.9.26
https://www.keidanren.or.jp/journal/times/2024/0926_08.html

을 통합·업데이트해, 광범위한 통일적이고 알기 쉬운 AI 사업자용 가이드라인을 검토, 2024년 4월에 ‘AI 사업자 가이드라인(제1.0판)’[67]을 정리하였다. 2024년 7월에는 방위성이 기술 발전에 따른 전투 방식의 변화에 대응하고 인구 감소에 따른 인적 효율화를 도모하기 위해 AI 활용을 추진하기 위한 기본방침을 처음으로 수립, 발표했다.

2024년 8월에는 AI 전략회의 산하에 ‘AI 제도연구회’를 설치해, 일본의 국가 AI 전략 방향을 검토하고, 2025년 2월에 ‘중간보고서’를 발표했다.

2025년에는 AI 기술의 연구개발 및 활용을 적극적으로 추진하여 산업 경쟁력을 강화하고 사회 문제를 해결하기 위한 일본 최초의 포괄적인 AI 기본법인 ‘AI 관련 기술의 연구개발 및 활용 추진에 관한 법률’을 공포, 시행하였다.[68] 이 법은 유럽연합(EU)의 AI법과 달리, 엄격한 처벌 조항을 두지 않고 자발적인 규제와 혁신을 유도하는 데 초점을 맞추고 있다.

또한 일본 정부는 「인공지능기본계획(초안)」[69](2025.9)을 마련하고, AI 활용과 개발을 국가 전략 차원에서 가속화해 경제성장과 사회혁신을 이끌어 내겠다는 비전을 제시했으며, 본 계획은 연내 수립 예정이다.

67) 経済産業省; 総務省, AI 事業者ガイドライン（第 1.0 版）, 2024.4.19
https://www.meti.go.jp/press/2024/04/20240419004/20240419004-1.pdf
68) 内閣府, AI法制（AI Act）に関する検討, 2024.3.25. https://www8.cao.go.jp/cstp/ai/ai_act/ai_act.html
69) 内閣府, AI戦略2025（案）, 2024. https://www8.cao.go.jp/cstp/ai/ai_plan/aiplan2025_draft1.pdf

2. 2024년 인공지능 정책 동향

1) AI 제도연구회, 국가 AI 정책·제도 방향 '중간보고서' 발표[70)]

(1) 개요

일본 정부가 AI 법규제의 기본 방향 등을 논의하는 전문가 회의인 'AI 제도연구회'를 'AI 전략회의' 산하에 설치했다. 연구회의 목적은 AI 기술의 진전과 그에 따른 사회적 영향을 고려해 적절한 법규제의 틀을 검토하는 것이다. 'AI 제도연구회'는 첫 회의를 2024년 8월 2일 수상관저에서 개최했으며, 전문가 논의와 의견 수렴 결과 등을 토대로 2025년 2월 국가 AI 정책·제도 방향에 대한 '중간보고서'를 발표했다.

(2) AI 제도의 기본 방침

AI 제도의 기본적인 방침으로 첫째, AI 혁신을 촉진하면서 적절히 위험에 대응해 나가는 것이 중요시되고 있다. 구체적으로는 ▲연구개발 지원, 인재 육성, 데이터 및 연산 자원의 정비 등을 통한 혁신 촉진 ▲법령과 가이드라인 등의 적절한 조합 ▲OECD 원칙, 히로시마 프로세스 국제 지침 등의 공통 지침과 개별 기존 법령 활용 등을 추진한다. 둘째, 글로벌 협력 추진을 통해 AI 거버넌스 형성을 향한 글로벌 논의를 주도하고, 국제적 정합성 및 상호 운용성을 확보한다.

(3) 구체적인 AI 제도·정책의 방향

구체적인 AI 제도·정책의 방향으로 ▲전반적인 사항 ▲정부의 활용 ▲기반 서비스 등에서의 활용으로 나눠 제시했다.

생성 AI와 같은 AI는 범용성이 높고 다양한 분야에서 이용되고 있으며, 위험에 대한 대응도 다양하다. 개인정보 및 저작권 처리, 허위·오정보 대응과 같은 위험에 대한 대응에서는 기존 법 등을 중심으로 한 대응이 전제되지만, AI는 횡단적인 대응이 필요한 경우도 있으므로, 전체를 조망하는 정부의 컨트롤 타워 기능 강화, 전략 수립, 또한 안전성 향상을 위한 투명성 및 적정성 확보 등이 요구되며, 필요에 따라 제도 정비가 필요하다. 전반적인 사항의 구체적 방향으로 ①정부의 컨트롤 타워 기능 강화 및 전략 수립 ②안전성 향상(AI 라이

70) 内閣府, AI制度研究会, 2024.2.4. https://www8.cao.go.jp/cstp/tougosenryaku/21kai/siryo.pdf

프사이클 전반에 걸친 투명성과 적정성 확보, 국내외 조직이 실천하는 안전성 평가 및 인증에 관한 전략적 촉진, 중대 사고 등에 관한 정부의 조사 및 정보 공개) 등을 추진한다.

일본에서 개인・기업의 AI 활용률은 타국과 비교해 현저히 저조한 상황이다. AI는 국민생활과 경제활동 발전의 기반으로서 그 활용의 중요성이 커질 것으로 보인다. 이러한 상황을 방치할 경우 일본의 국제 경쟁력이 약화될 우려가 있다. 따라서 우선 정부가 선도적으로 AI를 활용하고 국민의 활용을 촉진하는 방안이 고려될 수 있다. 정부 활용의 구체적 방향으로 ①정부 조달 ②정부・지자체의 AI 서비스 이용 등이 유용하다.

의료기기, 자율주행차, 기반 서비스 등 특히 국민 생활과 사회 활동에 미치는 영향이 큰 생명・신체 안전 및 시스템적 위험과 관련된 사항에 대해서는 특히 주의를 기울여 대응할 필요가 있다. 기반 서비스 등 활용의 구체적인 방향으로 ①업종별로 각 관할 부처가 기존 산업법을 바탕으로 대응하고, ②추가적인 대응이 필요한 사항에 대해서는 각 부처가 협력하여 대응해야 한다.

2) ‘AI의 이용・개발에 관한 계약 체크리스트’ 공개[71)]

(1) 개요

일본 경제산업성은 생성 AI의 보급을 비롯한 최근 시장 환경 변화를 감안하여 당사자 간 적절한 이익 및 위험 분배, 나아가 AI 활용 촉진을 목적으로 일본 사업자가 사용하기 쉬운 형식의 ‘AI 이용・개발에 관한 계약 체크리스트’(2025.2)를 마련했다.

사업 활동에서 AI 기반 서비스 활용을 검토하는 사업자가 급증하는 가운데, AI 기술이나 법률에 익숙하지 않은 사업자가 계약 실무에서 법적 위험을 충분히 검토하지 못하거나, 기밀 데이터나 정보가 제3자에게 제공될 우려가 제기되고 있다. 이에 경제산업성은 AI 기술이나 법률에 익숙하지 않은 일본 사업자들이 실무에서 쉽게 활용할 수 있도록 계약 체크리스트를 개발했다.

71) 経済産業省,「AIの利用・開発に関する契約チェックリスト」を取りまとめました, 2025.2.18
https://www.meti.go.jp/press/2024/02/20250218003/20250218003.html

⑵ 체크리스트의 주요 내용

체크리스트는 계약 당사자 간 이익과 위험을 적절히 분배하고 AI 활용을 촉진하기 위한 목적으로, 주로 사내 법무 부서나 고문 변호사의 계약 조항 검토나 사업 부문 담당자의 계약 초기 검토 시 활용하는 것을 상정하고, ▲AI 개발자 ▲AI 제공자 ▲AI 이용자로 대표되는 당사자 간 통상적인 계약 조건에서 확인이 필요한 사항을 정리하고 있다.

체크리스트는 AI 서비스의 사용자가 제공하는 입력(Input)과 서비스 제공자가 해당 입력을 통해 만들어 내는 산출물(Output)을 중심으로 구성되어 있다. 입력은 프롬프트와 학습용 원시 데이터를 포함하며, 산출물은 데이터 분석 결과를 비롯한 AI 생성물과 이를 가공한 결과물 등을 의미한다.

사용자의 입력 제공 시 입력의 정의와 제공 의무 및 조건, 입력한 정보의 이용 목적, 입력에 대한 관리 및 보안 체계, 보유 기간과 삭제 기간, 입력 정보의 제3자 제공 여부 등을 검토하고, 산출물 제공 시에는 산출물의 정의, 서비스 제공자의 산출물 완성 의무와 조건, 산출물의 이용 목적, 사용자에 의한 산출물의 제3자 제공 여부 등을 검토할 수 있다.

⑶ 유의사항

경제산업성은 체크리스트가 AI 사용 계약 체결 시 고려해야 할 사항을 제시할 뿐, AI 사용으로 인한 위험 발생 자체를 막을 수는 없으며 위험 가능성이 사업상 허용 가능한 범위에 있다면 계약 진행이 합리적일 수 있다고 설명한다.

체크리스트를 바탕으로 어떻게 대응하는 것이 적절한지는 개별 사용자의 구체적 상황에 따라 달라질 수 있으므로, 서비스 내용, 계약 형태, 계약상 의무의 이행 가능성 등을 종합적으로 고려해 판단할 것을 권고하고 있다.

3) ‘생성 AI의 조달・활용에 관한 가이드라인’ 수립[72)]

(1) 개요

일본 디지털청은 행정 업무에 생성 AI를 도입・활용하는 데 필요한 지침인 ‘행정의 진화와 혁신을 위한 생성 AI의 조달・활용에 관한 가이드라인’(2025.5)을 수립했다.

본 가이드라인은 생성 AI의 활용 촉진과 리스크 관리를 동시에 추진하기 위해 정부 내 생성 AI 거버넌스, 각 부처의 조달・활용 시 규칙을 규정하고 있으며, 경제산업성, 총무성 등 정부 각 부처가 각종 업무에서 생성 AI의 활용과 리스크 관리를 범정부적으로 추진하기 위해 마련되었다.

이에 따라 각 부처는 AI 총괄책임자(CAIO)를 설치해 생성 AI의 활용을 추진하고 거버넌스와 리스크 관리를 총괄해야 한다.

(2) 가이드라인의 주요 내용

가이드라인은 정부 내 생성 AI의 ▲추진・거버넌스 체제 구축 ▲조달・활용 규칙 등으로 구성되어 있다.

〈표 3-18〉 ‘생성 AI의 조달・활용에 관한 가이드라인’의 주요 내용

구분	주요 내용
생성 AI의 추진·거버넌스 체제 구축	• 각 부처에 신설되는 AI 총괄책임자(CAIO)가 생성 AI의 활용 파악·추진, 거버넌스, 리스크 관리를 총괄 • 비교적 고위험이 될 가능성이 있는 생성 AI의 이용이라도 ‘첨단 AI 활용 자문위원회’의 각 부처에 대한 조언과 상담 창구 등 시스템을 통해 안전하고 효과적인 AI 프로젝트 실시를 지원해, 첨단 생성 AI의 활용을 촉진
생성 AI의 조달·활용 규칙	• 각 부처의 생성 AI 시스템의 ①AI 총괄책임자(CAIO) ②기획자 ③제공자 ④이용자 등별로 규칙을 규정 – (CAIO) 각 부처의 이용자(직원)를 대상으로 생성 AI의 이용 규칙을 수립 – (기획자·제공자) 본 가이드라인의 ‘조달 체크 시트’ 및 ‘계약 체크 시트’*를 참고하여 사양서 작성 및 사업자와의 계약 등을 실시함으로써 안전하고 고품질의 생성 AI 시스템의 조달을 확보하고, 운용 개시 후에도 적절한 이용과 안전성·품질 확보를 정기적으로 검증

72) デジタル庁,「行政の進化と革新のための生成AIの調達・利活用に係るガイドライン」, 2025.5.27
https://www.digital.go.jp/assets/contents/node/basic_page/field_ref_resources/e2a06143-ed29-4f1d-9c31-0f06fca67afc/80419aea/20250527_resources_standard_guidelines_guideline_01.pdf

구분	주요 내용
	* ▲조달 체크 시트는 거버넌스, 개발·운용 프로세스 요건, 생성 AI 시스템 요건의 각 항목을 정리 ▲계약 체크 시트는 생성 AI의 조달 시 계약서에서 결정하는 항목을 정리 - (제공자·이용자) 리스크 발생 시 각 부처 CAIO에게 적절히 보고해 제공자가 필요한 대응을 하도록 하고, 또한 '첨단 AI 활용 자문위원회'는 각 사례에 대한 보고를 받고 필요에 따라 재발방지책 등을 검토

출처 : 総デジタル庁, 「行政の進化と革新のための生成AIの調達・利活用に係るガイドライン」, 2025.5.27.

4) 'AI 관련 기술의 연구개발 및 활용 추진에 관한 법률' 공포[73]

(1) 개요

일본 정부는 2025년 5월 28일 AI 개발 촉진과 안전 확보를 목표로 하는 일본 최초의 AI 기본법인 'AI 관련 기술의 연구개발 및 활용 추진에 관한 법률'을 제정하고, 6월 4일에 공포・시행했다.[74]

생성 AI를 비롯한 AI 기술의 발전은 국민 생활의 향상 및 국민 경제 발전에 기여하는 반면, 일본 내 AI 개발・활용은 뒤처져 있으며, 또한 많은 국민이 AI로 인해 발생하는 위험에 불안을 느끼고 있는 상황이다.

유럽과 미국에서는 AI 관련 법제도에 대한 논의와 검토가 활발히 진행되어 온 반면, 일본은 AI 기술의 급격한 변화와 리스크의 다양성을 감안하여 사업자 자율에 맡기는 가이드라인 등 법에 의존하지 않는 대응을 중심으로 진행되어 AI의 악용과 범죄에 대한 법적 대응을 강화해야 한다는 목소리가 고조되고 있었다.[75]

---

73) 参議院, 議案情報, 2025.5.28.
e-Gov, 人工知能関連技術の研究開発及び活用の推進に関する法律, 2025.6.4
https://laws.e-gov.go.jp/law/507AC0000000053

74) 법안 3장(AI 기본계획)과 4장(AI 전략본부)의 조항은 공포일로부터 3개월 이내 시행을 명시하고 있으며, 국제 동향 및 사회경제적 상황 변화를 감안하면서 AI 법안의 시행 상황을 검토하여 필요하다고 인정할 때에는 필요한 조치를 취할 수 있도록 규정

75) 유럽은 2024년 5월에 AI법(Artificial Intelligence Act)을 제정해 이른바 법적 구속력이 있는 하드 로우(Hard Law) 접근방식을 채택하고 있는 반면, 일본은 2024년 4월에 총무성·경제산업성에서 「AI 사업자 가이드라인」을 공표하는 등 이른바 법적 구속력이 느슨한 소프트 로우(Soft Law) 접근방식을 채택

이에 따라 일본 정부는 2024년 8월부터 내각부 산하에 설치된 ‘AI 제도연구회’를 통해 AI 제도와 관련된 법 정비의 필요성에 대한 논의를 시작하고, 논의 결과를 바탕으로 관련 법안을 제출(2025.2), 2025년 5월에 법안이 참의원을 통과했다.

일본 정부가 AI 법제화 필요성에 따라 AI 전략 방향성을 명확히 한 것으로, 규제 강화에 중점을 둔 유럽만큼 엄격하지는 않지만, 정부가 조사 권한을 가지고 AI를 둘러싼 개발 환경과 발전 속도에 적극 대응하겠다는 의지를 표명하며, 6월 4일에 AI법을 공포・일부 시행한 데 이어, 9월 1일에는 AI 전략본부 설치 관련 규정 등을 포함해 전면 시행하였다.

(2) 법안의 주요 내용

동 법안은 AI 관련 기술의 연구개발과 활용 시의 기본이념, 기본 시책, AI 기본계획 수립, AI 전략본부 신설 등을 규정하고 있다.

<표 3-19> 일본 AI 추진법의 개요

구분	주요 내용
목적	• 국민생활 향상, 국민경제 발전
기본이념	• 경제사회 및 안보상 중요 → 연구개발력 유지, 국제 경쟁력 향상 • 기초연구부터 활용까지 종합적·계획적으로 추진 • 적정한 연구개발·활용을 위한 투명성 확보 • 국제협력에서 주도적 역할 수행
AI 전략본부	• 본부장: 내각총리대신, 구성원: 전 국무대신 • 관계 행정기관 등에 필요한 협력 요청
AI 기본계획	• 연구개발·활용 추진을 위해 정부가 시행해야 할 시책의 기본 방침 등
기본 시책	• 연구개발 추진, 시설 등의 정비 및 공동이용 촉진 • 인력 확보 • 교육 진흥 • 국제적인 규범 수립에 참여 • 적정성 확보를 위한 국제규범에 부합하는 지침 정비 • 정보 수집, 권익 침해 사안의 분석 및 대책 검토, 조사 • 사업자·국민에 대한 지도·조언·정보 제공
책무	• 국가, 지방자치단체, 연구개발기관, 사업자, 국민의 책무 • 관계자 간 협력 강화 • 사업자는 국가 등의 시책에 협력
부칙	• 재검토 규정(필요한 경우 필요한 조치)

출처: 内閣府, 人工知能関連技術の研究開発及び活用の推進に関する法律案律（AI法）の概要, 2025.6.4

기본이념은 AI 관련 기술의 연구개발과 활용은 경제사회 발전의 기반이자 국가 안보 면에서도 중요하므로 AI 관련 연구개발 능력 유지와 국제 경쟁력 향상을 기본 취지로 삼을 것을 요구하고 있다. 또한 부정한 목적 또는 부적절한 방법으로 AI 관련 기술이 연구・개발 또는 활용되면 범죄에 이용되거나 개인정보 유출, 저작권 침해 등 국민 생활의 안정과 권익을 침해할 수 있다는 점에서 투명성 확보를 강조하고 있다.

기본 시책은 AI 기술의 연구개발, 시설・장비의 정비나 공유 촉진, 인재 확보, 교육 진흥, 국제협력을 규정하고 있다. AI 관련 기술의 기초 연구에서 실용 연구에 이르기까지 일관된 연구개발 추진 및 성과 이전을 위한 체제 정비 등의 시책을 마련하고, 데이터센터나 데이터셋의 공유를 활성화할 수 있는 조치도 추진한다. 또한 지방자치단체, 연구기관, 사업자와 긴밀히 협력해 기초 연구와 활용 등 각 단계에 필요한 다양한 인재 양성 및 자질 향상에 필요한 시책과, 국민의 AI 이해도를 높이기 위한 AI 관련 기술 교육과 학습 진흥안을 개발한다.

AI 기본계획은 기본이념과 기본 시책을 바탕으로 AI 관련 기술의 연구개발 및 활용 추진 시책에 관한 기본 방침과 종합적・계획적으로 마련해야 할 시책을 규정하고 있다. 내각총리대신은 AI 전략본부가 작성한 AI 기본계획안에 대하여 국무회의의 의결을 요청하여야 하며, 의결이 이루어지면 즉시 AI 기본계획을 공표하도록 규정하고 있다.

AI 전략본부 신설에 대해서는 내각에 설치해 AI 기본계획안의 작성 및 실시 등을 관장하며, 내각총리대신이 본부장, 내각 관방장관 및 AI 전략 담당 대신이 부본부장, 그 외의 국무대신이 구성원을 담당하도록 하고 있다. AI 전략본부는 AI 기본계획의 초안 작성과 이행 촉진에 관한 사항, 기타 AI 관련 기술의 연구개발과 활용 추진 관련 시책에서 중요 사안의 기획과 입안, 포괄적 조정 업무를 수행한다.

5) ‘AI를 활용한 무기의 연구개발 가이드라인’ 발표[76)]

(1) 개요

일본 방위성이 2025년 6월 6일 무기의 연구개발 시 AI 활용 방침을 안내하는 ‘장비품 등

76) 防衛装備庁, 装備品等の研究開発における責任あるＡＩ適用ガイドライン, 2025.6.6
https://www.mod.go.jp/atla/soubiseisaku_ai_guideline.html

의 연구개발에 있어 책임 있는 AI 적용 가이드라인’ 을 발표했다. 이 지침은 일본 방위성이 2024년 7월 발표한 ‘AI 활용을 위한 기본 방침’[77]을 토대로 무기 관련 연구개발 사업의 계획 수립과 시행에서 AI의 위험을 관리하고 이점을 극대화하기 위한 프레임워크를 제공하고 있다.

책임 있는 AI 무기 등의 연구개발을 위해서는 국제법 및 국내법을 준수하고, ‘자율형 치사 무기 시스템’(LAWS)이나 ‘AI와 자율성의 책임 있는 군사적 이용에 관한 정치 선언’으로 대표되는 국제적 논의와 조화를 이루는 방식으로 추진할 필요가 있다. 이를 통해 연구개발에 참여하는 사업자 등에 대한 예측 가능성을 확보하고, 무기 등에 대한 AI 활용을 강화한다.

(2) 가이드라인의 주요 내용

‘AI 무기 등의 연구개발에 관한 기본 요구사항’ 에서는 AI 무기 등이 준수해야 할 요건을 설정하고, 법적・정책적 요건과 기술적 요건의 개요를 제시하고 있고, ‘AI 무기 등의 연구개발에 관한 실시 사항’ 에서는 AI 무기 등의 분류, 법적・정책적 심사, 기술적 심사라는 3가지 개념을 제시하고, 심사 항목의 기준과 심사 체계를 제시하고 있다.

기본 요구사항은 국제인도법의 원칙이 모든 무기에 적용되어야 한다는 방침에 따라 법적 요건으로 AI를 활용한 무기 개발 시 국제법과 국내법의 준수를, 기술적 요건으로 인간의 참여와 책임을 요구하고 있다.

- (법적·정책적 요건) 국제인도법을 포함한 국제법과 국내법을 준수해야 하며, 인간의 개입 없이 AI가 자동으로 목표를 설정해 공격하는 자율살상무기(LAWS)의 연구개발을 금지
- (기술적 요건) AI 시스템 활용 시 인간의 참여와 통제를 가능하게 하고, 편향을 방지하고 검증 가능성과 투명성, 신뢰성을 보장하며, 오작동과 심각한 장애 발생 위험을 완화하는 체계를 마련해 안전성을 확보

77) ▲목표 탐지·식별 ▲정보 수집·분석 ▲지휘통제 ▲후방 지원 ▲무인기(드론) 등 무인 장비 ▲사이버 안보 ▲사무처리의 7개 핵심 분야를 중심으로 AI 활용과 인간 관여의 필요성을 명기
https://www.mod.go.jp/j/press/news/2024/07/02a_03.pdf

실시 사항은 AI 무기의 연구개발 시 ▲AI 무기의 분류 ▲법적·정책적 심사 ▲기술적 심사의 3단계 심사를 거쳐 사업을 추진하며, 법적·정책적 심사와 기술적 심사는 고위험 무기에만 진행한다.

- (AI 무기의 분류) 자율살상무기 관점에서 집중적으로 위험을 관리해야 할 무기는 고위험으로, 그 외의 AI 무기는 저위험으로 분류해 차등화된 관리를 실시[78]
- (법적·정책적 심사) 고위험으로 분류된 AI 무기를 대상으로 국제인도법을 포함한 국제법과 국내법 준수를 보장할 수 있는지, 인간의 개입이 필요 없는 자율살상무기가 아닌지를 검토
- (기술적 심사) 고위험 AI 무기 시제품이 실제 배치와 운용 단계에서 법적·정책적 요건을 충족할 수 있는 기능을 갖추고 있으며, 적절한 위험 완화 대책이 마련되어 있는지를 기술적 관점에서 확인

6) ‘AI 기본계획(초안)’ 수립[79]

(1) 개요

일본 정부의 ‘AI 기본계획(초안)’(2025.9)은 일본이 ‘세계에서 가장 AI를 개발·활용하기 쉬운 국가’가 되는 것을 목표로, AI 기술의 급속한 진보와 국제 경쟁 격화에 대응하기 위한 국가 전략을 제시하고 있다.

일본의 현 상황을 보면, AI 활용이 충분히 진전되지 못하고 있고 투자 규모도 경제 규모에 비해 적다는 문제의식이 있다. 이러한 인식을 바탕으로, ‘반전 공세’의 기회를 포착해 민관이 하나가 되어 AI 혁신을 추진하겠다는 강력한 의지를 표명하고 있다.

<표 3-20> 생성 AI의 이용 상황 변화와 AI에 대한 민간투자액의 변화

구분	2023년	2024년
생성 AI의 이용 상황 변화	• 개인 생성 AI 서비스 이용 경험 - 중국(56.3%), 미국(46.3%), 독일(34.6%), **일본(9.1%)** • 기업 업무에서의 생성 AI 활용률 - 미국(84.7%), 중국(84.4%), 독일(72.7%), **일본(46.8%)**	• 개인 생성 AI 서비스 이용 경험 - 중국(81.2%), 미국(68.8%), 독일(59.2%), **일본(26.7%)** • 기업 업무에서의 생성 AI 활용률 - 미국(90.6%), 중국(95.8%), 독일(90.3%), **일본(55.2%)**

78) 저위험 AI 무기는 법적·정책적 심사와 기술적 심사를 면제하고 사업 수행 담당자의 자체 점검을 중심으로 위험관리를 진행

79) 内閣府, 人工知能基本計画骨子（たたき台）, 2025.9.12
https://www8.cao.go.jp/cstp/ai/ai_plan/aiplan2025_draft1.pdf
日本DX地域創生応援団, 日本政府「人工知能基本計画骨子（案）」(2025.12.23. 방문)
https://digital-supporter.net/japan-ai-vision

구분	2023년	2024년
AI에 대한 민간투자액의 변화	• 1위: 미국(약 672억 달러) • 2위: 중국(약 78억 달러) • 3위: 영국(약 38억 달러) • 9위: 한국(약 14억 달러) • **12위: 일본(약 7억 달러)** • 13위: 아랍에미리트(약 4억 달러)	• 1위: 미국(약 1,091억 달러) • 2위: 중국(약 93억 달러) • 3위: 영국(약 45억 달러) • 8위: 아랍에미리트(약 18억 달러) • 11위: 한국(약 13억 달러) • **14위: 일본(약 9억 달러)**

출처 : 内閣府, 人工知能基本計画の骨子（たたき台）の概要について, 2025.9.12

본 계획안은 전문조사회의 검토를 거쳐 2025년 내 목표로 ‘AI 전략본부’ 회의 보고 후 확정될 계획이다.

(2) 기본 방침

기본 방침은 ‘AI 관련 기술의 연구개발 및 활용 추진에 관한 법률’의 기본이념을 바탕으로 3원칙과 4대 기본 방침을 제시하였다.

3원칙은 다음과 같다.

- 인간 중심의 AI 사회 원칙: 인간과 AI가 협력하여 인간이 창의력을 발휘할 수 있는 사회 실현
- 민첩한 대응: 변화에 즉각 대응하며 유연하고 신속하게 PDCA 사이클을 운영
- 국내외 정책 연계 대응: 국내 정책과 대외 정책을 유기적으로 결합하고, 적극적인 국제 협력을 통해 일본을 다양한 AI 혁신의 거점으로 조성

4대 기본 방침은 다음과 같다.

- AI 활용의 가속화(AI를 사용): 일본 사회 전체가 세계 최첨단 AI 기술을 적극적으로 활용하여 새로운 혁신을 창출하고, 데이터의 집적·활용, 특히 조직을 초월한 데이터 공유를 촉진
- AI 개발 역량의 전략적 강화(AI를 창조): 인프라부터 애플리케이션까지 AI 생태계 각 계층에서의 개발을 추진하여 「신뢰할 수 있는 AI」를 개발하고, 연구개발부터 사회 구현까지의 선순환을 실현
- AI 거버넌스의 주도(AI의 신뢰성 제고): 인간과 AI가 협업하여 혁신의 선순환을 실현하기 위해 AI의 적정성을 확보하는 거버넌스를 구축하고, 국제적 거버넌스를 일본이 주도
- AI 사회를 향한 지속적 변혁(AI와 협업): 인간과 AI의 협업을 위해 산업과 고용의 형태, 제도와 사회 구조를 선도적이며 지속적으로 변혁하고, AI 인재 양성·확보에 더해 창의력 등 ‘인간 고유의 능력’ 향상을 도모

(3) 강구해야 할 시책

AI 관련 기술의 연구개발 및 활용 추진에 대한 4대 기본 방침에 따라 정부가 종합적·계획적으로 강구해야 할 시책도 제시하고 있다.

<표 3-21> 'AI 기본계획(초안)'의 4대 기본 방침과 강구해야 할 시책

구분	주요 내용
AI 활용의 가속화	• 정부·지자체에서의 철저한 AI 활용 ① 정부 AI 추진, 정부에 의한 AI의 적정하고 선도적인 활용 ② 생성 AI 등을 적절히 활용한 정부 업무의 합리화·효율화 촉진 ③ 지자체에서의 AI 적정 활용 촉진 • 사회 과제 해결을 위한 AI 활용 추진 ① 의료·헬스케어, 요양, 금융, 교육, 방재, 환경 보전, 농림수산업, 식품 산업, 인프라 건설·관리, 조선·선박 산업 등에서의 AI 에이전트 및 피지컬 AI 등의 개발·실증·도입 촉진 ② 중소기업을 포함한 광범위한 지역 기반 산업에 대한 AI 도입 촉진 ③ 방위력의 근본적 강화를 위한 AI 활용 추진 ④ 국민의 안전·안심 확보를 위한 경찰 활동 고도화를 위한 AI 활용 추진 ⑤ 정보통신 분야의 보안 확보를 위한 AI 활용 추진 • AI 활용 촉진을 통한 새로운 사업 및 산업 창출 ① 피지컬 AI의 선도적 도입 지원 ② 산학 연구자 등에 대한 AI 활용 지원 ③ AI 관련 혁신 기술을 보유한 스타트업 등 지원 • AI 활용 확대를 위한 체계 구축 ① AI의 사회적 구현을 위한 규제·제도 점검 및 재검토 ② 통계 작성 등으로 분류 가능한 AI 개발 등에서의 본인 동의 방식 및 규정 준수 실효성 확보 등을 검토한 '개인정보보호법' 개정
AI 개발 역량의 전략적 강화	• 일본 내 AI 개발 역량 강화 ① 새로운 데이터셋, AI 개발에 필요한 멀티모달 데이터의 창출·제공 등 데이터 연계 기반 구축 ② 최고급 인재를 포함한 국내외 AI 개발자 확보를 위한 처우 및 생활환경 개선 등 종합대책 마련 ③ 선진적 지식을 도입하기 위한 대학·연구기관 및 국내외 민간 사업자 등과의 연계·협력 추진 ④ AI 모델의 고성능화 및 멀티모달화 촉진 ⑤ AI 평가 기반 및 테스트베드 구축 ⑥ 글로벌 사우스를 포함한 AI 산업의 해외 시장 진출 지원 • 일본의 경쟁 우위 AI 모델 등 개발 추진 ① AI 로봇의 공공 수요 창출을 포함한 피지컬 AI 개발·실증 촉진 ② 기반 모델 개발·활용, 연구 데이터 창출·활용 고효율화, 정보 기반 강화 등을 통한 AI for Science 추진 ③ 신약 개발 고도화를 위한 신약 개발 AI 추진 • 신뢰할 수 있는 AI 기반 모델 개발 ① 일본의 문화·습관 등을 고려한 신뢰할 수 있는 AI 개발·평가, 고품질 일본어 데이터 구축·확충(기존 축적 데이터 활용 포함) ② 일본이 신뢰할 수 있는 AI 개발 거점이 되어 오픈형 AI 모델을 포함한 생태계를 세계로 확산 • AI 개발·활용 기반의 증강·확보

구분	주요 내용
	① AI 데이터센터, 효율적인 전력·통신 인프라 구축(와트·비트 연계), 전광(All-Optical) 네트워크, 차세대 정보통신 기반 연구개발 ② 고성능 AI 반도체 등의 연구 및 개발, 후가쿠 NEXT 개발
AI 거버넌스의 주도	• 신뢰할 수 있는 AI 생태계 구축 ① AI법 제16조에 따른 조사 연구 실시 ② 사업자 등의 AI 연구 및 개발·활용에 있어 적정성 확보를 위한 자율적 노력을 촉진하기 위한 AI법 제13조에 따른 지침 및 각종 가이드라인 등 마련·철저한 홍보 ③ AI 관련 사이버 사건 대응 능력 향상 등, AI를 악용한 사이버 공격 및 사기를 비롯한 각종 범죄 등에 대한 대응 ④ AI 모델의 적정성 평가 기능 구축을 포함한 AI 안전연구소(AISI)의 근본적 강화 ⑤ 일본의 AI 평가 기능 향상에도 기여하는, AI 생성 콘텐츠 판별 기술 및 AI 제어 기능 등의 개발 지원 • 글로벌 사우스를 포함한 국제 협력 ① 히로시마 AI 프로세스 프렌즈 그룹 및 외교 기회를 활용한 글로벌 사우스 등과의 협력 강화 ② AI 관련 국제 규격 제정(ISO/IEC JTC1에서의 AI 분야 국제 표준화 활동 참여 등) ③ AISI 네트워크 등을 활용한 AI 거버넌스 주도 ④ 군사 분야 AI에 대해 인도적 고려와 안보 관점을 고려한 균형 잡힌 논의를 통한 국제적 논의에 적극 참여 ⑤ GPIA(Global Partnership on AI) 도쿄 전문가 지원 센터 등을 통한 프로젝트 기반 지원 ⑥ 동아시아·아세안경제연구센터(ERIA)를 활용한 글로벌 사우스와의 협력모델 구축
AI 사회를 향한 지속적 변혁	• AI를 기축으로 한 산업 구조 구축 ① 일본 기업 등에 AI 기축의 조직 경영 개혁(AI 트랜스포메이션)을 촉진하기 위해 해당 노력의 가시화, 노력이 진전되는 사업자에 대한 지원의 중점화 ② 구체적 모델 제시를 포함하여 지역 AI 인프라를 활용한 새로운 지역 산업·고용 창출 ③ 규제 샌드박스 제도, 스타트업 지원 제도 등을 포함한 AI 관련 산업의 입지 촉진 • AI 사회에서의 프레임워크 검토·실증 ① AI의 사회 구현을 위한 규제·제도 점검·재검토 ② AI 활용 시 민사 책임 등 방향성 검토 ③ 적절한 재산 보호와 활용으로 이어지는 투명성 확보 ④ AI에 의한 고용 영향(대체성과 보완성) 조사·분석, 이를 바탕으로 한 포괄적 대책 지속적 시행 • AI 인재 육성·확보 ① AI 및 차세대 반도체 등 활용·개발 관련 엔지니어, 연구자, 데이터 관리 인재 등 육성·확보 ② AI 활용·개발 관련 산관학 네트워크 및 커뮤니티 지원 ③ 종업원 및 근로자 대상 AI 재교육 지원 ④ 고급 필수 근로자 창출을 위한 재교육 지원 ⑤ 디지털 스킬 표준 개정 ⑥ 초등·중등 교육 및 일반 시민 대상 AI 리터러시 향상 지원 • AI 시대에서의 인간 역량 향상 지원 ① 인간과 AI가 협업하는 사회에서 인간과 AI의 역할 분담을 지속적으로 모색하고, 인간이 인간으로서 살아가는 역량 육성 ② AI 시대에서의 교육 추진 ③ AI 시대의 일하는 방식 검토

출처 : 内閣府, 人工知能基本計画骨子（たたき台), 2025.9.12.

제5절 영국

1. 그간의 인공지능 정책 추진 경과

<표 3-22> 영국의 주요 AI 정책 추진 현황

시기	발표기관	정책
2021년 9월	디지털문화미디어스포츠부(DCMS)	국가 AI 전략 - 10년 비전 제시
2022년 7월	정부과학실·DCMS	국가 AI 전략: AI 실행 계획 발표
2022년 9월	DCMS	AI 규제안(AI Rulebook) 공개
2023년 3월	과학혁신기술부(DSIT)	AI 규제 백서: AI 규제에 대한 친혁신적 접근법 발표
2023년 8월	영국 하원	AI 거버넌스 중간보고서 발표
2023년 11월	영국 정부	제1차 AI Safety Summit(Bletchley Summit) 개최 - 블레츨리 선언 주도, Frontier AI Taskforce 출범
2023년 11월	DSIT	AI안전연구소(AI보안연구소로 명칭·기능 변경) 설립 발표
2024년 1월	내각사무처·DSIT	중앙정부 부문을 위한 생성형 AI 프레임워크 공개
2024년 4월	경쟁시장청(CMA)	AI 기반 모델 시장 분석 보고서 업데이트
2024년 5월	영국 하원	AI 거버넌스 최종보고서 발표
2024년 11월	DSIT·교육부	AI 보증(AI assurance) 시장 분석 보고서 발표
2025년 1월	DSIT	AI 기회 실행계획 발표
2025년 2월	내각부	AI 플레이북 공개
2025년 5월	DSIT	AI 성장 구역(AIGZ) 요건·절차 공개
2025년 6월	교육부·DSIT·총리실	디지털·AI 역량 강화 위한 테크퍼스트(TechFirst) 정책 발표
2025년 6월	교통부	교통 AI 액션플랜 제시
2025년 7월	DSIT	AI 연구 자원(AIRR) 가이드 발표
2025년 7월	DSIT	영국 컴퓨팅 로드맵 공개
2025년 7월	법무부	사법체계를 위한 AI 액션플랜 발표
2025년 9월	금융감독청(FCA)	AI와 FCA: 우리의 접근 보고서 공개
2025년 9월	교통부	교통부 과학·혁신·기술 계획 2025 발표
2025년 11월	DSIT	AI 성장 구역 신청·운영 지침 발표
2025년 11월	DSIT	과학 전략을 위한 AI 정책 문서 공개

영국 정부는 2021년 10월, AI를 국가의 미래 경쟁력을 좌우할 핵심기술로 규정하며 글로벌 AI 선도국으로 도약하겠다는 비전을 공식화했다. 이에 따라 정부는 최초의 국가 차원 AI 비전인 ‘국가 AI 전략(National AI Strategy)’을 발표하고 2022년 12월 18일 영국 정부 공식 사이트(GOV.UK)를 통해 상세 자료를 공개했다.[80]

이 전략은 향후 10년 동안 영국이 AI 분야에서 초강대국 수준의 리더십을 구축하는 것을 목표로 한다. 이를 위해 정부는 정책 과제를 세 가지 큰 축으로 ①AI 생태계에 대한 지속적·장기적 투자, ②산업과 지역 전반으로의 기술 확산 추진, ③효율적이고 신뢰 가능한 거버넌스 체계 마련으로 구분하고, 각각을 단기(3개월 이내), 중기(6~12개월), 장기(12개월 이후) 단계로 나누어 실행 로드맵을 제시했다.

<표 3-23> 영국 국가 AI 전략(2021)의 부문별 정책 과제

구분	장기적 투자	산업·지역 확산	효과적 거버넌스
단기 (향후 3개월)	• 데이터 접근성과 활용 기반 강화 • 국가 사이버-물리 인프라 프레임워크 구축 • AI, 데이터사이언스, 디지털 기술 개발 지원 확대	• 보건 및 사회복지 분야 AI 기반 기술에 대한 국가 전략 개발 • 국방 AI에 대한 전략 수립 • 저작권 및 지식재산 분야 AI 연구 강화	• 데이터윤리혁신센터(CDEI) 인증 로드맵 준비 • AI 거버넌스를 위한 데이터베이스 구축 시작 • 국방부 AI 지침 개발 추진 • 국제 AI 활동 분야 범정부적 접근방식 개발
중기 (6~12개월)	• 민간기업 AI 자금 수요와 과제에 대한 평가 • 국립컴퓨팅교육센터 지원 • AI 관련 경력 개발 지원 • 미국과 AI R&D 협력 • 글로벌 인재 영입 추진	• 경제 및 산업 전반에서 AI 기회 연구 강화 • 에너지 등 주요 산업 AI 적용 확대 • 개발도상국의 기술 개발 지원	• AI 거버넌스 및 규제 국가 백서 발간 • 알고리즘 투명성 제고 위한 범정부 표준 개발 • AI 표준 허브에 대한 시범 운영 • 정부와 공공 분야 AI 안전 인식 제고
장기 (12개월 이후)	• 반도체 공급망에 대한 검토 • AI 모델 교육을 위한 정부 데이터셋 개발 • AI 연구혁신 프로그램 확대 • AI 인재 육성 및 투자 유치 • AI 무역 및 투자 거래 규정의 개발	• AI 사무국, UKRI 프로그램 출범 및 운영 • AI 신뢰성, 채택 가능성 등 안전에 대한 강화 • 정부기관간 협력 통해 전략 과제 개발	• AI 기술 표준 참여 툴깃 개발 • 공통 R&D 과제에 대한 글로벌 파트너십 확대 • 공공 부문 AI 윤리 및 안전 지침 개정 • 국가안보 및 국방 부문의 전문가와의 협력 확대

출처 : National AI Strategy(2021.10.), SPRi 정리

80) Gov.uk, National AI Strategy, 2022.12.18.

영국의 국가 AI 전략은 영국이 가진 강점을 토대로 민간과 공공 전반에 걸쳐 복원력과 생산성, 성장과 혁신을 강화하는 AI의 잠재력을 극대화하여 영국에서 AI를 통한 단계적 변화를 주도하겠다는 비전을 제시하고 있다.

영국 정부는 2022년 7월 '국가 AI 전략 - 실행계획(National AI Strategy - AI Action Plan)' 을 발표했다. 이 실행계획은 국가 AI 전략 발표 이후 1년간 성과를 점검하고, 전략 과제의 이행을 촉진시키기 위한 것으로서 '국가 AI 전략' 의 3대 핵심 영역별로 전략 발표 이전과 이후를 구분해 평가하고, 각 영역에서 향후 12개월간 강화할 중점 과제들을 명확히 제시했다.[81)]

<표 3-24> 영국 AI 실행 계획 주요 내용

영역	Before(2021년)	After(2022년)	향후 12개월(2023년)
AI 생태계의 장기적 투자	• 연구 역량/인재는 강점 • 데이터·컴퓨팅 접근성 부족 • 상용화 전환 경로 부족 • 인재 유치·유지 필요	• 데이터 가용성 프레임워크, Future of Compute 리뷰, NHS TRE 지원 • AI·데이터 석사(2,000명) • 글로벌 인재 비자 제도 개선, 미·영 공동 연구 선언	• 국가컴퓨팅 전략 • 컴퓨팅 인프라·용량 강화 • AI 연구 자원(AIRR) 확대
산업·지역 전반 확산	• 일부 산업 AI 활용 편중 • 중소기업·지역은 접근 제한 • 공공 활용은 초기 단계	• 상용화 경로 연구, IP 제도 개선 • Innovate UK(3,500만 파운드) • 국방·HMRC AI 전략 • 교통·농업·에너지 등 부처별 AI 적용 확대	• 중소기업·지역 지원 도구 강화 • 공공 조달 확대로 초기시장 창출 • 보건·교통·에너지 부문 대형 AI 프로젝트 추진
효과적 거버넌스	• AI 규제·윤리 논의는 분산되어 통합 기준 부재 • 국제 규범 형성 과정에서 영국 영향력 강화 필요	• 알고리즘 투명성 표준 파일럿 완료 • CDEI AI Assurance 로드맵 • AI Standards Hub 출범 • Pro-innovation 규제 접근 초안, Defence AI Centre • G7 협력, GPAI 투자, OECD, UNESCO 활동	• 국제 기술 전략 발표 • AI 규제 백서 발표 • 국제협력 확대, 다자 규범 형성 참여 강화

출처 : National AI Strategy - AI Action Plan(2022.7.18.), SPRi 정리

81) Gov.uk, National AI Strategy - AI Action Plan, 2022.7.18.

또한 영국 정부는 2022년 7월, 혁신을 활성화하면서도 국민 신뢰와 데이터 보호를 강화하기 위한 방향성을 담은 ‘AI 규제안(AI Rulebook)’을 공개했다. 이 규제안은 AI를 책임 있게 활용하도록 유도함과 동시에, 기업이 감당해야 하는 규제 준수 부담을 완화하는 조치들을 포함하고 있다.[82)]

이어 2023년 3월에는 과학혁신기술부(DSIT: Department for Science, Innovation and Technology) 장관 명의로 ‘친혁신적 AI 규제 접근(A pro-innovation approach to AI regulation)’ 백서가 발표되었는데, 이 문서는 AI의 발전을 촉진하면서도 AI가 초래할 수 있는 위험을 효과적으로 관리하는 방안을 제시했다. 특히, 위험과 편익의 균형을 고려하여 일정 수준의 위험을 수용하는 대신 혁신을 가속할 수 있는 유연하고 비례적인 규제 프레임워크를 제안한 점이 특징이다.[83)]

<표 3-25> 영국 AI 규제에 대한 친혁신 프레임워크 핵심 요소

핵심 요소	주요 내용
규제기관 조정 기반 마련 - AI 특성에 따른 정의 설정	• AI를 효과적으로 규제하기 위해서는 규제기관 간 명확한 역할 정의와 공통된 개념 정립이 필요하다고 판단하여, AI의 기술적 특성을 반영한 적응성(adaptivity), 자율성(autonomy)을 핵심 기준으로 제시함 • 이러한 기준을 바탕으로 규제체계를 조정하고 기관 간 협업 구조를 강화
상황 중심의 적용 방식 구축	• 일률적 규제 적용이 아니라, 기술이 사용되는 환경·목적·맥락에 따라 규제 강도와 적용 방식이 달라질 수 있도록 상황별 접근 모델을 제안 • 규제가 AI 활용에 과도한 부담이 되지 않도록, 규제 준수 비용과 수행 가능성을 함께 고려하는 방향성을 강조
부문 간 공통 원칙(cross-sectoral principles) 제시	• ▲안전·보안·견고성 ▲투명성·설명 가능성 ▲공정성 ▲책임·거버넌스 ▲이의 제기·구제의 5개 범부문 원칙을 마련하여 다양한 규제기관들이 공통 기준으로 활용할 수 있도록 조치
규제의 일관성과 안정성 보장	• 규제기관이 원칙을 적절히 고려하도록 새로운 의무를 도입하고 규제기관의 권한을 강화하여 규제 환경 전반에 걸쳐 일관성을 유지 • 규제기관, 국제 파트너, 업계, 노동조합과 시민사회, 학계, 일반 대중을 포함한 광범위한 이해관계자들이 AI 규제를 보다 잘 이해하고 협업하도록 하는 방안 촉진 • 규제 체계 운영 상황을 중앙에서 모니터링하고 평가, 평가·검토·정책 지원·국제 협력 등 후속 관리 기능을 제도화

출처 : A pro-innovation approach to AI regulation(2023.3.), SPRi 정리

82) Gov.uk, Establishing a pro-innovation approach to regulating AI, 2022.7.18.
83) Gov.uk, A pro-innovation approach to AI regulation, 2025.8.3.(업데이트)

영국 정부는 백서 발표 이후 12주에 걸쳐 백서에 대한 공개 의견 수렴을 진행했으며, 수렴된 의견들을 반영하여 2024년 2월 정부 ‘친혁신적 규제 접근: 정부 답변(A pro-innovation approach to AI regulation: Government Response)’ 문서를 의회에 제출했다. 이 문서에 백서의 5대 원칙 ▲중앙 기능 가동 ▲리스크 등록부 추진 ▲규제자 역량 강화 ▲정부 거버넌스 개편 등에 대한 추진 방침이 담겼다.

<표 3-26> 영국 친혁신적 AI 규제 프레임워크의 5대 원칙

원칙	주요 내용
안전·보안·견고성 (Safety, security & robustness)	• AI 시스템은 안전하게 설계·운영되어야 하며, 외부 위협에 대해 강인하고 신뢰성 있게 작동해야 함
투명성과 설명 가능성 (Appropriate transparency & explainability)	• AI 작동 방식과 의사결정 과정은 이해관계자들이 납득할 수 있게 적절히 공개되고 설명할 수 있어야 함
공정성(Fairness)	• AI는 차별적 편향 없이 공정하게 설계·활용되어야 하며, 사회적 불평등을 심화시키지 않아야 함
책임성과 거버넌스 (Accountability & governance)	• AI 개발자·제공자·사용자는 결과에 책임을 지며, 명확한 감독 체계와 관리 구조로 운영해야 함
이의제기 및 구제 가능 (Contestability & redress)	• AI로 불리한 영향을 받은 개인이나 집단은 이의를 제기할 수 있어야 하고, 구제 절차를 보장받아야 함

출처 : A pro-innovation approach to AI regulation(2023.3.), SPRi 정리

2023년 8월 영국 하원의 과학혁신기술위원회(Science, Innovation and Technology Committee)는 AI의 급속한 발전 속도에 비해 AI 백서의 규제 대응이 늦다는 지적을 담은 ‘AI 거버넌스 중간보고서’를 발표했다.[84] 이 보고서는 AI 백서의 원칙을 법률화하는 ‘AI 법안(AI Bill)’ 도입 필요성을 강조하고, AI 정책을 설계할 때 검토해야 할 쟁점들로 ▲규제기관 역량 강화 ▲AI 샌드박스 구축 ▲중앙의 조정기능 강화 ▲중소기업 AI 채택 지원 ▲글로벌 협력·기준 주도 등을 제시했다. 이러한 내용들은 2024년 5월에 추가로 발표된 ‘AI 거버넌스 최종 보고서’[85]에 포함되었다.

84) UK Parliament, The governance of artificial intelligence: interim report, 2023.8.31.
85) UK Parliament, Third Report - Governance of artificial intelligence (AI), 2024.5.28.

영국은 2023년 11월 ‘AI 안전 정상회의(AI Safety Summit)’를 개최하여 글로벌 AI 위험 논의를 위한 글로벌 협력을 시작했다. 이 행사에는 미국, 중국, 캐나다, 독일, 인도, 대한민국을 비롯한 28개국과 EU가 참여했으며, 영국은 이 자리에서 AI 안전연구소(AISI: AI Safety Institute) 설립을 공식 발표하고, 국제 차원의 AI 안전기구 설립을 제안하는 등 글로벌 거버넌스 논의의 중심적 역할을 수행했다.

이어 2024년 1월 영국 내각부는 정부 조직이 생성 AI를 보다 책임감 있고 안전하게 활용하도록 지원하는 지침인 ‘생성 AI 프레임워크(Generative AI Framework for HMG)’를 발표했다. 이 프레임워크는 공공 부문이 생성형 AI의 특성과 위험을 이해하고 적절히 도입할 수 있도록 설계된 것으로, 정부 기관이 반드시 고려해야 할 10가지 공통 원칙을 제시하여 책임 있는 생성형 AI 활용 기준을 마련하고자 했다.[86)]

<표 3-27> 영국 정부 조직을 대상으로 하는 생성형 AI 프레임워크의 10대 원칙

구분	주요 내용
원칙 1	생성형 AI의 개념과 그 한계·제약을 정확히 파악한다.
원칙 2	생성형 AI 기술을 사용할 때 법적·윤리적 책임을 준수하며 적절하게 활용한다.
원칙 3	생성형 AI 도구를 안전하게 관리하고 유지하는 절차를 숙지한다.
원칙 4	AI 적용 과정에서 인간의 판단과 통제가 필요한 지점을 분명히 인식한다.
원칙 5	전체 생성 AI 활용 과정(수명주기)을 체계적으로 관리하는 방식을 이해한다.
원칙 6	수행하려는 작업에 맞는 적절한 도구와 기능을 선택하여 사용한다.
원칙 7	지속적으로 협력하고 열린 태도로 생성형 AI를 도입·운영한다.
원칙 8	초기 단계부터 민간 기업과의 협업 구조를 마련하여 활용성을 높인다.
원칙 9	생성형 AI 시스템 구축·운영에 필요한 기술적 전문성을 확보한다.
원칙 10	기관의 정책과 연계하여 해당 원칙을 실천하고 필요한 보완조치를 실시한다.

출처 : Generative AI Framework for HMG(2024.1.), SPRi 정리

86) Gov.uk, Generative AI framework for HM Government, 2024.1.18.

영국 정부가 발표한 생성형 AI 프레임워크는 생성형 AI가 공공 부문에서 활용될 수 있는 구체적인 업무 사례를 제시하고 있다. 여기에는 디지털 문의 응대 지원, 문서・서신・통화 기록 분석을 통한 시민 요구 파악, 고도화된 검색 기능, 복잡한 문서 요약, 문서 초안 및 서신 작성 자동화, 소프트웨어 개발 보조, 텍스트・오디오 요약 등이 포함된다. 동시에 프레임워크는 생성형 AI의 구조적 한계도 명확히 지적하며, 부정확한 판단과 잘못된 추론 가능성, 환각 현상(hallucination), 편향적 또는 공격적 결과물 생성, 투명성 부족 등을 위험 요소로 제시하였다. 또한 정부는 생성 AI를 조달할 때 고려해야 할 지침도 제안했는데, 공급업체 요구사항을 명확히 문서화하고, 데이터 품질 관리 및 편향 완화 조치, 데이터 윤리 원칙 준수, 특정 공급업체에 종속되지 않도록 하는 전략, 그리고 지속적인 기술 지원・유지관리 필요성을 강조하였다.

2024년 8월 28일에는 DSIT와 교육부가 공동으로 교원의 업무 부담을 줄이기 위한 ‘콘텐츠 스토어(Content Store)’ 프로젝트에 400만 파운드(약 70억 원)를 투입한다고 발표했다. 이 프로젝트는 정부의 커리큘럼 지침, 수업 자료, 익명화된 학생 평가 데이터를 통합하여 AI 기업이 활용할 수 있는 학습 세트를 구축하고, 이를 기반으로 학교 현장에서 안정적으로 사용할 수 있는 맞춤형 고품질 교육 콘텐츠를 개발하는 것이 핵심 목표다. 콘텐츠 스토어는 과제 채점・교재 제작・행정 자동화 등 교육 현장의 다양한 작업을 지원할 수 있는 AI 기업을 주요 대상으로 하며, 학부모들이 “교사가 학생과 대면하는 시간 확보를 위해 생성 AI를 적극 활용하길 바란다”는 의견을 제시한 점을 반영하였다.[87]

2024년 11월 6일 DSIT는 AI 보증(AI assurance) 시장 분석 보고서를 발표하며 중소기업과 스타트업이 AI 관련 위험을 평가・관리할 수 있도록 돕는 중앙 집중형 AI 보증 플랫폼을 구축하겠다는 계획을 내놓았다. 보고서에 따르면 영국 내 AI 보증 관련 제품・서비스 제공 기업은 약 524개이며, 이 중 전문 AI 보증 기업은 84개로 추산된다. 관련 산업의 연 매출은 약 10억 1천만 파운드(약 1조 8,210억 원), 종사자는 약 12,572명에 이르는 것으로 나타났다. 경제 규모 대비 영국의 AI 보증 시장은 미국・독일・프랑스보다 큰 수준이며, 현재의 수요・공급은 잠재력에 비해 아직 제한적이지만, 정부는 적절한 성장 기반이 마련될 경우 2035년까지 시장 규모가 65억 3천만 파운드(약 11조 7,750억 원) 이상으로 확대될 것으로 전망하고 있다.[88]

87) Gov.uk, Teachers to get more trustworthy AI tech, helping them mark homework and save time, 2024.8.28.

한편 영국 정부는 2023년 AI 안전 정상회의에서 공식적으로 설립을 발표한 영국 AI안전연구소(AISI)를 중심으로 글로벌 안전 프레임워크 구축을 추진하고 있다. 연구소는 2024년 5월 10일, 범세계적 안전 평가를 지원하기 위한 AI 안전성 테스트 플랫폼 '인스펙트(Inspect)'를 오픈소스로 공개했다. AISI는 이 플랫폼을 통해 다양한 연구 주체가 보다 용이하게 안전성 테스트를 수행하고 상호 협력할 수 있을 것이라고 설명했다. 또한 2024년 5월 20일에 발표한 제4차 경과 보고서를 통해 지금까지 공개된 5개 모델을 대상으로 ▲사이버 공격 위험도 ▲화학·생물학적 악용 가능성 ▲자율성 수준 ▲안전장치의 적절성 등을 평가한 결과를 공개했다. 모델명은 밝히지 않았으나, 일부 모델은 화학·생물 분야에서 전문가에 근접한 수준의 지식을 보유한 것으로 나타났다.

2024년 4월 11일 영국 경쟁시장청(CMA: Competition and Markets Authority)은 AI 기반 모델(foundation models)에 대한 업데이트 보고서를 발표했다. CMA는 이미 2023년 9월 1차 분석을 통해 기반모델의 혁신적 가치와 잠재 위험을 검토한 바 있으며, 이번 보고서에서는 이해관계자의 의견과 기술 발전 현황을 반영해 내용을 보완했다. CMA는 보고서를 통해 기반모델이 소비자·기업·국가 경제에 긍정적으로 기여하기 위해 따라야 할 AI 원칙을 제안하며, 시장 경쟁을 유지하면서 기술 발전을 촉진할 수 있는 규제 방향을 제시했다.[89]

<표 3-28> 영국 경쟁시장청(CMA)의 AI 이행 원칙

구분	주요 내용
접근성	기업과 개발자가 데이터·컴퓨팅 자원·전문 인력·필수 기술에 안정적으로 접근할 수 있도록 환경을 보장한다.
다양성	수요자와 기업이 필요에 맞춰 여러 종류의 AI 모델 또는 방식을 선택할 수 있는 다양성을 확보한다.
선택 가능성	자체 구축, 파트너십 활용, API 및 플러그인 연동 등 여러 선택 경로를 제공하여 도입 방식의 유연성과 선택권을 강화한다.
공정성	제품·서비스의 품질에 따른 정당한 경쟁 환경을 유지하며, 경쟁을 저해하는 불공정 행위를 금지한다.
투명성	소비자와 기업이 기반 모델의 위험성과 한계를 이해할 수 있도록 필수 정보 공개를 보장한다.
책임성	기반 모델 개발자와 배포자는 해당 모델이 초래하는 결과와 영향에 대한 책임을 부담한다.

출처 : AI Foundation Models: Update paper(2024.4.), SPRi 정리

88) Gov.uk, Assuring a Responsible Future for AI, 2025.11.6.
89) Gov.uk, AI Foundation Models: Update paper, 2024.4.17.

2. 2025년 인공지능 정책 동향

1) 과학혁신기술부(DSIT)

(1) AI 기회 실행계획[90]

영국 과학혁신기술부(DSIT)는 2025년 1월 13일 AI를 통한 국가 성장과 공공 서비스 혁신을 구체화하는 것을 목표로 하는 ‘AI 기회 실행계획(AI Opportunities Action Plan)’을 발표했다. AI 기회 실행계획을 발표한 이유는 2021년 발표한 AI 국가 전략과 2022년 발표한 AI 국가전략 실행계획을 통해서 큰 틀에서 도입과 활용을 위한 비전과 원칙을 제시한 이후에 실제로 산업·공공 분야에서 실행을 촉진할 필요가 있기 때문이다.

‘AI 기회 실행계획’은 기존에 발표된 국가 전략을 정책 실행 중심으로 전환한 것으로서, ▲AI 활성화를 위한 기반 마련 ▲AI를 통한 생활 변화 ▲AI를 통한 미래 보장 등 세 가지 정책 영역으로 구성되어 있으며, 총 50개 AI 도입과 활용 활성화를 위한 권고사항을 제시하고 있다.

우선 AI 활성 기반 마련은 세계적 수준의 컴퓨팅 및 데이터 인프라 확보, 인재·혁신가 유치, AI 신뢰성·안전 확보 방안들을 제시하고 있다. 두 번째로 AI를 통한 생활 변화는 민간 협력을 통해 공공 부문 AI를 신속하게 탐색(Scan), 시범 적용(Pilot), 확장(Scale)하는데 초점을 두고 있다. 세 번째로 AI를 통한 미래 보장은 영국을 AI 소비자가 아닌 생산자로 위치시키고 핵심적인 AI 분야에서 국가 챔피언을 육성할 수 있는 혁신 지원 방안을 제시했다.

DSIT는 ‘AI 기회 실행계획’ 발표 당일 ‘AI 기회 실행계획: 정부 응답(AI Opportunities Action Plan Government Response)’ 자료를 공개하고 실행계획에서 제시된 권고사항을 전반적으로 수용할 의사를 밝혔다.[91]

90) Gov.uk, AI Opportunities Action Plan, 2025.1.13.
91) Gov.uk, AI Opportunities Action Plan: government response, 2025.1.13.

<표 3-29> 영국 AI 기회 실행계획과 정부 답변

구분	권고사항	정부 답변(시기)
AI 활성화 기반 마련 (인프라)	AI 인프라 수요 장기 계획 마련	컴퓨팅 전략 및 10개년 로드맵 제시('25, 봄)
	2030년까지 AI 연구 자원 20배 확충	실행계획 수립, 최소 2배 이상 확대('25~'31)
	국가컴퓨팅 자원 전략적 배분	미션 중심으로 자원 배분 계획 제시('25, 봄)
	AI 성장 특구(AI Growth Zones) 조성	AIGZ 조성지 확정, 추가 후보지 선정('25. 봄)
	컴퓨팅 솔루션 제공자로서 기회 선점	컴퓨팅 전략에 관련 내용 포함('25, 봄)
	국제 컴퓨팅 협력 체계 구축	컴퓨팅 전략에 관련 내용 포함('25, 봄)
AI 활성화 기반 마련 (데이터)	최소 5개 공공 데이터셋 발굴	국가 데이터 라이브러리(NDL) 구축 포괄적 데이터 접근 정책 수립 권고사항 추진 방법 검토 세부 내용 추후 발표('25 여름)
	수집이 필요한 데이터를 전략적 설계	
	AI 공공 데이터셋 공개 지침/사례 개발	
	컴퓨팅 자원·데이터셋 접근권 패키지 제공	
	고부가가치 데이터셋 구축 투자	
	인센티브 체계 및 보상 제도 마련	
	학습 데이터셋, 국제 라이선스 체계 확립	외부 협력을 통해 정부 역할 검토 ('25 봄)
AI 활성화 기반 마련 (인재)	AI 인력 격차의 정확한 규모 파악	기업·교육 기관·노동조합·정부 협력조사('25 봄)
	AI 전공 졸업생수 확대, 맞춤형 교육 지원	학생청 및 이해관계자 협력('27 가을)
	인재풀의 다양성 확대	성별 불균형 해소 계획 수립('26 가을)
	AI 분야로의 교육 경로 다양화	기업·교육 기관·노동조합·정부 협력('26 가을)
	AI 장학 프로그램 신설로 영국 유학 지원	권위 있는 신규 장학제도 마련('26 가을)
	기술 교육 프로그램의 AI 대응 역량 확보	교육과정 평가 체계 재검토 작업('25 가을)
	최상위 수준의 헤드헌팅 역량 구축	권고 실현 방안 검토('26 봄)
	기존 이민 제도 활용 방안 모색	경쟁력 있는 비자 제도 운용('25 여름)
	튜링 AI 펠로우십 프로그램 확대	UKRI와 협력해 AI 펠로우십 확대('26 가을)
AI 활성화 기반 마련 (신뢰성/안전)	'AI 안전연구소'의 역량 유지/확장 지원	연구소 예산 확정 및 법정 기구화('25 봄)
	텍스트·데이터 마이닝 제도 경쟁력 제고	권고 관련 공개 협의 절차 착수('24)
	규제기관의 AI 역량 확충 재정 지원	위험 대응 및 역량 화보 방안 수립('25 봄)
	소관 부처 안전한 AI 혁신 촉진 관점 포함	규제혁신사무국 주도로 지침 제공('25 봄)
	AI 도입 가속화/규제 샌드박스 도입	고성장 잠재력 분야 식별, 전략 개발('25 봄)
	규제기관의 연례 보고서 발간 의무 부여	활동 보고, 취합 내용 종합 보고('25 봄)
	AI 보증 생태계 지원	연구 유지/추가 고려('26 봄)
	앨런 튜링 연구소의 역량 극대화	권고 동의('25 가을 업데이트)
AI 통한 생활 변화(공공 AI 탐색)	AI 활용 가능 영역 식별	권고 동의('25 가을 업데이트, 이후 지속 추진)
	AI 기술/활용 사례 이해와 전문성 강화	권고에 동의('25 가을 업데이트)
	AI 발전 예측 및 공공 부문 수요 제시	권고에 동의('25 여름 업데이트)
AI 통한 생활 변화(공공 AI 파일럿)	일관된 프레임워크를 활용과 지속 개선	AI 조달 위한 프레임워크 개발('25 여름)
	신속한 프로토타이핑 역량 구축	권고에 동의 : 정부 디지털 센터 착수 시
	외부 AI 인재 채용 위한 특화 지원 체계	AI 및 GDS 활동 확대('25 가을 업데이트)

구분	권고사항	정부 답변(시기)
	데이터 중심 실험 환경 조성	AI의 실험 환경 기반 기능 강화('25 업데이트)
	신속하고 유연한 AI 조달 프로세스 도입	AI 조달 절차 개선 방안 검토/이행 ('25 가을
AI 통한 생활 변화(확장)	고위급 지원 및 중앙의 자금 지원	시범 사업 확산 지원 서비스 검토('25 가을)
	미션 중심 국가 AI 공모 시행	공모 방안 검토 ('25 가을 업데이트)
	확장할 수 있는 기술 스택 개발·조달	성공 모델 검토로 방안 모색('25 가을 업데이트)
	인프라 상호운용·코드 재사용, 오픈소스 의무화	권고 동의('25 가을 업데이트)
AI 통한 생활 변화(공공·민간 협력)	시장 형성하는 주체로서 전략적 조달 시행	전략적 조달 시행('25 가을 업데이트)
	디지털 정부 인프라 활용	권고에 동의 : 기회 확대('25 가을 업데이트)
	'AI 지식 허브'를 구축	AI 지식 허브 시범 사업 추진('25 여름)
	'정부 디지털 센터' 구축	단기성과 창출을 위한 과제 이행('25 여름)
AI 통한 생활 변화(장애물 해소)	AI 도입 지원 공동 노력 기회 창출	재무부, 기업 통상부 등과 기회 발굴('25 봄)
	AI 부문 챔피언을 임명	업계 리더가 AI 도입 주도 역할 발굴('25, 여름)
	국가 전역에 걸친 AI 도입 촉진	지역 정부와 협력해 기회 발굴('25 여름)
AI 통한 미래 보장(AI 국가 챔피언 육성)	'영국 소버린 AI' 기관 신설	권고에 동의 : 신규 기능 신설

출처 : AI Opportunities Action Plan Government Response(2025.1.), SPRi 정리

(2) AI 연구 자원(AIRR) 가이드[92]

DSIT는 2025년 7월 17일에 연구자・산업계・학계가 활용할 수 있는 AI 슈퍼컴퓨팅 자원의 구조와 이용 방법 등을 명확히 제시한 'AI 연구 자원(AIRR: AI Research Resource) 가이드'를 발표했다.

이 가이드는 DSIT와 영국연구혁신기구(UKRI)[93]가 공동 작성한 것으로, 영국의 AI 특화 슈퍼컴퓨팅 인프라의 구조, 목적, 이용 절차 등을 설명하고 있으며, AIRR을 AI에 특화된 고성능 컴퓨팅(capacity and infrastructure) 자원 모음으로 규정하고, 연구자・학계・산업계가 AI 모델을 학습・실험할 수 있는 컴퓨팅 자원 및 저장 공간, 소프트웨어 지원, 전문가 지원 등을 포함한 포괄적 기술 기반 인프라로 정의하고 있다.

92) Gov.uk, AIRR advanced supercomputers for the UK, 2025.7.17.
93) UK Research and Innovation(UKRI) : 2018년 설립된 공공 연구·혁신 지원 기관으로서, 영국의 7대 연구위원회(AHRC, BBSRC, ESRC, EPSRC, MRC, NERC, STFC)와 Innovate UK, Research England 등을 통합하여 만들어졌고, 연간 수십억 파운드 규모의 연구 자금을 관리

AIRR은 연구자나 기업이 AIRR 포털을 통해 신청해 사용할 수 있고 프로젝트 개요, 필요한 연산자원 규모, 보조금 규정 등에 따라 접근방식을 ‘게이트웨이 방식’과 ‘신속 접근 방식’으로 구분하고 있다.

<표 3-30> AIRR 이용 가이드 주요 내용

구분	세부	설명
시스템	AIRR	• 브리스톨대학(Isambard-AI)과 케임브리지대학(Dawn)의 두 개의 슈퍼컴퓨터 클러스터로 구성(suite of advanced computers). • DSIT, UKRI가 공동 운영·지원하며, HPE, Nvidia, Intel, Dell이 파트너로 참여
접근방식	게이트웨이 방식 (Gateway route)	• 영국 등록 조직(학계, 산업계, 중소기업 등)을 대상으로 초기 사용자가 코드 및 알고리즘 테스트, 워크플로 벤치마킹 등에 활용 • 3개월 내 최대 10,000 GPU 시간 제공
	신속 접근 방식 (Rapid Access route)	• 마이크로·중소 기업(MSME)을 대상으로 시장 진입 전 단계의 실현 가능성 검토나 실험적 개발(experimental development) 등에 활용 • 프로젝트 시작 후 3개월 내 최대 20,000 GPU 시간 제공
향후 계획		• 2030년까지 AIRR 용량을 최소 20배 이상 확충 • 2030년까지 10억 파운드 추가 투자(이중 Isambard-AI와 Dawn 클러스터에 약 3억 5천만 파운드 이상 투자)

출처 : AIRR advanced supercomputers for the UK(2025.7.), SPRi 정리

(3) 영국 컴퓨팅 로드맵

DSIT와 영국연구혁신기구(UKRI)는 2025년 7월 17일 세계적 수준의 컴퓨팅 생태계 구축을 통한 국가 AI 역량 강화를 목적으로 ‘영국 컴퓨팅 로드맵(UK Compute Roadmap)’을 발표했다. 이 로드맵은 컴퓨팅 파워를 에너지·교통 등과 같은 핵심 기반 시설로 간주하고 ▲현대적 공공 컴퓨팅 생태계 구축 ▲컴퓨팅 활용 경제 전반 혁신 촉진 ▲최첨단 AI 인프라 구축 ▲주권적이고 안전하며 지속가능한 역량 확보 등 4대 전략을 제시했다. 또한 목표 달성을 위해 2030년까지 공공 투자를 단계적으로 확대하고, 국립 슈퍼컴퓨팅 센터(NSC) 설립, AIRR의 대대적인 확충 추진 등을 발표했다.

특히 영국 정부는 이번 로드맵에서 AI 전용 대형 인프라와 산업 지원 프로그램을 결합하여 지역 거점 단위의 클러스터 활성화와 민간투자를 활성화하는 방안에 중점을 두었다. 구체적으로는 2030년까지 최대 20억 파운드의 공공 투자를 통해 세계적 수준의 컴퓨팅 생태계를 구축하고 AIRR을 2030년까지 20배 가량 늘리기 위해 10억 파운드 이상을 투자한다고 밝혔는데, 실제로 에든버러에 첫 번째 국립 슈퍼컴퓨팅 센터를 설립하여 지역·분야의 핵심

거점 컴퓨팅 센터로 기능하도록 했다. DSIT의 지역 기반 인프라 정책 프로그램인 ‘AI 성장 구역(AIGZ: AI Growth Zone)’ 프로그램과 접목시켜 거점별 대형 AI 컴퓨팅 인프라와 지원 생태계를 구축하는 방안도 제시했다.

(4) AI 성장 구역

‘AI 성장 구역(AIGZ)’은 AI 산업 성장을 영국 전역으로 확산하기 위해 특정 지역에 대규모 컴퓨팅 자원과 산업·학계 협력 생태계를 집중적으로 구축하는 전략으로, DSIT는 2025년 5월 AIGZ의 구체적 신청 요건과 절차를 공개했다. 이 전략은 특정 지역에 첨단 AI 연구와 기업 혁신이 가능한 물리적·디지털 기반 시설을 마련해 지역 균형 발전을 촉진하는 동시에, 해당 인프라를 기반으로 기업, 대학, 스타트업이 협력하는 허브 조성을 동반 추진함으로써 지역별 특화 산업과 연계해 AI 기술을 적용·상용화해 경제적 파급효과를 창출하고자 한다.

특히 DSIT는 AI 성장 구역을 발표하면서 데이터센터 및 지원 인프라 투자 촉진과 접근성 개선을 통해 영국의 혁신 역량과 고숙련 일자리 창출을 효과를 유도하는 데 중점을 두고 있다. 이를 위해, 영국 전역에 걸쳐 지방·지역 당국과 산업계(데이터센터 개발사, 에너지 기업)를 대상으로 ▲기술적 타당성 ▲이행 가능성 ▲지역적 영향 ▲정부 지원 요청 등 4개 항목을 투자 평가 기준으로 제시했다.

AIGZ의 구체적인 실행 전략은 2025년 6월 ‘관심 표명(EOI)’ 공고와 정식 신청 지침을 발표한 뒤, 11월 정책 패키지와 함께 컬렉션 형태로 재정비한 ‘AI 성장 구역 신청·운영 지침(AI Growth Zones - Collection&Application Guidance)’을 발표하면서 좀 더 상세하게 드러났다.[94]

이 지침은 AI 성장 구역 정책을 실제로 집행하기 위해 지방·지역 당국과 민간 사업자가 따라야 할 신청 절차와 선정 기준을 체계화한 실행용 가이드로서, AI 성장 구역이 AI 데이터센터와 연계 인프라에 대한 투자를 촉진하고, 전력 접근성·계획 인허가를 개선해 혁신을 가속한다는 정책 목표를 전제로, 어떤 지역이 성장 구역으로 적합한지 판단하기 위한 자격

94) Gov.uk, Collection - AI Growth Zones, 2025.11.13.

요건과 증빙 수준을 상세히 제시했다.

또한 이번 정책 패키지에는 ‘AI 성장 구역의 제공(Delivering AI Growth Zones)’이라는 제목의 정책 보고서가 포함되었는데, 이 보고서에서는 영국 전역에 AI 산업 클러스터를 조성하여 국가 AI 경쟁력을 지역 단위로 확장하기 위한 전략을 제시하고 있다. 대규모 데이터센터, 전력·냉각·네트워크 등 AI 인프라가 집적된 지역을 AI 성장 지역으로 지정하고, 기업·대학·연구기관을 중심으로 고도 AI 연구·개발이 지속적으로 이루어질 수 있는 물리적 기반을 제공하는 데 초점을 설정하고 있다.

특히 전력망 연결 지연, 토지 이용 규제, 대규모 설비 인허가 등 AI 인프라 확충을 가로막던 구조적 장애를 제거하는 데 중점을 둔다. 이를 위해 정부는 지방정부와 민간 사업자를 대상으로 기술 타당성, 인프라 이행 가능성, 지역적 파급효과 등을 기준으로 성장 지구 지정을 평가하고, 전력 지원·계획 인허가·투자 유치 패키지 등 차별화된 지원 프로그램을 운영하기로 했다.

(5) 과학 전략을 위한 AI[95)]

‘과학 전략을 위한 AI(AI for Science Strategy)’는 AI를 과학 연구의 핵심 기반으로 삼아 국가 연구 경쟁력을 대폭 강화하기 위한 전략으로, 2025년 11월 DSIT가 정책 보고서 형태로 발표했다. 이 전략은 AI를 활용해 연구 속도를 혁신적으로 단축하고, 대규모 실험·시뮬레이션을 자동화하며, 기초과학과 응용과학 전반에서 발견·발명 과정을 가속하는 것을 목표로 한다. 특히 국가 슈퍼컴퓨터(AIRR 컴퓨트 포함), 데이터 공유 인프라, 연구자 기술 역량 강화를 핵심 축으로 삼아 공공 연구기관·대학·산업계가 공동으로 활용가능한 과학용 AI 생태계를 구축하겠다는 비전을 제시했다.

또한 이 전략은 AI 기반 과학 연구의 확산이 가져올 위험을 최소화하기 위해 연구 무결성, 데이터 품질, 안전한 모델 사용 기준 등 거버넌스 요소도 병행 강조한다. 정부는 연구자들이 AI를 안전하고 투명하게 활용할 수 있도록 평가·검증 체계를 마련하고, 국제 공동 연구와 컴퓨팅 협력 네트워크를 강화하며, 장기적으로는 영국을 AI 기반 과학 혁신의 글로벌 허브로 도약시키는 것을 주요 성과 지표로 삼고 있다.

95) Gov.uk, Policy Paper – AI for Science Strategy, 2025.11.20.

2) 내각부

(1) AI 플레이북(AI Playbook)[96]

2025년 2월 영국 정부의 디지털 서비스 혁신을 총괄하는 중앙 조직으로 영국 내각부(Cabinet Office) 산하에 있는 정부 디지털 서비스(GDS: Government Digital Service)가 공공부문의 AI 활용 촉진을 위한 종합 지침인 ‘AI 플레이북(AI Playbook)’을 공개했다.

이 플레이북은 2024년 발표된 생성 AI 프레임워크를 더욱 확장한 것으로서, 생성 AI를 넘어 다양한 AI 기술 전반으로 확장된 것이 특징이다. 여기서 내각부는 AI를 안전하고 책임감 있게 적용하기 위한 10대 원칙과 ▲AI 정의와 적용 분야 ▲비즈니스 사례 ▲사회적 웰빙 증진 ▲보안 기회 ▲거버넌스 구조 ▲품질 보증 및 위험관리 등에 대한 가이드라인을 제시했다. 또한 정부 조직이 AI를 도입할 때 다양한 학제 팀을 구성해 데이터・보안・윤리・거버넌스를 아우르는 전문성을 확보할 것을 권고하며, 의사결정자의 이해 제고와 리더십 교육도 강조했다. 더불어 공정성, 투명성, 설명 가능성, 인간 감독 유지 등을 통해 신뢰성 있는 AI 운용을 보장해야 한다는 점도 명시했다.

3) 교육부

(1) 테크퍼스트(TechFirst)[97]

2025년 6월 교육부는 DSIT 및 총리실과 공동으로 청년과 노동자의 디지털・AI 역량 강화를 위한 ‘테크퍼스트(TechFirst)’ 정책을 발표했다. 총 1억 8,700만 파운드를 투자하여 학교・지역사회에 디지털 및 AI 학습 기회를 제공하고, 산업계와 협력하여, 장기적으로 AI 스킬 격차를 해소하는 것이 핵심 목표이다. 구체적으로는 테크유스(TechYouth) 프로그램을 통해 향후 3년간 영국 내 모든 중학교 학생 100만 명에게 AI・기술 교육과 진로 교육을 제공하고, 2030년까지 산업계와 협력을 통해 750만 명 노동자에게 필수 AI 기술 교육을 제공할 계획이다. 정부는 이를 위해 온라인 플랫폼을 구축해 교육 자료・진로 정보를 제공하고, 지역 파트너를 통해 학교・대학 단위로 프로그램을 제공한다는 계획도 발표했다.

테크유스 외에 3종의 맞춤형 지원 프로그램도 마련되었다. 학부 대상의 ‘테크그라드

96) Gov.uk, Artificial Intelligence Playbook for the UK Government, 2025.2.10.
97) Gov.uk, PM launches national skills drive to unlock opportunities for young people in tech, 2025.6.8.

(TechGrad, 9,680만 파운드)’ 프로그램은 매년 1,000명 이상 국내 학생에게 학부 장학금을 제공하고, 핵심기술 분야 석사(100명), AI 부문(100명) 학생에게도 장학금을 지원한다. 박사 대상인 ‘테크엑스퍼트(TechExpert, 4,840만 파운드)’ 프로그램은 박사과정 학생 500명에게 최대 1만 파운드를 지원해 첨단 혁신 연구를 촉진한다. 지역 기업 대상으로 제공되는 ‘테크 로컬(TechLocal, 1,800만 파운드)’ 프로그램은 지역 혁신 기업과 중소기업을 대상으로 기술 개발과 AI 도입에 초기 자금을 지원한다.

4) 법무부

(1) 사법 체계를 위한 AI[98)]

‘사법체계를 위한 AI 액션 플랜(AI Action Plan for Justice)’은 영국 사법체계에서 AI를 활용해 사건 처리 효율성, 접근성, 공정성을 향상하기 위해 마련된 전략으로, 2025년 7월 법무부가 발표했다. 이 액션 플랜은 법원·교정·법률 서비스 전반에서 AI가 수행할 수 있는 역할을 구체화하며, 사법부와 행정부가 공통으로 활용할 수 있는 데이터·기술 기반을 강화하는 데 중점을 두고 있다. 또한 서류 검토 자동화, 절차 간소화, 예측 분석 등 현실적인 도입 사례를 통해 사법적 판단과 지원 업무가 보다 신속하고 정확하게 운영되도록 하는 방향을 제시했다.

아울러 법무부는 사법 분야에 특화된 위험기관 기준을 명확히 하고, 편향·불투명성·부적절한 자동화 등 AI 관련 위험을 방지하기 위한 안전장치를 병행 도입하기로 했다. 특히 인간 중심의 최종 의사결정 원칙, 설명 가능성 요구, 공정성 검증, 민감 정보 보호 등을 핵심 원칙으로 설정해, AI가 사법 체계의 신뢰성과 민주적 가치에 부합하도록 운영될 수 있게 하는 데 정책적 중점을 두고 있다.

5) 교통부

(1) 교통 AI 액션플랜[99)]

교통부는 2025년 6월 ‘교통 AI 액션플랜(Transport Artificial Intelligence Action Plan)’을 발표했다. 이는 도로·철도·항공 등 교통 전 분야에서 AI가 제공하는 혁신 기회를 체계적

98) Gov.uk, AI action plan for justice, 2025.7.31.
99) Gov.uk, Policy Paper - Transport artificial intelligence action plan, 2025.6.10.

으로 정리한 문서로, 교통 운영 효율성 제고, 혼잡 완화, 안전성 확보, 탄소 감축과 같은 과제 해결을 위해 AI 기술을 단계적으로 도입하는 로드맵을 제시한다. 이를 통해 교통 인프라 및 서비스 제공 과정에서 AI를 활용한 예측・자동화・상황 인식 기술을 확대해 공공서비스 품질을 높이는 데 초점을 맞추고 있다.

특히 이 계획은 교통 시스템의 복잡성과 공공 안전성을 고려해 AI 적용 시 높은 수준의 검증・감독 기준을 마련하고, 산업계・지방정부・연구기관과의 협력을 기반으로 표준화・데이터 공유・기술 테스트베드 구축을 병행 추진할 것을 명시했다. 이를 통해 교통 분야의 디지털 전환을 가속화하는 동시에, 안전성과 책임성을 담보하는 균형 잡힌 AI 도입 모델을 구현하는 것을 핵심 목표로 제시했다.

(2) 과학・혁신・기술 계획 2025[100]

2025년 9월 발표된 ‘교통부 과학・혁신・기술 계획 2025(DfT Science, Innovation and Technology Plan 2025)’은 교통부의 과학・기술・혁신 전략으로, 교통 시스템에서 AI와 같은 첨단기술을 폭넓게 활용하는 방향을 제시하고 있다. 이 계획은 대규모 교통 데이터의 분석・모델링・예측에 AI를 도입하고, 교통 서비스 운영 최적화 및 정책 결정 정밀화를 지원하는 것을 핵심 목표로 삼고 있는데, 특히 교통부 내부 R&D 체계를 현대화해 기술 기반 정책 기능을 고도화하는 데 중점을 두었다.

또한 이 계획은 기술 도입을 위한 인력 역량 강화, 연구기관・산업계와의 공동 프로그램, 실증 기반 정책 개발 등 종합적인 혁신 구조를 제안한다. AI・디지털 트윈・자율 시스템 등 신기술을 체계적으로 통합해 교통 부문의 지속가능성을 높이고, 교통안전과 시민 편익을 개선하는 동시에, 영국의 교통 과학기술 생태계를 국제 경쟁력을 갖춘 수준으로 발전시키는 것이 목표이다.

100) Gov.uk, DfT science, innovation and technology plan 2025, 2025.9.15.

6) 금융감독청

(1) AI와 규제 접근[101)]

2025년 9월에 영국 금융감독청(FCA: Financial Conduct Authority)은 ‘AI와 FCA: 우리의 접근(AI and the FCA: our approach)’이라는 보고서를 발표하고 금융 시장에서 AI가 확산하는 상황에 대응하여 감독 당국이 취할 규제 원칙과 지원 전략을 제시했다. 이 보고서는 금융 서비스 전반에서 AI가 초래할 혁신 가능성을 인정하면서도, 소비자 보호·시장 무결성·금융범죄 예방이라는 FCA의 핵심 책무가 흔들리지 않도록 규제 틀 안에서의 AI 활용 기준을 구체화하는 데 목적이 있다. FCA는 혁신 촉진과 위험관리 사이의 균형을 강조하며, 기업이 AI를 안전하고 책임 있게 활용하도록 감독·가이드라인·샌드박스를 연계한 정책 패키지를 제시했다.

특히 FCA는 모델 위험관리, 설명 가능성, 데이터 윤리, 운영 탄력성 등 금융 AI에 특화된 주요 위험 항목을 명확히 제시하며, 필요 시 ‘라이브 테스트(현장 실험)’ 방식의 감독 모델도 도입할 수 있음을 밝혔다. 또한 디지털 금융 범죄 대응, 소비자 취약성 문제 완화, 금융 접근성 개선 등 공익 지향적 결과를 AI 활용 핵심 목표로 설정해, 금융 생태계 전반에서 AI의 책임 있는 확산을 지원하겠다는 의지를 표명했다.

101) FCA, AI And FCA: our approach, 2025.9.9.

제6절 캐나다

1. 그간의 인공지능 정책 추진 경과

<표 3-31> 캐나다의 주요 AI 정책 추진 현황

시기	발표기관	정책(주요 내용)
2017년	연방정부·CIFAR	범캐나다 AI 전략 1단계(Pan-Canadian AI Strategy)
2019년	재무위원회·TBS	자동화된 의사결정 지침(DADM) - 공공부문 알고리즘 활용 시 AIA(영향평가) 의무화
2022년 6월	ISED·CIFAR	범캐나다 AI 전략 2단계
2023년 5월	연방정부	고급 생성형 AI 자율강령(Voluntary Code of Conduct)
2023년 12월	개인정보보호위원회	AI윤리가이드라인(책임 있고 신뢰할 수 있으며 개인정보를 보호하는 생성형 AI 기술 원칙)
2024년 6월	연방정부	Securing Canada's AI Advantage 패키지
2024년 하반기	ISED	캐나다 소버린 AI 컴퓨팅 전략(AI Sovereign Compute Strategy)
2024년 11월	연방정부	캐나자 AI안전연구소 설립
2025년 6월	연방정부	연방 공공서비스 AI 전략 - 공공부문 AI·데이터 활용 3개년 로드맵 수립
2025년 7월	AI 안전 연구소	AI 정렬 프로젝트(영국 AI보안연구소 공동)
2025년 9월	연방정부	청정에너지 AI 인재·연구 투자 이니셔티브 추진
2025년 10월	연방정부	캐나다 소버린 AI 컴퓨팅 전략

캐나다는 2017년 세계 최초의 국가 인공지능 전략을 발표하며 AI 분야의 선도적 입지를 구축했다. 「범캐나다 AI 전략(Pan-Canadian Artificial Intelligence Strategy) 1단계」는 인재 확보와 기초연구 역량 강화에 집중했다. 캐나다 고등연구원(CIFAR)이 전략 실행을 주도했으며, 3대 국립 AI 연구소인 △Amii △Mila △Vector 를 핵심 거점으로 지정했다. 「범캐나다 AI 전략」을 통해 캐나다는 'CIFAR 인공지능 연구석좌(Canada CIFAR AI Chairs)' 프로그램

을 마련하고, 전 세계의 우수한 연구자들을 유치하며 강력한 인재 기반을 확보했다.

2022년 6월, 캐나다 정부는 「범캐나다 AI 전략 2단계」를 발표하며 정책의 무게 중심을 이동시켰다. 1단계에서 축적된 세계적 수준의 연구 역량을 바탕으로, 'AI 기술 상용화'와 '책임 있는 AI 채택'을 가속하는 것을 새로운 목표로 제시했다. 정부는 약 4억 4,300만 CND 이상의 예산 투입을 발표했으며, 예산은 △3대 AI 연구소 연구 성과의 상용화 프로젝트 지원 △캐나다 글로벌 혁신 클러스터(Global Innovation Clusters)를 통한 산업 적용 확대 △캐나다 표준위원회(SCC)의 AI 표준 개발 등에 사용되었다. 1단계가 '연구'에 집중했다면, 2단계는 이 연구 성과를 캐나다 경제의 실질적인 성장 동력으로 전환하는 데 초점을 맞추었다고 볼 수 있다.

2024년에는 생성 AI 기술이 급부상, 이로 인한 막대한 컴퓨팅 파워 확충 요구가 급증하자, AI 경쟁력이 인재뿐만 아니라 인프라에 의해 좌우되는 모습을 띠기 시작했다. 캐나다 정부는 이러한 패러다임 변화에 대응하여 2024년 6월, 총 24억 CND 규모의 「캐나다 AI 우위 확보 패키지(Securing Canada's AI Advantage)」를 발표했다. 「캐나다 AI 우위 확보 패키지」는 캐나다 AI 생태계의 지속가능한 발전을 위한 포괄적인 지원책을 담고 있다.

「캐나다 AI 우위 확보 패키지」의 가장 큰 비중을 차지하는 것은 AI 컴퓨팅 인프라 구축이다. 전체 예산 중 20억 달러가 인프라 확충에 할당되었으며, 지역개발청을 통해 AI 스타트업이 신기술을 통해 시장에 진출할 수 있도록 돕는 데 2억 달러를 지원한다. 캐나다 국립연구위원회(NRC)의 산업 연구 지원 프로그램(Industrial Research Assistance Program)을 통해서는 중소기업의 AI 솔루션 도입을 촉진하는 데 1억 달러를 배정했다. 더불어, AI 기술의 안전한 개발과 배치를 감독하기 위한 '캐나다 AI 안전 연구소(AI Safety Institute of Canada)' 설립에도 5천만 달러를 투입했다.

이러한 예산 배정은 '캐나다 소버린 AI 컴퓨팅 전략(Canadian AI Sovereign Compute Strategy)'을 통해 구체적인 실행 계획으로 제시되었다. '캐나다 소버린 AI 컴퓨팅 전략'은 20억 달러의 예산을 활용하여 캐나다의 AI 주권을 확립하는 것을 최종 목표로 한다. 정부는 2024년 하반기에 △AI 개발자 △학계 △기업 등 광범위한 이해관계자들의 의견을 수렴하여 전략을 설계했으며, 캐나다의 연구자와 기업이 외부가 아닌 자국 내 인프라를 통해 안정적

으로 컴퓨팅 자원에 접근할 수 있도록 보장하는 것에 전략을 집중시켰다. 이를 위해 'AI 컴퓨팅 접근 기금(AI Compute Access Fund)'을 통한 단기적 접근성 지원과 함께, 캐나다의 풍부한 청정에너지 자원을 활용한 지속가능한 데이터센터 구축 등 장기적인 인프라 확보를 동시에 추진했다.

캐나다의 AI 정책은 범캐나다 AI 전략 1단계의 '인재 확보'에서 범캐나다 AI 전략 2단계의 '상용화' 모색으로, 그리고 2024년 이후 '소버린 AI 구축'이라는 방향으로 진화해왔다고 평가할 수 있다. 이는 캐나다가 자국의 강력한 AI 인재풀을 기반으로, AI 기술의 경제적 성과를 극대화하고 기술 주도권을 확보하기 위한 전략적 행보를 가속해온 것으로 보인다.

2. 2025년 인공지능 정책 동향

1) AI 및 데이터법(AIDA) 입법 절차 폐기

① AIDA 입법 과정[102)]

「인공지능 및 데이터법(AIDA)」의 입법 시도는 2022년 6월 공식적으로 시작되었다. 캐나다 혁신과학경제개발부(ISED)는 「디지털 헌장 이행법」의 일부로 AIDA 법안을 의회에 상정했다.

법안 발의의 배경에는 캐나다 내 AI 기술의 급속한 확산이 존재했다. AI가 캐나다 국민의 일상과 경제 활동 전반에 미치는 영향력이 커짐에 따라, 이에 따른 잠재적 위험을 관리할 국가 차원의 규제 기반이 시급했다. 캐나다는 AI 연구 분야의 선도국이자 2020년 'AI에 관한 글로벌 파트너십(GPAI)'의 공동 창립국으로서, 책임 있는 AI 거버넌스에 대한 국제적 논의를 주도해 온 국가적 입지도 고려되었다. △EU △영국 등 주요 파트너국들이 AI 규제 마련을 가속화하는 상황에서, 캐나다 역시 국제 무대에서의 상호운용성을 확보하고 자국 데이터 경제의 지속적인 성장을 위한 신뢰 기반을 구축할 필요가 있었다.

AIDA의 목적은 AI 기술의 잠재적 위험으로부터 캐나다 국민을 보호하는 동시에, 책임 있

102) ISED, The Artificial Intelligence and Data Act (AIDA), 2025.1
https://ised-isde.canada.ca/site/innovation-better-canada/en/artificial-intelligence-and-data-act-aida-companion-document

는 혁신을 촉진하는 데 있었다. 법안은 AI 시스템의 안전한 설계·개발·배포를 위한 명확한 규칙을 수립하고자 했다. 또한, AI가 초래할 수 있는 편향과 차별 등 유해한 결과를 방지하는 것을 핵심 목표로 삼았다. 이를 통해 캐나다는 AI 기술에 대한 국민의 신뢰를 확보하고, 캐나다의 가치가 반영된 AI 생태계를 조성하려 했다.

AIDA 법안의 가장 큰 특징은 '위험 기반 접근법(risk-based approach)'을 채택한 점이다. 이는 모든 AI 시스템을 일률적으로 규제하는 대신, 위험 수준에 따라 차등화된 의무를 부과하는 방식이다. 법안은 연구 혁신 활동에 대한 과도한 부담을 피하고자, 규제 대상을 '고영향 AI 시스템(high-impact AI systems)'에 집중시켰다.

고영향 AI 시스템은 개인의 건강, 안전, 또는 인권에 중대한 부정적 영향을 미칠 위험이 있는 시스템으로 정의되었다. 정부는 고용 심사, 대출 및 서비스 접근성 평가, 생체 인식을 통한 식별 및 추론, 법 집행, 보건 및 안전과 직결된 시스템 등을 고영향 시스템의 주요 사례로 제시했다.

법안은 이러한 고영향 시스템의 생애주기 전반에 걸쳐 관련 주체들에게 명확한 책임을 부여했다. AI 시스템을 설계·개발·출시하는 사업자는 잠재적 위험을 식별하고 완화하기 위한 조치를 취해야 했다. 또한, 시스템 운영을 관리하는 사업자는 지시된 방식으로 시스템을 사용하고, 지속적인 모니터링과 위험 평가를 수행할 의무를 부과했다.

이러한 의무는 △인간 감독 및 모니터링 △투명성 △공정성과 형평성 △안전 △책임 △타당성 및 견고성 6대 핵심 원칙을 기반으로 했다. 특히 AIDA는 AI 시스템의 체계적 편향(systemic bias)으로 인해 발생하는 집단적 피해와 편향적 결과물(biased output)을 인권법에 근거한 차별적 영향으로 다루고자 했다. 강력한 집행 수단도 법안에 포함되었다. 법규 위반 시 막대한 규모의 '행정적 벌금(Administrative Monetary Penalties)'을 부과할 수 있었으며, 고의로 유해한 AI를 사용하여 심각한 피해를 야기할 경우 형사처벌까지 가능하도록 규정했다.

② AIDA 입법 절차 폐기[103)]

AIDA는 2025년 1월을 기점으로 사실상 폐기되었다. 2022년 6월 발의된 이후 2년 넘게 논의가 이어졌으나, 최종적으로 입법화에 이르지 못하고 의회 절차가 중단되었다.

AIDA의 입법 절차 폐기 배경에는 여러 복합적인 요인이 작용했다. 법안은 초기부터 캐나다 산업계와 시민사회로부터 적지 않은 비판에 직면했다. 산업계는 '고영향 AI 시스템'의 정의가 모호하고 규제 범위가 광범위하여, 자칫 캐나다 AI 기업의 혁신을 저해하고 과도한 규제 부담으로 작용할 수 있다는 우려를 지속적으로 제기했다. 또한 법안의 거버넌스 구조 역시 핵심 쟁점이었다. AIDA는 법안의 집행과 감독에 대한 대부분의 권한을 캐나다 혁신과학경제개발부(ISED) 장관에게 부여했다. 이에 대해 시민사회와 법률 전문가들은 독립적인 감독 기구가 부재한 상태에서 특정 부처에 과도한 권한이 집중되는 구조적 문제를 비판했다. 이들은 AI 규제의 공정성과 투명성을 담보하기 어렵다고 지적했다.

이러한 기술적, 구조적 논쟁이 해결되지 못한 상황에서, 2025년 1월 트뤼도 총리의 사임으로 인한 정치적 불안정은 의회에 계류 중이던 주요 법안들의 추진 동력의 상실로 이어졌고, AIDA 역시 공식적으로 폐기 수순을 밟게 되었다. AIDA 법안이 폐기됨에 따라 2025년 현재, 캐나다의 민간 부문 AI 활동을 포괄적으로 규율하는 연방 법률은 공백 상태에 놓이게 되었다. AIDA의 대안이 즉각적으로 마련되지 않았기 때문에, 캐나다의 AI 거버넌스는 기존의 부문별 규제에 의존하게 되었다.

대표적으로는 △자동 의사결정 지침(Directive on Automated Decision-Making)과 △알고리즘 영향 평가 도구(Algorithmic Impact Assessment tool)가 있으나, 이는 연방 공공 부문에 한정되어 적용될 뿐 민간 부문에 대해서는 법적 구속력이 없다.

캐나다는 향후 AIDA와 같은 포괄적 AI 규제 체계를 다시 마련하기 위해서는 새로운 의회에서 법안을 원점부터 재발의해야 한다. 이는 상당한 시간이 소요될 것으로 예상되며, 그동안 캐나다는 공공 부문 지침과 자발적 표준을 중심으로 AI 거버넌스 공백을 관리해 나갈 것으로 전망된다.

---

103) Montreal AI Ethics Institute, The Death of Canada's Artificial Intelligence and Data Act: What Happened, and What's Next for AI Regulation in Canada?, 2025.01
https://montrealethics.ai/the-death-of-canadas-artificial-intelligence-and-data-act-what-happened-and-whats-next-for-ai-regulation-in-canada

2) 연방 공공서비스 AI 전략 발표[104)]

「연방 공공서비스 AI 전략(Strategy for the Federal Public Service on the Use of Artificial Intelligence and Data)」은 2025년부터 2027년까지 이행되는 3개년 계획으로, 캐나다 공공 부문이 수립한 최초의 포괄적 인공지능 전략이다.

「연방 공공서비스 AI 전략」 수립 배경에는 공공서비스 품질을 향상하고 내부 운영의 효율성을 극대화하려는 요구가 있었다. 캐나다 정부는 「자동화된 의사결정에 관한 지침」 등을 통해 AI 사용의 위험을 관리해왔으나, 기술이 급속도로 발전함에 따라 방어적 지침을 넘어 AI를 적극적으로 활용하기 위한 체계적인 청사진이 필요해졌다. 이에 따라 캐나다 정부는 「연방 공공서비스 AI 전략」을 통해 AI와 데이터를 서비스 개선에 활용하고 AI 상용화의 핵심 동력으로 삼고자 했다.

「연방 공공서비스 AI 전략」은 △중앙 AI 역량(Central AI capacity) △정책, 법률 및 거버넌스(Policy, legislation and governance) △인재와 교육(Talent and training) △참여와 투명성 및 국민가치(Engagement, transparency, and value to Canadians) 네 가지 주요목표를 겨냥하고 있다.

<표 3-32> 연방 공공서비스 AI 전략 4대 목표 및 주요 내용

구분	주요 내용
중앙 AI 역량	• AI전문 지원센터(AI Centre of Expertise) 설립 - 고부가가치 AI 사용 사례 발굴부터 데이터 준비, 평가, 조달, 거버넌스 등 정부 서비스 관련 전 주기 지원 제공 - 부처 간 지식 공유와 협력 장려, AI 솔루션 확산을 위한 촉진 • 공동 인프라 활성화 - AI 프로젝트 수행 시 필요한 고성능 컴퓨팅 및 클라우드 인프라 제공 - 공통 데이터 관리 체계와 공급업체용 표준화된 플랫폼 구축 • 시범사업 선정·추진 - AI 적용 과정의 장벽과 실무 필요 사항에 대해 테스트로 사전 확인 - AI 확산 및 맞춤 지원을 위한 과제 도출과 거버넌스 모델 설계 활용

104) Government of Canada, AI Strategy for the Federal Public Service 2025-2027: Overview, 2025.6. https://www.canada.ca/en/government/system/digital-government/digital-government-innovations/responsible-use-ai/gc-ai-strategy-overview.html

구분	주요 내용
정책, 법률 및 거버넌스	• 공동 AI거버넌스와 위험관리 체계 확립 - AI도입에서 폐기까지 연방 공동 거버넌스 및 리스크관리, 운영 지침 제공 - 윤리 검토, 데이터 주권, 시스템 신뢰성 등을 반영한 정부 차원의 거버넌스 구조를 통해 전략 실행과 고위험 프로젝트를 체계적으로 감독 • 정책 및 법적 장벽 해결 - AI 도입을 가로막는 정책·법령·조달 도구 등을 검토하여, 개인정보 보호, 데이터 주권, 외부 기관의 AI 활용 등 정책 격차 보완 - 기술·사회 변화에 민첩하게 대응할 수 있도록 지침과 도구 정비, AI 관련 국내외 전략 및 법적 기준과의 정합성 제고 방안 마련 • Think AI 접근 문화 확립 - 정책·서비스 업무 기획 단계부터 AI를 고려하는 접근법 - AI적용성과 데이터 연계 가능성을 평가하고 관련 요구 사항을 통합하도록 예산과 전략을 사전에 고려
인재와 교육	• 역량 강화 계획 수립 - 공무원 대상 역할별 맞춤형 AI 교육과 리더십 프로그램을 통한 역량 강화 - 재교육과 협업으로 AI변화에 유연하게 대응할 수 있는 인재 기반 마련 • 인재 역량 기준 설정 - AI를 위한 인재 요구사항과 직원들의 기존 보유 지식 기준점 마련 - 기준점을 통해 학습자 유형 개발 및 추가 교육 대상 선발 • 인재 확보 계획 - AI 인재 유치를 위한 유연한 경력 경로 마련과 채용 유지 전략 수립 - 인턴십, 산학협력, 과제 기반 운영 등과 같은 AI인재 파이프라인 구축
참여와 투명성 및 국민가치	• AI공공 등록부 운영 - 정부 AI시스템의 목적, 책임 주체, 개발 단계 등의 표준화 및 데이터화 - 데이터 출처, 모델 정보, 위험 평가, 인간 감독, 개인정보 보호 등 AI 사용 주기 전반에 걸친 정보 제공 및 공유 • AI도입 성과 지표 개발 - AI 프로젝트의 재무적·비재무적 요소를 아우르는 성과 지표 체계 마련 - AI 도입 현황, 협업·공유 성과, 부서별 자원 투자 등을 종합적으로 추적할 수 있는 통합 프레임워크 구축 • AI도입에 대한 국민 참여 - 공공에 중대한 영향을 미치는 AI정책에 대해 초기 단계부터 이해관계자 및 국민의 실질적인 참여 약속 - AI 시스템의 접근성·형평성 확보를 위해 취약계층, 공무원, 사용자 참여를 확대하고, 지속적인 피드백 채널을 운영

캐나다 정부는 「연방 공공서비스 AI 전략」을 통해 AI 기술을 책임감 있게 활용하여, 신뢰할 수 있는 현대적 공공서비스 환경 구축을 겨냥하고 있다. AI를 통해 데이터 기반의 의사결정 체계를 정착시키고, 행정의 생산성을 높이는 것이 핵심이다. 동시에 투명한 거버넌스와 공공 참여를 보장함으로써, 기술 도입 과정에서 발생할 수 있는 위험을 관리하고 국민적 신뢰를 확보하는 것을 목표로 한다.

3) 캐나다 소버린 AI 컴퓨팅 전략[105)]

캐나다는 AI 연구자와 스타트업 및 스케일업 기업의 경쟁력 유지를 위해 필수적인 컴퓨팅 인프라에 대한 접근성을 개선하기 위해 2025년 10월 '캐나다 소버린 AI 컴퓨팅 전략(Canadian Sovereign AI Compute Strategy)'을 발표했다. 캐나다 정부는 2024년부터 5년간 총 20억 캐나다 달러를 투자하여 자국 AI 산업 및 연구자들에게 최첨단 컴퓨팅 인프라에 대한 접근성을 확보하고 글로벌 AI 경쟁에서 캐나다의 경쟁력 강화를 도모한다.

동 전략은 3대 핵심 요소들을 통해 캐나다의 컴퓨팅 역량과 AI 생태계 지원을 강화하하며 경제 성장을 견인하는 데 초점을 두고 있다. 우선 민간 부문 투자 독려로, 캐나다 정부는 에너지, 토지, 기후 측면에서 캐나다의 자연적 경쟁 우위를 활용하여 국내 AI 컴퓨팅 역량 증대를 통해 캐나다 AI 생태계를 지원하기 위해 최대 7억 달러를 투자한다. 이를 위해 'AI 컴퓨트 챌린지(AI Compute Challenge)'를 통해 기업, 컨소시엄, 산학 협력체들의 프로젝트 제안을 중심으로 상업적 배치가 가능한 완전 통합형 AI 데이터센터 솔루션를 구축할 계획이다. 둘째, 공공 슈퍼컴퓨팅 인프라 구축으로, 연구자, 정부, 산업의 요구를 충족시키기 위한 공공 슈퍼컴퓨팅 인프라에 최대 10억 달러의 투자를 추진한다. 이에 따라 'AI 주권 컴퓨트 인프라 프로그램(AI Sovereign Compute Infrastructure Program, SCIP)'을 통해 최대 7억 500만 달러를 새로운 AI 슈퍼컴퓨팅 시스템에 투자하고, 소규모 보안 컴퓨팅 시설을 국가 안보 목적 및 공공과 산업의 연구 개발을 위해 구축될 예정이다.

셋째, 3억 달러 규모의 'AI 컴퓨트 액세스 펀드(AI Compute Access Fund)'를 형성하고, 캐나다 혁신 기업들의 AI 컴퓨팅 자원 구매를 지원할 계획이다. 이 펀드를 통해 캐나다 정부는 컴퓨팅 자원의 높은 비용과 제한된 국내 역량 가용성 문제를 해결하고 고성능 컴퓨팅 역량이 필요한 잠재 분야의 AI 장벽 해소를 도모하며 AI 생태계 조직들과 긴밀히 협력할 계획이다.

105) ISED, Canadian Sovereign AI Compute Strategy, 2025.10

4) 청정에너지 AI 인재・연구 투자[106)]

캐나다 정부는 2025년 9월, 자국의 AI 역량과 핵심 산업 정책을 결합하는 새로운 청정에너지 AI 인재 및 연구 투자 계획을 공개했다. 청정에너지 AI 인재 및 연구 투자는 「범캐나다 AI 전략」으로 다져진 AI 연구 분야의 우위를 지키고, '넷제로(Net-zero)'로 대표되는 청정에너지 경제 확충을 목표로 한다.

청정에너지 AI 인재 및 연구 투자는 △숙련 기술 인력의 공백과 △핵심 연구기관의 지속적인 성장 지원, 두 가지 과제를 동시에 해결하기 위한 전략적 결정이다. 캐나다 고용사회개발부(ESDC)와 혁신과학경제개발부(ISED)가 공동으로 주관하며, 총 2,870만 CND 규모의 투자가 앨버타 대학교와 앨버타 기계 지능 연구소(Amii)에 집중된다.

앨버타 대학교는 1,990만 달러를 지원받으며, 「청정에너지 분야 인력 양성(Upskilling for Clean Energy)」 프로젝트에 사용된다. 「청정에너지 분야 인력 양성」 프로젝트는 「분야별 인력 솔루션 프로그램(Sectoral Workforce Solutions Program)」의 일부로 운영되며, 최대 1,775명의 인력을 대상으로 새로운 기술 교육을 제공한다. 특히 기존 산업 종사자나 경력이 단절된 인력의 재취업을 돕는다. 넷제로 경제에 필요한 인력 수요에 선제적으로 대응하는 것이 목적이다.

Amii는 880만 캐나다달러를 지원받으며, Amii의 AI 연구 능력 확장을 목표로 한다. 이는 「범캐나다 AI 전략」의 핵심 목표와 일치하는 것으로, 세계적 수준의 AI 인재를 유치하고 캐나다의 AI 리더십을 공고히 하는 데 사용된다. 캐나다 정부는 이를 통해 AI 기술혁신을 촉진하고, 다양한 산업에 적용하여 경제 성장을 견인하는 기반을 마련하고자 한다.

106) Employment and Social Development Canada, Government of Canada invests $28.7 million to Train Clean Energy Workers and Expand AI Research Capacity, 2025.09.
https://www.canada.ca/en/employment-social-development/news/2025/09/government-of-canada-invests-287-million-to-train-clean-energy-workers-and-expand-ai-research-capacity.html

5) AI 지원 이니셔티브[107]

「AI 지원 이니셔티브(AI-Assist Initiative)」는 캐나다 내 중소기업의 기술 도입 확산에도 초점을 맞추고 있는 사업으로, 캐나다 국립연구위원회 내 산업연구 지원 프로그램(NRC Industrial Research Assistance Program)이 본 이니셔티브를 주관한다. 캐나다 정부는 우수한 AI 연구 역량에도 불구하고, 자본과 전문성이 부족한 중소기업들이 AI 기술을 실제 비즈니스에 적용하는 데 어려움을 겪고 있다고 판단하고, 「AI 지원 이니셔티브」를 통해 중소기업이 직면한 기술적·재정적 장벽을 낮춰 AI 도입을 가속시키고자 노력하고 있다.

「AI 지원 이니셔티브」는 정부가 중소기업에 직접 자금을 지원하지 않는 대신 NRC IRAP 공모를 통해 AI 도입 지원 역량을 갖춘 비영리 단체들을 선정 후, 선정된 비영리 기관이 운영 자금을 지원하는 방식으로 간접 지원된다. 이후, 비영리 단체들은 정부와 기업 간 중개자를 수행하는 동시에, AI 기술 도입을 희망하는 중소기업을 발굴한다. 더불어 중소기업을 대상으로 맞춤형 자문 서비스를 제공해야 한다. 여기에는 기업의 현 비즈니스를 진단하고, 가장 적합한 AI 솔루션을 기획하는 과정이 포함된다. 비영리 단체는 중소기업당 최대 25,000 캐나다달러의 자금을 지원하여 초기 프로젝트 실행을 돕는다.

「AI 지원 이니셔티브」는 중소기업이 큰 초기 투자 부담 없이 AI 기술을 시험하고 내재화할 수 있도록 지원하며, 이를 통해 중소기업은 생산성을 향상해 새로운 성장 동력 확보와 캐나다 산업 생태계 전반의 경쟁력을 강화하는 것을 겨냥하고 있다.

6) 'AI 정렬 프로젝트' 참여[108]

2025년 7월, 캐나다는 영국 AI보안연구소(UK AI Security Institute)와 함께 AI의 안전한 개발을 목표로 하는 국제 연합 연구 「AI 정렬 프로젝트(Alignment Project)」를 발표했다.

「AI 정렬 프로젝트」는 AI가 개발자의 의도대로만 작동하도록 제어하는 기술을 다룬다.

---

107) National Research Council Canada, AI Assist call for proposals from organizations, 2025.08. https://nrc.canada.ca/en/support-technology-innovation/ai-assist-call-proposals-organizations

108) UK Department for Science, Innovation and Technology, AI Security Institute launches international coalition to safeguard AI development, 2025.07. https://www.gov.uk/government/news/ai-security-institute-launches-international-coalition-to-safeguard-ai-development

캐나다는 「AI 정렬 프로젝트」의 핵심 파트너로, 캐나다 AI 안전연구소(CAISI)가 UK AISI와 직접 협력한다. 또한, CIFAR는 「AI 정렬 프로젝트」를 위해 100만 달러의 기금을 지원했으며, 이는 AI의 동작 제어 기술을 확보하려는 양국의 공동 목표를 반영한다.

7) AI 표준 및 인증 개발[109)]

캐나다의 AI 표준화 노력은 2022년 「범캐나다 AI 전략 2단계」의 일환으로 시작되었다. AI 기술의 상용화가 가속화되면서, 기술의 신뢰성과 책임성을 입증할 명확한 기준이 필요해졌다. 이에 캐나다 표준위원회(SCC)가 주도하여 2022년 6월 「AI 관리시스템 인증 파일럿 프로그램(Artificial Intelligence Management Systems Accreditation Pilot)」을 착수했다. 「AI 관리시스템 인증 파일럿 프로그램」은 향후 공식 인증 프로그램을 운영하기 위한 요구사항을 정의하고 검증하는 것을 목표로 했다.

파일럿 프로그램은 2024년 9월 「AI 인증 교훈 보고서(AI accreditation lessons learned report)」를 발간하며 성공적으로 마무리되었다. 보고서는 파일럿 운영을 통해 얻은 시사점을 정리하고, 후속 인증 체계 설계의 근거를 마련했다. 다만, 당초 AIMS 프로그램은 2022년 발의된 AIDA의 원활한 이행을 지원하는 핵심 도구로 기획되었으나, 2025년 1월 AIDA의 입법 절차가 최종 폐기되면서, 캐나다의 민간 AI 규제에 법적 공백이 발생했다. AIDA의 부재는 역설적으로 AIMS 프로그램의 중요성을 더욱 부각시켰다. 법적 강제 수단이 사라진 상황에서, 조직이 스스로 책임 있는 AI 거버넌스를 입증할 수 있는 신뢰할 수 있는 경로가 필요해졌기 때문이다.

이러한 맥락을 바탕으로, SCC는 2025년 9월 공식 「AIMS 인증 프로그램」을 시장에 도입했다. 「AIMS 인증 프로그램」의 핵심 목적은 AI를 개발하거나 활용하는 조직이 국제적으로 인정받는 표준에 따라 AI 관리 시스템의 △효과성 △안전성 △신뢰성을 검증받을 수 있도록 하는 것이다. 이는 조직에게 제3자 인증을 통해 AI 기술에 대한 이해관계자들의 신뢰를 구축할 수 있는 명확한 경로를 제공하는 것을 목표로 한다. 궁극적으로 AIMS 인증은 법적 규제 대신, 표준 기반의 자발적 거버넌스를 통해 캐나다 AI 생태계 전반의 신뢰를 구축하는 핵심 제도로 기능을 할 것으로 기대된다.

---

109) SCC, SCC launches accreditation pilot for AI Management System, 2022.6. https://scc-ccn.ca/resources/news/scc-launches-accreditation-pilot-ai-management-systems

8) AI 윤리가이드라인[110]

캐나다 개인정보보호위원회(OPC)는 생성형 AI 기술의 급속한 발전에 대응하여 2023년 말 '책임 있고 신뢰할 수 있으며 개인정보를 보호하는 생성형 AI 기술 원칙(Principles for responsible, trustworthy and privacy-protective generative AI technologies, 이하 'AI 윤리 가이드라인') '을 발표했다. 동 가이드라인은 캐나다가 2020년 AI에 관한 글로벌 파트너십(GPAI) 창립 회원국으로서 책임 있는 AI 개발에 대한 국제적 규범을 형성하기 위한 노력의 일환으로 이뤄졌다. 캐나다 개인정보보호위원회는 생성형 AI가 개인정보를 포함한 방대한 데이터를 학습하여 콘텐츠를 확률적으로 생성한다는 점에서 새로운 개인정보보호 위험을 제기한다고 판단했다. 이 가이드라인에 따라 AI 시스템을 제공하는 모든 관련 조직은 캐나다의 적용 가능한 개인정보보호 법규를 준수할 의무를 지닌다.

9) 캐나다 디지털규제포럼 '합성 미디어' 보고서[111]

2025년 9월, 캐나다의 주요 디지털 규제기관들이 '합성 미디어(Synthetic Media)'에 대한 공동 분석 보고서 「디지털 환경 안에서의 합성 미디어(Synthetic media in the digital landscape)」를 발표했다. 보고서는 생성 AI 기술의 발전으로 AI가 생성하거나 변형한 콘텐츠가 디지털 환경 전반에 급속히 확산하는 상황에 대응하기 위한 것이다.

보고서 발간 배경에는 "AI를 활용해 인위적으로 생성 및 조작된 △이미지 △영상 △텍스트 △오디오 등"을 의미하는 합성 미디어가 캐나다의 디지털 경제와 사회에 미치는 복합적인 영향을 파악해야 할 필요성이 있었다. 특히 합성미디어는 한 개의 규제 영역이 아닌 △AI △콘텐츠 △프라이버시 △노동 △소비자보호 등을 아우르는 다층적 규제 대상이라는 점에서 다양한 기관 간 협의가 필수적이다. 이에 따라 △캐나다 경쟁국(Competition Bureau) △캐나다 개인정보보호위원회(Office of the Privacy Commissioner) △캐나다 저작권위원회(The Copyright Board of Canada) 등 핵심 규제기관들이 참여하는 협의체 △'캐나다 디지털 규제포럼(Canadian Digital Regulators Forum)'가 논의를 주도해, AI 기반의 합성 미디어가 야기하는 문제가 개별 기관의 소관을 넘어선다는 데 공감했으며, 범정부 차원의 공동 분석을 통해 현황을 진단하고 향후 정책 방향을 모색하는데 의견을 모았다. 대표적으로는 △합성

110) OPC, Principles for responsible, trustworthy and privacy-protective generative AI technologies, 2023.7
111) Competition Bureau, Canadian Digital Regulators Forum: Synthetic media in the digital landscape, 2025.9. https://competition-bureau.canada.ca/en/how-we-foster-competition/education-and-outreach/publications/canadian-digital-regulators-forum-synthetic-media-digital-landscape

미디어의 투명성 확보 △규제기관 간 협력과 역할 분담 △창작자・소비자・플랫폼・AI 개발사 간 이해관계 조정 및 거버넌스 설계의 중요성에 공감대를 형성하며 공통의 과제로 인식했다.

보고서의 주요 내용은 합성 미디어가 가져오는 기회와 위험을 균형 있게 다루고 있다. 우선, 합성 미디어는 창의성과 혁신을 촉진하는 긍정적 잠재력을 지닌다. 예술, 엔터테인먼트, 교육 등 다양한 분야에서 새로운 형태의 콘텐츠 제작을 가능하게 한다. 또한 시장의 경쟁을 활성화하고 새로운 경제적 이익을 창출할 기회 요인으로 평가되었다.

반면, 보고서는 합성 미디어가 초래할 수 있는 심각한 위험성 역시 명확히 경고했다. 가장 큰 위협은 정교하게 조작된 콘텐츠를 통한 소비자 기만 및 사기 행위의 증가이다. 또한, AI가 학습 과정에서 기존 저작물을 무단으로 사용하거나 변형하여 발생하는 저작권 침해 문제도 핵심 위험으로 지적되었다. 개인의 동의 없이 음성이나 초상을 합성하여 발생하는 심각한 개인정보 침해 가능성도 강조되었다. 이 보고서는 향후 캐나다 규제 당국이 합성 미디어 관련 정책과 규제를 마련하는 데 중요한 기초 자료로 활용될 전망이다.

<표 3-33> 캐나다 디지털규제포럼 기관별 합성미디어 쟁점 정리

기관	역할	쟁점	기관별 주요 과제
저작권 위원회	• 저작권 사용에 관한 경제 규제 및 준사법적 기능 수행	• 합성미디어 시대에 누가 저작자이며 어떤 창작물이 저작권 보호 대상인지 • AI 훈련 및 합성콘텐츠 이용에 대해 어떤 라이선스를 설정할 것인지	• 저작자성·요율·레퍼토리·법경제적 불확실성 관리
방송통신 위원회	• △방송 △통신 △온라인 미디어 규제 및 콘텐츠의 △접근성 △공정성 보장	• 합성미디어가 △뉴스 △콘텐츠 제작 △유통 △접근성에 미치는 영향과 콘텐츠의 정의 • △저널리즘 △창작 노동 △이용자 보호 규범을 어떻게 조정할지	• 미디어·저널리즘·접근성 규범 재설계
캐나다 경쟁국	• △경쟁촉진 △담합 및 독점 △기만적 마케팅 △스팸 규제 등 집행	• △합성미디어 △딥페이크 활용 기만 사기 및 스팸이 어떤 위험을 만드는지 • 어떤 법집행 전략이 필요한지	• 기만 마케팅·딥페이크 사기·라벨링 기준 마련
개인정보보호 위원회	• 연방 개인정보보호법 관련 조사 및 집행, 권고	• 합성미디어 시스템에 입출력되는 개인정보 보호 문제를 어떻게 다룰지 • 연방 프라이버시법 적용 범위와 권리구제를 어떻게 설계할지	• 훈련·출력 단계의 개인정보 보호, 법 적용·권리구제 경로 설정

출처 : Synthetic media in the digital landscape, SPRi 정리

10) '콘텐츠 자격증명' 공동 지침 발간[112)]

2025년 1월, 캐나다 사이버보안센터(CCCS)는 △미국 국가안보국(NSA) △영국 국립사이버보안센터(NCSC) △호주 신호국(ASD)과 공동으로 디지털 콘텐츠의 무결성 위협에 대응하기 위한 관련 지침 「콘텐츠 자격 증명: 생성적 AI 시대의 멀티미디어 무결성 강화(Content credentials: Strengthening multimedia integrity in the generative AI era)」를 통해 공동의 기술적 대응 방안을 제시했다.

지침은 '콘텐츠 자격증명(Content Credentials)' 기술의 도입을 권고하는 것으로, '콘텐츠 자격증명'은 디지털 콘텐츠의 출처와 이력(provenance)에 대한 검증 가능한 정보를 제공하는 기술 표준이다.

'콘텐츠 자격증명'은 콘텐츠 파일에 위변조가 어려운 메타데이터를 첨부하는 방식으로 작동한다. 메타데이터에는 △콘텐츠의 최초 생성자 △생성 일시 △사용 도구 △편집 및 수정 이력 등이 상세히 기록된다. 특히, 콘텐츠 제작 과정에서 AI가 사용되었는지 여부 역시 명시해야 한다.

공동 지침은 '콘텐츠 자격증명'을 '디지털 콘텐츠의 영양 성분표'에 비유하며 △사용자가 정보를 소비하기 전에 해당 콘텐츠가 어떻게 만들어졌는지 확인하고 △신뢰 여부를 스스로 판단할 수 있는 근거를 제공할 수 있어야 한다고 강조했다. 4개국 기관은 '콘텐츠 자격증명' 표준의 광범위한 채택을 촉구하며, 이를 통해 조직은 투명성을 높이고, AI가 생성한 조작 콘텐츠에 효과적으로 대응할 수 있을 것으로 기대했다.

112) Australian Government, Content credentials: Strengthening multimedia integrity in the generative AI era, 2025.01.
https://www.cyber.gov.au/business-government/secure-design/artificial-intelligence/content-credentials-strengthening-multimedia-integrity-in-the-generative-ai-era

제7절 프랑스

1. 그간의 인공지능 정책 추진 경과

<표 3-34> 프랑스의 주요 AI 정책 추진 현황

시기	발표기관	정책(주요 내용)
2018년 3월	대통령실·정부	'AI for Humanity' 연설 및 1단계 국가 AI 전략 발표 - 연구·인재 중심 투자, 3IA 연구소와 슈퍼컴퓨터 등 인프라 구축
2021년 10월	프랑스 정부	France 2030 혁신투자계획 - AI를 10대 전략 분야 중 하나로 선정
2022년	정부·Bpifrance 등	국가 AI 전략 2단계 추진 개시 - 산업·공공서비스에 AI 확산, 사회·환경 과제 해결 중심
2023년 9월	총리실	AI 혁신위원회(Commission de l'IA) 설치 - 프랑스 AI 전략 재설계를 위한 자문기구
2024년 3월	AI 혁신위원회	'AI: 프랑스에 대한 우리의 야망'(IA : Notre ambition pour la France) 보고서 발간 - 인프라·인재·산업·신뢰·프랑스어·지속가능성 6대 축 25개 권고
2025년 1월	정부·연구기관 연합	국립 AI 평가·안전 연구소(INESIA) 출범 - AI 성능·안전 평가, 규제 지원 및 테스트베드 역할
2025년 1월	개인정보보호 위원회	2025-2028 전략 계획 발표 - 책임 있는 AI 기술 활용을 핵심과제로 설정
2025년 2월	범정부 AI 위원회	국가 AI 전략 3단계 발표 - AI 인프라·인재·산업·행정 활용·신뢰성·친환경 AI 6대 정책 축과 1,090억 유로 민간 투자 유치 목표 설정
2025년 2월	대통령	AI Action Summit 개최 - InvestAI·CurrentAI 등 국제 이니셔티브와 연계한 글로벌 AI 허브 비전 천명
2025년 7월	프랑스 정부	Osez l'IA 전략 발표 - 프랑스 기업들에서 AI 활용 확산 도모
2025년 7월	보건복지부	AI와 보건의료 데이터 전략(Stratégie IA et données de santé) 1차 발표 - '데이터' 파트 전략 발표
2025년 9월	프랑스 정부	AI 선도 프로젝트(Pionniers de l'intelligence artificielle) 공모 사업 발표
2025년 11월	보건복지부	AI와 보건의료 데이터 전략 2차 발표 - 'AI' 파트 전략 발표

1) 개요

2018년 3월 에마뉘엘 마크롱(Emmanuel Macron) 대통령은 「AI 포 휴머니티(AI for Humanity)」 정책 비전 연설에서 「국가 AI 전략(SNIA: Stratégie nationale pour l’ intelligence artificielle)」[113]을 발표하며, AI를 미래 성장과 국가 경쟁력 확보의 핵심 동력으로 규정했다. 「국가 AI 전략」은 프랑스 정부의 공식 AI 정책 프레임워크로, 「AI 포 휴머니티」에서 제시된 정책 비전(인간 중심의 AI, 윤리・신뢰, 연구・혁신・산업 강조)을 구체화한 실행 전략이다. 동 전략은 ‘인간 중심의 AI’를 국가 비전으로 설정하고, 기술혁신과 윤리적 규범의 조화를 통해 지속가능한 디지털 사회를 구현하는 것을 목표로 한다.[114]

프랑스의 「국가 AI 전략」은 2021년 이후 국가 미래 전략인 「프랑스 2030(France 2030)」[115]과 연계되어 추진되고 있다. 2021년 10월 12일 공식 발표된 「프랑스 2030」은 총 540억 유로 규모의 국가 기술・산업 혁신 투자계획으로, 이 중 AI는 반도체, 로봇, 바이오, 에너지, 모빌리티, 친환경 기술 등과 함께 10대 혁신 전략 분야 중 하나로 포함되어 100억 유로 규모의 투자가 계획되어 있다.[116]

프랑스의 AI 정책은 「프랑스 2030」과 연계되면서 2021년부터 전략 범위를 확장하고, 연구개발 중심의 제1단계(Phase 1)에서 산업 전반의 AI 확산과 사회 통합(Integration sociale)을 목표로 하는 제2단계(Phase 2)로 전환했다. 사회 통합은 AI 기술을 산업・연구개발에 국한하지 않고 시민 생활, 교육, 행정, 공공서비스 등 전 사회 영역에 확산하여 디지털 격차를 완화하려는 정책 목표이다. 이후 2025년 2월, 범정부 AI 위원회(Comité interministériel de l’ intelligence artificielle)는 「국가 AI 전략」 제3단계(Phase 3)를 발표하며, 「프랑스 2030」 계획과 연계해 전 산업 및 사회로의 AI 활용 확산과 디지털 격차 해소, 공공 행정의 효율성 강화 등을 주요 과제로 삼고 국가 자율성, 기술 주권, 혁신 리더십 확보를 위한 정책을 본격화하고 있다.[117]

113) Gouvernement de la République française, stratégie nationale pour l’intelligence artificielle (IA), 2018.
114) Élysée, Discours du Président de la République sur l'intelligence artificielle, 2018.3.29.
115) 프랑스 정부, France 2030 전략 개요(Understanding France 2030) (2025.10.30. 방문)
https://www.info.gouv.fr/grand-dossier/france-2030-en/understanding-france-2030
116) Service d'information du Gouvernement, Understanding France 2030, 2024.5.8.
117) entreprises.gouv.fr, La stratégie nationale pour l’intelligence artificielle, 2025.2.7.

2) 「프랑스 2030」과 「국가 AI 전략」

2021년 10월 발표된 「프랑스 2030」 계획은 에마뉘엘 마크롱 대통령이 직접 주도하는 범정부 프로젝트로, 기후 변화 대응, 에너지 전환, 디지털 주권, 산업 혁신을 총괄하는 국가 미래 전략 프로그램이다. AI를 포함한 10대 핵심 전략 산업을 중심으로 국가 혁신 주권 확보를 목표로 하는 중장기 투자 계획이라 할 수 있다.

「프랑스 2030」의 비전은 ▲혁신 주권 확보(AI·반도체·에너지 등 전략 기술의 자립성 강화), ▲지속가능한 성장(저탄소·절약형 AI, 에너지 효율성, 친환경 인프라), ▲사회·산업 통합(기술혁신을 산업 경쟁력 및 공공서비스 개선으로 확산)이라는 세 가지 축을 중심으로 한다. 「프랑스 2030」의 목표는 프랑스의 경제적·기술적·정책적 주권을 강화하면서, AI를 사회와 경제에 유익하게 활용하는 것이다.

이러한 비전과 목표 아래 「국가 AI 전략」은 「프랑스 2030」 계획의 일부로 포함되어 3단계에 걸쳐 추진되고 있다. 2021년에는 「프랑스 2030」을 통해 AI를 국가 혁신의 중심축으로 재정의하였고, 2025년 2월부터는 「국가 AI 전략」 제2단계(2022~2025)를 마무리하고, 제3단계로 전환했다. 제2단계는 산업 현장과 공공서비스에서의 AI 활용 확산에 중점을 두었고, 제3단계는 '신뢰할 수 있는 AI'와 '저탄소·절약형 AI'를 중심으로 국가 차원의 AI 자립 생태계 구축을 목표로 하고 있다. 각 단계는 시대적 환경과 기술 발전 수준에 따라 목표와 접근 방식이 진화해 왔으며, 연구 중심 → 산업 확산 → 사회 통합 및 기술 주권 강화로 이어지는 점진적 전환을 꾀하고 있다.[118]

<표 3-35> 프랑스 「국가 AI 전략」 단계 및 중점 내용

구분	주요 내용	
제1단계 (2018~2022)	시점	• 2018년 3월 에마뉘엘 마크롱 대통령이 「AI 포 휴머니티」 국가 AI 정책 비전 연설에서 「국가 AI 전략」 발표
	목표	• 프랑스를 유럽의 AI 연구 중심국으로 육성하고, 이를 위한 학문적 기반 조성
	투자 규모 /분야	• 총 15억 유로 규모의 예산 • 연구·교육 중심 • 알고리즘·데이터 과학 분야 우선

118) Service d'information du Gouvernement, L'intelligence artificielle au service de nos quotidiens, 2025.2.

구분	주요 내용	
	주요 내용	• 4대 AI 연구 허브 설립 : 3IA(Instituts Interdisciplinaires d'Intelligence Artificielle) 클러스터 : 파리, 니스, 그르노블, 툴루즈) 설립 • 쟝 제이(Jean Zay) 슈퍼컴퓨터 구축, AI 연구자·박사급 인재 양성 • 국립디지털과학기술연구소(INRIA: Institut national de recherche en informatique et en automatique), 국립과학연구원(CNRS: Centre national de la recherche scientifique), 프랑스 원자력 및 대체에너지 위원회(CEA: Commissariat à l'énergie atomique et aux énergies alternatives) 등 국립연구기관을 중심으로 AI 알고리즘·데이터 과학 연구 강화 • 개인정보보호위원회(CNIL: Commission Nationale de l'Informatique et des Libertés), 국가디지털윤리위원회(CNPEN: Comité Consultatif National d'Ethique)을 통한 AI 윤리·규제 논의 착수
	주요 집행 주체	• 국립디지털과학기술연구소, 3IA 클러스터, 공공투자은행(Bpifrance: Banque publique d'investissement), 고등교육연구부(MESR: Ministère de l'Enseignement Supérieur et de la Recherche)
	키워드	• AI 연구 중심, 국가 클러스터 조성, 기초 역량 확보
제2단계 (2022~2025)	시점	• 2022년 11월, 「프랑스 2030」 계획의 일환으로 「국가 AI 전략」 제2단계 발표
	목표	• 산업·경제 전반으로 AI 확산, 신뢰성 확보 및 친환경 기술 개발
	투자 규모 /분야	• 총 25억 유로 규모 • 임베디드 AI, 신뢰할 수 있는 AI, 친환경·저전력 AI, 생성형 AI 등 등 우선 분야에 집중 투자
	주요 내용	• 3대 추진축 : ① 인재 양성 및 연구 강화, ② 최첨단 기술 공급 육성, ③ 수요-공급 연결 및 상용화 촉진 • AI 연구·교육 클러스터 구축(5억 6천만 유로) • 산업계 AI 도입 지원(중소기업 400개 지원 목표) • AI 전문인력 양성(매년 학사·석사 3,500명 육성)
	주요 집행 주체	• 경제재정산업디지털주권부(BERCY: Ministère de l'Économie, des Finances et de la Souveraineté industrielle et numérique), 고등교육연구부, 공공투자은행, 3IA 클러스터 ※ 국립디지털과학기술연구소, 국립과학연구원, 프랑스 원자력 및 대체에너지 위원회는「프랑스2030」 연구 연계 틀 안에서 각 단계 지속 관여
	키워드	• 산업 확산, 생성형·친환경 AI, 중소기업 지원

구분	주요 내용	
제3단계 (2025~)	시점	• 2025년 2월 6일, 범정부 AI 위원회에서 「프랑스 2030」과 연계한 「국가 AI 전략」 제3단계 발표
	목표	• 전 사회·산업 분야로 AI 활용 확대, 기술 주권과 신뢰성 강화
	투자 규모 /분야	• 민간 1,090억 유로 규모 AI 투자 유치 계획 발표 : 2025년 2월 파리 AI 액션 서밋(AI Action Summit) 직전 대통령 발표 • EU의 인베스트AI(InvestAI) 펀드(2천억 유로 규모) 및 공익을 위한 AI를 목표로 하는 국제협력 프로젝트인 커런트 AI(Current AI) 등 국제 이니셔티브와 연계해 프랑스 AI 생태계를 세계적 수준으로 견인
	주요 내용	• 프랑스를 글로벌 AI 강국으로 도약시키기 위한 4대 우선 과제 제시 ① AI 인프라 확충 : 전국 35개 부지에 초대형 데이터센터 구축, 고성능 컴퓨팅(HPC) 인프라 확대 ② 인재 양성·유치 : 세계 최고 수준 AI 인재 프랑스 유치, 교육·산학 협력 강화 ③ 산업·행정 AI 활용 가속화 : 공공행정서비스 혁신, 산업별 AI 통합 촉진 ④ 신뢰할 수 있는 AI 구축 : 윤리·투명성·책임성을 기반으로 한 '신뢰할 수 있는 AI(Trustworthy AI)' 구현 • 프랑스 전역에 초대형 데이터센터 구축부지 35곳 선정 • 공공행정의 AI 혁신 적용 • 최고 수준 인재의 프랑스 유치
	주요 집행 주체	• 범정부 AI 위원회가 중심적 거버넌스 역할, INESIA(Institut national pour l'evaluation et la securite de l'intelligence artificielle, 국립 AI 평가 및 안전 연구소) 출범, 개인정보보호위원회, 공공투자은행 등
	키워드	• 디지털 주권, AI 인프라, 신뢰성, 국제 리더십

출처 : 프랑스 2030- 국가 AI 전략 외 정부 발표 및 기사 종합(2025.11.9. 검색[119)120)]), SPRi 정리

3) AI 혁신위원회 권고안[121)]

프랑스 정부는 2023년 9월 AI 혁신위원회(Commission de l' intelligence artificielle)를 설치하고, 프랑스가 AI 선도국으로 도약하기 위해 실행가능한 정책·규제 권고를 마련하도록 위임했다. AI 혁신위원회는 6개월의 조사 끝에 2024년 3월, 25개 권고를 담은 최종 보고서 'AI : 프랑스에 대한 우리의 야망(IA : Notre ambition pour la France)'을 제출했다. 동 권고안은 프랑스의 경제·주권·경쟁력, 공공서비스, 윤리·신뢰를 아우르는 국가 차원의 권고를 제시하고 있다.

119) entreprises.gouv.fr, France 2030 : stratégie nationale pour l'intelligence artificielle, 2024.9.20./2025.2.12.
120) Service d'information du Gouvernement, IA : une nouvelle impulsion pour la stratégie nationale, 2025.2.12.
121) economie.gouv.fr, IA : Notre ambition pour la France, 2024.3.15.

<표 3-36> AI 혁신위원회 25개 권고 요약

번호	주요 내용	
1	AI의 사회적 수용 기반 조성	사회 전체가 AI를 이해하고 활용할 수 있도록 교육, 공공 토론, 시민참여를 체계화하여 '집단적 수용(appropriation)' 문화 형성
2	AI 영향 연구 및 관측체계 강화	AI가 고용, 생산성, 환경, 정보 생태계에 미치는 영향을 지속적으로 관측하고 정책 설계에 활용할 연구 인프라 구축
3	노사 및 사회적 대화 체계화	산업별·직능별로 AI 도입 과정에서 사회적 대화와 공동 설계를 제도화하여 신뢰 기반의 AI 도입 모델 정착
4	공공부문 AI 활용 역량 강화	공무원과 공공기관 인력의 AI 이해·활용 능력을 높여 행정 효율성과 시민 서비스 품질 제고
5	의료 분야의 증강의료 실험	의료진·환자 협업 구조에 맞춘 실용형 AI 도구(예: 개인 메모 보조)를 시험·확대해 진료 품질과 효율성 제고
6	교육 분야 AI 도입 및 평가 체계 구축	교사·학생이 함께 생성형 AI를 실험적으로 도입하고 학습효과·공정성·안전성에 대한 평가 프로토콜 마련
7	민간투자 및 혁신펀드 연계 강화	AI·딥테크(Deep Tech) 기업 지원을 위한 '프랑스 & AI(France & IA)' 등 투자 펀드를 조성하여 민간 자본을 혁신산업으로 유도
8	컴퓨팅 주권 확보 및 연산 허브 구축	대규모 AI 학습을 지원할 유럽 내 자립형 컴퓨팅 인프라를 구축해 디지털 주권 강화
9	데이터 접근 및 활용 절차 간소화	보건·문화 등 공공데이터의 접근성과 상호운용성을 높여 혁신 활용과 개인정보 보호 조화
10	AI 학습데이터 투명성·표준화 강화	데이터 출처 공개, 저작권 보호, 표준화된 학습 데이터 구축을 통해 신뢰할 수 있는 데이터 시장 조성
11	공공·민간 데이터 거버넌스 구축	데이터 공유와 보호를 병행하는 협력형 거버넌스 모델을 도입해 공공이익 데이터 활용 확대
12	AI 전문·융합인재 양성 및 유치	다학제 기반의 AI 교육, 국제 연구자 유치 프로그램을 확충해 인재 생태계 강화
13	AI 산업화 및 중소기업 전환 촉진	산업별 AI 도입을 가속하고, 중소기업의 AI 활용 역량을 높이는 지원정책 추진
14	글로벌 AI 거버넌스 구상 추진	프랑스가 주도하는 국제 AI 거버넌스 이니셔티브를 구축해 국가·기업·시민사회 간 협력 제도화
15	국제 공익 AI 연구기금 설립	'국제 공익 AI 연구기금(International Fund for Public Interest AI)'을 설립해 비영리·공익 목적의 AI 연구와 서비스 개발 지원
16	글로벌 AI 연대 메커니즘 도입	전 세계적 디지털 격차 해소를 위해 컴퓨팅 투자 중 일정 비율을 개도국에 재투자하는 '1% AI 연대기금' 도입
17	AI 평가·감사 체계 구축	국립 AI 평가 및 안전 연구소를 중심으로 한 평가 생태계를 조성해 AI 시스템의 성능, 안전성, 신뢰성을 공공·민간 협력으로 검증
18	시장 경쟁정책의 AI 대응 강화	플랫폼 집중을 감시하고, 시장의 변화에 맞춰 실시간으로 규제와 경쟁정책을 업데이트하는 동적 경쟁정책(dynamic competition policy)으로 공정한 AI 시장 질서 유지
19	오픈소스·투명한 AI 생태계 육성	연구자와 시민이 접근가능한 개방형 AI 모델·도구를 지원해 신뢰성 있는 기술 생태계 조성
20	공공연구의 자율성 확대	AI 연구 관련 행정절차를 간소화하고, 공공연구의 자율적 실험 환경 조성
21	지속가능한 AI 개발 촉진	에너지·자원 효율적 AI 모델과 저전력 하드웨어를 지원해 친환경적 AI 생태계 구축
22	국제협력 및 다자 플랫폼 강화	'글로벌 AI 파트너십(GPAI: Global Partnership on AI)' 등 국제협력체를 통한 규범·윤리·기술 교류를 활성화해 프랑스의 글로벌 리더십 강화
23	AI 윤리·책임·법제 정비	AI 투명성, 책임성, 저작권 보호를 반영한 법제 개정을 통해 사회적 신뢰 기반 강화
24	정기적 AI 정책·기술 평가 체계화	AI 기술 발전과 사회 영향에 대한 정기 평가를 실시해 정책을 연 1~2회 갱신·보완
25	국가 AI 투자계획 실행	향후 수년 동안 약 50억 유로 규모의 국가 AI 투자계획을 통해 디지털 주권과 혁신 역량 확보

출처 : IA : Notre ambition pour la France(2024.3.), SPRi 정리

AI 혁신위원회의 25개 권고는 기술·산업·사회·윤리·문화 전반에 걸친 입법 및 정책 제안을 담고 있으며, 프랑스가 AI 혁신에서 주도적 위치를 확보하기 위해 필요한 제도적 기반을 제시하고 있다. 이는 정책·규제 관점에서 ▲연산 인프라 및 기술 주권, ▲인재 및 기술 기반 강화, ▲경제 전반의 대규모 도입, ▲신뢰 및 책임 프레임워크 구축, ▲프랑스어·문화·저작권 보호, ▲지속가능한 AI라는 여섯 가지 전략 축으로 분류할 수 있다. 특히 지속가능성은 모든 정책 축에 걸쳐 적용되는 횡단적·전략적 원칙으로 강조되고 있다.

AI 혁신위원회의 권고 발표(’24.3) 이후, 프랑스 정부는 「국가 AI 전략」 제3단계를 발표(’25.2)하며 권고의 골자를 정책화했다. 권고의 주요 내용은 「국가 AI 전략」 제3단계의 4대 우선 과제와도 연결된다. 「프랑스 2030」의 재원으로 추진하는 제3단계는 ▲연산 인프라를 포함한 AI 인프라 확충 ▲인재 양성·유치 ▲산업·행정 AI 활용 가속화 ▲신뢰성 있는 AI 구축이라는 4대 우선 과제를 제시하고 있는데, 이는 AI 혁신위원회 권고의 주요 내용을 반영하고 있다.

<표 3-37> 주요 권고와 「국가 AI 전략」 제3단계 연계

AI 혁신위원회 권고 6가지 축('24.3)	「국가 AI 전략」 제3단계 우선 과제('25.2~)
1. 연산 인프라 및 기술 주권	4대 우선 과제 중 ① AI 인프라 확충 : 전국 35개 부지에 초대형 데이터센터 구축, 고성능 컴퓨팅 등 연산 인프라 강화
2. 인재 및 기술 기반 강화	② 인재 양성·유치 : 교육·재교육 및 산학 협력 강화, 해외 인재 유치 재원 확대
3. 경제 전반의 대규모 도입	③ 산업·행정 AI 활용 가속화 : 공공행정서비스 혁신, 산업별 AI 통합 촉진, 「프랑스 2030」 기반 30개 프로젝트 가동
4. 신뢰 및 책임 프레임워크 구축	④ 신뢰성 있는 AI 구축 : 윤리·투명성·책임성을 기반으로 한 신뢰할 수 있는 AI 구현
5. 프랑스어·문화·저작권 보호	제3단계 문서 내 문화·저작권 연계 방향 언급
6. 지속가능한 AI	저탄소·절약형 AI, 에너지 효율성, 친환경 인프라 언급

출처 : IA : Notre ambition pour la France(2024.3.) & 프랑스 2030-국가 AI 전략(2025.2.)

2. 2025년 프랑스 인공지능 정책

1) 관련 조직 신설·개편

(1) 국립 AI 평가 및 안전 연구소(INESIA)[122]

국립 AI 평가 및 안전 연구소(INESIA: Institut national pour l’evaluation et la securite de l’intelligence artificielle)는 2025년 1월 프랑스 정부가 새롭게 출범시킨 AI 평가·안전 전문기관으로, AI 기술의 신뢰성과 안전성 확보 및 국가 차원의 리스크 평가 전략을 실현하기 위해 설립되었다. 별도 법인 설립이 아닌, 기존 연구 네트워크의 연합형 공공연구 네트워크 형태로, 기술 평가 및 규제 지원 기능을 수행한다. 국가사이버보안국(ANSSI: Agence nationale de la securite des systemes d’information), 국립디지털과학기술연구소(INRIA: Institut national de recherche en informatique et en automatique), 국립 측정·시험연구소(LNE: Laboratoire national de métrologie et d’essais), 디지털플랫폼 규제전문센터(PEReN: Le Pole d’Expertise de la Regulation Numerique) 등 AI 신뢰성·안전성 평가를 담당하는 기존 주요 기관들과의 협업을 통한 AI 평가 및 안전 연구를 수행하며, AI 모델 및 시스템의 성능, 안전성, 신뢰성 조사·검증, AI 규제 및 거버넌스 정책 지원 및 기술적 자문을 제공한다. 프랑스가 추진하는 ‘안전하고 포용적인 AI’ 프레임워크 및 거버넌스 전략과 맞물리며, 향후 AI 거버넌스 체계 내에서 기술 평가·감사·안전성 검증을 지원하는 역할을 할 것으로 예상된다.

(2) AI·디지털위원회(CIAN)[123]

2025년 7월 프랑스 정부는 기존 국가디지털자문기구(CNNum: Conseil national du numérique)를 개편해 AI·디지털위원회(CIAN: Conseil de l’Intelligence Artificielle et du Numérique)를 출범시켰다. AI·디지털위원회는 총리실 산하의 국가 자문기구로, 경제재정산업디지털주권부가 협력 지원·운영 형태로 참여하여 정책·산업 관련 자문을 공동 수행한다. AI·디지털위원회로의 개편은 AI·디지털 의제를 통합적으로 다루고, 공익·윤리·산업 정책 간 균형을 강화하기 위해서이다.

AI·디지털위원회는 정부 대표 외에 학계·산업계·시민단체를 포함한 민간 전문가가 참

122) economie.gouv.fr, La France se dote d’un Institut national pour l’évaluation et la sécurité de l’intelligence artificielle (INESIA), 2025.2.3.
123) economie.gouv.fr, Un nouveau Conseil national de l’IA et du Numérique pour éclairer les décisions publiques, 2025.7.24.

여해 시민사회 참여를 제도화했다는 점이 특징이다. 주요 기능으로 ▲AI 및 디지털 정책 관련 정부 자문 및 권고안 제시, ▲AI 윤리·데이터 보호·공정성 관련 사회적 합의 도출, ▲AI 산업정책 및 교육·고용 정책과의 연계 검토, ▲공공부문 AI 활용 가이드라인 자문 및 국제협력 정책 조언 등을 진행한다.

2) 파리 'AI 액션 서밋' 주최

글로벌 AI 거버넌스 주도를 위한 노력의 연장선상에서, 프랑스 정부는 2025년 2월 10~11일 파리 그랑팔레(Grand Palais)에서 파리 'AI 액션 서밋(AI Action Summit)'을 주최하며, 글로벌 AI 거버넌스 논의의 중심 국가로 부상하였다. 이번 서밋은 엘리제궁(Elysee Palace) 및 외교부(Ministere de l' Europe et des Affaires etrangeres) 주도로 개최되었으며, 인도와의 공동의장국 체제로 진행되었다. 100여 개국 정부 대표, 국제기구(OECD, UNESCO, UNDP 등), 주요 기술기업, 시민사회단체가 참여하여 '공익을 위한 AI(AI for the Public Interest)'를 주제로 논의했다.[124] 파리 'AI 액션 서밋'은 2023년 영국에서 개최된 'AI 안전 정상회의(AI Safety Summit)'와 달리, 안전(safety) 중심의 위협 관리에서 행동(action) 중심의 실행·혁신 모델로의 전환을 선언했다. 에마뉴엘 마크롱 대통령은 개막 연설에서 "AI는 인류 전체의 자산이며, 공공의 이익을 위해 관리되어야 한다"고 강조하며 '혁신과 책임의 균형'을 프랑스 AI 개발의 핵심으로 제시했다.[125] 이번 서밋에서의 주요 논의 및 발표는 아래와 같다.

〈표 3-38〉 파리 'AI 액션 서밋' 주요 논의

구분	주요 내용
포용적·지속가능한 AI 선언(Statement on Inclusive and Sustainable AI) 채택	인권, 기후, 교육, 보건 등 공익 영역에서 AI를 윤리적이고 지속가능하게 활용하기 위한 다자간 원칙을 확정
글로벌 AI 거버넌스 플랫폼 설립 제안	프랑스는 OECD, UNESCO 등과 연계하여, 각국의 AI 규제·표준을 조정하는 'AI 거버넌스 글로벌 파트너십(Global Partnership on AI Governance)' 출범을 제안
혁신 프로젝트 및 투자 촉진	프랑스 정부는 민간 및 국제 투자 유치 규모를 1,090억 유로로 발표하고, 데이터센터·AI 반도체·신뢰할 수 있는 AI 시스템 분야를 핵심 우선 영역으로 제시
AI 민주화·교육 확대	공공부문, 지방정부, 시민사회와 연계한 AI 리터러시 교육 및 시민참여형 AI 거버넌스 추진 의지 표명

출처 : 파리 AI 액션 서밋 발표자료(2025.2.), SPRi 정리

124) diplomatie.gouv.fr, AI Action Summit (10 & 11 February 2025), 2025.2.11.
125) Élysée, Statement on Inclusive and Sustainable Artificial Intelligence for People and the Planet, 2025.2.11.

파리 ‘AI 액션 서밋’은 프랑스 AI 전략의 국제적 확장판으로 볼 수 있다. 자국 내에서는 「프랑스 2030」 및 ‘프랑스 AI 강국 조성’ 전략을 추진하며 AI 거버넌스 논의를 확장시켰고, 외교적으로는 EU 규범 중심의 신뢰할 수 있는 AI 모델(European Trustworthy AI)을 국제표준으로 확산시키는 역할을 했다. 프랑스는 향후 OECD가 운영하는 AI 정책 분석 및 데이터 플랫폼(OECD AI Policy Observatory) 및 ‘글로벌 AI 파트너십’ 등과의 연계를 강화하여, 산업 경쟁력·윤리·공공성의 조화를 이루는 글로벌 AI 거버넌스 모델을 발전시켜 나갈 계획이다.

3) 커런트 AI 출범[126)]

‘커런트 AI(Current AI)’는 2025년 2월 파리 ‘AI 액션 서밋’에서 공식 출범한 공익 지향적 글로벌 AI 협력 이니셔티브이다. 정부, 재단, 기술기업 등이 참여하여 ‘공익을 위한 신뢰할 수 있는 AI(Open and Trustworthy AI for the Public Interest)’를 실현하는 것을 목표로 한다. 초기 조성 규모는 약 4억 달러이며, 프랑스 정부를 비롯해 맥아더재단(MacArthur Foundation), 맥거번재단(Patrick J. McGovern Foundation), 구글 딥마인드(Google DeepMind), 세일즈포스(Salesforce) 등이 핵심 창립 파트너로 참여했다.

‘커런트 AI’는 비영리 글로벌 네트워크로 조직되어 있으며, 이사회, 위원회, 사무국 구조로 운영된다. ▲이사회(Board)-독립된 전문가로 구성되어 전략 결정 및 자원 배분 관리, ▲위원회(Council)-정부·시민사회 대표로 구성되어 정책 자문 및 윤리 감독 기능 수행, ▲사무국(Secretariat)-조직 운영 및 프로그램 실행을 총괄하는 역할을 한다. 또한 프랑스 및 캐나다 정부, OECD, UNESCO 등 국제기구와 연계된 협력 플랫폼을 통해 각국의 AI 공익 프로젝트를 조정한다.

‘커런트 AI’는 AI 기술을 사회적 가치 창출 수단으로 활용하기 위해 아래의 중점 영역을 추진한다. 이러한 프로그램은 2025년 2월 11일 발표된 ‘파리 AI 공익 헌장(The Paris Charter on AI in the Public Interest)’의 핵심 가치에 따라 운영되며, AI가 인간과 지구의 이익을 위해 개발·활용되어야 한다는 원칙을 명시하고 있다.

126) Élysée, The Paris Charter on Artificial Intelligence in the Public Interest, 2025.2.11. (The Paris Charter on Artificial Intelligence in the Public Interest)

<표 3-39> 커런트 AI의 중점 영역

구분	주요 내용
데이터 및 개방성(Data & Openness)	공익 데이터 공유 및 오픈소스 AI 도구 확산
책임 및 신뢰성(Accountability & Safety)	투명성 기준, 감사 체계 및 책임 있는 거버넌스 구축
언어 및 문화 다양성(Linguistic Diversity)	다언어 AI 데이터셋 개발 및 비영어권 접근성 확대
보건·기후 응용(Health & Climate)	AI를 활용한 보건의료 및 기후 지속가능성 프로젝트 추진
시민참여(Public Engagement)	시민사회 참여형 AI 정책 모델 및 윤리 교육 프로그램 운영

출처 : 커런트 AI 웹사이트(2025.11.9. 방문)[127]

4) 개인정보보호위원회 '2025-2028 전략 계획'[128]

2025년 1월, 프랑스 개인정보보호위원회(CNIL: Commission Nationale de l'Informatique et des Libertes)는 '2025-2028 전략 계획(Plan stratégique 2025-2028)'을 발표했다. 동 계획은 디지털 환경의 급속한 진화에 따라 책임감 있는 AI 기술 활용과 데이터 보호의 균형 확보를 핵심 과제로 설정하고, AI 전담 부서 설립 및 생성형 AI 규제 가이드라인 마련 등과 같은 계획을 포함하고 있다. 이와 관련해 CNIL은 AI 기술의 복잡성과 사회·경제적 영향력 증대에 대응하기 위해 조직 내 '서비스 AI' 부서를 신설하고, 경제분석팀을 별도로 두기로 했다. 관련 AI 부서는 특히 AI 시스템의 개인정보 침해 위험, 투명성 기준, 규제 대응 등을 분석·감독하는 역할을 맡는다.

2025-2028 전략 계획은 ▲윤리적이고 권리를 존중하는 AI 육성 ▲아동청소년 대상 디지털 보호 강화 ▲개개인의 사이버보안 인식 확산 및 디지털 신뢰도 제고 ▲일상적인 디지털 서비스의 이용자 보호 조치 강화라는 4대 축을 중심으로 설계되었다.

127) Current AI (2025.11.9. 방문) https://www.currentai.org
128) CNIL, 2025-2028 PLAN STRATÉGIQUE, 2025.1.6.

<표 3-40> 프랑스 개인정보보호위원회 ‘2025-2028 전략 계획’ 4대 축

4대 축	주요 내용
윤리적이고 권리를 존중하는 AI 육성	• AI 개발 및 사용에 대한 지식·전문성 공유 • AI 관련 법률 체계 명확화 및 균형 잡힌 규제 적용 • AI 시스템의 준수성 감시 및 통제 강화 • 공정·투명한 AI 설계와 개인정보보호 강화 기술 장려
아동청소년 대상 디지털 보호 강화	• 아동·청소년 대상 온라인 서비스 감독 강화 • 이용자 연령 및 권리 행사 지원 시스템 구축 • 맞춤형 개인정보보호 인터페이스 도입 • 청소년 대상 디지털 교육 및 인식 제고
개개인의 사이버보안 인식 확산 및 디지털 신뢰도 제고	• 개인·기관 대상 보안 인식 제고 및 대응 능력 강화 • 보안 위협 예방·대응을 위한 기술·교육 지원 • 보안 규범 준수 감시 및 제도적 제재 강화 • 개인정보보호와 사이버보안의 통합적 접근
일상적인 디지털 서비스의 이용자 보호 조치 강화	• 모바일 앱 및 디지털 플랫폼의 이용자 보호 준수 강화 • 개인 데이터와 사용자 인터페이스 투명성 확보 • 디지털 ID 시스템의 프라이버시 준수 보장 • 모바일 앱과 개인정보처리 방식에 대한 감독 및 감시

출처 : 개인정보보호위원회(CNIL) 2025-2028 전략 계획(2025.1.), SPRi 정리

CNIL 2025-2028 전략 계획에는 향후 AI 시스템의 라이프사이클별 준수성 검토 기준과 감시 기술 아키텍처 개발을 명시하고 있다. 아울러 2025년 하반기에 AI 가치사슬 내 참여 주체(모델 설계자, 통합자, 재사용자 등)의 책임 범위를 명확히 규정하는 권고안 발표를 계획 중이며, 이는 공개 토론을 거칠 예정이다. 또한, EU 개인정보보호이사회(EDPB: European Data Protection Board) 및 타 유럽기관과의 협력 강화를 통해 유럽 차원의 AI 관련 규제 기준을 수립하고, 데이터 보호 기준을 국제적으로 확산하는 역할에도 주목하고 있다.[129]

동 전략 계획의 일환으로 2025년 6월 공개된 ‘PANAME(Privacy AuditiNg of AI Models)’ 프로젝트는 AI 모델의 개인정보 처리 적합성을 점검하는 감사 도구 개발을 목표로 한다. CNIL은 디지털플랫폼 규제전문센터(PEReN), 국가사이버보안국(ANSSI) 등과 협력해 GDPR 규정 준수를 위한 감사 모델 개발 프로젝트를 지휘하고 권장사항을 마련할 계획이다.[130]

129) CNIL, CNIL publishes its European and international strategy for 2025-2028, 2025.4.14.
130) CNIL, IA : la CNIL finalise ses recommandations sur le développement des systèmes d’IA et annonce ses futurs travaux, 2025.7.22.

5) ‘프랑스 과학/AI 연구진 선택’ 프로그램[131)]

프랑스 정부는 2025년 5월 발표한 ‘프랑스 과학/AI 연구진 선택(Choose France for Science / AI Faculty)’ 프로그램을 통해 세계적인 연구자를 프랑스로 유치하고 있다. 동 프로그램은 국립디지털과학기술연구소가 주관기관이며, 참여하는 대학·연구소가 공모를 통해 해외 연구자를 채용하면 해당 연구자에게 연구팀 구축, 연구비, 정착 지원 등을 제공하는 방식으로 운영된다. 정착 지원에는 가족 동반 이주 지원, 배우자 취업 지원, 생활 정착 서비스 등이 포함되어 연구자 유치 후 장기적인 활동이 가능하도록 설계되어 있다. 연구 주제는 AI, 사이버보안, 양자컴퓨팅, 컴퓨터과학 등 프랑스가 전략적으로 강화하고자 하는 분야와 연계되어 있다. 단일 연구자 유치를 넘어 프랑스 내 연구기관이 국제적으로 경쟁력 있는 기관으로 성장하기 위한 정책이기도 하다.

6) AI 확산 전략 ‘Osez l’IA’[132)]

2025년 7월 1일 프랑스 디지털·AI담당 국무장관 클라라 샤파즈는 「국가 AI 전략」의 일환으로 프랑스 기업들에서 AI 활용을 확산하기 위한 전략인 ‘Osez l’IA’를 발표했다. 동 전략은 2030년까지 프랑스 대기업의 100%, 중견·중소기업의 80%, 초소형기업의 50% 이상이 AI를 활용하도록 만드는 것을 목표로 한다. 프랑스는 이를 위해 ▲인식 제고 ▲역량 강화(교육) ▲지원의 세 가지 축을 중심으로 기업에서의 AI 확산을 장려할 방침이다.

<표 3-41> 프랑스 ‘Osez l’IA’ 전략 주요 내용

3대 축	주요 내용
인식 제고	• 300명의 AI 홍보대사로 구성된 국가 네트워크를 운영해 인식 제고 활동 수행 • AI 홍보대사는 ‘일반 대중’ 홍보대사(기업인, 운동선수, 예술가 등)와 산업 및 지역 홍보대사들로 구성
역량 강화	• 2025년 말까지 맞춤형 콘텐츠를 제공하는 공공 교육 플랫폼인 ‘AI 아카데미’ 출범 • ‘AI 아카데미’ 등을 통해 2030년까지 1,500만 명의 직업인을 대상으로 AI 교육 제공 목표
지원 (기업 AI 실행 지원)	• 중소기업 대상 진단 지원, 보조금, 국가지원 보증 대출 제공 • AI 솔루션 및 부문별 사용 사례 카탈로그 제공 • 새로운 프랑스 솔루션 개발에 대한 지원 강화 프로그램 운영

출처 : Osez l’IA(2025.7.), SPRi 정리

131) INRIA, Researchers in AI, quantum, cybersecurity and computer science: choose France for your projects!, 2025.6.10.

132) Gouvernement de la République française, Osez l'IA : le plan pour diffuser l'IA dans toutes les entreprises, 2025.7.1.

7) AI 선도 프로젝트 공모사업[133)]

2025년 9월에는 「프랑스 2030」 전략의 일환으로 AI 혁신의 전환점이 될 기술을 육성하기 위해 ‘AI 선도 프로젝트(Pionniers de l’intelligence artificielle)’ 공모 사업이 발표되었다. 이 사업은 정부의 「국가 AI 전략」 제3단계 추진과 연계되며, 획기적인 혁신 기술을 조기에 발굴하고 확장하는 것을 목표로 한다.

운영 주체는 총리실 산하의 투자총괄사무국(SGPI: Secrétariat général pour l’investissement), 경제재정산업디지털주권부 산하의 기업청(DGE: Direction générale des entreprises), 고등교육연구부 산하의 국가연구·혁신청(DGRI: Direction générale de la recherche et de l’innovation) 및 환경·에너지부 산하의 지속가능발전총감독청(CGDD: Commissariat général au développement durable) 등이며, 실제 운영은 공공투자은행과 국립디지털과학기술연구소가 디지털 프로그램 기관으로 참여한다.

동 사업은 기업 또는 연구센터 단독 또는 컨소시엄 형태로 신청할 수 있으며, 기술 파괴적 특성을 가진 프로젝트를 우대한다. 중점 적용 분야는 산업, 에너지, 정보 시스템 보안, 바이오메디컬 및 헬스케어, 생태전환 등 전략성이 높은 영역이다. 기술 유형 측면에서는 새로운 AI 아키텍처, 혁신적 학습 방법, 모델과 환경의 상호작용, 특히 로보틱스와 통합되는 AI를 포함하는 응용 기술 개발을 목표로 하고 있다.

프로젝트 운영은 3단계(선택적 승급 방식) 구조로 구성되며, 각 단계는 성과 기준에 따라 다음 단계 진입이 결정된다. 각 단계 간 이동은 중간 평가를 통해 결정되며, 운영 기관들 간 교차 전문성 평가를 거쳐 우수 프로젝트가 선정된다. 1단계의 지원자는 공공 투자은행 또는 디지털 프로그램 기관 중 하나와 계약을 체결하며, 이후 2단계, 3단계는 공공 투자은행이 국가를 대신하여 관리한다. 유럽 파트너가 포함된 컨소시엄은 커런트 AI 재단에서 추가 자금을 받을 자격이 있다.

AI 선도 프로젝트는 지원 규모가 상당하여 대형 자본이 필요한 AI 프로젝트가 참여 가능하며, 리스크가 높은 연구형 성격의 프로젝트도 도전할 수 있는 진입 통로를 제공한다. 또

133) INRIA, France 2030 | Lancement de l’appel à projet « Pionniers de l’intelligence, 2025.9.18.

한, 운영 구조는 혁신 기술을 조속히 시장화할 수 있는 메커니즘을 제공하며, 프랑스의 AI 주권 확보 전략과도 연결된다.

8) AI와 보건의료 데이터 전략[134)]

프랑스 보건복지부는 2025년 7월 1일 'AI와 보건의료 데이터 전략(Stratégie IA et données de santé)'을 발표했다. 이는 AI 개발과 보건의료 데이터의 2차적 활용을 연계하는 국가 전략으로, 궁극적으로 프랑스를 AI 기반 의료기기, 디지털 의료 도구, 임상 연구, 데이터 기반 제약 혁신을 위한 선도적인 환경으로 자리매김할 수 있도록 만드는 것을 목표로 하며 보건의료 데이터 활용을 가로막는 장벽을 제거하는 데에 초점을 맞추고 있다. 동 전략은 이를 위한 범부처 협력 및 다년간의 전략 계획을 담고 있으며, 특히 2028년까지 프랑스에서 유럽건강데이터공간(EHDS: European Health Data Space)을 구현하기 위해 필요한 준비 작업을 포함하고 있다. EU 이사회는 2025년 1월 21일 '유럽건강데이터공간(EHDS)에 관한 규정(EU 2025/327)'을 채택했는데, 동 규정은 EHDS 체계 아래, EU 시민이 모국이 아닌 다른 회원국에서도 의료기록 열람 등과 같은 자신의 건강데이터에 보다 신속하고 용이하게 접근할 권리를 가질 수 있도록 ▲EU 회원국 간 건강 데이터의 원활한 교환 ▲ 개인의 건강 데이터에 대한 접근성 및 통제권 강화 ▲연구 및 혁신 목적의 데이터 활용 촉진 ▲EU 내 디지털 건강 서비스와 제품의 단일시장 형성 등을 목표로 각 회원국에 관련 작업을 지시했다.

동 전략은 1차와 2차에 걸쳐 발표되었는데, 먼저 7월 공개된 1차 전략은 '데이터'에 초점을 맞춰 ▲신뢰와 투명성 강화 ▲데이터베이스 구축 ▲데이터 공유 및 접근성 증진 ▲데이터 접근 간소화의 4가지 축을 중심으로 보건의료 데이터에 대한 접근과 활용을 현대화하기 위한 토대를 마련하는 데에 중점을 두었다.

134) Gouvernement de la République française, Stratégie IA et données de santé, 2025.7.1.

〈표 3-42〉 프랑스 ‘AI와 보건의료 데이터’ 1차 전략 주요 내용

4대 축	주요 내용
신뢰와 투명성 강화	• 프랑스 건강 데이터 생태계에 새로운 거버넌스 구조 구축 • 프랑스 건강데이터허브(HDH)[135]를 통한 EHDS 구축 사업(HealthData@EU, TEHDAS 1·2) 추진 • 국민들이 프랑스 국가건강데이터시스템(SNDS) 데이터에 대한 권리를 행사할 수 있는 기반 마련
데이터베이스 구축	• 데이터 재사용을 위한 데이터베이스의 보강, 상호운용성 확보 • 데이터 공유에 대한 원칙 명확화
데이터 공유 및 접근성 증진	• 프랑스의 국가적인 데이터베이스 및 메타데이터 디렉토리와 EU 규정에 부합하는 경제 모델 개발을 통해 전체 생태계에 도움이 되는 방식으로 데이터 공유 촉진·규제
데이터 접근성 간소화	• 데이터 접근에 대한 행정절차 간소화 및 속도 개선 • 조화로운 원칙 구현

출처 : Stratégie IA et données de santé(2025.7.1.), SPRi 정리

이어서 보건복지부는 2025년 11월 12일 동 전략의 두 번째 파트인 ‘AI’ 전략을 발표했다. 이는 AI를 통해 의료 시스템을 혁신하고 개선하기 위한 것으로, 구체적으로 ▲전문성을 바탕으로 한 명확한 구현 프레임워크 구축 ▲AI 시스템에 대한 적절한 평가 및 분석 ▲AI가 전체 보건의료 생태계에 미치는 영향 예측 등을 목표로 설정했다. 동 전략은 프랑스가 의료 분야에 AI를 도입하기 위해 이미 시작한 여러 구체적인 사업들을 기반으로 하고 있는데, 대표적으로 ‘SHAIPED 프로젝트’가 있다. 이는 AI 기반 의료기기의 개발 및 배포를 위한 유럽 공간을 구축하는 것을 목표로 하는 EU 지원 프로젝트로, 프랑스 건강데이터허브(HDH: Health Data Hub)의 주도로 11개 회원국의 30개 파트너 기관이 참여하는 컨소시엄에 의해 2028년까지 추진된다. 의료 데이터 활용 협력을 활성화하고, AI 기반 의료기기의 시장 출시 기간을 단축하고, 이러한 기기의 임상 효과를 검증해 유럽의 의료 시스템에 통합하는 데에 방점을 두고 있다.

135) 국가건강데이터시스템(SNDS)의 일부로, 각 기관 데이터를 카탈로그화해 중앙서버에 복사·저장한 뒤, 승인된 연구자가 데이터에 접근할 수 있도록 지원

제8절 독일

1. 그간의 인공지능 정책 추진 경과

<표 3-43> 독일의 주요 AI 정책 추진 현황

시기	발표기관	정책
2018년 11월	연방정부	국가 AI 전략 발표 - “AI Made in Germany” 슬로건과 연구·기업·법제도 12개 분야 종합 계획으로 구성
2019년 11월	연방 교육연구부(BMBF) 등 주요 부처	AI 전략 중간 보고 - AI 전략 첫 번째 업데이트
2020년 12월	연방정부	AI 전략 두 번째 업데이트(Fortschreibung 2020) - 전문가·연구·지식이전·규제·사회 5대 우선 분야 제시
2023년 11월	연방 교육연구부	AI 행동계획 - 11개 실행 분야 제시
2025년 5월	연방정부	디지털국가현대화부(BMDS) 신설
2025년 6월	연방 정보보안청(BSI)	‘AI를 위한 SBOM’ 개념 문서 발표
2025년 7월	연방정부	하이테크 어젠다 발표 - AI 등 6대 하이테크 전략 분야 선정. 2030년까지 부가가치의 10%를 AI 기반으로 창출하고 AI 기가팩토리 구축 목표
2025년 9월	연방정부	EU AI법 시행령(독일 국내법) 장관 초안 발표 및 의견 수렴 - 연방네트워크청(BNetzA)을 독일 시장감시 기관으로 지정
2025년 10월	연방정부	연방정부 현대화 어젠다 발표 - AI 활용을 통한 범부처 및 공공행정 현대화·효율화 강조

1) 개요

2018년 독일은 “AI Made in Germany”라는 국가 슬로건 하에 최초의 AI 국가전략(KI-Strategie Deutschland)을 발표했다. 이 전략의 AI 지원 정책은 ▲연구 강화 ▲혁신 챌린지 ▲기업역량 강화 ▲스타트업 지원 ▲노동시장 변화 ▲인재양성 및 유치 ▲공공수요 및 정부서비스 혁신 데이터 활용 ▲촉진(Promotion) ▲법제도 ▲표준 ▲네트워크 ▲사회적 논

의 등 12개 분야로 구성되어 있는데 이를 통해 AI 생태계 조성을 목표로 하고 있다.[136]

또한 2019년 11월 정부는 AI 국가전략에 대한 실행 경과를 점검하고 향후 방향을 제시하는 중간보고서(Zwischenbericht)를 발표했다. 이 문서는 전략 시행 첫 해의 성과와 도전과제를 정리한 것으로, 당시에는 세 개 부처로 나뉘어 있었던 연방교육연구부(BMBF), 연방경제에너지부(BMWi), 연방노동사회부(BMAS)가 공동으로 작성했다.[137] 한편 독일 연방정부는 2020년 6월 발표한 '코로나19 경기부양책'에 따라 2025년까지 인공지능 분야 투자를 기존 30억 유로에서 50억 유로로 증액한 '인공지능 전략 2020 개정안(Fortschreibung KI-Strategie)'을 발표하면서 AI 정책을 지속적으로 지원하고 있다.[138]

2023년 11월 연방교육연구부(BMBF)는 지금까지 실행했던 AI 정책을 좀 더 확장하고 집중하도록 하기 위한 AI 행동계획을 발표했는데 이 계획은 현재 독일 AI 산업에서 가장 시급한 11개 실행 분야(▲연구 기반 지속 강화 ▲새로운 관점 및 접근법을 위한 연구 아젠다 설계 ▲목표 지향적 AI 인프라 확장 ▲AI 전문성 강화 ▲유럽 및 국제 파트너십 강화 등)를 제시하고, 이미 시행한 조치와 향후 단기·중기적으로 시행해야 할 조치도 함께 포함하고 있다.[139] 연방교육연구부(BMBF)는 16억 유로를 AI에 투자했으며 특히 현재 진행 중인 50개 AI 관련 진흥 프로그램을 더해 최소 20개의 새로운 이니셔티브를 추가로 시작했다.[140]

136) German GOV, Fünf Jahre KI-Strategie – wo steht Deutschland? (2025.10.25. 방문) https://www.ki-strategie-deutschland.de
137) die Bundesregierung, Zwi-schen-be-richt: Ein Jahr KI-Stra-te-gie, 2019.11.15.
138) BMWK, Strategie Künstliche Intelligenz der Bundesregierung Fortschreibung 2020. 2020.12.2.
139) BMWK, BMBF-Aktionsplan "Künstliche Intelligenz", 2023.11.
140) German GOV, Artificial intelligence: a key technology (2025.10.25. 방문) https://www.tatsachen-ueber-deutschland.de/en/research-and-innovation/research-and-innovation

2. 2025년 인공지능 정책 동향

2025년 7월 독일은 하이테크 어젠다(Hightech Agenda Deutschland)를 채택하면서 2030년까지 AI가 독일 전체 경제(GDP)의 10%를 담당하도록 하겠다는 가속화 정책을 발표했다.[141] 이 발표에는 2027년까지 운영될 유럽연합(EU) 내 대용량 처리 센터 건설 입찰에 독일 정부가 도이치텔레콤(Deutsche Telekom)과 협력해 최소 하나의 AI '기가팩토리'를 독일 내에 건설하는 계획이 포함되어 있다. 또한 2024년 8월부터 EU 「AI 법(AI Act)」이 발효된 가운데 독일은 2025년 그 법적 이행을 준비하는 작업에 착수했고 「AI 법」 적용을 위한 국내법을 마련하기 위한 노력을 지속 중이다.

1) 관련 조직 신설·개편[142]

(1) 디지털국가현대화부(BMDS)

독일의 새 연방정부는 디지털과 국가현대화를 중심축으로 삼으며 연방정부 내에 연방디지털국가현대화부(BMDS: Bundesministerium für Digitales und Staatsmodernisierung)를 신설하고 디지털 행정, 데이터 정책, AI·디지털 인프라 등을 여러 기존 부처(연방 내무부, 연방 경제·에너지부, 연방 사법 및 소비자보호부 등)에 흩어져 있던 기능을 한 곳으로 통합했다. BMDS는 총리실의 디지털 전략, 내무부의 연방 IT 시스템, 디지털 서비스 및 사이버보안, 경제부의 데이터 정책 및 디지털 경제 감독, 재무부의 정부 중앙 IT 공급업체(ITZBund) 및 연방 클라우드 운영, 그리고 법무부의 EU 「AI 법」 시행을 감독하는 역할을 수행한다. 이는 단일 컨트롤타워(BMDS)를 통해 부처 간 중복, 불일치, 책임소재 불명확 문제를 해소하고, 각종 AI 및 디지털 프로젝트의 기획-예산-집행까지 일관된 체계를 마련하기 위함이다. 또한 독일 연방정부는 디지털·AI 관련 책임과 권한이 집중됨으로써 정책 조정력과 실행 속도가 향상될 것으로 전망하고 있다.

(2) 독일 新내각에 따른 부처 변화

한편 주요 부처의 변화를 살펴보면 ▲연방 경제·에너지부(BMWE)에는 에너지 관련 기능이 추가됐고 ▲연방 연구·기술·우주항공부(BMFTR)는 우주항공 ▲연방 환경, 기후보호,

141) 독일 정부는 AI 공세(AI Offensive)라는 표현을 쓰는 등 AI 활성화에 강한 의지를 피력

142) euronews, Merz's government decree gives German digital ministry extensive powers, 2025.5.7.
Center for Data Innovation, Germany's New Digital Ministry Will Make or Break The Government's AI Ambitions, 2025.6.2.

자연보호 및 원자력 안전부(BMUKN)는 기후보호(Klimaschutz) 및 기후 관련 국제정책(internationale Klimapolitik) ▲연방 농업・식량・지역부(BMLEH)는 고향/지역 정체성(für Heimat) 기능 ▲연방 사법 및 소비자보호부(BMJV)는 BMUKN로부터 이관받은 소비자보호(Verbraucherschutz) 기능을 각각 추가했다.

<표 3-44> 독일 연방 정부 조직 변화

舊 연방부처명	설립일	新 연방부처명	新 연방부처의 특성
신설	2025년 5월 6일	BMDS(연방 디지털국가현대화부)	• AI 정책 핵심 부처
BMWK (연방 경제기후보호부)	2021년	BMWE(연방 경제·에너지부)	• 에너지 관련 기능 추가
BMBF (연방 교육연구부)	1998년	BMFTR(연방연구·기술·우주항공부)	• 우주항공(raumfahrt) 기능 추가
BMUV(연방 환경·자연보호·핵안전 및 소비자보호부)	2021년	BMUKN(연방 환경, 기후보호, 자연보호 및 원자력 안전부)	• 소비자보호(Verbraucherschutz) 부문 BMJV 이관 • 기후보호(Klimaschutz) 및 기후 관련 국제정책(internationale Klimapolitik) 기능 추가
BMEL(연방 식량·농업부)	2013년	BMLEH(연방 농업·식량·지역부)	• 고향/지역 정체성(für Heimat) 기능 추가
BMAS (연방 노동사회부)	2005년	유지	• 유지
BMI (연방 내무 및 지역부)	2021년	BMI(연방 내무부)	• 고향/지역 정체성 기능 BMLEH 이관
BMJ(연방 법무부)	2021년	BMJV(연방 사법 및 소비자보호부)	• 소비자보호 기능 추가

출처 : BMDS, BMWE, BMFTR 등 발표자료 종합(2025.10.), SPRi 정리

2) AI를 위한 SBOM 개념 문서[143)]

연방정보보안청(BSI: Bundesamt für Sicherheit in der Informationstechnik)은 AI 시스템의 복잡해진 공급망(supply chain) 구조가 새로운 보안·거버넌스 리스크를 초래하고 있다고 진단하고, 특히 단순히 소프트웨어 컴포넌트를 나열하는 전통적인 '소프트웨어 자재명세서(SBOM: Software Bill of Materials)' 만으로는 AI 시스템에 내재한 위험을 충분히 제어할 수 없다고 판단했다. 이에 BSI는 AI 시스템을 구성하는 요소들을 투명하게 기록하고 추적할 수 있게 하기 위해 'AI를 위한 SBOM(SBOM for AI)' 개념을 제시했다. 전통적인 SBOM은 소프트웨어를 구성하는 코드·라이브러리·패키지 등 기술적 요소를 식별하는 도구였으나, AI 시스템의 특수성 즉, 데이터 기반 학습, 모델 구조의 불투명성, 외부 컴포넌트 의존성 등으로 인해 기존 SBOM으로는 AI 공급망 리스크를 충분히 관리할 수 없다는 점에서 BSI는 전통적인 SBOM을 확장한 개념을 새로이 제안한 것이다.

BSI는 AI 시스템의 신뢰성 확보는 소프트웨어 수준이 아니라 데이터-모델-인프라 전체 공급망 수준에서 이뤄져야 한다고 강조하며, 'AI를 위한 SBOM'을 다음과 같이 정의했다. 먼저 'AI를 위한 SBOM'은 AI 시스템 개발과 운용에 관여하는 모든 구성요소의 출처·변경이력·보안속성을 명확히 하여 AI 공급망 전반의 투명성(transparency)과 추적가능성(traceability)을 확보하기 위한 것이다. 또한 AI 모델의 오염된 데이터셋, 비인가 학습방법, 취약 인프라 등으로부터 발생할 수 있는 보안 리스크를 조기에 식별하기 위한 것이다. 즉, 'AI를 위한 SBOM'은 AI 시스템의 신뢰 사슬(trust chain)을 문서화하는 일종의 디지털 이력서로 볼 수 있다. BSI는 'AI를 위한 SBOM'이 반드시 포함해야 할 최소한의 정보 항목을 7개 클러스터(cluster)로 구분했다. 각 항목은 기계판독(machine-readable) 방식으로 기록되고 디지털 서명(digital signature)을 통해 무결성을 보장한다. 이러한 구조를 통해 AI 시스템의 모든 구성요소를 문서화하고, 향후 문제가 발생했을 때 해당 요소의 출처와 변경이력을 추적할 수 있다.

143) Bundesamt für Sicherheit in der Informationstechnik, BSI-leitet G7-Arbeitsgruppe: Gemeinsames Konzept für eine "SBOM for AI" veröffentlicht, 2025.6.16.

<표 3-45> AI 시스템의 필수 기술정보 클러스터 7가지

구분	주요 내용
모델 정보(ModelInformation)	• 모델 이름, 버전, 구조, 제작자, 소유권 정보
학습 방법(TrainingMethods)	• 알고리즘, 하이퍼파라미터, 학습환경 등
데이터셋(Datasets)	• 데이터 출처, 품질, 전처리, 사용 목적
보안 속성(SecurityProperties)	• 검증 결과, 취약점 목록, 보안 업데이트 내역
시스템 아키텍처(SystemArchitecture)	• 하드웨어, 네트워크 구조, 연결된 API 등
성능 지표(PerformanceMetrics)	• 정확도, 편향성, 공정성, 설명가능성 등
인프라(Infrastructure)	• 배포 환경, OS, 라이브러리 버전 등

출처 : BIS 홈페이지(www.bsi.bund.de, 2025.10.25. 방문)

3) 독일 하이테크 어젠다(Hightech Agenda Deutschland)[144]

2025년 7월 30일 독일 정부는 AI, 양자 기술, 마이크로일렉트로닉스, 생명공학, 핵융합, 기후 중립적 에너지 생산, 그리고 기후 중립적 이동성 기술 등 6대 핵심기술 분야에서 선도국으로 도약하기 위한 기술혁신 전략인 '독일 하이테크 어젠다(Hightech Agenda Deutschland)'를 채택했다. 핵심기술과 관련한 연구, 개발, 응용 분야 가속화를 통해 독일 및 유럽의 기술 역량을 확대하겠다는 것이다. 특히 이번 발표는 AI 분야에서 독일 경제 및 산업이 갖는 위상을 강화하기 위해, 2030년까지 전체 경제생산(Bruttowertschöpfung)의 약 10%를 AI 기반으로 달성하고, 작업생산성(Arbeitsproduktivität)을 향상시키며, AI를 핵심 연구·응용 분야에서 중요한 도구로 만들겠다는 목표가 설정되어 있다. 이를 위해 독일은 2030년까지 국내총생산(GDP)의 최소 3.5%를 연구개발에 투자할 예정으로, 2029년까지 이 목적을 위한 연구에만 20억 유로 이상이 투자될 예정이다.[145]

특히 AI 기술을 통해 독일이 다시 기술적 자립성(Technologiesouveränität)과 경쟁력(Wettbewerbsfähigkeit)을 확보해야 한다는 방향성에 따라 구체적인 조치들을 설정했는데,

144) die Bundesregierung, Hightech Agenda Deutschland, 2025.7.30.
145) 최소 3.5%를 연구 개발에 투자, 2029년까지 20억 유로 투자의 경우 AI에만 국한된 것이 아니고 6개 분야에 발표이므로 해석에 주의 필요

다음은 독일 하이테크 아젠다 중 AI(Künstliche Intelligenz) 분야의 주요 내용 및 핵심 사항을 정리한 것이다.

<표 3-46> 독일 하이테크 어젠다 핵심 내용

구분	핵심 내용
AI-로보틱스 부스터(KI-Robotikbooster)	• 2026년부터 다목적 로봇(Multipurpose Robots)을 위한 리드 프로젝트(Leitprojekte) 출범 예정 • 이를 통해 '구현된 인공지능(embodied AI)'의 응용 가능성 제시, 과학·산업을 위한 연구 인프라(Forschungsinfrastruktur) 확충, 실험·훈련·응용센터(Erprobungs- und Weiterbildungszentren) 구축 등을 추진
핵심 산업 및 연구분야를 위한 전환(Transfer) 리드프로젝트	• 2025년 가을 자동차, 화학, 바이오테크, 클린테크, 의료, 식품농업(Agrifood) 등과 같은 핵심 업종 및 소재, 기후, 생물다양성, 에너지 및 지속가능성 연구 분야 등에 AI 응용을 강화하기 위한 프로젝트를 업계·연구자 인사들과의 대화(Fachgespräche)를 통해 시작 • 2026년부터 본격적 경쟁 프로젝트 및 사업화(Transfer) 착수
유럽 내 주요 산업AI 모델 개발 참여	• 유럽 차원의 중요 프로젝트에 독일이 참여해 고도로 전문화된 산업용 AI 모델을 개발하고, 동시에 독일·유럽 내 기술 자립성을 강화
의료·헬스케어 분야에서의 AI 활용 확대	• 예방부터 치료, 돌봄 및 사후관리까지 의료 영역 전반에서 AI를 활용한 예측·진단·치료 기술 확산을 위한 포트폴리오 구축
스타트업 및 사업화 확대	• 기업화(Spin-outs)를 촉진하고, 전국적 인큐베이션 구조를 구축하며, 특히 AI 스타트업의 스케일업(확장)을 위해 'EXIST 창업프로그램' 등을 활성화
AI칩 및 하드웨어 개발 지원	• 스마트하고 에너지 효율적인 AI칩 개발을 위한 디자인 이니셔티브(Design-Initiative) 개시 • 2025년부터 첫 리드 프로젝트 추진

출처 : BMFTR 홈페이지(www.bmftr.bund.de, 2025.10.27. 방문)

2025년 10월 29일 프리드리히 메르츠(Friedrich Merz) 연방 총리는 하이테크 어젠다를 공식 출범한다고 밝혔다. 베를린에서 열린 출범식에서 메르츠 총리는 우수한 연구를 제품과 서비스로 전환하는 데는 상당한 개선의 여지가 있으며, 특히 AI 분야에서 진전이 필요하고

서 강조했다. 카르스텐 빌트베르거(Karsten Wildberger) 연방 디지털국가현대화부(BMDS) 장관도 “AI는 인류 역사상 가장 혁신적인 동력”이라고 밝히며 독일이 AI에 적극적으로 참여하고 선두에 서야 한다고 역설했다. 또한 AI 산업의 가장 확실한 접근 방식은 산업 데이터를 활용하고 에이전트 AI를 통해 유럽과 전 세계를 위한 솔루션을 구축하는 것이라고 설명했다.[146]

4) AI 시장 감독 및 혁신 촉진법(KI-MIG)[147]

EU AI법에 따르면 회원국들은 2025년 8월 2일까지 ‘시장감시 기관(notifying authority 및 market surveillance authority)’을 지정해야 하지만 독일은 이 마감일을 지키지 못했다. 연방 디지털국가현대화부(BMDS)는 독일 내에서 EU 「AI 법」을 시행하기 위한 국내 규정을 마련하기 위해 「AI 법」 시행령에 관한 독일 연방주 및 관련 단체들과의 협의 절차를 공식적으로 시작했는데, 해당 협의는 2025년 9월 11일 장관급 법안 초안(KI-MIG: Künstliche Intelligenz Marktüberwachung und Innovationsförderungsgesetz)을 발표하면서 시작됐다.

초안에 따르면 독일 정부는 연방네트워크청(BNetzA)을 중앙 감독기관 및 EU 집행위원회와의 접점 역할을 하는 시장감시 기관으로 지정할 예정이다. 또한 2026년 8월 2일까지 최소 한 곳 이상의 ‘국가 AI 규제 샌드박스/연구소’를 설립할 계획이다. BMDS는 초안을 발표하면서 독일에서 EU 「AI 법」은 과도한 관료주의를 지양하고 조율된 시장감독을 보장하는 ‘혁신 친화적’ 방식으로 시행될 것이라고 설명했다. 초안에 대한 의견 수렴 절차는 2025년 10월 10일까지 진행되었고, 독일 정부는 차후 의견 수렴 내용을 바탕으로 최종안을 마련할 방침이다.[148]

<표 3-47> AI 시장 감독 및 혁신 촉진법 초안 내용

구분	주요 내용
중앙 감독기관 지정	• 연방 네트워크청(BNetzA)이 중앙 시장 감시 및 신고기관으로 지정 • AI 관련 행정집행과 기업 지원을 위한 전담부서(‘KoKIVO’ - Koordinierungs- und Kompetenzzentrum für die KI-Verordnung) 신설

146) BMDS, Gesetz zur Durchführung der KI-Verordnung, 2025.9.12.
heise online, Hightech Agenda: Merz calls for solidarity for technological sovereignty, 2025.10.30.
147) Heuking, New draft bill for the implementation of the AI Regulation, 2025.9.25.
148) White & Case, AI Watch: Global regulatory tracker - Germany, 2025.5.20.
Germany's Implementation Act for the EU AI Act, 2025.9.25.

구분	주요 내용
부문별 감독	• 제품안전·금융서비스 등 기존 감독기관이 분야별 AI 응용을 지속 관리 • 고위험 분야(생체인식, 치안 등)의 경우 독립적 감독기구(UKIM: Unabhängige KI-Marktüberwachungskammer) 설립이 명시
AI 샌드박스 도입	• 혁신 지원을 위해 BNetzA 산하 AI 규제 샌드박스(선도사업 실증 장소)를 2026년까지 설립해 실험적 도입 및 기업 지원에 초점
기업 지원 서비스 데스크	• AI 시장 진입과 실무 준수 지원을 위한 서비스 센터를 운영해, 복잡한 규제 환경 속 기업의 원스톱 지원체계를 목표로 함
분산 감독과 협업	• 각주별 데이터 보호기관(German DPA) 등은 EU 일반 개인정보 보호 규정(GDPR) 등 기존 법률에 따라 AI 관련 업무 병행 • '중앙-지방', '감독-지원'의 복합 구조가 특징
제재와 집행	• 법 위반 시 벌금 등 집행 조항 도입, EU 및 독일 현행 경쟁법, 개인정보보호법 등과 유사한 수준의 제재 마련

출처 : KI-MIG(2025.9.11.), SPRi 정리

5) 연방 정부 현대화 어젠다[149)]

2025년 10월 1일 독일 메르츠 내각은 연방정부 현대화를 골자로 하는 '연방정부 현대 어젠다(Modernisierungsagenda - für Staat und Verwaltung)'를 채택하고, 정부 디지털화를 위한 핵심 수단으로 AI의 활용을 강조했다. 이는 독일을 선도적인 경제 강국으로 되돌리기 위한 범부처 로드맵으로 ▲관료주의로 인한 비용 절감 ▲규제 개선 ▲시민기업 중심의 행정서비스 ▲공무원 인적자원 현대화 ▲연방 행정 효율화 등 다섯 가지 실행 영역에 초점에 맞춰 80가지의 조치를 담고 있는데, 독일 정부는 전(全) 영역에 걸쳐 특히 AI와 같은 첨단기술을 활용한 효율화를 강조하고 있다. 이 계획은 앞서 2025년 7월 발표된 '하이테크 어젠다'에 이어 발표된 것으로, 독일의 국가 행정을 근본적으로 변화시킬 것으로 기대를 모으고 있다.

구체적으로 이 계획에 따르면, 독일 정부는 향후 각 부처의 입법안 작성을 지원하는 AI 기반 도구를 도입할 예정이다. 또한 수출 지향 기업을 위한 AI 기반 온라인 플랫폼을 구축하여 대외 무역 진행 및 개발 협력에 관한 정보를 종합적으로 제공할 계획이다. 더불어, 비자 신청자의 서류 검토에서 AI를 활용하여 비자 처리 속도를 개선할 방침이다.

149) BMDS, Modernisierungsagenda für Staat und Verwaltung (Bund), 2025.10.1.

제9절 싱가포르

1. 그간의 인공지능 정책 추진 경과

<표 3-48> 싱가포르의 주요 AI 정책 추진 현황

시기	발표기관	정책
2017년 5월	국립연구재단	AI Singapore(국가 AI 프로그램) 개시
2019년	싱가포르 정부	국가 AI 전략 NAIS 1.0 발표
2019년	개인정보보호위원회(PDPC)	모델 AI 거버넌스 프레임워크(v1.0) - 기업 대상 AI 윤리·거버넌스 원칙 제시
2019년	정보통신미디어개발청(IMDA)·PDPC	AI Verify 및 AI Verify Foundation 출범
2020년 1월	PDPC	모델 AI 거버넌스 프레임워크(v2.0) - 설명가능성·투명성 강화, 실무 적용 가이드 추가
2020년대	국립연구재단	AI Singapore 확대
2023년 12월	싱가포르 정부	국가 AI 전략 NAIS 2.0 발표
2025년 7월	IMDA·AI Verify Foundation	글로벌 AI 보증 샌드박스 - 생성형 AI 서비스의 글로벌 테스트·검증 샌드박스 운영

1) 개요

싱가포르는 2019년 국가 AI 전략(National AI Strategy, NAIS 1.0)을 발표하면서 경제・사회 핵심 분야(보건・금융・교통・도시 등)에 실용적・확장 가능한 AI 도입을 목표로 삼았다. 그 이후 2019년 싱가포르 개인정보보호위원회(PDPC)는 모델 AI 거버넌스 프레임워크(Model AI Governance Framework) 초판을 공개하면서 AI를 책임감 있게 개발하고 배치하기 위한 지침을 제공했다. 또한 싱가포르 정부는 지침 제시에서 멈추지 않고, 윤리원칙 기업・기관이 실제로 지침을 따르는지 검증할 수 있는 AI Verify 개발 도구와 세부 가이드를 마련하면서 지침의 실효성을 높이고자 했다.[150]

150) Allen & Gledhill,Releases Asia's First Model AI Governance Framework for Public Consultation, Pilot Adoption and Feedback (2025.11.17. 방문)

2020년 1월에는 ‘모델 AI 거버넌스 프레임워크’의 두 번째 판[151]을 선보였는데, 민간 기업이 이를 자발적으로 채택하는 본래의 취지는 변함이 없으나 2019년 출간된 지침서와의 차이점이라면 AI 시스템이 내린 결정은 설명할 수 있고(explainability) 투명하며(transparency) 공정(fair)해야 한다는 원칙과 AI 시스템이 인간 중심이 되어야 한다는 두 원칙을 채택했다.

한편 2023년 12월 싱가포르 정부는 국가 AI 전략(National AI Strategy 2.0, NAIS 2.0)을 발표하면서 기존 NAIS 1.0을 확장해 장기적·공공선 중심의 AI 전략으로 확장했다. 이 전략은 ‘싱가포르와 세계를 위한 공공선으로서의 AI(AI for the Public Good, for Singapore and the World)’라는 비전 아래 AI 연구·인재 육성, 컴퓨팅 인프라 확보, 규제 및 윤리 지침 마련 등이 포함된 싱가포르의 전체 AI 정책을 포괄하는 핵심 전략이라고 할 수 있다. 끝으로 2025년 싱가포르는 하이브리드 양자-클래식 컴퓨팅 이니셔티브(Hybrid Quantum-Classical Computing, HQCC 1.0)를 발표함으로써 AI 역량의 계산적 기반 강화를 위한 차세대 계산 패러다임(quantum computing)인 양자 기술 인프라 확충에 나섰다.

2) 싱가포르의 인공지능 정책 변화(NAIS 1.0과 2.0 비교 중심)

2019년 싱가포르는 AI를 통해 스마트 네이션(Smart Nation)으로 도약하겠다는 국가 AI 전략(National AI Strategy 1.0, NAIS 1.0)을 발표했다. 그전까지는 디지털 전환이나 데이터 기반 행정이 중심이었는데, 이 전략부터는 AI를 경제성장 동력이자 사회 혁신의 핵심 도구로 규정했다. 또한 국가 규모의 한계를 극복하기 위해, AI를 활용하여 지리적 제약을 뛰어넘는 성장(frontier)을 강조했다.[152]

이 전략은 △싱가포르를 AI의 글로벌 허브로 발전 △새로운 비즈니스 모델 창출과 생활 개선을 위한 혁신 서비스 제공 △AI 경제의 미래 적응을 위한 인적 자원 정비의 3대 비전 하에 의료, 교육, 물류, 안전·보안 등 국가 주요 부문에서 AI 프로젝트를 수립하며 싱가포르 AI 생태계 기초 강화를 도모했다. 싱가포르 최초의 NAIS 1.0 전략은 국가 AI 사무국을 구심점으로 한 거버넌스 체계 구축과 산업 분야별 AI 프로젝트 개발을 통해 AI 확산을 위한 기초를 마련했다. 즉, AI를 단순히 연구만 하는 게 아니라, 국민이 체감할 수 있는 실생활

https://www.allenandgledhill.com/sg/publication/articles/9803/releases-asia-s-first-model-ai-governance-framework-for-public-consultation-pilot-adoption-and-feedback

151) PDPC, Singapore's Approach to AI Governance,2020.1. https://www.pdpc.gov.sg/help-and-resources/2020/01/model-ai-governance-framework

152) Singapore, Smart Nation 1.0, 2025.4.1. https://www.smartnation.gov.sg/about/our-vision/sn1

중심의 프로젝트를 추진하겠다는 것이다.

4년 뒤인 2023년에 발표된 국가 AI 전략(National AI Strategy 2.0, NAIS 2.0)은 이전 전략의 연장선이지만, 범위와 기준이 확대됐다. NAIS 1.0이 AI를 활용해 경제와 행정을 혁신하는 것이라면, NAIS 2.0은 AI를 모든 사회와 산업 시스템의 핵심 기반으로 활용함으로써 진일보한 형태이다. NAIS 2.0의 핵심 비전은 AI를 공공선(public good)을 위해 활용하고 그 경험을 세계와 나누겠다(AI for the Public Good – for Singapore and the World)는 것이며 이를 위해 정부는 세 가지 전략적 전환(Three Key Shifts)을 제시했다.

① 기회에서 필수로(From Opportunity to Necessity): AI는 더 이상 선택이 아니라 모든 산업과 사회 시스템의 필수 기반이다.
② 지역에서 세계로(From Local to Global): 싱가포르의 경험을 아시아 및 세계와 연결하고, 글로벌 AI 거버넌스 논의에 선도적으로 참여한다.
③ 개별 부문에서 시스템 전체로(From Sectoral to Systemic): 특정 분야 중심의 프로젝트를 넘어서, 사회·경제 전반에 AI를 체계적으로 통합한다.

NAIS 2.0은 AI 기술이 성숙함에 따라 국가 전반의 시스템 혁신으로 전환해야 한다는 인식에서 출발했다. 특히 생성형 AI(Generative AI)가 확산되고 신뢰·윤리·보안에 대한 국가적 관리 능력이 중요해지면서 싱가포르 정부는 NAIS 2.0 전략을 3가지 범주인 △Activity Driver(정부, 산업, 연구), △사람 및 커뮤니티(인재, 역량, 장소 조성), △인프라 및 환경(컴퓨팅, 신뢰할 수 있는 환경, 데이터) 과 이를 가능하게 하는 10개 조력자 역할로 규정했다.[153] 따라서 NAIS 1.0은 AI를 통한 경제 혁신과 효율성 제고를 목표로 한 실용적 출발점이었다면, NAIS 2.0은 AI를 사회적 가치, 신뢰, 국제 협력의 핵심 축으로 발전시킨 업그레이드 전략이라고 볼 수 있다.

153) Singapore, National AI Strategy 2.0, 2023. https://file.go.gov.sg/nais2023.pdf

<표 3-49> NAIS 1.0와 NAIS 2.0 비교

구분	NAIS 1.0(2019년)	NAIS 2.0(2023년)
정책 비전	AI를 통한 경제 성장과 사회 혁신 실현	AI를 공공선(Public Good)으로 발전시켜 지속가능한 사회 구축
슬로건 / 핵심 문구	AI as an Engine for Growth and Progress	AI for the Public Good – for Singapore and the World
정책 성격	기술 도입 중심의 실행형 전략	사회·경제 전반을 아우르는 시스템 혁신 전략
주요 목표	① 생산성 향상 ② 삶의 질 개선 ③ 새로운 산업 창출	① 사회적 신뢰 기반 구축 ② 공공·민간 부문 AI 통합 ③ 글로벌 협력 및 책임 있는 AI 확산
중점 추진 분야	5대 분야 집중 -의료(Healthcare) -스마트 시티(Smart Estates), -교통·물류(Transport & Logistics), -교육(Education), -공공안전(Safety & Security)	범사회적 통합 -산업 전반의 AI 내재화 -공공 서비스 전면 디지털화 -신뢰·윤리·보안 체계 구축
정책적 의미	AI를 경제혁신의 수단으로 본 정책적 출발점	AI를 사회 기반 인프라이자 글로벌 공공재로 인식하는 정책적 성숙 단계

출처: NAIS 1.0, NAIS 2.0 보고서 재정리.

2. 2025년 인공지능 정책 동향

1) 하이브리드 양자-고전 컴퓨팅(Hybrid Quantum-Classical Computing 이니셔티브, HQCC 1.0)

(1) 개요

2025년 3월 11일, SupercomputingAsia 2025 행사에서 조세핀 테오(Josephine Teo) 디지털개발정보부 장관(Minister for Digital Development and Information)은 하이브리드 양자-고전 컴퓨팅(Hybrid Quantum-Classical Computing 이니셔티브, HQCC 1.0)의 출범을 발표했다.

싱가포르 정부는 이 이니셔티브에 약 2,450만 달러(싱가포르 달러) 규모의 자금을 투입하여 양자 컴퓨팅과 고성능 컴퓨팅(HPC)[154]/AI 인프라의 통합을 가속화하겠다고 했으며 미들

웨어 개발, 하이브리드 알고리듬(Algorithm)·도구 개발, 인력 양성 및 산업 연계에 중점을 둘 것이라고 했다. 국립양자사무소(NQO)가 주도하는 이 프로젝트는 3~5년 간 ASPIRE 2A+[155]와 같은 고전 슈퍼컴퓨터의 강력한 연산 능력과 양자 컴퓨터의 고유한 장점을 결합한 혼합 연산 기술 개발을 목표로 한다. 이를 통해 양자와 고전 컴퓨팅의 매끄러운 통합을 가능케 하는 미들웨어 플랫폼과 하이브리드 알고리즘을 개발하고, 향후 금융, 물류, 생물연산 등 분야에 특화된 양자응용을 촉진할 계획이다.[156]

한편 이 이니셔티브는 AI 모델 규모가 폭발적으로 증가하면서 연산 에너지, 비용, 환경 부담이 커지고 있는데 HQCC의 하이브리드 연산을 통해 효율적·에너지 절감형 인프라로서의 역할을 할 것으로 보이며 NAIS 2.0의 핵심 중 하나인 신뢰 가능한 AI(Trusted AI)를 실현시키기 위한 알고리듬 투명성, 데이터 무결성, 연산 검증 가능성을 시험함으로써 AI 거버넌스 모델을 확장하는 실험적 장(場) 역할로도 활용될 것으로 보인다.

2) 신뢰할 수 있는 AI 생태계를 위한 이니셔티브 발표(Singapore launches new tools to help businesses protect data and deploy AI in a trusted ecosystem)[157]

(1) 글로벌 AI 보증 샌드박스(Global AI Assurance Sandbox)

2025년 7월 7일 싱가포르 정보통신미디어개발청(Infocomm Media Development Authority, IMDA)과 AI Verify Foundation(AIVF)은 Global AI Assurance Sandbox라는 이름으로 공식 출범했다. 이 샌드박스는 기업이 실제로 운영되거나 운영 예정인 생성형 인공지능(GenAI) 애플리케이션을 대상으로 기술적 테스트 및 검증을 할 수 있는 안전한 환경을 제공하는 글로벌 이니셔티브이다. 단순한 원칙 기반의 프레임워크를 넘어서, 실제 테스트·검증(assurance) 단계로 나아가려는 싱가포르의 전략이 반영된 것이다. 대상은 빌더(Builder) 또는 생성형 AI 애플리케이션을 개발·배포하는 조직(Deployers of GenAI applications) 그리고

154) 고성능 컴퓨팅(HPC: High-Performance Computing)은 일반 컴퓨터보다 월등히 빠른 속도로 대규모 데이터 처리 및 복잡한 계산을 수행하는 컴퓨팅 기술을 의미
155) 싱가포르 국립 슈퍼컴퓨팅 센터(NSCC)에서 구축한 최첨단 고성능 컴퓨팅(HPC) 시스템
156) The Quantum Insider, Singapore Invests $24.5 Million in Quantum and Supercomputing Integration Initiative,2025.3.14., https://thequantuminsider.com/2025/03/14/singapore-invests-24-5-million-in-quantum-and-supercomputing-integration-initiative
157) IMDA, Singapore Launches New Tools to Help Businesses Protect Data and Deploy AI in a Trusted Ecosystem, 2025.3.14 https://www.imda.gov.sg/resources/press-releases-factsheets-and-speeches/press-releases/2025/singapore-launches-new-tools-to-help-businesses-protect-data-and-deploy-ai-in-a-trusted-ecosystem

전문 기술 테스트 벤더(Specialist Technical Testing Vendors)이다. 기존 파일럿 단계에서는 17개의 애플리케이션이 전 세계 9개국에서 16개의 테스트 벤더와 함께 진행되었다. 테스트 대상은 기반모델(foundation model)이 아니라, 실제로 배포되거나 배포될 생성형 AI 애플리케이션이다. 테스트 기간은 최대 약 3개월이며, 테스트 결과는 공개적인 사례연구(case-study) 또는 보고서 형식으로 발행된다.[158] IMDA는 LLM 기반 애플리케이션의 안전 테스트를 위한 스타터 키트(Starter Kit for Safety Testing of LLM-based applications)라는 가이드킷을 발행했고 샌드박스 참여 기업들이 이를 활용하도록 설계되어 있다. 또한 이 샌드박스를 통해 얻은 인사이트는 향후 테스트 표준 마련에 활용될 예정이다.

(2) 개인정보보호 강화 기술(Privacy-Enhancing Technologies, PET)도입 가이드

IMDA와 개인정보 보호 위원회(Personal Data Protection Commission, PDPC)는 새로운 개인정보보호 강화 기술(Privacy-Enhancing Technologies, PET)도입 가이드를 통해 기업이 PET를 도입하여 더 큰 비즈니스 가치를 창출하도록 지원하는 것을 목표로 하고 있다. PET는 개인정보 노출 없이 데이터 활용(분석・AI 학습・협업)을 가능하게 해 새로운 서비스로 확장할 수 있다. 여기에는 AI 및 머신러닝을 통해 데이터에 대한 접근성을 높이고 여러 이해관계자(예: 조직 내 여러 팀 또는 여러 부문)간에 통찰력을 공유하는 것이 포함된다. 싱가포르는 PET Sandbox 운영 경험을 바탕으로 PETs Adoption Guide를 발행해 실무 도입 장벽을 낮추고자 한다. 또한 이 가이드는 실무 템플릿・체크리스트・사례를 담고 있어 도입 초기 문서화에 바로 활용 가능하다.

(3) 데이터 보호를 위한 새로운 국가 인증(New National Certification for Data Protection)

데이터 보호 인증 마크(Data Protection Trustmark, DPTM)는 IMDA와 PDPC가 공동 개발했으며 싱가포르 내 조직의 개인정보 보호・데이터 거버넌스 수준을 인증하는 자발적 인증 제도이다. IMDA는 Enterprise SG 및 싱가포르 인증 위원회와 협력하여 DPTM를 새로운 싱가포르 표준(SS 714:2025)으로 격상시켰다. 이는 글로벌 데이터 보호 벤치마크 및 국제 모범 사례와 동등한 수준을 의미하는데 조직이 책임감 있는 데이터 보호 관행(accountable data protection practices)을 갖추었음을 고객・파트너・규제기관에 보여주는 것이다.

158) AI Verify Foundation, AI Assurance Pilot (2025.11.17. 방문)
https://aiverifyfoundation.sg/ai-assurance-pilot

제10절 이스라엘

1. 그간의 인공지능 정책 추진 경과

<표 3-50> 이스라엘의 주요 AI 정책 추진 현황

시기	발표기관	정책
2021년	정부·TELEM 포럼	국가 AI 프로그램 Phase 1 권고안 - AI를 전략기술로 지정
2022년	프라이버시보호국 등	AI 관련 개인정보 가이드라인·AI 규제 백서 초안
2023년 12월	혁신과학기술부	이스라엘 AI 정책 2023 - 책임 있는 혁신을 중심 원칙으로 하는 첫 종합 AI 정책 문서
2024년 5월	정부	공공부문 AI Open Call 프로그램
2024년 9월	정부	국가 AI 프로그램 Phase 2 출범
2025년 7월	이스라엘 혁신청(IIA)	고급 AI 연구 데이터 자산 프로그램
2025년 10월	총리실	국가 AI 본부 - 정책 및 거버넌스 통합 추진

1) 개요

2021년은 이스라엘이 AI 정책의 출발 선상에 놓여있는 해라고 할 수 있다. 이스라엘은 정부결정 제212호(Resolution 212)를 통해 혁신 촉진, 하이테크 성장 장려 및 기술·과학 리더십 강화 프로그램을 채택하고, AI를 국가 핵심 혁신 분야로 공식 지정하였다.[159] 이 결정에 따라 혁신과학기술부(Ministry of Innovation, Science & Technology)가 국가 AI 정책의 주관 부처로 지정되었으며, 데이터 정책, 윤리, 국제협력, 민간 부문 적용을 총괄하도록 했다. 또한 2021년에는 텔렘(TELEM) 포럼[160]이 AI 연구역량 강화를 위한 권고안을 제시하였다. 여기에는 AI 슈퍼컴퓨터 센터 설립, 히브리어·아랍어 자연어처리(NLP) 기술 강화, 인적자원

159) https://innovationisrael.org.il/wp-content/uploads/2023/10/Governmnet-Resoluion-No.-173.pdf#:~:text=Further%20to%20Resolution%20212%20and%20the%20work,will%20be%20formulated%20by%20the%20aforementioned%20team.

160) 이스라엘의 TELEM 포럼은 '연구 개발을 위한 국가 기반 시설 포럼(The National Infrastructure Forum for Research and Development)'의 약자로, 국방부를 포함한 여러 국가 기관 간의 연구 개발(R&D) 자원을 조정하고 통합하기 위한 자발적인 협의체임, 참여 기관으로는 이스라엘 국방부 산하 국방 연구 개발국(DDR&D/MAFAT), 이스라엘 혁신청(Israel Innovation Authority), 고등교육위원회 기획예산위원회 위원장, 과학기술부 사무총장, 재무부 예산과 부국장 등 여러 고위급 인사들이 참여

양성, 규제 환경 정비 등이 포함되었다. 초기 예산은 약 5억 뉴 이스라엘 셰켈(NIS) 규모[161]로 책정되었으며, 이 시기를 통해 이스라엘은 본격적인 국가 단위 AI 추진 체계를 갖추기 시작했다.[162]

2022년에는 AI의 사회적 책임과 데이터 보호를 중심으로 한 규제·윤리 논의가 본격화되었다. 2022년 10월 이스라엘 프라이버시보호국(Privacy Protection Authority)은 AI 기반 알고리즘이 개인 정보를 수집하거나 처리하는 경우, 정보주체에게 목적과 수신자를 명시해야 한다는 가이드라인을 발표했다. 이는 개인정보보호법(PPL)을 AI 시대에 맞게 해석한 최초의 사례였다.[163] 이어 11월에는 정부가 AI 규제·윤리 백서(White Paper) 초안을 공개하고 공공 의견 수렴을 시작하였다. 이는 향후 국가 차원의 책임 있는 AI 정책 수립을 위한 기초 문서로, 투명성·설명가능성·인간 감독의 원칙을 명확히 제시했다는 점에 의미가 있다.[164]

2023년은 이스라엘 AI 정책이 제도적으로 구체화된 해이다. 2023년 2월 24일 정부결정 제173호(Resolution 173)를 통해 이스라엘의 기술적 리더십 강화 프로그램을 승인하고, AI가 핵심 전략 분야라는 것을 재확인했다. 이 결정에서는 △기초 및 응용 AI 연구 가속화 △산업용 AI 인프라 구축 △공공 및 민간 분야의 AI 적용 확대를 주요 원칙으로 명시했다. 이를 위해 2023년~2026년 동안 최대 5억 NIS의 예산이 배정되었다. 2023년 12월에는 혁신과학기술부 장관이 이스라엘의 AI 규제 및 윤리 정책(Israel’s AI Policy 2023)을 공식 발표하였다. 이는 이스라엘 최초의 종합 AI 정책 문서로, 책임 있는 혁신(Responsible Innovation)을 핵심 철학으로 삼았다. 또한 정부는 히브리어·아랍어 AI 응용기술 개발에 약 3천만 NIS를 투입하며, 언어·문화적 포용성을 강조했다.[165]

한편 2024년은 이스라엘 AI 정책이 실행 중심 단계로서 AI 정책의 가속도를 내기 시작한 시점이라고 할 수 있다. 9월 정부는 국가 AI 프로그램(Phase2)[166]을 공식 출범하며, 2024~2027년 동안 약 5억 NIS를 추가 투자하기로 결정했다. 이 프로그램은 연구개발뿐 아니

161) 뉴 이스라엘 셰켈(New Israeli Shekel)이라는 공식 통화의 약자로서, 5억 NIS는 한화로 약 2,273억 원(2025년 12월 5일 기준)
162) https://techpolicy.org.il/blog/israeli-court-ruling-border-controllers/
163) https://www.loc.gov/item/global-legal-monitor/2022-10-20/israel-privacy-protection-authority-announces-scope-of-informed-consent-for-collection-of-personal-information-by-ai/
164) https://www.gov.il/BlobFolder/news/most-news20231218/en/Israels%20AI%20Policy%202023.pdf
165) https://www.gov.il/en/pages/ai_2023
166) 이스라엘 정부는 Telem Forum의 권고사항을 바탕으로, 2021년부터 첫 번째 단계(Phase 1)를 공식 시작

라 공공부문 혁신, 산업별 적용, 데이터 인프라 구축을 포함한다.[167] 2024년 5월에는 정부 주도로 9개 공공기관이 참여한 AI 시범사업(Open Call for AI in Government)이 추진되어 약 4,500만 NIS가 배정되었다. 이는 행정 효율성, 예측 분석, 공공 서비스 품질 향상을 목표로 한 최초의 대규모 공공부문 AI 프로젝트였다.[168]

2025년은 이스라엘이 AI를 사회 전반으로 확장한 시기로 평가된다. 데이터 인프라 강화 차원에서 약 4,400만 NIS 규모의 AI 데이터 저장소 구축 지원 사업을 추진하였다. 또한 2025년 10월에는 총리실 산하에 국가 AI 본부(National AI Headquarters)를 신설하는 방안이 공식화되었다. 이는 부처 간 분산된 정책을 통합하고, 국가 차원의 전략 조정 기능을 강화하기 위한 조직적 변화이다.[169]

2. 2025년 인공지능 정책 동향

1) 2025년 AI 정책 강화를 위한 조직 확대

(1) 국가 AI 본부 설립 승인

이스라엘 정부[170]는 2025년 10월 19일, 국가 차원의 AI 전략을 이행하기 위해 국가 AI 본부(National Artificial Intelligence Headquarters / Directorate)를 설립하기로 결정했다. 지금까지 이스라엘 정부는 2023년 및 2024년 여러 정책과 프로그램 등이 실행된 바 있으며, 이를 바탕으로 보다 중앙 집중적이고 강력한 거버넌스 체계를 마련하고자 했다.[171] 이러한 맥락에서 정부는 본부 설립을 승인하였으며, 이는 AI를 경제・안보・사회 전반의 핵심 역량으로 전환하기 위한 전략적 조치라고 볼 수 있다.[172]

---

167) https://innovationisrael.org.il/en/press_release/second-phase-ai-program/#:~:text=Israel%20Launches%20Second%20Phase%20of%20National%20AI,Investment%20in%20Research%20and%20Development%20Infrastructure.%2017/09/2024.

168) https://aiisrael.org.il/press_release/council-of-europe-opens-first-ever-global-treaty-on-ai-for-signatureranked-7th-in-the-world-in-ai-development-israel-invests-133m-in-phase-2-of-national-program/

169) https://www.ifi.today/news/1447-Israel-to-Launch-National-AI-Headquarters-Under-Prime-Ministers-Office.html

170) 이스라엘 재무부와 총리실의 공동 사업

171) 지금까지 펼친 정책의 실효성 문제점에 대한 논의가 존재하고 있으며 이스라엘이 전쟁 중이라는 특수한 상황에서 군사 기술력 강화라는 해석이 함께 공존하고 있음

172) https://www.ifi.today/news/1447-Israel-to-Launch-National-AI-Headquarters-Under-Prime-Ministers-Office.html

이 본부는 총리실(Prime Minister’s Office)산하 부처로 설치되는데 정부는 본부를 총리실 내부에 두고 주요 정책 조정 및 실행의 허브로 삼기로 했다. 설립 승인 당시, 총리실은 본부 책임자(Director)를 2주 내에 임명[173]하고, 본부는 승인 후 약 45일 이내에 기능을 시작할 것이라는 일정으로 제시했다. 승인된 초기 예산은 2025년~2026년 기간 동안 약 1억 2천만 NIS로 책정되었으며 주요 승인 내용은 다음과 같다.[174]

<표 3-51> 국가 AI 본부 관련 주요 승인 내용

구분	주요내용
우수 AI 인재 확보	• 국내외 AI 전문가 유치 및 산학·연 협력 강화
고성능 컴퓨팅 인프라 구축	• 슈퍼컴퓨터 및 대규모 AI 모델 훈련을 위한 인프라 확충 계획이 포함되어 있으나, 구체적 비용이나 규모는 조정 중
정부 시스템 내 AI 응용 확대	• 공공 서비스 및 행정 업무에 AI를 통합하여 효율성·혁신성 제고
사회·경제적 영향 평가	• AI 도입이 고용, 에너지 소비, 규제 등에 미치는 영향을 연구하고 대응

출처 : www.ifi.today

(2) AI 및 자율성 관리국(AI and Autonomy Administration) 설립

2025년 1월 1일자로 이스라엘 국방부(Israel Ministry of Defense, IMoD)는 국방연구개발국(Directorate of Defense Research and Development, DDR&D)산하에 AI 및 자율성 관리국(AI and Autonomy Administration)을 공식 설립했다고 발표 했다. 이 조직은 AI 및 자율시스템 분야 등 기존의 여러 부서·프로젝트로 분산 운영해 왔던 구조에서 벗어나 집중적인 연구·개발·운용 통합체계를 마련하기 위한 조치이다.[175]

이 조직의 설립은 단순한 기술 부서의 신설이 아니라, AI 기반 자율시스템이 미래 전장의 중심축으로 부상함에 따라 이스라엘이 군사적·산업적 경쟁력을 유지하기 위한 전략적 전환점을 의미한다. 국방부 차관인 에얄 자미르(Eyal Zamir)는 미래의 전장은 병사와 자율시스

173) 2025년 9월 25일, 총리실은 공개 모집 없이 에레즈 아스칼 준장을 총리실 산하에 신설된 국가인공지능(AI)국 국장으로 임명. 아스칼 준장은 이전에 C4I 및 사이버 방위국 산하 디지털 전환국 국장, 시각 정보 부대 9900 사령관, 공군 및 군사 정보국 연구부에서 정보 업무를 담당한 이력이 있음

174) https://www.gov.il/en/pages/press_21092025a

175) https://jewishpress.com/news/israel/government-israel/defense-ministry-launches-new-ai-and-autonomy-administration/2025/01/01/

템이 통합된 팀으로 작전하게 될 것이라고 강조했다. 이는 단순히 무기체계의 자동화 수준을 높이는 것이 아니라, AI가 전술·전략적 의사결정에 참여하는 수준으로 확장되는 것을 목표로 한다.[176]

AI 및 자율성 관리국은 국방연구개발국(DDR&D)이 관리하는 여러 프로젝트 중에서도 이스라엘 방위군의 작전 우위를 유지하기 위해 모든 군부의 전투 능력을 변화시키는 역할을 담당한다. 특히 육상, 해상, 공중, 우주 등 모든 전장 영역에서 AI 기술을 적용할 수 있는 플랫폼을 개발하며 국방부, 산업계, 학계, 스타트업 간 협력의 허브 역할을 수행한다. 이스라엘은 세계적인 기술 스타트업 생태계를 갖추고 있기 때문에, DDR&D는 이를 적극적으로 활용하여 민군 겸용 기술을 국방 분야로 이전하는 동시에 군사 연구성과를 민간산업으로 환류시키는 기술 선순환 구조를 강화하고 있다.[177]

2) AI 연구 개발을 위한 데이터 저장소 구축(Establishing and Enabling Access to Data Assets for Advanced AI Research)[178]

(1) 개요

2025년 7월 23일 이스라엘 정부는 AI 연구 생태계 강화를 위해 AI 연구 개발을 위한 데이터 저장소 구축(Establishing and Enabling Access to Data Assets for Advanced AI Research) 프로그램을 공식 추진하고 있다. 이 사업은 이스라엘 국가 AI 전략(Phase2)의 핵심 과제 중 하나로, AI 연구·개발(R&D)에 필수적인 고품질 데이터 자산(Data Assets)을 체계적으로 구축하고, 이를 산업계와 학계가 폭넓게 활용할 수 있도록 개방하는 것을 목표로 한다.

이 프로그램은 이스라엘 혁신청(Israel Innovation Authority, IIA)이 주관하며, 공모(Call for Proposals)형식으로 운영된다. 즉, 국내 기업이나 연구기관이 데이터 저장소를 구축하거나, 기존 데이터를 통합·개방하는 사업 계획서를 제출하면 심사를 거쳐 정부의 보조금 지원을 받을 수 있다. 선정된 프로젝트는 데이터 자산의 집계(Aggregation), 수집(Collection), 개발(Development), 접근성 확보(Accessibility) 등 데이터 전 주기 관리 활동을 수행하며, AI 연구에 필요한 데이터 인프라를 실질적으로 강화하게 된다.[179]

176) https://euro-sd.com/2025/01/major-news/41924/imod-ai-and-autonomy-admin/
177) https://iaeai.org/tpost/i1m7dsg2r1-israels-ministry-of-defense-and-idf-laun
178) https://innovationisrael.org.il/en/press_release/data-for-ai/
179) https://innovationisrael.org.il/en/calls_for_proposal/data-for-ai/

지원 대상 분야는 광범위하다. 생명과학・의료(Life Sciences & Health), 첨단 제조(Advanced Manufacturing), 농업기술(AgriTech), 푸드테크(FoodTech), 사막 및 기후기술(DeserTech & Climate), 방위산업 및 국가안보(Defense & Homeland Security) 등 이스라엘의 주요 전략산업들이 모두 포함된다. 이는 단순한 데이터 저장소 구축을 넘어, 각 산업군의 특성에 맞춘 데이터 활용 기반을 마련함으로써 산업별 AI 융합 연구의 가속화를 의도한 것이다.

<표 3-52> 프로그램 지원 대상 분야

구분	주요내용
생명과학·의료(Life Sciences & Health)	• 약물 개발, 의료 기기, 진단 기술, AI 를 통해 의료 시스템에 적용 가능한 의료, 게놈 또는 임상 데이터
첨단 제조(Advanced Manufacturing)	• 반도체와 ICT를 포함한 첨단 산업, 생산 라인, 공급망, 감지 시스템, 자동화에서 발생하는 데이터
사막 및 기후기술(DeserTech & Climate)	• 기후 데이터 저장소, 재생 에너지 데이터, 수질 모니터링, 날씨 예보, 환경 감지 및 기후 관련 애플리케이션
농업기술(AgriTech)	• 센서 기반 농업 데이터, 위성 추적, GIS 시스템, 양식업, 포장 공장, 식품 공급망
푸드테크(FoodTech)	• 대체 단백질, 육류 대체 식품, 새로운 생산 공정, 영양 추적을 포함한 혁신적 식품 분야의 연구 저장소 및 기술.
방위산업 및 국가안보(Defense & Homeland Security)	• 사이버 보안, 방위 및 이중 용도 기술, 우주 산업, 방위 AI, 블록체인, 핀테크, 스마트 모빌리티, 소프트웨어와 관련된 데이터

출처 : www.innovationisrael.org.il/

이 프로그램의 재정 지원 규모는 약 4,400만 이스라엘 NIS에 달한다. 정부는 선정된 프로젝트에 대해 총 사업비의 55%에서 최대 66%까지 보조금을 제공하며, 신청 자격은 이스라엘에 등록된 법인이어야 하며, 해외 기관이나 기업이 참여할 경우 이스라엘 현지 법인과의 협력(Consortium) 또는 하도급(Sub-contractor)형태로만 가능하다.

제11절 인도

1. 그간의 인공지능 정책 추진 경과

<표 3-53> 인도의 주요 AI 정책 추진 현황

시기	발표기관	정책
2015년	전자정보기술부 (MeitY) 등	디지털 인도 전략 발표
2018년 6월	NITI Aayog 국책연구소	AI를 위한 국가 전략 - 'AI for All' 비전과 5대 우선 분야(보건·농업·교육·스마트시티·교통) 제시
2024년 3월	MeitY	인도AI 미션(IndiaAI Mission) 발표 - "Making AI in India & Making AI Work for India" 슬로건, 7대 전략 축 제시
2024년~	MeitY·IndiaAI	IndiaAI 사업부 운영 개시 - 미션 추진을 위한 전담 조직
2025년 1월	MeitY	인도 AI안전연구소 설립 발표
2025년 3월	MeitY	공공부문 리더 대상 AI 역량 프레임워크 발표
2025년 3월	MeitY	선진 인도를 위한 AI 로드맵 발표
2025년 10월	NITI Aayog	포용적 사회 발전을 위한 AI 보고서 발표
2025년 11월	MeitY	AI 거버넌스 가이드라인 발표

인도는 ▲2015년 '디지털 인도(Digital India)' ▲2018년 '국가 AI 전략(National AI Strategy)' ▲2024년 '인도AI미션(IndiaAI Mission)'으로 이어지는 정책 개발을 통해 AI 혁신을 위한 기반을 다지고 AI 정책 목표와 주요 사업을 명확화·구체화하였다.

'디지털 인디아'는 정보통신 기술을 활용해 행정, 교육, 금융, 보건 등 사회 전반의 디지털화와 포용적 성장을 목표로 하는 국가 정책이다. AI를 전면에 내세운 정책은 아니지만 디지털 인프라 구축, 디지털 접근성 향상, 정부 디지털화 등을 견인하여 AI 정책 추진을 위한 디지털 기반을 수립하는 데에 지대한 영향을 미쳤다. 예로, '디지털 인디아'를 통해

구축된 생체 정보(지문, 홍채) 기반 디지털 신원인증 플랫폼(‘Aadhaar’)과 같은 디지털 공공 인프라(DPI: Digital Public Infrastructure)는 현재 ‘인도AI미션’ 하에서 AI 솔루션과 통합되고 있다.[180)]

‘국가 AI 전략’ 은 인도 총리가 주도하는 인도 정부의 공공 정책 싱크탱크인 NITI아요그(NITI Aayog)에서 발표한 전략문서로, 최초로 국가적인 수준에서 AI 발전 비전 및 목표를 광범위하게 설정하고 있다는 점에서 의의가 있다. ‘AI 전략’ 은 ‘모두를 위한 AI(AI for All)’ 라는 AI 정책 비전을 제시하고, 역량 강화와 포용을 위한 도구로서의 AI 활용을 강조하였다. 특히 인도가 직면하고 있는 여러 도전과제를 AI를 활용해 해결할 것을 권고했고, 그 결과 인도에서는 보건의료, 농업, 교육, 스마트도시, 교통 등 AI 활용의 효과가 특히 더 기대되는 분야를 중심으로 실용적인 AI 애플리케이션 및 서비스 개발이 추진되었다.[181)]

가장 핵심적인 AI 정책인 ‘인도AI미션’ 은 기존의 AI 정책 비전인 ‘모두를 위한 AI’ 와 더불어 “인도에서 AI 개발하고 인도를 위해 AI를 활용하자(Making AI in India and Making AI Work for India)” 를 슬로건을 내걸고 마련된 포괄적인 AI 생태계 구축 정책으로, 2024년 3월에 발표되었다. 인도에서 AI 기술혁신을 촉진하고 국가적 AI 역량을 강화하는 것이 목표로, 이를 위해 정책적으로 지원할 핵심 활동 영역을 크게 7개 필라(Pillar)로 구분하고, 각각의 필라에 따라 세부적인 목표들을 수립하였다. 주요 내용은 ▲AI 컴퓨팅 용량 확보 ▲토종 AI 파운데이션 모델 개발·배포(AI 혁신센터) ▲AI 데이터셋 플랫폼 구축 ▲AI 애플리케이션 개발 ▲AI 인재 개발 ▲AI스타트업 지원 ▲안전한 AI 거버넌스 구축 등으로, 인도 정부는 관련 활동에 2024년부터 5년 동안 1,037억 1,920만 루피(약 12억 5,000만 달러)의 예산을 투입할 계획이다. ‘인도AI미션’ 은 AI 발전을 위한 정부 활동의 근간으로, 인도 정부 수립 100주년이 되는 2047년까지 인도를 선진국으로 만든다는 장기적인 국가 비전인 ‘선진 인도(Viksit Bharat)’ 및 기존의 디지털 정책과 결합되어 추진된다. 주무부처인 전자정보기술부(MeitY)는 ‘인도AI미션’ 수행을 위해 2024년부터 독립 사업부인 인디아AI(IndiaAI)를 운영하고 있다.[182)]

180) Digital India, About Us (2025.11.3. 방문) www.digitalindia.gov.in/about-us
181) NITI Aayog, National Strategy for Artificial Intelligence, 2018.6.
182) indiaai.gov.in (2025.11.3. 방문)

<표 3-54> ‘인도AI미션’ 7개 필라 및 주요 내용

필라	세부 목표 및 내용
AI 컴퓨팅 용량 확보	• AI 스타트업과 연구 커뮤니티를 지원하는 확장 가능한 AI 컴퓨팅 생태계 구축 • 민관 협력을 통해 GPU 1만 8,000개 이상 규모의 AI 컴퓨팅 인프라 구축 추진 • 적격 사용자로 선정된 AI스타트업·연구자 등에 최대 40% 저렴한 비용으로 AI 컴퓨팅 성능 제공
AI 혁신센터 (토종 AI 파운데이션 모델 개발)	• 인도 AI 혁신센터(IAIC: IndiaAI Innovation Centre)를 설립해 자체 AI 대규모멀티모달모델(LMM)과 ‘특정 영역에 특화된(domain-specific)’ 파운데이션 모델의 개발·배포 수행 • 글로벌 표준에 부합하는 동시에 인도가 직면한 고유의 과제를 해결하는 토종 AI 모델의 구축과 배포 촉진
통합 데이터셋 플랫폼 구축	• 비개인(익명화된) 데이터에 대한 접근성을 확대해 인도 AI 스타트업과 연구자들이 AI 혁신을 주도할 수 있도록 지원 • 양질의 통합 데이터셋 플랫폼을 마련해 데이터셋에 대한 접근성을 개선·간소화
AI 애플리케이션 개발	• 현실 세계 과제 해결에 효과적인 AI 솔루션의 개발·확산·홍보 지원 • 의료, 기후변화·재난관리, 거버넌스, 농업, 학습장애보조 등 사회·경제적으로 중요한 분야에서 효과적인 AI 솔루션 개발을 촉진하기 위해 개발 챌린지 프로그램 등 운영
AI 인력 개발	• AI 교육 프로그램에 대한 진입장벽을 완화하여 AI 인재 파이프라인 구축 • 대학, 대학원, 박사 교육과정에서 AI 과정을 증설하고, 인도 내 2·3선 도시(인구 10만 명 미만의 소도시)에 데이터/AI 랩을 설치해 AI 교육 접근성을 확대
AI 스타트업 자금 지원	• 혁신적인 AI 프로젝트에 대한 재정지원을 통해 딥테크 AI 스타트업의 성장 지원 • AI 스타트업에 대한 지원 절차를 간소화하고, 효과적인 지원 메커니즘 마련
안전한 AI 거버넌스 구축	• AI의 책임 있는 개발·배포·도입을 보장하는 가드레일 설계·개발 • 인도AI안전연구소(AISI: IndiaAI Safety Institute)의 주도로 책임 있는 AI를 위한 지침, 거버넌스 프레임워크, 체크리스트, 기타 도구 개발 추진

출처 : IndiaAI 홈페이지(indiaai.gov.in, 2025.11.3. 방문), SPRi 정리

이와 더불어 인도는 2023/24년도부터 2027/28년도까지 총 990억 루피(약 11억 2,000만 달러)를 투자하는 AI 우수성 센터(CoE: Centres of Excellence for AI) 이니셔티브를 시행 중이다. CoE는 뉴델리 올인디아의료과학원(AIIMS), 델리공과대학(IIT Delhi), 로파르공과대학(IIT Ropar), 칸푸르공과대학(IIT Kanpur) 등 명문대학들이 주도하고 정부와 산업계 파트너 및 스타트업이 협력하여 의료 부문 등 여러 중점 분야에서 혁신적인 AI 솔루션 개발을 추진하는 협업 공간으로, 궁극적으로는 인도에서 AI 생태계를 더욱 발전시키고, AI의 잠재력을 최대한 활용하여 인도가 직면한 주요 사회적 과제들을 해결하는 것을 목표로 운영된다. 인도는 2024년까지 ‘의료’, ‘농업’, ‘지속가능한 도시’ 분야에서 3개의 CoE를 설립하였다.

2. 2025년 인공지능 정책 동향

2025년 인도는 4번째 CoE 설립 계획을 발표하는 한편, ‘인도AI미션’을 위한 다양한 후속 활동을 전개하였다. 특히 AI 컴퓨팅 인프라 확대, 데이터셋 플랫폼 구축, 토종 AI 모델 육성 등에서 상당한 진전을 이루면서 AI 생태계 구축을 가속화하였다. 또한 AI 거버넌스를 위한 정책 문서 및 가이드라인을 공개하며 향후 AI 정책 추진 방향을 구체화적으로 제시하였다. 이러한 노력으로 인도는 세계적인 AI 선도국가로 부상하고, 2047년 ‘선진 인도’ 비전을 실현할 수 있는 견고한 기반을 다지고 있다.

1) 교육 분야 CoE 설립 추진[183]

인도 니르말라 시타라만(Nirmala Sitharaman) 재무장관은 2025년 2월 교육 분야 CoE를 설립하기 위해 2025년 예산안에서 50억 루피(약 5,900만 달러)를 배정했다고 밝혔다. 이로써 인도는 기존에 설립된 3개의 CoE에 더해 총 4개의 CoE를 구축하게 되었다. 이들 CoE는 인도 내 AI 생태계 구축을 촉진하고 우수 인재 양성 등에 기여할 전망이다.

2) 선진 인도를 위한 AI 로드맵 발표[184]

(1) 개요

MeitY는 2025년 3월 6일 ‘선진인도를 위한 AI 로드맵(India’s AI Revolution - A Roadmap to Viksit Bharat)’에서 ‘인도AI미션’의 주요 추진 성과를 검토하고 향후의 추진 방향을 제시하였다. 이전까지 인도는 언어 장벽 극복, 생산성 개선 등과 같은 사회적인 문제를 해결할 수 있는 AI 솔루션 개발에 주력해왔으나, ‘인도AI미션’을 통해 자체 AI 인프라 구축과 토종 파운데이션 모델 개발까지 아우르는 AI 생태계 전반의 발전을 추구하는 방향으로 정책적 초점을 전환하였다. 이 가운데 2025년 1월 중국 딥시크가 저비용으로 구축한 AI 모델의 전 세계적인 성공사례를 목격한 인도는 동 로드맵에서 특히 AI 컴퓨팅 자원 확대 및 토종 AI 모델 개발 등에 주력하여 글로벌 AI 개발 분야에서 경쟁력과 기술 주권을 확보하겠다는 의지를 강조하였다.

---

183) AIM India, Govt Announces 4th AI Centre of Excellence for Education, Expansion in Skilling & IIT Infrastructure, 2025.2.1.

184) MeitY, India’s AI Revolution - A Roadmap to Viksit Bharat, 2025.3.6.
PIB, Transforming India with AI - Over ₹10,300 crore investment & 38,000 GPUs powering inclusive innovation, 2025.10.12.

⑵ ‘인도AI미션’ 추진 성과 및 향후 계획

2025년에 인도는 컴퓨팅 인프라 분야에서 괄목할만한 성과를 달성하였다. ‘인도AI미션’ 초기 단계에서 인도 정부는 1만 4,517개의 GPU를 확보했고, 이를 1만 8,963개까지 확대할 계획이었다. 그러나 현재는 당초의 목표를 크게 상회해 3만 8,000여 개의 GPU를 확보했고, 이렇게 확보한 AI 리소스를 ‘AI 컴퓨팅 포털(AI Compute Portal)’을 통해 AI 연구자와 기업 등에 시간 당 65루피(약 0.7달러)라는 저렴한 가격으로 제공할 수 있게 되었다. ‘AI 컴퓨팅 포털’은 인도의 AI 연구자, 기업가, 스타트업 등의 최신 AI 애플리케이션 및 솔루션 연구와 개발을 지원하기 위해 구축되었으며, AI 개발에 필요한 컴퓨팅 자원과 데이터셋, 저장소, 클라우드 서비스 및 각종 도구들을 저렴하게 제공한다. 동 포털은 2025년 3월 6일 AI 데이터셋 플랫폼 ‘AIKosh’와 함께 공개되었다.

‘AIKosh’는 데이터셋에 대한 접근성을 개선하기 위해 마련된 통합형 데이터셋 플랫폼으로, 데이터셋은 물론 여러 모델, 용례 등을 함께 제공한다. 정부/비정부 기관들의 데이터를 통합해 AI 모델 학습을 위한 대규모 데이터셋을 개발해 제공함으로써 개발자가 기본 모듈을 구축하는 대신 AI 솔루션 개발에만 집중할 수 있도록 지원한다. ‘AIKosh’는 20여 개 분야에 걸쳐 3,000개 이상의 데이터셋과 243개의 AI 모델을 제공 중으로, 2025년 7월을 기준으로 플랫폼 방문객수는 26만 5,000회, 등록 사용자수는 6,000명을 달성했고, 다운로드 횟수는 13,000건을 기록하였다.

또한 2025년에는 인도의 니즈(needs)에 부합하는 자체 AI 파운데이션 모델 개발 사업이 본격화되었다. 인도 정부는 2024년 말부터 글로벌 표준에 부합하는 동시에 인도 고유의 과제를 해결할 수 있는 토종 AI 파운데이션 모델을 구축하는 것을 목표로 ‘인도AI 혁신 이니셔티브(IndiaAI Innovation Initiative)’를 추진하였다. 인도 고유의 데이터셋을 활용하는 대규모 멀티모달모델(LMM), 대규모언어모델(LLM), 또는 특정 도메인이나 애플리케이션을 지원하는 소규모언어모델(SLM) 개발 프로젝트를 선정해 지원하는 것이 골자이다. 2025년 1월 30일부터 제안 모집을 시작해 인도 전역의 기관, 스타트업, 연구기관 및 AI 혁신가들로부터 총 506개의 제안을 접수하고, 1단계에서 토종 LLM 개발업체로 사르밤AI(Sarvam AI), 소켓AI(Soket AI), 간AI(Gan AI), 지나니닷AI(Gnani.AI) 등 4개 스타트업을 지원 대상으로 선정하였다. 이 중 인도의 소버린 LLM 생태계 구축을 위해 선정된 사르밤AI는 1,200억 파라미터 규모의 오픈소스 LLM을 개발 중으로, 향후 해당 모델은 거버넌스 및 공공 서비스 개선에 활

용될 예정이다. 소켓AI는 국방・의료・교육 부문에 초점을 맞춰 인도의 언어 다양성에 최적화된 1,200억 파라미터 규모의 오픈소스 모델을 개발한다. 지나니닷AI는 고급 추론 기능을 갖춘 140억 파라미터 규모의 음성 AI 모델을 구축하고 있으며, 간AI는 700억 파라미터의 다국어 텍스트-음성 변환 모델을 만드는 데 주력하고 있다. 이어서 2025년 9월에도 8개 파운데이션 모델 프로젝트를 추가 지원 대상으로 선정했는데, 여기에는 농업・의료・거버넌스 특화 AI 아바타, STEM 및 의학적 문제해결을 위한 대규모 추론 모델, 22개 인도어에 대한 다수의 SLM 개발 프로젝트 등이 포함되었다.

이 밖에도 ‘인도AI미션’ 하에서 다음과 같은 진전이 이루어졌다.

① (AI 애플리케이션 개발 이니셔티브) 2024년부터 의료, 농업, 기후변화, 거버넌스, 학습장애보조기술 등의 분야에서 인도 고유의 과제 해결을 위한 AI 애플리케이션 개발 지원사업을 추진해 2025년 7월까지 30건의 애플리케이션을 승인

② (AI 인력 양성) AI 전문 인력 양성을 위해 박사과정 연구자 500명, 대학원생 5,000명, 학부생 8,000명에게 재정지원을 제공하는 것을 목표로 ‘인도AI 퓨쳐스킬(IndiaAI FutureSkills)’ 이니셔티브를 시행 중으로, 2025년 7월까지 총 200명 이상에게 장학금을 제공

- 또한 인도 내 2·3선 도시에 데이터·AI 랩을 설립을 추진하면서, 총 27개 데이터 랩 설립 대상지를 선정

③ (AI 스타트업 재정 지원) 2025년 3월부터 ‘인도AI 스타트업 글로벌(IndiaAI Startups Global)’이라는 AI 스타트업 재정지원 프로그램을 개시

- 세계적인 스타트업 인큐베이터 스테이션F(Station F) 및 프랑스 경영대학원 HEC파리(HEC Paris)와 협력해 인도 AI 스타트업 10곳에 스테이션F에서 HEC파리가 설계한 스타트업 프로그램(자원·멘토링·네트워킹 기회 제공)에 참여할 수 있는 기회를 제공

④ (안전하고 신뢰할 수 있는 AI) AI의 책임 있는 개발·배포·채택을 위한 가드레일을 마련하기 위해 관련 프로젝트를 선정해 지원

- 2024년 진행된 책임 있는 AI 프로젝트 1차 지원 사업에서 8개 프로젝트를 선정해 지원했고, 2차 지원 사업(2024년 12월 10일~2월 28일)에서는 400건 이상의 의향서를 접수

- 2025년 1월 30일 인도AI 안전연구소(AISI) 설립을 발표하고, 학계, 스타트업, 산업계, 정부 부처/부서를 포함한 모든 관련 이해관계자와 협력해 AI의 안전·보안·신뢰 확보 및 인도 고유의 데이터셋을 활용한 AI 연구개발을 활성화할 수 있는 협력 모델을 추진

3) 공공 부문을 위한 AI 역량 프레임워크 발표[185]

(1) 개요

공공 부문에서의 AI 활용은 국가적 개발과 경제적 목표 달성에 중요한 역할을 한다. 이에 MeitY는 2025년 3월 6일 '공공 부문 리더들을 위한 AI 역량 프레임워크(A Competency Framework for AI Integration in India)'를 발표하였다. 이는 공공 부문에서 AI 도입 및 활용을 책임지고 주도하는 공무원들이 AI 기술과 그 사회·윤리적 함의를 이해하고, AI 정책 및 프로젝트를 설계·개발·적용·모니터링하는 데 필요한 지식과 능력을 체계적으로 강화할 수 있도록 지원하는 문서이다.

(2) 주요 내용

'공공 부문 AI 역량 프레임워크'는 공무원들에게 요구되는 행동적 역량, 기능적 역량, 도메인별 역량을 명시하고 있다. 공무원들이 AI 환경을 효과적으로 탐색하는 데 필요한 기술과 지식, 전문성, 윤리적 기반을 갖출 수 있도록 안내한다는 점에서 공공 부문에 AI를 통합하기 위한 가이드북이자 청사진이라 할 수 있다.

<표 3-55> 공무원을 위한 AI 역량 프레임워크

핵심 역량 범주	개요	주요 내용
행동적 역량 (Behavioural Competencies)	AI 이니셔티브 관리에 필요한 기술 및 태도	• 혁신적 사고, 체계적 사고, 시민 중심적 사고 강조 • AI 솔루션이 사회적 가치에 부합하고 현실 세계의 과제를 효과적으로 해결할 수 있도록 보장
기능적 역량 (Functional Competencies)	AI 프로젝트를 추진하기 위해 필요한 실무 기술 및 지식	• AI 리터러시, 데이터 거버넌스, 이해관계자 관리 등과 관련 • 규제 준수 등 AI 시스템의 성공적인 구현에 필요한 조치 강조
도메인별 역량 (Domain Competencies)	특정 영역에서 공무원의 역량 개발을 지원하는 교육 및 프로그램	• 농업, 의료, 교육, 스마트시티, 교통 등 국가적으로 중요한 분야에서의 역량 강화 방법 안내

출처 : Empowering Public Sector Leadership: A Competency Framework for AI Integration in India(2025.5.6.), SPRi 정리

185) MeitY, Empowering Public Sector Leadership: A Competency Framework for AI Integration in India, 2025.5.6.

4) 포용적 사회 발전을 위한 AI 보고서 공개[186)]

(1) 개요

2025년 10월 NITI아요그 연구소는 AI 기술이 인도의 비공식 노동자들의 삶을 어떻게 변화시킬 수 있는지를 체계적으로 탐구한 연구보고서 ‘포용적 사회 발전을 위한 AI(AI for Inclusive Societal Development)’를 발표하였다. 인도에는 대략 4억 9,000만 명의 비공식 노동자들이 존재하지만 이들은 제한된 노동 시장 접근성, 부족한 기술능력, 불충분한 사회보장제도, 낮은 생산성과 같은 문제들에 직면해 있다. 보고서는 AI, 블록체인, 몰입형 학습 등과 같은 최신 기술이 비공식 노동자들이 직면한 이러한 여러 장벽을 제거하는 데에 기여한 실제 사례들을 소개하면서 신기술을 통해 비공식 노동 부문을 개혁할 수 있음을 설명하였다. 그러나 AI 등 기술 그 자체만으로는 비공식 노동 부문을 변화시킬 수 없다는 점을 지적하고, 인도 내 비공식 인력을 디지털화하고 역량을 강화하기 위한 새로운 국가적 이니셔티브로 ‘디지털 슈람세투 미션(Digital ShramSetu Mission)’을 제안하였다.

(2) 주요 내용

힌디어로 ‘슈람세투’는 ‘노동의 다리’를 의미한다. NITI아요그는 인도의 모든 노동자가 미래 디지털 경제에서 성공하는 데 필요한 기술, 도구, 그리고 기회를 가질 수 있도록 보장하는 데에 중점을 두고 ‘디지털 슈람세투 미션’을 구상하였다. 비전은 AI와 같은 최첨단 기술을 활용해 인도의 비공식 노동자들을 ‘미래에 준비된’, ‘권한과 역량을 가진’ 공식적인 노동자로 전환하는 것으로, 이를 실현하기 위해 저렴하고, 접근성이 뛰어나며, 비공식 노동자들이 직면한 여러 과제에 대응할 수 있는 기술 솔루션에 대한 연구개발(R&D)을 추진할 것을 제안하고 있다. 대표적인 예는 ▲디지털 ID 시스템을 통한 급여 적시 지급, 자격 인증, 사회보장혜택 제공 ▲AI 기반 교육 프로그램 제공 ▲자격증 발급 및 검증을 위한 분산형 신뢰 모델 구축 ▲블록체인 기반의 스마트 계약 시스템 운영 등이다.

NITI아요그는 이러한 목표를 실현하기 위해 2025년부터 2029년 및 그 이후까지 취해야 할 조치들을 총 4단계의 로드맵으로 제시하였다.

〈표 3-56〉 ‘디지털 슈람세투’ 구현 로드맵

단계		주요 내용
1단계 (2025~2026년)	미션 방향성 설정	• 명확한 목표, 일정, 성과측정 방법 등을 담은 정책 선언문 작성 • 정부, 산업계, 학계, 시민사회 이해관계자들이 함께 참여해 우선순위와 목표 설정

186) NITI Aayog, AI for Inclusive Societal Development, 2025.10.8.

단계		주요 내용
2단계 (2026~2027년)	제도 구축 및 거버넌스 설계	• 다부처형 거버넌스 구조, 리더십 역할, 실행을 위한 청사진 수립 • 관련 법규제, 디지털 인프라 준비도 등 고려
3단계 (2027~2029년)	시범 프로그램 등 실시	• 시범 프로그램을 통해 기술 솔루션을 실제 세계에서 테스트 • 모니터링 및 평가 실시
4단계 (2029년~)	전국 시행 및 통합	• 검증된 솔루션을 전국적으로 배포

출처 : AI for Inclusive Societal Development(2025.10.8.), SPRi 정리

5) AI 거버넌스 가이드라인 발표[187]

(1) 개요

MeitY는 ‘인도AI미션’에 따라 2025년 11월 5일 AI 기술의 안전하고 책임 있는 사용을 보장하는 것을 목표로 ‘AI 거버넌스 가이드라인’을 발표하였다. 가이드라인은 AI 혁신을 촉진하는 동시에 AI 기술에 개인과 사회에게 제기하는 위험을 완화하기 위한 강력한 거버넌스 프레임워크를 제안한다. AI 개발에 있어 ‘책임 있는 AI’ 원칙을 강조하지만, 이를 강제하기보다는 기업의 자발적인 위험 평가와 투명성 확보 등을 권장하면서 규제보다는 ‘책임 있는 혁신’에 초점을 맞추고 있다.

(2) 주요내용

‘AI 거버넌스 가이드라인’은 AI 시스템의 신뢰성 확보, 인간의 감독, AI 역량 강화, 관련 표준 설정의 필요성 등을 설명하고, 정부, 학계, 산업계, 시민사회 등 다양한 이해관계자들에게 지속적인 협력을 기반으로 강력하고 포용적인 AI 생태계를 구축할 것을 강조한다. 주요 내용은 ▲윤리적이고 책임감 있는 AI를 위한 7가지 가이드 원칙 ▲AI 거버넌스의 6가지 축(pillar) 및 권장사항 ▲단기·중기·장기적 액션플랜 ▲산업계 및 규제기관을 위한 실용적 지침 등이다. 이는 AI의 혁신적인 힘을 포용적이고 지속가능하며 공평한 성장을 위해 활용하겠다는 인도 정부의 AI 개발 의지를 보여주는 것으로, 이를 통해 인도는 AI 생태계 구축 노력에 있어 중요한 이정표를 수립하였다.

<표 3-57> AI 거버넌스 가이드라인 주요 내용

구분		주요 내용
7가지 가이드 원칙	‘신뢰’는 기본원칙	• 신뢰는 AI 혁신·도입·발전과 위험 완화를 위해 필수적
	인간 중심	• 인간 중심 설계, AI 시스템에 대한 인간의 감독, 인간의 AI 역량 강

187) MeitY(India AI), India AI Governance Guidelines: Empowering Ethical and Responsible AI, 2025.11.5.

구분		주요 내용
	(People First)	화 등 보장
	제약(규제)보다는 혁신	• 다른 모든 조건이 동일한 경우, 제약(규제)보다는 혁신 우선
	공정성과 형평성	• '포용적 개발'을 촉진하고 차별을 방지
	책임성	• 명확한 책임 분담 및 규정 집행
	이해가능성을 내재한 설계 (Understandable by Design)	• AI 시스템의 사용자와 규제기관들이 해당 시스템의 작동 방식, 사용자에 대한 영향, 의도한 예상 결과 등을 이해할 수 있도록 관련 내용을 명확히 공개
	안전성·복원력·지속가능성	• 피해 위험을 최소화하기 위한 안전장치를 갖춘 설계 및 복원력 • 환경적으로도 책임감 있고 자원 효율적인 AI 개발 추진('경량' 모델 도입 장려)
6가지 필라와 권장사항	인프라	• 데이터, 컴퓨팅 등 기본적인 리소스에 대한 접근성 확대
	역량 구축	• AI에 대한 신뢰·역량 강화 및 위험과 기회에 대한 인식증진을 위한 교육과 기술능력 개발, 훈련 프로그램 개시
	정책 및 규제	• 혁신을 지원하고 AI 관련 위험을 완화하는 균형적이며 유연한 프레임워크 채택 • 현행 법률을 검토하여 AI 시스템 관련 규제 격차를 파악하고 법률 개정 추진
	위험 완화	• 인도 고유의 위험 평가 프레임워크 개발 • (취약계층 보호 등) 필요 시 위험 완화를 위한 추가 의무 도입
	책임성	• AI의 기능, 위험도 등에 따른 차등적인 책임 시스템 도입 • AI 가치사슬 내 다양한 행위자들의 의무 준수를 지원하고 의무 준수에 대한 투명성 강화
	기관/기구	• 기술정책위원회(TPEC: Technology & Policy Expert Committee) 및 AI 거버넌스 그룹(AIGG: AI Governance Group) 설립 추진 • 신설된 AI안전연구소(AISI)를 통해 안전하고 신뢰할 수 있는 AI 개발 및 활용 촉진
액션플랜	단기	• 주요 거버넌스 기관/기구 설립 • 인도 특화 위험 프레임워크 개발 • 자발적 서약 채택 • 법률 개정안 마련 • 명확한 책임 체계 구축 • 인프라 접근성 확대 • 인식증진 프로그램 시행 • AI 안전 도구에 대한 접근성 확대
	중기	• 공통 표준 발표 • 법률/규정 개정 • AI 사고 대응 시스템 운영 • 규제 샌드박스 시범 운영 • '디지털 인디아'를 통해 구축한 핵심 기반인 디지털 공공 인프라(DPI)와 AI 통합 확대
	장기	• 역량 구축, 표준 설정, 접근성 개선 등에서 지속적인 노력 전개 • 거버넌스 프레임워크 검토 및 업데이트 • 새로 부상한 위험과 역량을 기반으로 신규 입법 추진

구분		주요 내용
산업계 및 규제기관 지침	산업계	• 모든 법률 준수 및 자발적 프레임워크 채택 • 투명성 보고서 발표 • 불만처리 메커니즘 제공 • 기술-법률 융합(techno-legal) 솔루션을 통한 위험 완화
	규제기관	• 피해를 완화하는 동시에 혁신 지원 • 준수가 까다로운 제도는 회피 • 기술-법률 융합 접근방법 장려 • AI 규제 프레임워크의 유연성을 보장하고 정기적으로 검토

출처 : India AI Governance Guidelines: Empowering Ethical and Responsible AI(2025.11.5.), SPRi 정리

MeitY는 2025년 12월 중 AI 거버넌스 그룹(AIGG)과 기술정책위원회(TPEC)를 구성할 계획이다. AIGG는 MeitY와 내무부, 통신부 등 다수의 부처와 인도경쟁위원회(CCI: Competition Commission of India)와 같은 규제기관들이 참여하는 소규모 기구로, 부처 간 AI 거버넌스 및 정책 개발을 조율하는 역할을 수행할 예정이다. TPEC는 AIGG에 AI 정책 시행 및 전략에 대한 자문을 제공하는 전문가 기구로, 신기술, 데이터 사이언스, 법률, 공공 정책, 사이버 보안 분야 전문가들로 구성될 전망이다.

제12절 UAE

1. 그간의 인공지능 정책 추진 경과

<표 3-58> UAE의 주요 AI 정책 추진 현황

시기	발표기관	정책
2017년 9월	연방정부	UAE 4차 산업혁명 전략 - AI·로봇·블록체인 등 첨단기술로 경제 다각화·정부혁신 추진 선언
2017년 10월	연방정부	국가 AI 전략(UAE Strategy for Artificial Intelligence 2031)
2017년 10월	연방정부	AI 담당 국무장관 신설 - 세계 최초 AI 전담 장관직 임명
2018년	아부다비 정부	G42(Group 42) 설립 - 대규모 데이터센터·클라우드·자체 LLM 개발 등 국가 AI 인프라·산업화를 주도하는 기술지주회사(왕실 소유 공기업)
2019년 10월	아부다비	MBZUAI(무함마드 빈 자이드 AI 대학) 설립 - AI 석·박사 특화 대학원, 전액 장학을 통한 글로벌 인재 허브 구축
2020년 7월	연방정부	AI·디지털경제·원격근무 애플리케이션부 신설 - 기존 AI 부처를 확대 개편하여 디지털 전환과 통합
2020년대	두바이미래재단·연방정부	'스마트 두바이'·'아부다비 AI 허브' 구도 형성 - 두바이는 스마트시티·서비스 혁신, 아부다비는 인프라·산업 중심 전략
2025년 5월	연방정부	AI 기반 규제 지능 생태계 승인
2025년 5월	연방정부	미국-UAE AI 가속화 협약, 스타게이트 파트너십 체결
2025년 10월	두바이 정부	두바이 AI 가속화 이니셔티브 발표

아랍에미리트연합(UAE)의 AI 정책 추진을 이해하기 위해서는 UAE의 정치체계를 간소하게나마 이해할 필요가 있다. UAE는 7개의 토후국이 모인 연방국으로, 실질적인 국정 운영은 전제군주제와 유사하며 아부다비와 두바이가 UAE 정치와 경제의 핵심축을 이루고 있다. 따라서 아부다비와 두바이의 전략 및 대규모 투자를 중심으로 정책 경과를 살펴보되, UAE와 두바이, 아부다비가 추진하는 정책에 대해 개별적이면서도 서로의 연관성 상에서 살펴보아야 한다.

아부다비는 UAE의 수도로, UAE의 석유 매장량 대부분이 아부다비에 집중되어 있다. 아부다비는 석유 경제를 바탕으로 연방 전체의 정치와 경제를 실질적으로 주도하고 있다. UAE 대통령 무함마드 빈 자이드(Mohammed bin Zayed)는 대통령이자 아부다비의 통치자(Ruler)에 자리하며 UAE의 대내외 정책 전반을 총괄하는 최고 실권자로 활동하고 있다.

두바이의 경우, 20세기부터 일찌감치 비석유 부문으로 눈을 돌려 부동산 개발을 바탕으로 물류와 관광 산업을 통해 비약적인 경제 성장을 달성했다. 그 결과 중동 지역의 비즈니스 및 금융 중심지로 자리매김했으며, 두바이의 통치자 모하메드 빈 라시드(Mohammed bin Rashid)은 UAE의 부통령 겸 총리직을 겸임하고 있다.

UAE의 AI 정책은 아부다비와 두바이를 중심으로 추진된다. 아부다비와 두바이는 UAE의 경제 성장이라는 큰 틀 안에서 협력하나, 각자의 강점과 목표에 따라 AI 정책을 차별화하여 전개하는 양상을 보인다.

국가적 차원의 AI 전략과 대규모 투자는 아부다비가 주도하고 있다. 특히 MBZ의 친동생이자 국가안보보좌관인 타흐눈 빈 자이드 알 나히얀(Tahnoun bin Zayed)은 아부다비의 국부펀드를 통제하며, UAE의 AI 정책 및 전략에 있어 핵심 역할을 하는 △G42 △MGX의 소유주이기도 하다.

연방 정부 차원에서는 2017년 오마르 술탄 알 올라마(Omar Sultan al-Olama)를 AI 담당 국무장관으로 임명하며 미래 성장 동력으로서 정책적 의지를 분명히 했다. 오마르 술탄 알 올라마는 현재 'AI, 디지털 경제 및 원격 근무 애플리케이션부'를 이끌고 있으며, 두바이 미래 재단(Dubai Future Foundation)과도 밀접하게 연관되어 있어 두바이의 혁신 정책 수립과 추진에도 기여하고 있는 것으로 보인다.

1) ‘UAE 4차 산업혁명 전략’ ‘국가 AI 전략’ 발표와 AI 장관 임명[188)189)190)]

UAE는 2017년 하반기, 탈석유 시대를 대비한 국가 차원의 청사진을 제시하며 AI 정책의 핵심 기틀을 집중적으로 마련했다. 이는 미래 경제 패권을 선점하기 위한 목표에서 시작되었다.

2017년 9월, 모하메드 빈 라시드 부통령 겸 총리는 ’UAE 4차 산업혁명 전략(UAE Strategy for the Fourth Industrial Revolution)’을 발표했다. 전략의 최우선 목적은 UAE를 4차 산업혁명 기술의 글로벌 중심지로 변모시키는 것으로, 혁신 기술을 통해 국가 경제를 다각화하고, 미래 도전에 신속히 대응하는 정부 시스템 구축을 목표로 삼았다. AI, 로봇공학, 나노기술, 생명공학 등 첨단 기술을 경제 전반에 도입하고, 관련 규제 환경을 선제적으로 정비하는 것이었다.

2017년 10월, UAE 정부는 ’국가 AI 전략(UAE Strategy for Artificial Intelligence)’을 연이어 발표했다. ’UAE 센테니얼 2071(UAE Centennial 2071)’ 비전의 첫 번째 핵심 프로젝트로, 4차 산업혁명 전략을 AI 분야에서 구체화하는 실행계획이었다. ’국가 AI 전략’의 목적은 △AI를 활용해 정부 서비스와 행정 효율성을 100% 수준까지 향상 △AI를 새로운 핵심 경제 성장 동력으로 육성하여 2031년까지 AI 분야의 글로벌 리더로 도약하는 것이다. 이를 위해 △운송 △보건 △우주 △재생 에너지 △물 △기술 △교육 등 8개 핵심 분야를 선정하고 AI 기술의 우선 적용을 계획했다.

UAE 정부는 국가 전략의 실행력을 담보하기 위해 같은 해 10월, 내각 개편을 단행했다. ’AI 담당 국무장관’ 직책이 신설하였으며, 당시 27세의 오마르 술탄 알 올라마(Omar Sultan Al Olama)가 초대 장관으로 임명되었다. 이는 AI를 국가 핵심 의제로 격상시키고 젊은 리더십을 통해 변화를 주도하겠다는 상징적인 조치로, ’AI 담당 국무장관’의 핵심 역할은 ’국가 AI 전략’ 이행을 총괄하는 것이다. ’AI 담당 국무장관’은 모든 정부 부처와 기관의 AI 관련 이니셔티브를 조정 및 통합하고, △AI 인재 양성 △데이터 거버넌스 구축 △인프라 확충 등 AI 생태계 전반을 조성하는 책임을 맡았다.

188) UAE Government, UAE's Fourth Industrial Revolution Strategy, 2017.9.
189) UAE Government, UAE Strategy for Artificial Intelligence, 2017.10.
190) Business Insider, The United Arab Emirates is the first country in the world to hire a minister for artificial intelligence, 2017.12.

2) ‘G42’와 ‘무함마드 빈 자이드 AI 대학’ 설립[191][192]

UAE는 2017년 ’국가 AI 전략’을 실현하기 위한 구체적인 실행 체계 구축에 즉시 착수했다. 이는 기술 개발을 주도할 상업적 주체와 인재를 양성할 학술적 기반을 동시에 확보하는 방향으로 전개되었다.

먼저 2018년, 아부다비는 AI 및 클라우드 컴퓨팅 전문 기업인 ’G42(Group 42)’를 설립했다. G42의 핵심 설립 목적은 UAE 국가 AI 전략의 실질적인 이행을 주도하고, 기술 주권을 확보하는 것이었다. G42는 셰이크 타흐눈 빈 자이드 국가안보보좌관가 이끌고 있으며, 아부다비의 핵심기술 투자 포트폴리오의 중추이다.

G42는 ’국가 챔피언’으로서 정부, 보건의료, 금융, 석유 및 가스, 항공 등 국가 핵심 산업 전반의 AI 도입 및 디지털 전환을 총괄하는 역할을 맡았다. 또한, 국가 차원의 대규모 데이터센터와 클라우드 인프라를 구축하고, 자체 LLM을 개발하는 등 UAE를 기술 생산국으로 변모시키는 임무를 수행하기 시작했다. 이는 민간 기업의 형태를 띠지만, 사실상 아부다비의 AI 전략을 실행하는 가장 강력한 정책 수단이다.

AI 기술 개발과 더불어, UAE는 이러한 기술을 지속해서 연구하고 운영할 전문 인력의 확보를 위해 2019년 10월, 세계 최초의 AI 전문 대학원 ’무함마드 빈 자이드 AI 대학(MBZUAI)’의 설립을 공식 발표했다.

MBZUAI는 AI 분야의 글로벌 인재를 UAE로 유치하고, 국내외 최고 수준의 AI 연구자를 양성하여 지속가능한 AI 생태계를 구축하는 것을 목적으로 설립되었다. MBZUAI은 △컴퓨터 비전 △머신러닝 △자연어 처리 등 AI 핵심 분야의 석사 및 박사 과정을 제공하며, 전 세계 학생들에게 학비와 생활비 전액을 지원하는 파격적인 조건을 제시하며 AI 인재 허브로의 도약을 꾀했다.

191) Wikipedia, G42 (company), 2025.11.
192) WAM, Abu Dhabi announces establishment of the Mohamed bin Zayed University of Artificial Intelligence, 2019.10.

3) ‘UAE AI 디지털경제 원격 근무 애플리케이션부’ 설립[193)]

UAE는 2020년 7월, 내각 개편을 통해 AI 담당 부처의 역할과 범위를 대폭 확대했다. 기존의 ’AI 담당 국무장관’ 조직은 ’AI, 디지털 경제 및 원격 근무 애플리케이션부(Ministry of State for Artificial Intelligence, Digital Economy, and Remote Work Applications)’로 공식 명칭을 변경하였으며, 기능이 역시 강화되었고, 초대 AI 장관이었던 오마르 술탄 알 올라마가 확대된 부처의 장관직을 그대로 맡았다.

조직 확대 개편의 직접적인 배경에는 코로나 19 팬데믹으로 인해 비대면 활동이 일상화되었고, 디지털 경제로의 전환을 가속하는 계기가 만들어졌기 때문이었다. 또한, 원격 근무가 전 산업 영역에서 새로운 업무방식으로 급부상하며, UAE 정부는 이러한 급격한 환경 변화를 국가 시스템 혁신의 기회로 판단했다.

’AI, 디지털 경제 및 원격 근무 애플리케이션부’ 설립 목적은 AI를 독립적인 미래 기술로만 다루는 것이 아니라, ’디지털 경제’ 및 ’원격 근무’와 유기적으로 통합 관리하는 것으로, AI 기술을 활용해 디지털 경제의 성장을 촉진하고, 원격 근무 환경을 효율화하여 국가 전반의 회복탄력성과 경쟁력을 높이는 것을 목표로 삼았다.

부처의 역할 역시 구체화 되었다. △기존의 ’국가 AI 전략’ 이행 △AI 기술의 공공 및 민간 도입 총괄 △AI를 기반으로 한 ’디지털 경제’ 육성 전략을 수립 △핀테크, 전자상거래, 스마트 서비스 등 관련 산업 생태계 조성 △정부 및 민간 부문에서 효율적이고 안전한 원격 근무가 정착될 수 있도록 관련 정책, 인프라 및 애플리케이션 개발을 지원한다.

4) ‘MGX 펀드 매니지먼트’ 설립과 ‘두바이 AI 범용 계획’ 발표[194)195)]

아부다비와 두바이는 2024년 각자의 강점을 바탕으로 한 대형 AI 이니셔티브를 동시에 추진했다. 2024년 3월, 아부다비는 AI 분야의 글로벌 기술 패권을 확보하기 위한 강력한 투자 수단을 확보했다. 아부다비 정부는 ’MGX 펀드 매니지먼트(MGX Fund Management Limited)’

193) Wikipedia, Ministry of State for Artificial Intelligence, Digital Economy and Remote Work Applications. 2025.11.
194) MGX, Abu Dhabi launches comprehensive global investment strategy on artificial intelligence, 2024.5.
195) UAE Government, Dubai Universal Blueprint for Artificial Intelligence, 2024.12.

의 설립을 공식화했다. MGX의 설립 목적은 AI 및 첨단 기술 분야에 대한 전략적 투자를 집행하는 아부다비의 핵심 투자사 역할을 수행하는 것이다.

MGX는 아부다비의 대표 국부펀드 무바달라(Mubadala)와 G42의 파트너십으로 탄생했다. MGX의 핵심 역할은 △AI 인프라(반도체, 데이터센터) △파운데이션 모델 △AI 애플리케이션 등 AI 가치 사슬 전반에 걸쳐 글로벌 투자를 단행하는 것이다. MGX는 운용 자산 1,000억 달러 이상 확보를 목표로 하며, 이는 막대한 자본력을 바탕으로 AI 기술을 선점하려는 아부다비의 전략을 보여준다.

2024년 4월, 두바이는 AI 기술의 실용적 도입과 생태계 조성을 위한 청사진을 제시했다. 함단 빈 모하메드(Hamdan bin Mohammed) 두바이 왕세자는 '두바이 AI 범용 계획(Dubai Universal Blueprint for AI)'을 발표했다. '두바이 AI 범용 계획'의 목적은 △AI 기술 적용의 글로벌 선도 도시 △AI를 통한 공공 서비스의 획기적으로 향상으로, AI를 경제 성장의 핵심 동력으로 삼아 GDP에 기여하는 것을 목표로 한다.

'두바이 AI 범용 계획'는 AI 기술을 도시 운영과 산업 전반에 신속하게 통합하는 데 초점을 맞추고 있다는 특징을 가지고 있다. 이를 위해 △정부 부처별 '최고 AI 책임자(CAIO)' 임명 △AI 스타트업 육성을 위한 'AI 가속화 프로그램' 운영 △AI 기업들이 활동할 수 있는 '상용 플랫폼' 구축 등을 포함했다. 이는 AI를 공공 서비스와 산업 현장에 즉시 적용하려는 두바이 특유의 실용적 접근 방식을 보여준다.

5) Microsoft-G42-아부다비 전략적 파트너십[196)]

2024년 4월, Microsoft는 G42에 15억 달러를 투자하는 전략적 파트너십을 체결했다. Microsoft-G42 파트너십의 목적은 UAE의 AI 기술 개발을 가속화하고, G42를 MS의 최첨단 AI 기술을 활용하는 글로벌 선도 기업으로 육성하는 것으로, 미국 정부의 안보 기준을 준수하며 AI 기술을 안전하게 발전시키려는 양국 간의 전략적 합의를 기반으로 한다. MS는 G42를 거점으로 △중동 △중앙아시아 △아프리카 전역에 MS Azure 클라우드 서비스를 확장하는 것을 목표로 한다.

196) G42, Microsoft Invests $1.5 Billion in Abu Dhabi's G42 to Accelerate AI Development and Global Expansion, 2024.4.

파트너십을 통해 G42-MS는 △MS의 G42의 소수 지분에 투자를 통한 15억 달러 조달 △G42의 자사 AI 애플리케이션과 서비스를 MS 애저 클라우드 플랫폼으로 이전 △이전한 자사 AI 애플리케이션 기반으로 AI 솔루션을 개발 △MS의 브래드 스미스(Brad Smith) 사장의 G42 이사회 합류 △UAE와 주변 지역의 AI 개발자 및 인재 육성을 위해 10억 달러 규모의 양사 '개발자 펀드' 공동 조성에 합의했다.

G42-MS-아부다비 파트너십은 미국 상무부와의 '보안 및 보증'에 관한 구속력 있는 협정을 포함하며, G42가 MS의 첨단 AI 기술에 접근하는 대가로 미국과 동맹국의 안보 표준을 준수하겠다는 것을 의미한다는 점에 의의가 있다.

G42와 아부다비는 MS의 선진 AI 기술에 대한 안정적인 접근성을 확보하며 막대한 기술적, 자본적 동력을 얻게 되었다. 이를 통해 아부다비는 글로벌 AI 허브로서의 입지를 다지게 되었으며, MS는 급성장하는 중동 시장에서 애저 클라우드의 지배력을 강화하고, 지정학적으로 민감한 AI 기술 이전에 대한 선례를 마련했고 볼 수 있다.

2. 2025년 인공지능 정책 동향

1) AI 기반 규제 지능 생태계 승인[197)]

2025년 5월, UAE 정부는 'AI 기반 규제 지능 생태계(Regulatory Intelligence Ecosystem)'의 도입을 공식 승인했다. 이는 모하메드 빈 라시드 부통령의 지시에 따라 내각 사무국이 주도하는 프로젝트이다.

'AI 기반 규제 지능 생태계'는 AI 기술을 입법 과정 전반에 통합하는 내용을 담고 있다. 'AI 기반 규제 지능 생태계'를 통해 UAE의 AI 모델은 전 세계의 △법률 데이터 △정책 보고서 △학술 자료 등을 실시간으로 수집하고 분석하여 세계 최초의 스마트 입법 지도를 개발한다. 또한, 새로운 법률안이 국가 경제·사회·환경에 미칠 다각적인 영향을 사전에 시뮬레이션한다. 입법 과정의 일부를 자동화하여 효율성을 높이고, 의사 결정자에게는 데이터에 기반한 최적의 규제 시나리오를 제공한다.

'AI 기반 규제 지능 생태계'는 AI와 같은 첨단 기술의 발전 속도를 기존의 규제가 따라가지 못하는 현상을 극복하는 것을 목적으로 내세웠다. 사후 대응적인 입법 체계에서 벗어나, 미래의 규제 수요를 예측하고 선제적으로 대응하는 '적극적 거버넌스'로의 전환을 목표로 하며, 궁극적으로 UAE의 입법 환경을 세계에서 가장 민첩하고 효율적으로 만들고자 한다.

'AI 기반 규제 지능 생태계'의 승인은 2025년 UAE AI 정책의 방향성을 제시한다. 이는 AI 기술을 '규제의 대상'으로만 보는 것이 아니라, '규제의 도구'로 활용하는 접근이다. MS-G42 파트너십과 같은 대규모 기술 투자가 안정적으로 운영되기 위해서는 예측가능하고 신속한 규제 환경을 마련할 필요가 있으며, UAE는 AI를 활용한 입법을 통해 '규제 불확실성'을 최소화하는 장치로 여기고 있는 것으로 보인다. 이는 UAE가 글로벌 AI 허브로 도약하기 위한 핵심적인 '소프트 인프라'를 구축하는 조치로 평가할 수 있다.

197) UAE Government, The General Secretariat of the Cabinet introduces its New Regulatory Intelligence Ecosystem to Government Leaders and Experts, 2025.5.

2) 미국-UAE AI 가속화 협약[198]

2025년 5월, 미국과 UAE 정부는 AI 분야의 협력을 가속하기 위한 포괄적인 정부 간 파트너십을 체결했다. '미국-UAE AI 가속화 협약(U.S.-UAE AI Acceleration Partnership)'의 목적은 양국을 AI 기술의 개발과 배포에 있어 핵심 전략적 파트너로 공식화하는 것이다. 미국은 UAE의 자본력과 AI 도입 의지를, UAE는 미국의 원천 기술과 글로벌 표준 주도권을 상호 인정했다. 이는 MS-G42 파트너십에서 강조된 '안전하고 신뢰할 수 있는 AI' 개발 원칙을 정부 차원에서 보장하고, 중국을 포함한 제3국의 기술적 영향력을 배제하려는 안보적 합의를 배경으로 두고 있다.

협약에는 △양국의 AI 연구 개발(R&D)을 위한 공동 투자를 촉진과 핵심 인력 교류 확대 △AI의 안전성, 보안, 개인정보 보호와 관련된 기술 표준 및 규제 프레임워크의 긴밀한 조율 △AI 기술이 공공 서비스, 보건, 기후 변화 등 공통의 과제를 해결하는 데 사용될 수 있도록 협력한다는 내용이 담겼다.

UAE는 '미국-UAE AI 가속화 협약'을 통해 후술할 'Stargate 파트너십'을 미국에 약속했으며, 동시에 AI 기술 발전에 있어 미국의 공식적인 핵심 파트너로서의 지위를 확보했다. 이는 G42를 비롯한 UAE 기업들이 미국의 최첨단 AI 기술과 반도체에 안정적으로 접근할 수 있는 외교적 보증을 얻었음을 의미한다.

3) Stargate 파트너십[199]

△무바달라 △G42 △OpenAI △Oracle △NVIDIA △SoftBank Group △Cisco는 UAE의 핵심 자산들을 전략적으로 통합하여, AI 및 고성능 컴퓨팅(HPC) 분야에서 글로벌 수준의 혁신을 주도하기 위한 'Stargate 파트너십'을 체결했다.

'Stargate 파트너십'은 미국과 UAE 정부가 새롭게 구축한 '미국-UAE AI 가속화 협약' 프레임워크를 기반으로, △안전하고 △보안성이 높으며 △책임감 있는 AI를 개발하고 AI 및 첨단 기술 분야에서 협력을 강화하는 것을 목표로 한다. 더불어 UAE 기업들 역시 '미국 우선

198) UAE Government, Statement by Ambassador Yousef Al Otaiba on the U.S.-UAE AI Acceleration Partnership, 2025.5.
199) G42, Global Tech Alliance Launches Stargate UAE, 2025.5.

투자 정책'에 따라 미국의 디지털 인프라에 대한 투자를 확대할 예정이다.

'Stargate'는 참여 기업 간의 '자원 공동 활용'과 '공동 연구개발'을 골자로 하고 있다. G42가 주도하는 기술 컨소시엄으로, G42가 AI 인프라 클러스터를 건설하고, OpenAI와 오라클이 운영한다. 엔비디아는 블랙웰 GB300 시스템을 공급하며, 시스코는 제로 트러스트 보안을 기반하는 AI 연결을 지원한다. 참여 기업들은 각자가 보유한 인프라, 자본, 인력, 데이터를 공동 투입하여 AI 연구 및 상용화 프로젝트에 집중적으로 투자하며, 구체적으로는 △차세대 HPC 기술 개발 △산업별 특화 AI 애플리케이션 개발 △AI 스타트업 육성을 위한 조인트 벤처 설립 등을 우선적으로 추진한다.

이번 파트너십을 통해 건설되는 UAE-미국 AI 인프라 캠퍼스는 약 25.7㎢으로, 5기가와트의 AI 데이터센터 용량과 지역 컴퓨팅 리소스를 제공할 예정이다. △원자력 △태양광 △천연가스를 에너지원으로 사용하며, 지속 가능한 컴퓨팅 인프라를 추진하는 과학 단지가 들어설 예정이다. 최종적인 AI 클러스터는 2026년에 가동될 예정으로, 최초의 200메가와트급 컴퓨팅 인프라를 목표로 한다.

'Stargate 파트너십'은 UAE가 외부에서 수입한 기술들을 내부에서 효과적으로 소화하고 활용하기 위한 체계로, 아부다비가 UAE의 핵심 국영기업들을 AI라는 단일 목표 아래의 협력체계로 정렬하고 있다고 볼 수 있다.

4) 두바이 AI 가속화 이니셔티브[200]

2025년 10월, 함단 빈 모하메드 두바이 왕세자는 '두바이 AI 가속화 이니셔티브(Dubai AI Acceleration Initiative)'를 발표했다. 함단 빈 모하메드 두바이 왕세자는 '두바이 경제 어젠다(D33)'에 발맞춰, 정부 기관들이 AI 도구를 활용해 더 효율적이고 안전하며 혁신적인 서비스를 제공할 수 있도록 준비 태세를 강화하는 것이 목표라고 밝히며, 모하메드 빈 라시드의 비전에 따라 지식과 혁신에 기반한 통합 디지털 경제를 구축하고 글로벌 첨단 기술 허브로서의 입지를 공고히 할 것이라고 강조했다. '두바이 AI 가속화 이니셔티브'는 크게 3가지 방향으로 추진된다.

200) Government of Dubai Media Office, Hamdan bin Mohammed approves new AI initiatives to fast-track Dubai's digital transformation, 2025.10.

<표 3-59> 연방 공공서비스 AI 전략 4대 목표

구분	내용
AI 인프라 권한 강화 플랫폼	• 두바이 정부 기관들의 신속한 AI 기술 도입 및 확장을 지원하는 통합 디지털 환경 제공 • 고급 인프라와 즉시 사용 가능한 스마트 서비스를 결합
두바이 AI 가속화 태스크포스	• 정부 기관 간의 제도적 통합 강화 및 AI 도입 의사결정 가속 • 두바이 디지털 전환을 지원하는 통합 프레임워크 내에서 전략적 비전 정렬
유니콘 30 프로그램	• 두바이에서 탄생한 30개의 유망 신생 기업의 글로벌 유니콘 기업화 • 10개 핵심 이니셔티브를 통한 스타트업의 글로벌 확장 지원

제13절 대만

1. 그간의 인공지능 정책 추진 경과

<표 3-60> 대만의 AI 정책 추진 현황

시기	발표기관	정책
2018년 1월	행정원	대만 AI 행동계획 1.0(2018~2021)
2019년 9월	과학기술부	AI 기술 연구개발 지침
2023년 4월	행정원	대만 AI 행동계획(2023~2026)
2023년 8월	행정원	행정원 및 소속 기관의 생성형 AI 사용 참고 지침
2023년 11월	국가과학기술위원회	반도체 칩 주도의 대만 산업 혁신 방안
2024년 4월	디지털발전부	AI 제품과 시스템 평가 지침 초안
2024년 7월	국가과학기술위원회	AI 기본법(초안) 마련
2025년 4월	입법원	산업혁신조례 개정안 통과
2025년 6월	경제부	AI 신진 인재 양성 계획
2025년 6월	국가과학기술위원회	스마트 로봇산업 육성계획
2025년 7월	행정원	10대 AI 인프라 프로젝트
2025년 8월	행정원	AI 기본법(정부안) 승인
2025년 12월	입법원	AI 기본법 3차 독회 통과

대만은 2010년대 후반부터 국가 차원에서 AI 역량 강화를 위한 중장기 전략을 추진해 왔다. 대만 행정원은 2018년 1월 ‘대만 AI 행동계획 1.0(AI Taiwan Action Plan)’(2018~2021년)을 발표하고 AI를 활용한 산업 혁신을 본격화했다. 이 계획은 ▲AI 인재 양성 ▲반도체 칩 등 핵심기술 강화 ▲AI 혁신 허브 구축 ▲규제 완화 및 테스트베드 제공 ▲산업 디지털 전환 등 5대 분야를 중점 추진했다. ‘AI 대만 행동계획 1.0’이 마무리되는 2021년에는 구체적인 성과를 내기 시작했다. 정부와 연구기관은 대규모 투자를 기반으로 ‘타이와니아

(Taiwania)’ 슈퍼컴퓨터 시리즈를 구축해 고성능 연산 인프라를 확대하고, 실시간 객체 인식 알고리즘인 ‘YOLOv4’를 개발하는 등 AI 핵심기술 연구에서도 성과를 거두었다. 또한 대만 AI 아카데미(Taiwan AI Academy)와 AI 재단(Artificial Intelligence Foundation)이 설립되어 차세대 인재 교육과 민간 생태계 조성을 위한 기반을 마련했다. 2022년에는 과학기술부를 국가과학기술위원회(NSTC)로 개편하고 디지털발전부(MODA)를 신설하는 등 정부 조직을 정비하여 AI 및 디지털 정책 추진체계를 한층 강화했다.

대만 정부는 장기적인 AI 역량 강화를 위해 2023년 ‘대만 AI 행동계획2.0’(2023~2026년)을 새롭게 수립했다. ‘행동계획 1.0’이 컴퓨팅 인프라, 핵심기술, 인재 양성 등 역량 확충에 초점을 두었다면, ‘행동계획 2.0’은 이러한 기조를 유지하면서도 AI 윤리와 거버넌스, 사회적 수용성, 위험기관에 대한 비중을 크게 높인 것이 특징이다.

AI 기술 확산과 함께 부작용에 대비하기 위한 법·제도 논의도 병행되었다. 과학기술부는 2019년 9월 AI 기술 연구개발 지침을 마련했으며, 2023년에는 행정원이 AI 챗봇 등 생성 AI를 업무에 활용할 때 발생할 수 있는 기밀 유출 및 개인정보 침해 위험을 우려하여, ‘행정원 및 소속 기관의 생성형 AI 사용 참고 지침’을 마련했다. 이 지침은 기밀문서나 민감한 정부 데이터를 AI로 작성하거나 입력하는 행위를 금지하고, 업무상 취득한 개인정보나 내부 자료 등을 외부 AI 서비스에 올리지 못하도록 규정하는 등, 공공 데이터 보호와 안전한 활용을 위한 기본 원칙과 책임 범위를 구체적으로 제시하였다. 대만 디지털발전부는 2024년 3월 AI 제품과 시스템의 개발과 관련된 잠재적 위험을 제시해 업계에 참고 자료로 제공하기 위한 ‘AI 제품과 시스템 평가 지침’ 초안을 마련하였다. 한편, AI 서비스의 확산으로 개인정보 침해, 차별·편향, 안전·보안 및 책임소재 문제가 대두됨에 따라, NSTC는 관계 부처, 학계·산업계 전문가와의 논의를 거쳐 「AI 기본법」 초안을 마련해 2024년 의견 수렴과 수정을 거쳐 2025년 8월 행정원에서 법안이 최종 승인되었다.

대만 정부는 AI 산업 진흥 및 전 산업의 디지털 전환을 뒷받침하기 위한 정책과 법제 마련도 병행하고 있다. 대만 입법원은 2023년 1월 「산업혁신조례 개정안」을 통과시켜 2029년까지 시행하기로 했다. 이 법안은 반도체 기업 등 대만 기술혁신과 글로벌 공급망의 핵심이 되는 기업이 R&D와 첨단 생산 공정 설비 구입에 투자할 경우 각각 투자 비용의 25%와 5%를 당해연도 영업세에서 공제받을 수 있도록 했다. 대만은 2025년 4월에도 산업 전반에

AI와 클라우드 컴퓨팅 등 디지털 기술의 도입을 장려하기 위해 산업혁신조례 일부 조항 개정안을 통과시켰다.

또한 행정원은 2023년 11월 전체 회의에서 NSTC가 제출한 ‘반도체 칩 주도의 대만 산업 혁신 방안’을 통과시켰다. 이 계획은 반도체 설계·제조·패키징·테스트 등에서 대만이 보유한 세계 최고 수준의 경쟁력을 차세대 AI 기술 및 고성능 컴퓨팅과 결합하여 제조업뿐 아니라 금융, 통신, 에너지, 의료 등 전 산업의 혁신과 디지털 전환을 견인하겠다는 전략이다. 2024년부터는 1단계(향후 5년) 사업이 본격적으로 시행되면서, AI와 반도체 중심의 국가 혁신 생태계를 강화하고, 글로벌 공급망 재편 속에서 대만의 전략적 지위를 공고히 하는 것을 목표로 하고 있다.

2. 2025년 인공지능 정책 동향

1) 입법원

(1) 산업혁신조례 개정

대만 입법원은 2025년 4월 투자 세액 공제 항목에 ‘AI’를 추가하고, 더 많은 자금이 스타트업에 유입되도록 자금 조달 기회를 확대하는 한편, 핵심기술의 외부 유출을 방지하기 위해 특정 국가나 지역, 산업 또는 기술에 대하여 일정 금액 이상 투자할 경우 승인을 받도록 하는 내용을 골자로 하는 「산업혁신조례」 개정안을 통과시켰다.

개정 조례는 산업 전반에 AI와 클라우드 컴퓨팅 등 디지털 기술의 도입을 장려하고, 기업의 위기 대응 능력을 강화하며, 산업의 탄소 절감 효율 가속화를 목적으로 한다. 이를 위해 기존 투자 세액공제 항목의 ‘지능형 기계’ 항목을 유지하고, ‘지능형 기술 요소’의 정의에서 사물인터넷, 대량 정보 부분을 삭제하며, ‘AI’ 및 ‘에너지·탄소 절감’ 항목을 추가하였다. 또한, ‘지능형 기계’ 등 항목에 대해 투자한 금액이 100만 대만달러(한화 약 4,600만 원) 이상 10억 대만달러(한화 약 460억 원) 이하인 경우 법인세 공제 혜택이 적용되는데, 이번 세액공제 항목이 추가되면서 투자액의 상한선이 20억 대만달러(한화 약 920억 원)로 상향 조정되었다.

2) 행정원

(1) AI 기본법

대만 행정원은 2025년 8월 AI 개발과 응용에 우호적인 환경을 조성하고 대만을 AI 개발의 핵심 거점으로 만들기 위한 「AI 기본법(人工智慧基本法)」을 통과시켰다. 대만 정부는 2024년 7월 AI 기본법 초안을 공개한 뒤 의견 수렴을 거쳐 개정안을 마련해 이번에 행정원 회의에서 통과시켰으며, 2025년 12월에 입법원의 3차 독회(심의)를 통과해 최종 공포만을 남겨두고 있다. 대만 디지털개발부는 법안 통과 시 AI 법에 따라 AI 연구와 개발, 응용 및 거버넌스의 기본 원칙을 수립하고 대만을 글로벌 AI 개발의 핵심 축으로 만들겠다는 비전에 따른 AI 발전 방향을 확립한다는 구상이다.

법안은 AI 연구개발과 응용의 7대 원칙으로 ①지속가능성 ②인간의 자율성 ③개인정보보호와 데이터 관리 ④안전성 ⑤투명성과 설명 가능성 ⑥공정성 ⑦책임성을 제시했으며, AI 인프라와 혁신 지원, 인재 양성과 교육, 위험관리와 규제방안을 제시했으며, 기술혁신과 국제 통합 촉진에 중점을 두는 동시에 지식재산권과 개인정보 보호를 강조하였다.

AI 인프라와 혁신 지원 측면에서는 정부에 AI 연구개발과 응용 인프라 및 혁신 환경 구축을 요구하였으며, AI의 연구·개발·활용 및 인프라 구축을 촉진하기 위해 자금 지원과 세제 혜택, 보조금 등을 제공하고, 각 산업 주무 부처는 AI 기반 혁신 제품이나 서비스에 대한 규제 샌드박스를 도입하기로 하였다. 또한 AI의 혁신적 활용을 촉진하기 위해 공공-민간 파트너십을 통해 민간 부문의 참여를 확대하고, AI 인재와 기술, 인프라 관련 국제협력과 공동 연구개발을 촉진한다는 내용을 포함하였다.

인재 양성과 교육 측면에서는 국민의 AI 리터러시 향상을 위해 학교, 산업계, 공공기관 등에서 AI 교육을 꾸준히 추진하고 AI 도입으로 인한 실직자들에게 직업 상담과 재취업을 지원할 것을 요구하였다. 위험기관과 규제 측면에서는 디지털개발부 주도로 AI 위험 분류 체계를 마련하고, 각 부처에 이를 토대로 필요에 따라 위험에 상응하는 단계별 관리 기준을 수립하도록 하였다. 또한 AI 기반 의사결정의 검증 가능성과 인간의 통제 가능성을 강화하기 위해 법률이나 지침을 제정해 검증과 추적 또는 책임 메커니즘을 확립하고, 고위험 AI 시스템으로 인한 피해 구제 또는 보상 제도를 마련할 것을 규정하였다. AI 개발과 활용 과정에서 불필요한 개인정보 수집, 처리 또는 이용을 방지하고, 다문화적 가치를 반영할 수

있는 AI 학습용 데이터를 충분히 확보하며, AI 학습용 데이터 활용 시 지식재산권 침해를 방지할 것을 요구했다.

<표 3-61> 대만 10대 AI 인프라 프로젝트의 구성

분야	주요 내용
3대 스마트 응용	• AI 소프트웨어 산업 발전과 식품, 의류, 주택, 교통, 교육, 엔터테인먼트 분야에서 국제적 경쟁력 갖춘 AI 응용 제품 개발 • 제조, 서비스, 농업 등 다양한 산업에서 AI 도입으로 효율성과 생산성 개선 및 신규 시장 창출 지원 • 전국민 스마트 생활권 구축 및 교통, 외식, 의료, 공공 서비스 분야에서 AI 적용을 촉진하고 스마트 서비스 제공 추진
3대 핵심기술	• 실리콘 포토닉스(차세대 AI 컴퓨팅을 뒷받침할 광기반 반도체 기술) • 양자 기술과 AI 결합을 통한 혁신 추진 • AI 로보틱스(AI 로봇과 자동화 생태계 육성으로 AI 로봇의 글로벌 거점 목표)
4대 인프라	• 소버린 AI로 자주적 AI 기술 개발과 데이터 보안 강화 • AI 애플리케이션을 지원하는 고성능 컴퓨팅 플랫폼 구축 • 대만 전역의 AI 자원과 인프라의 공평한 분배 촉진 • 운송, 스마트시티, 제조, 금융, 기타 부문에서 AI의 심층 통합 추진

출처 : 행정원, 10대 AI 인프라 프로젝트(2025.7)

대만 정부는 동 계획에 따라 실리콘 포토닉스·양자·로봇 3대 기술 개발 프로그램과 함께 산업계 컨소시엄, 지역 AI 센터 구축 등 세부 사업을 단계적으로 전개하고 있다. 특히 실리콘 포토닉스 분야에서는 TSMC가 '실리콘 포토닉스 산업연맹(SiPhIA)'을 설립하고 엔비디아, 브로드컴 등 주요 고객과 협력하여 기술을 개발해 2026년 대량 생산을 목표로 하고 있으며, 양자 기술에서는 완전한 산업사슬 구축을 추진 중이다.

3) 디지털발전부

(1) AI 산업 발전 정책 현황[201]

대만 디지털발전부(MODA)의 황옌난(黃彦南) 장관은 2025년 6월 입법원(국회) 교통위원회의 요청에 따라 현재 추진하는 AI 산업 발전 정책에 대한 특별 보고를 진행하였다. 황옌난 장관에 따르면 디지털발전부는 △컴퓨팅 파워 △데이터 △인재 △마케팅 △자금 지원의 5대 정책 도구를 통해 대만 AI 산업 발전을 위한 전반적 전략을 추진하고 있다. 그는 AI 기술혁신이 민간 기업의 자유로운 경쟁과 혁신 동력에서 시작되어야 하며, 정부의 역할은 혁신을 주도하는 것이 아니라 정책적 수단을 통해 AI 산업의 건전한 발전을 지원하는 혁신

201) 立法院,「推動A I 之產業發展政策」專題報告, 2025.06.18.

생태계를 구축하는 것이라고 강조하였다.

우선 컴퓨팅 파워 영역에서는 AI 혁신의 문턱을 낮추기 위한 컴퓨팅 파워 지원 계획의 일환으로, 제품 개발의 초기 단계에서 아이디어를 검증할 수 있도록 AI 스타트업들에게 무료 GPU를 제공하기로 하였다. 이를 통해 AI 스타트업들이 하드웨어 투자 비용을 절감하고 혁신적인 아이디어를 실용적인 기술 솔루션으로 빠르게 전환할 수 있도록 지원한다는 방침이다.

데이터 영역에서는 방대한 고품질 데이터에 크게 의존하는 AI 개발 특성을 고려해 대만의 소버린(Sovereign) AI 기반을 강화할 수 있도록 정부 데이터의 개방을 추진하고 있다. AI 모델에 대만 언어와 문화를 반영하기 위해 중앙집중식 코퍼스(Corpus) 관리와 공유 플랫폼으로 '주권 AI 훈련 코퍼스'를 구축하고 정부 기관이 저작권을 소유한 데이터를 플랫폼에서 단계적으로 공개할 계획이다.

인재 영역에서는 AI 인재 양성을 위해 AI 전문기관과 협력해 인재 양성 목표, 인재 유형, 관련 교육과정을 개괄하는 'AI 인재 인증 가이드라인'의 개발을 추진 중으로, 정부와 학계, 민간 교육기관 간 협력을 통해 AI 도구 활용, AI 모델 개발, AI 연구의 3개 측면에서 인재 집단을 양성할 방침이다.

마케팅 측면에서는 기업에 가장 필요한 것은 보조금이 아니라 대규모의 안정적인 시장 기회라는 관점에 따라 AI 산업 대상의 네트워킹과 매칭 행사를 개최하는 한편, 제품 품질 향상과 시장 검증을 촉진하기 위한 경쟁 체계를 도입하겠다는 계획이다.

마지막으로 자금 지원 측면에서는 AI 스타트업의 성장을 지원하기 위해 행정원 국가 발전기금을 활용해 2024년 8월 수립한 'AI 혁신 투자 강화 계획'에 따라, 10년간 100억 대만 달러(한화 약 4.650억 원)를 투자할 계획이다. 이 프로그램은 기술력을 갖추고 초기 시장 검증이 완료된 AI 스타트업을 대상으로 기업 규모 확대와 신속한 제품 상용화를 지원한다.

4) 경제부

(1) AI 인재 육성 계획

대만 경제부는 2025년 3월 타이페이에서 열린 ‘AI EXPO Taiwan’ 개막식에서 국가 경쟁력 강화를 위해 20만 명의 AI 인재를 육성하겠다고 발표했다.[202] 대만의 AI 인재 육성 정책은 대만이 직면한 저출산 심화와 노동력의 급속한 고령화로 인한 구조적 인재 부족 문제에 대응하기 위한 국가 차원의 전략적 조치로서, 향후 4년간 AI 전문가 20만 명 이상 양성을 목표로 10억 대만달러를 투자할 계획이다. 이 계획은 세제 혜택, 산학협력 강화, 젊은 층 대상 경쟁 프로그램 등 다양한 인센티브 정책과 함께, 대만 중소기업 지원을 위한 산업별 AI 시범 사이트 구축 계획을 포함한다.

대만 경제부는 2025년 6월에는 국립 타이베이대학, 타이베이이공대학, 정쿵대학, 포모사대학 등 4개 대학과 협력해 ‘AI 신진 인재 양성 계획[203]’을 발표했다. 이 계획은 다양한 산업의 AI 도입 및 산업 고도화와 스마트 혁신을 촉진하기 위함으로, 대만의 신규 졸업생을 대상으로 설계된 1년 기간의 AI 응용 교육과정으로 구성되어 산업에 즉시 기여할 수 있는 실무 역량을 갖춘 AI 응용 전문가 양성을 목표로 한다. 실무 중심의 3단계로 구성되며, 1단계(4개월)는 다양한 산업 분야의 배경지식을 쌓고 실무과제와 응용 요구사항을 이해하는 산업 전문 역량 개발 과정으로 진행된다. 2단계(4개월)는 시뮬레이션 환경을 통한 학습으로 AI 응용 실무 역량을 강화한다. 3단계(4개월)는 기업 인턴십 단계로, 기업에서 AI 기술을 활용해 실제 문제 해결에 도움이 되는 실무 교육을 받는다. 교육 수료 후 참가자들은 2년 이상 기업에서 근무하며 AI 구현을 지원하게 되며, 교육생의 재정 부담을 덜기 위해 정부에서 전액을 지원한다.

5) 국가과학기술위원회(NSTC)

(1) 스마트 로봇산업 육성계획

대만 NSTC는 고령사회 진입과 노동력 인구 감소에 대응해 스마트 로봇산업의 전략적 육성을 통한 노동력 부족 보완, 고령화 의료 서비스의 지속 제공과 함께 반도체와 ICT 분야에서 대만의 우위를 로봇산업으로 확장을 위해 2025년 6월 스마트 로봇산업 육성계획을 발표

202) Taiwan News, Taiwan aims to produce 200,000 AI professionals in 4 years, 2025.3.26.
203) MOEA, 經濟部啟動「AI新秀計畫」, 培育AI應用即戰力人才, 2025.6.17.
https://www.moea.gov.tw/MNS/populace/news/News.aspx?kind=1&menu_id=40&news_id=119662

했다. 부처간 협력과 산-학-관-연 협력을 통해 자율형 스마트 로봇을 가속화하여, 대만의 세계적 스마트 로봇산업 경쟁력을 종합적으로 강화하는 동시에 다양한 미래사회 요구에 대응할 계획이다.

이 계획은 3대 목표로 시스템 개발 역량을 갖춘 로봇 스타트업 육성과 국내 생산량을 5년 내 40억 대만달러에서 500억 대만달러로 확장, 로봇 보급률 확대를 제시하였다. 또한 4대 전략으로 핵심기술 배치 확대와 산업 생태계 구축, 사이버보안 및 표준과 법규 제정, 인재 양성과 활용 촉진을 수립하였다.

<표 3-62> 대만 스마트 로봇산업 육성계획

분야	주요 내용
3대 목표	• 시스템 개발 역량을 갖춘 로봇 스타트업 육성: 정부와 민간의 공동 투자로 100억 대만달러를 관련 기업 육성에 투입 • 국내 생산량 확대: 공급망과 산업 생태계 개선으로 전문 서비스 로봇 생산가치를 5년 내 40억 대만달러에서 500억 대만달러로 확대 • 보급률 확대: 개발된 서비스 로봇을 노동 집약적 응용 분야에 도입해 국내 노동력 부족 문제를 해결
4대 전략	• 핵심기술 배치: 산학연 연계로 기초기술과 연계한 스마트 로봇 연구센터와 로봇 기술을 개발하는 스마트 로봇 혁신 및 응용 연구개발센터 설립 • 산업 생태계 구축: 스마트 로봇 연구센터 및 스마트 로봇 혁신 및 응용 연구개발센터와 대만 남부 기술창업기지를 결합한 로봇산업 클러스터를 건설해 로봇산업 생태계를 구축 • 사이버보안, 표준, 법제: 서비스형 스마트 로봇 제품의 표준 시험 추진, 사이버보안 지침 강화 등을 통해 국가 정보통신 보안을 확립 • 인재 양성과 활용 촉진: 대학 교육과정 조정, 직업훈련과 현장실습 활성화, 국내외 인재 채용 강화, 석박사급 인재 양성, 실증 현장 구축 및 판로 개척 등을 추진

출처 : NSTC, 스마트 로봇산업 육성 계획(2025.6)

제14절 호주

1. 그간의 인공지능 정책 추진 경과

<표 3-63> 호주의 주요 AI 정책 추진 현황

시기	발표기관	정책
2019년	산업부	AI 윤리원칙(8대 원칙)
2019년	CSIRO Data61	국가 AI 로드맵
2020년	산업부	글로벌 AI 파트너십(GPAI) 창설 참여
2020년	(감독기관) e세이프티위원회	온라인안전법 제정(2021년 1월 시행)
2021년	산업부	AI 액션플랜 수립
2022년	산업부	AI를 국가 중요기술로 분류하고 새로운 국가 AI 전략개발 착수
2024년	산업부	자율 AI 안전기준(Voluntary AI Safety Standard) 도입
2025년 12월	산업부	호주 국가 AI 계획(National AI Plan) 공개

호주는 2019년 말 '인공지능 윤리 원칙(AI Ethics Principles)'을 수립하며 AI 정책 기반을 마련했다. 이어 CSIRO Data61 주도로 국가 AI 로드맵(Artificial Intelligence: Solving problems, growing the economy and improving our quality of life)을 발표하여 보건·의료, 스마트 인프라, 자연자원 관리 등 3대 분야에서 AI 활용 전략을 제시했다. 이러한 초기 청사진과 함께 2019년 11월 '전국 AI 서밋(Techtonic)'을 개최하여 산·학·연 전문가들과 국가 AI 생태계 발전 방향을 논의했다. 2020년에는 글로벌 AI 거버넌스 참여를 본격화하여, 6월 캐나다·EU 등과 함께 '글로벌 AI 파트너십(GPAI)'의 창립 회원국으로 참여했다. 이를 통해 국제 무대에서 책임있는 AI 개발 논의에 적극 기여하고 호주 AI 역량을 보여줄 발판을 마련했다.[204]

204) Australian Department of Industry, AI the next frontier to solving national challenges, 2016.11.15.

한편, 국내 법·제도 정비 측면에서는 「온라인 안전법(Online Safety Act 2021)」 제정을 준비하여 AI가 생성한 온라인 불법·유해콘텐츠에 대응할 수 있는 규제 틀을 마련했다. 2021년 1월 발효된 이 법률은 e세이프티위원회(eSafety)가 딥페이크 등 AI 기반 유해 콘텐츠도 신속히 조치할 수 있는 권한을 부여하여, AI 기술 악용에 대한 초기 규제 기반을 확보했다.[205]

2021년에는 연방정부 차원의 종합 AI 지원 전략이 구체화되었다. 5월 모리슨 정부는 디지털 경제 전략의 일환으로 'AI 액션플랜'을 수립하고 1억 2,410만 호주달러 규모의 투자를 발표했다.[206] 이를 통해 국가 인공지능 센터(National AI Centre) 설립, AI 및 디지털 역량센터 4곳 신설, AI 분야 장학생 지원(약 460명, 4,730만 호주달러 규모) 등이 추진되었다.[207] 특히 국가 AI 센터는 CSIRO 산하 Data61에 설치되어 호주 전역의 AI 전문성 및 산업 역량을 결집·조율하는 역할을 맡았다. 2021년 12월 AI 센터가 공식 출범하면서 마이크로소프트 출신 스텔라 솔라(Stela Solar)를 초대 센터장으로 영입하였고, 이를 통해 중소기업의 AI 도입 지원과 산업별 AI 활용 촉진에 본격 착수했다. 이러한 정부 주도의 AI 역량 강화 노력은 2030년까지 AI가 호주 GDP에 연 6,000억 호주달러 규모로 기여할 것이라는 전망 아래, 호주를 신뢰성 있는 글로벌 AI 선도국으로 도약시키겠다는 비전을 담고 있었다. 2022년 5월 출범한 알바니지 노동당 정부는 이전 정부의 AI 진흥 기조를 계승하면서도 안전하고 책임 있는 AI 활용에 초점을 맞춘 새로운 전략 수립을 예고했다. 과학기술산업부(구 산업부)를 중심으로 AI를 국가 중요기술로 지정하고 정책 우선순위를 재편했으며,[208] 국가 AI 전략을 새로이 마련하기 위한 준비 작업에 들어갔다. 또한 2022~2023년 사이 정부는 '안전하고 책임 있는 AI' 활용 방안에 대한 공개 의견 수렴을 진행하여, 급속한 AI 기술 발전에 따른 규제 공백과 거버넌스 개선 필요성을 점검했다.[209]

이와 병행하여 개인정보 보호법 개정 검토, 산업별 기존 규제 정비 등 AI 확산에 대비한 법·제도 점검 작업도 이루어졌다. 예를 들어 재무당국과 개인정보보호청(OAIC)은 AI에도

---

205) White & Case, AI Watch: Global regulatory tracker - Australia, 2025.11.19.
206) Australian Department of Industry, Launch of National Artificial Intelligence Centre to back Australia's AI future, 2021.12.14.
207) 앞의 글
208) Australian Department of Industry, Science & Resources, List of Critical Technologies in the National Interest, 2023.5.19.
209) Australian Department of Industry, Launch of National Artificial Intelligence Centre to back Australia's AI future, 2021.12.14.

기존 법령이 기술중립적으로 적용됨을 강조하면서도, 2024년에는 생성 AI 개발·활용 시 프라이버시 가이드라인을 발행하는 등 보완조치를 내놓았다.[210] 2023년 정부는 본격적인 국가 AI 정책 전환기를 맞아 다각적인 움직임을 보였다. 우선 3월경 '호주에서의 안전하고 책임 있는 AI'에 관한 공청회를 열어 '고위험 AI 분야의 규제 옵션(이른바 AI '가드레일') '에 대한 의견을 수렴했다. 이후 2024년 1월 정부는 해당 의견 수렴에 대한 중간결과로서, 기존 법체계만으로는 일부 AI 리스크를 충분히 다루기 어렵다는 인식을 표명하고 추가 대응을 시사했다. 이에 따라 2024년 2월 초 산·학·연 전문가 12인으로 구성된 ' AI 전문가 그룹(AI Expert Group) '을 출범시켜, 투명성, 안전성 검증, 책임성 확보를 위한 구체 정책 및 고위험분야 AI 규제방안을 4개월간 자문하도록 했다.[211]

한편 산업 현장의 AI 도입 확산을 지원하기 위해, 2023년 말 정부는 AI 도입 지원센터(AI Adopt Centres) 구축을 추진했다. 그 결과 2024년 10월까지 제조업, 농업, 서비스 등 분야별로 특화된 4개의 AI Adopt Center가 개소되어 중소기업을 대상으로 현장 맞춤형 AI 적용 컨설팅과 윤리적 AI 활용 교육을 제공하기 시작했다.

국제협력 면에서도 호주는 2023년 영국 주도의 블레츨리 선언(Bletchley Declaration) 및 일본 주도의 히로시마 AI 프로세스 등에 서명함으로써, 글로벌 AI 안전성 및 거버넌스 논의에 적극 동참하였다. 이를 통해 글로벌 스탠다드 수립에 기여하고 국내 AI 정책에 국제적 책임원칙을 수용하려는 움직임을 보였다.

2024년에는 앞선 논의들을 바탕으로 자율적 가이드라인 정립과 제도 인프라 보강이 이루어졌다. 2024년 3월, 호주 정부는 업계와 공공부문이 참고할 수 있는 ' 자발적 AI 안전 기준(Voluntary AI Safety Standard) '을 최초로 발표하여, 윤리 원칙(2019년에 제정된 8대 원칙)과 국제 논의를 반영한 10대 안전 기준을 권고했다. 이 지침은 법적 구속력은 없지만, 기업들이 AI 시스템의 투명성·공정성·책임성을 높이기 위한 모범관행을 제시한 것으로 평가된다. 또한 앞서 구성된 AI 전문가 그룹의 권고를 토대로, 고위험 AI 활용에 대한 규제 필요성을 지속 검토하는 한편, 당분간은 현행 법체계 내에서 부처별 감독기능을 활용하는 영국식 분권 규제 접근을 취하기로 가닥을 잡았다. 이러한 방향은 과도한 선제 규제가 혁신

210) White & Case, AI Watch: Global regulatory tracker – Australia, 2025.11.19.
211) Australian Computer Society, Govt appoints AI Expert Group – Will guide introduction of 'guardrails'., 2024.2.15.

을 저해하지 않도록 한다는 생산성위원회 및 경쟁당국(ACCC)의 조언과 맥을 같이 한다.

정부 조직 및 예산 측면에서는, 2024년 5월 예산안을 통해 향후 5년간 약 3,990만 호주달러를 AI 정책에 배정하였다. 이 중 2,160만 호주달러는 기존 국가 AI 센터와 자문기구 운영(과학기관 CSIRO에서 산업부로 이관) 및 AI 거버넌스 조율 기능 강화에 책정되었다. 또 1,570만 호주달러는 산업별 AI 영향 분석과 규제정비 프로젝트(예: 의료, 소비자보호, 저작권 분야 점검)에 투입하고, 260만 호주달러는 국가안보 맥락의 AI 위험 대응에 배정되었다. 아울러 연방정부 디지털전환청(DTA)을 중심으로 공공부문이 AI 활용 모범 사례가 되도록 가이드라인 정비와 시범사업도 추진하기로 했다. 이러한 일련의 조치는 새로운 국가 AI 전략 수립을 앞두고 법적·제도적 기반을 보강하면서, 민간의 자율적 혁신을 유도하려는 정책 기조를 반영한 것이다.[212)]

적으로 민감한 AI 기술 이전에 대한 선례를 마련했고 볼 수 있다.

212) Australian Computer Society, 'Profound disappointment' over Budget's AI funding - Australia risks becoming uncompetitive and unproductive., 2024.5.15.

2. 2025년 인공지능 정책 동향

1) 국가 AI 계획(National AI Plan 2025)

2025년 12월, 호주 정부는 최초의 범정부 종합 「국가 AI 계획(National AI Plan 2025)」[213]을 공식 발표하였다. 이 전략 로드맵은 AI 기술을 통해 보다 공정하고 강건한 호주를 건설하고 모든 국민이 혜택을 공유하도록 하겠다는 비전을 제시한다. 계획은 ① AI 선도기술 투자 및 활용 촉진, ② AI 혜택의 사회 전반 확산, ③ 안전하고 신뢰할 수 있는 AI 구축의 세 가지 목표 아래 다양한 정책과제를 담고 있다. 호주 정부는 동 계획을 통해 AI가 의료, 돌봄, 교육, 고용 등 분야에서 국민 생활의 품질 향상과 생산성 제고에 기여할 것이라고 강조했다.[214]

특히 경제적 기회 포착과 산업경쟁력 강화가 핵심으로 부각되었다. 정부는 첨단 데이터센터 인프라에 대한 투자를 유치하고 AI 분야 인재 양성 및 재교육을 강화하여, AI로 인한 '일자리 변화에 대비한 노동력 재스킬링(reskilling)'을 지원할 계획이다. 이를 통해 AI 도입이 노동자를 대체하기보다는 근로자의 역량을 향상시키는 방향으로 나아가도록 유도하고, 기술 진보의 생산성 이익이 노동자와 기업 모두에 공정하게 배분되도록 한다는 방침이다. 또한 정부와 민간이 보유한 대규모 데이터셋 개방·공유를 촉진함으로써 AI 모델 개발을 가속화하고 혁신을 견인하겠다는 구상도 포함되었다.

규제 및 거버넌스 접근 측면에서, 호주는 이번 전략에서 새로운 AI 특별법을 제정하지 않기로 결정하여 큰 방향 전환을 명확히 했다. 이는 일부 고위험 AI 활용에 대해 엄격한 규제를 검토하던 이전 입장에서 선회한 것으로, 대신 견고한 기존 법·규제체계에 의한 관리를 지속하기로 한 것이다. 정부는 "호주의 기존 법률과 규제 프레임워크가 AI로 인한 신흥 위험을 다루는 기본 토대가 될 것"이라고 명시하며, 부처별 기존 규제기관이 자기 분야에서 AI로 인한 피해를 식별·관리하도록 역할을 부여했다.[215]

---

213) Australian Department of Industry, Science and Resources, Artificial Intelligence 홈페이지, Australia's plan for AI.
https://www.industry.gov.au/science-technology-and-innovation/technology/artificial-intelligence

214) The Guardian, 'Enable workers' talents': no need for AI legislation in Australia, Labor says, 2025.12.1.

215) Reuters, Australia rolls out AI roadmap, steps back from tougher rules, 2025.12.2.

예를 들어 금융에서는 ASIC(증권투자위원회), 개인정보는 OAIC(정보위원) 등이 현재 법령 하에서 AI 관련 이슈를 감독하게 된다. 이처럼 부문별 현행법의 기술중립적 적용을 통해 AI를 규율하되, 추후 보완입법이 필요하면 도입을 검토하는 탄력적 접근을 취하고 있다. 이는 영국 등 일부 국가와 유사한 노선으로서, 정부는 "기술 발전 속도에 맞추어 향후 규제를 지속적으로 다듬어갈 것"이라고 밝혔다. 한편 이러한 결정을 뒷받침하기 위해 정부는 2024년 논의된 고위험분야 AI 가드레일에 대한 의사결정을 유보하고, 우선 산업계 자율지침과 기존 법 집행으로 위험을 관리하는 데 집중하고 있다.[216)]

이를 위해 2025년 10월 호주 산업부 산하 ' 국가 AI 센터(NAIC) '가 발표한 「AI 도입 지침(Guidance for AI Adoption)」이 중요한 역할을 한다. 이 지침은 2024년 도입된 자율 안전 기준(VAISS)을 발전시킨 실용 지침서로서, 기업들이 AI를 도입할 때 따라야 할 6대 핵심 실천사항을 제시하고 개발자와 사용자 모두를 위한 구체적 방안을 담고 있다. 정부는 이를 통해 법적 강제 없이도 업계 전반에 신뢰성 있는 AI 활용 문화를 정착시키고자 한다.

2) 호주 AI안전연구소 설립 계획 발표

또한 정부는 AI 위험 대응 역량 강화를 위해 전문 전담기구 설립을 공식화했다. 2025년 말 국가 AI 계획 발표와 함께, 2026년까지 산업부 산하에 ' 호주 AI 안전연구소(AI Safety Institute) '를 신설할 것을 발표한 것이다. 정부는 이 연구소에 3,000만 호주달러의 예산을 투입하여 최첨단 AI 기술의 위험을 모니터링·검증하고, 안전성 평가 및 정책 권고를 수행하게 할 예정이다. 이를 통해 생성형 AI로 인한 허위정보 확산, AI 악용범죄, AI로 인한 사이버·사회적 위험 등에 선제적으로 대응하고, 필요한 경우 향후 추가 규제 도입 여부를 자문받겠다는 계획이다. 이처럼 전문기관을 통한 유연한 대응은 급변하는 AI 기술 환경에서 기민하게 위험에 대처하면서도 섣부른 입법으로 혁신을 저해하지 않으려는 정부 의도가 반영된 것이다.[217)]

호주의 2025년 AI 정책 동향은 ' 혁신 촉진과 위험기관의 균형 '으로 집약할 수 있다. 2025년 새로 수립한 「국가 AI 계획」은 경제적 가치 창출과 사회적 혜택 극대화를 목표로 대규모 투자와 인프라 강화를 추구하는 한편, 별도 입법 없이 기존 제도를 활용하는 유연한

216) White & Case, AI Watch: Global regulatory tracker – Australia, 2025.11.19.
217) The Guardian, 'Enable workers' talents': no need for AI legislation in Australia, Labor says, 2025.12.1.

규제 접근을 공식화하였다. 대신 전문 연구소 설립과 가이드라인 보급을 통해 비법령적 수단으로 AI 거버넌스를 보완하고자 하며, 국제사회와 보조를 맞춰 책임있는 AI 사용 원칙을 확산시키고 있다. 또한 호주 정부는 “계획을 지속적으로 보강하여 새 기회를 포착하고 신속하게 대응함으로써 호주인을 보호하겠다”는 입장을 밝히고 있어, 향후 기술 발전에 따라 정책을 계속 업그레이드해나갈 것으로 전망된다. 결론적으로 2025년 호주 AI 정책은 공격적인 AI 산업육성과 신중한 규제 접근을 병행함으로써, 혁신 선도국 지위 확보와 국민 신뢰 제고라는 두 가지 목적을 달성하기 위한 균형 잡힌 행보를 보여주고 있다고 할 수 있다.

제15절 우리나라

1. 그간의 인공지능 정책 추진 경과

〈표 3-64〉 우리나라의 주요 AI 정책 추진 현황

시기	발표기관	정책
2019년 12월	과학기술정보통신부·관계부처	인공지능 국가전략 - 3대 목표·9대 전략·100대 과제 제시, AI를 국가 혁신 인프라로 규정
2021년 5월	과기정통부·방통위·개인정보위	신뢰할 수 있는 인공지능 실현전략 - 기술·제도·윤리 3대 축, 10대 정책과제 제시
2022년 9월	과기정통부	대한민국 디지털 전략 - 데이터·네트워크·AI 기반 초연결·초지능 사회 비전
2023년 1월	과기정통부	제1차 데이터산업 진흥 기본계획 - 데이터·AI·클라우드 융합 생태계 조성 청사진
2023년 4월	과기정통부	초거대 AI 경쟁력 강화 방안 - 데이터·컴퓨팅·알고리즘 측면의 초거대 AI 국가전략 발표
2024년 4월	과기정통부·관계부처	AI·디지털 혁신성장 전략 - AI 3대 강국(G3) 도약을 위한 민·관 합동 전략
2024년 7월	개인정보보호위원회	AI 개발·서비스를 위한 공개된 개인정보 처리 안내서 - 공개 개인정보의 AI 학습·서비스 활용 기준과 안전조치 제시
2024년 9월	과기정통부	제1차 국가 AI 위원회 - 국가 AI전략 정책방향 발표 - AI G3 도약 비전, 4대 플래그십 프로젝트와 4대 분야 정책방향 제시
2024년 12월 / 2025년 1월	국회 / 정부	인공지능 발전과 신뢰 기반 조성 등에 관한 기본법(AI기본법) 제정 -2025.1.21 공포, 2026년 1월 22일부터 효력
2025년 2월	과기정통부·관계부처	국가 AI 역량 강화 방안 - AI 컴퓨팅 인프라 확충·차세대 모델 개발·AI 전환 가속화(AI+X) 3대 전략과 6대 핵심 과제 제시
2025년 2월	방송통신위원회	생성형 인공지능 서비스 이용자 보호 가이드라인 발표 - 안전하고 신뢰할 수 있는 생성형 AI 서비스 이용 환경을 구축하고 이용자의 권익을 보장하기 위한 가이드라인

2025년 2월	산업통상자원부	K-휴머노이드 연합 출범
2025년 4월	과기정통부·관계부처	AI 바이오 확산전략 수립
2025년 8월	개인정보보호위원회	생성형 AI 개인정보 처리 안내서 공개 - AI 개발과 활용의 전 과정에 걸쳐 개인정보보호법 적용의 불확실성을 해소하고, 기업·기관의 자율적 법률 및 규제 준수 지원
2025년 8월	서울시	행정 분야 자체 LLM 도입 계획 발표 - 공무원의 행정업무에 생성형 AI를 적용하는 '챗봇 2.0' 사업에 착수
2025년 9월	과기정통부·관계부처	대통령 직속 국가인공지능전략위원회 공식 출범 - 국가 AI 정책 제반 사항 심의 및 의결 등
2025년 10월	중소벤처기업부	AI 기반 스마트제조혁신 3.0 전략 발표
2025년 11월	교육부	모두를 위한 AI 인재양성 방안 발표 - 국민 모두의 AI 역량 함양을 지원
2025년 11월	과기정통부	AI기본법 시행령 제정안 마련 - 인재·인프라·데이터·안전 등 위임 사항을 구체화하여 법체계 안착 도모
2025년 12월	국가인공지능전략위원회	대한민국 AI 액션플랜(안) 공개 - 2030년까지 피지컬 AI 분야 세계 1위 달성을 위한 98개 핵심 과제와 300개 정책 권고사항

2019년 12월, 우리 정부는 대통령 주재로 열린 제53회 국무회의를 통해 부처 합동 「인공지능(AI) 국가 전략」을 발표하고 AI 분야 최상위 국가로의 도약을 위한 AI 정책의 기반을 마련하고 AI 정책 추진을 본격화하였다. 이를 통해 정부는 3대 목표와 함께 9대 전략 및 100대 실행과제를 제시하고, 2030년까지 AI를 통한 경제효과 455조원을 창출하고 삶의 질을 세계 10위 수준으로 도약한다는 종합적인 국가 비전과 전략을 구상하였다.[218)]

- (3대 목표) ▲세계를 선도하는 AI 생태계 구축 ▲AI를 가장 잘 활용하는 나라 ▲사람 중심의 AI 구현)

218) 과학기술정보통신부, 인공지능(AI) 국가전략발표, 2019.12.17.

산업 및 사회 전반에 걸쳐 AI 기술의 도입 및 확산이 가속화되며 예상치 못한 사회적 이슈와 우려가 대두됨에 따라 정부는 2021년 5월, 「신뢰할 수 있는 AI 실현 전략」을 수립하고, 주요 추진전략으로 제시하였다. 정부는 추진전략을 중심으로 2025년까지 단계적으로 추진할 10대 실행과제를 설정하고 AI 신뢰성을 확보하기 위한 정부 차원의 노력을 강화하였다.[219]

- (주요 추진전략) ▲(기술) 신뢰 가능한 AI 구현 환경 조성 ▲(제도) 안전한 AI 활용을 위한 기반 마련 ▲(윤리) 사회 전반의 건전한 AI 의식 확산

또한, 정부는 「대한민국 디지털 전략(’ 22.9)」과 「신성장 4.0 전략(’ 22.12)」을 발표하고 이에 대한 후속조치를 본격화하며, AI 등 신산업 필요 데이터의 전략적 생산과 새로운 데이터 공유 기반 구축을 위한 「제1차 데이터산업 진흥 기본계획(’ 23.1)」을 수립하고, AI로 인한 국민의 혜택 증진 및 산업혁신을 위한 대규모 AI 수요창출을 위한 「인공지능 일상화 및 산업 고도화 계획(’ 23.1)」을 발표하였다.[220]

2023년 3월, 정부는 관계부처 및 산업계가 함께 참여하는 초거대 AI 산업 발전을 위한 ‘인공지능 최고위 전략대화’를 개최하고, 초거대 AI 경쟁력 강화 및 산업생태계 조성 방안, 챗GPT 등 초거대AI 확산에 따른 사회적·문화적 영향력, 인공지능 관련 규제 개선 및 윤리·신뢰성 제고 방안 등 AI 관련 주요 현안에 대해 논의하였다.[221] 이후 4월, 정부는 관계부처 합동으로 「디지털플랫폼정부 실현계획」과 「초거대 인공지능 경쟁력 강화 방안」을 마련하고[222], 7월 초거대 AI 개발·고도화를 지원하는 기술·산업 인프라 확충, 민간·공공 초거대 AI 융합 등 초거대 AI 혁신 생태계 조성, 범국가 AI 혁신 제도·문화 정착 등 초거대 AI 활용지원 사업에 본격 착수하였다.[223]

2024년 4월, 정부는 생성 AI의 등장과 함께 AI가 산업을 넘어 인문·사회 분야까지 영향력이 확대되며 다층적이고 복합적인 문제를 촉발함에 따라 기존 분야별로 운영되던 포럼·자

219) 과학기술정보통신부, 신뢰할 수 있는 인공지능 구현 전략 발표, 2021.5.13
220) 과학기술정보통신부, 한덕수 총리, 세계 최고의 데이터 생태계 구축으로 인공지능이 일상화 되는 대한민국의 미래를 그리다. 2023.1.26.
221) 과학기술정보통신부, 과기정통부, 초거대 인공지능(AI) 산업 발전을 위한 인공지능 최고위 전략대화 개최, 2023.3.8.
222) 과학기술정보통신부, 디지털플랫폼정부 추진 본격 시동(초거대 인공지능(AI) 경쟁력 강화 방안 포함), 2023.4.14.
223) 과학기술정보통신부, 중소·벤처기업, 공공기관·지자체 대상 초거대 인공지능 활용지원, 첫 발 뗀다, 2023.7.3.

문위원회·협의체 등의 한계를 인식하고, 상호 연계되고 통합된 시각에서 국가 전체 AI 혁신의 방향을 이끌 거버넌스가 필요하다는 범정부적 공감대가 형성됨에 따라 민·관 AI 최고위거버넌스인 'AI전략최고위협의회'를 출범하였다.[224] 또한, 정부는 'AI전략최고위협의회' 출범식을 통해 AI 3대 강국(G3) 도약을 위한 민·관 합동 전략으로써 「AI·디지털 혁신성장 전략」을 발표하였다. 동 전략은 2022년 9월 「대한민국 디지털 전략」이 발표된 이후 생성 AI 등장에 따른 잠재적 산업·경제적 기회를 포착하고 활용하는 것에 중점을 두며, 정부는 AI를 통한 혁신이 저성장, 저출산, 고령화 등 한국이 직면한 구조적 한계를 극복할 수 있는 재도약의 기회로 인식하고, 이를 돌파하기 위한 4대 전략 분야별(혁신·산업·사회·제도) 프로젝트를 추진한다.[225]

- (전략 분야별 프로젝트) ▲(혁신) 글로벌 선도수준의 AI 기술혁신 ▲(산업) 제조 및 서비스 전반의 AI 대전환 가속화 ▲(사회) 국민 일상에 선도적 AI 도입 추진 ▲(제도) 온전한 수용을 위한 새로운 디지털 질서 정립

이와 더불어 정부는 'AI전략최고위협의회' 출범식을 통해 「AI 일상화를 위한 '24년 국민·산업·공공 프로젝트 추진계획(안)」을 발표하였다. 정부는 국가 전반의 AI 활용도 제고와 국민체감 확산을 동 추진계획(안)의 비전으로 설정하고 ▲국민의 AI 서비스 경험률을 51%('23)에서 60%('24)까지 ▲기업의 AI 도입률을 28%('23)에서 40%('24)까지 ▲공공부문의 AI 도입률을 55%('22)에서 80%('24)까지 높이는 것을 목표로 총 7,102억 원 규모의 69개 과제를 추진한다고 발표했다.[226]

<표 3-65> AI 일상화 주요 추진과제

구분	주요 내용
①국민생활	(18개 과제 / 755억 원 투입) • (건강·질병) AI 서비스 확산으로 국민의 건강 향상 기여 • (사회·복지) 장애인·어르신 일상생활 보조 및 사각지대 없는 복지 기여 • (보육·교육) 맞춤형 돌봄·교육 제공 지원으로 부모의 걱정 완화 • (문화·주거) AI 콘텐츠 전국민 이용 지원, AI 홈서비스 제공 등
②산업 현장	(24개 과제 / 2,881억 원 투입) • (법률·미디어 등 민간 전문영역) 전문가 업무보조 및 국민 접근성 향상 • (바이오·제약·의료) 신약·항체개발 첨단화 및 지역 의료 AI 확산 기여

224) 과학기술정보통신부, 민.관이 머리를 맞대고 인공지능 3대 강국 도약 방안을 모색하기 위한 인공지능 최고위 협치 체계 『인공지능 전략최고위협의회』 출범, 2024.4.4.
225) 관계부처 합동, AI G3 도약을 위한 AI·디지털 혁신성장 전략, 2024.4.4.
226) 관계부처 합동, AI 일상화를 위한 '24년 국민·산업·공공 프로젝트 추진계획(안), 2024.4.4.

구분	주요 내용
	• (소비·유통물류) 소상공인 영업지원·노동감소 및 스마트 물류관리 촉진 • (제조·공정) 공정효율 향상·비용절감 기여, AI로 인한 새로운 서비스 창출 기여 • (농·축·수산) 무인화·자동화·안전화로 전통적 지배환경의 진화 촉발
③정부 행정	(14개 과제 / 1,157억 원 투입) • (재난·안전) 화재·홍수 등 국가적 재난 상황의 위기대응 능력 향상 • (공공·행정) 대국민 서비스의 질은 향상, 행정 업무의 양은 감소
④기반 인프라	(13개 과제 / 2,309억 원 투입) • (리터러시·인재) 국민 일상의 AI 접근 향상, 현장형·실무형 인재 양성 • (격차 해소) 지역·계층 등 소외 없는 맞춤형 AI 교육 제공 • (신뢰·안전) 민간자율 AI 신뢰·안전성 확보 지원과 생성 AI로 인한 위험 대응

출처 : 관계부처 합동, AI 일상화를 위한 ’24년 국민・산업・공공 프로젝트 추진계획(안), 2024.4.4.

2024년 5월, 산업통상자원부는 AI 보편화 시대를 앞두고 산업의 변화상을 전망하고 AI 활용 산업 혁신을 위한 정책과제를 도출하기 위해 「AI 시대의 신산업정책」 수립을 위한 ‘AI산업정책위원회’를 출범하고, 「AI 시대의 신산업정책」의 6대 중점 분야(①자율제조 ②디자인 ③R&D ④유통 ⑤에너지 ⑥AI 반도체) 중 첫 번째 과제인 자율제조 관련 「AI 자율제조 전략 1.0」을 발표하였다. 동 전략은 3대 중점 분야인 ①AI 자율제조 도입 확산 ②AI 자율제조 핵심역량 확보 ③생태계 진흥을 중심으로 연간 1천억 원 이상의 예산을 투입하여, 9% 수준에 불과한 AI 자율제조 확산율을 ‘30년까지 30% 이상, 제조 생산성을 20% 이상 높이는 것을 목표로 한다.[227]

2024년 7월, 개인정보보호위원회는 7월 17일 AI 개발에 필수적인 공개 데이터가 현행 개인정보 규율체계 내에서 적법하고 안전하게 처리되도록 지원하기 위해 「인공지능(AI) 개발・서비스를 위한 공개된 개인정보 처리 안내서」를 발표하였다. 동 안내서는 공개된 개인정보 수집·활용의 법적 기준을 명확화하고, 기업이 AI 개발 및 서비스 단계에서 취해야 할 안전조치의 적정성과 관련하여 고려해야 할 최소한의 기준을 제시함으로써 공개된 개인정보를 활용하는 기업들이 개인정보 침해 이슈를 최소화하고 법적 불확실성을 해소해 혁신성장을 이룰 수 있도록 지원한다. 특히, 개인정보보호법 제15조에 따른 ‘정당한 이익’ 조항에 의해 공개된 개인정보를 AI 학습 및 서비스 개발에 활용할 수 있음을 설명하고, 해당 조항이 적용되기 위해서는 세 가지 요건을 충족해야 함을 강조한다.[228]

- (세 가지 요건) ▲AI 개발 목적의 정당성 ▲공개된 개인정보 처리의 필요성 ▲구체적 이익형량)

227) 산업통상자원부, 인공지능(AI) 시대 우리 산업이 나아갈 길을 그린다, 2024.5.8.
228) 개인정보보호위원회, 인공지능(AI) 개발·서비스에 이용되는 ‘공개 데이터’ 처리 기준 제시, 2024.7.17.

2024년 9월, 과학기술정보통신부는 대통령 주재로 개최된 제1차 국가 AI 위원회에서 AI G3 도약을 위한 국가 AI혁신 비전을 담은 「국가 AI전략 정책방향」을 발표하고, AI G3 국가 도약을 통한 글로벌 AI 중추국가 실현이라는 비전하에 민·관이 함께 범국가적으로 추진해 나갈 핵심과제인 '4대 AI 플래그십 프로젝트'와 AI 생태계의 핵심인 4대 분야의 정책 추진방향을 제시하였다.[229)]

<표 3-66> 4대 AI 플래그십 프로젝트

구분	주요 내용
프로젝트 1	• '국가AI컴퓨팅 센터' 구축 등 최신 GPU 15배 확충, 국산 AI반도체 상용화 지원
프로젝트 2	• 4년간('24~'27) 민간 총 65조 원 규모 AI분야 투자, 정부는 투자 활성화 지원
프로젝트 3	• 국가 전반 AI 大전환으로 '30년까지 AI 도입률 산업 70% + 공공 95% 달성
프로젝트 4	• AI 안전·안보역량 조기확보로 글로벌 AI 거버넌스 주도

출처 : 과학기술정보통신부, 인공지능 3대 강국(AI G3) 도약을 위한 대한민국 AI 혁신의 청사진 제시, 2024.9.26.

<표 3-67> 4대 분야 정책 추진 방향

구분	주요 내용
스타트업·인재 확충	(국가 AI 혁신을 뒷받침할 기초체력 확충을 위해 AI 스타트업·인재 육성) • 전문·특화분야 AI 스타트업 성장기반 조성과 M&A 활성화 환경 조성으로 '30년까지 AI 유니콘 기업 10개 육성 • AI인재 양성의 수월화·글로벌화·보편화를 통해 '30년까지 20만 AI인재('23, 5.1만명) 확보
기술·인프라 혁신	(AI 핵심·원천기술 확충 및 AI 인프라 혁신 추진) • AI-반도체 공동연구 협력거점 구축·운영 등을 통해 최고기술선도국과 기술협력 확대 • 미래 AI 신시장인 온디바이스 AI 경쟁력 확보 • 개인정보 규제혁신을 통해 AI 신산업 혁신 뒷받침 • 양질의 저작물 활용을 촉진하는 제도 개선으로 생성 AI에 적합한 데이터 활용기반 조성 • 저지연·저전력·Edge 네트워크 혁신 등 AI 네트워크 고도화 • 「국가기간전력망확충 특별법」 제정 등을 통해 AI 트래픽·전력 수요 증가에 대응
포용·공정기반 조성	(지속 가능한 AI발전·확산 기반 조성) • 전국민의 AI 접근·활용 보장을 위한 「디지털포용법」 제정 추진 • AI 혁신을 뒷받침할 공정경쟁 질서와 정당한 권리보호 기반 확립 • AI 개발·활용 과정에서의 부당한 차별 또는 프라이버시 침해 방지 방안 마련
글로벌 리더십 확보	(AI시대의 새로운 질서 정립 및 글로벌 AI 규범·거버넌스 주도) • AI 책임·권리 귀속관계 등 기본법리를 정립 • AI 자동화 결정시대의 개인·취약계층 보호 방안 마련 • 글로벌 AI 중추국가로서 선도국·국제기구와 AI 협력 지속 확대 • 글로벌 AI거버넌스 논의 주도

출처 : 과학기술정보통신부, 인공지능 3대 강국(AI G3) 도약을 위한 대한민국 AI 혁신의 청사진 제시, 2024.9.26.

229) 과학기술정보통신부, 인공지능 3대 강국(AI G3) 도약을 위한 대한민국 AI 혁신의 청사진 제시, 2024.9.26.

2. 2025년 인공지능 정책 동향

1) 인공지능 기본법(AI기본법) 제정[230][231]

2020년 7월 국회에서 법안으로 처음 발의된 「인공지능 발전과 신뢰 기반 조성 등에 관한 기본법(AI기본법)」은 4년이 넘는 기간 동안 다양한 의견을 수렴 및 논의의 과정을 거쳐 2024년 12월 국회 본회의에서 의결된 후, 2025년 1월 21일 법률로 제정되어 내년 1월부터 시행될 예정이다. AI기본법은 AI에 관한 국가 차원의 거버넌스 체계를 정립하고, AI 산업을 체계적으로 육성하며, AI로 인해 발생가능한 위험을 사전에 예방하기 위한 내용을 담고 있다.

2) 과학기술정보통신부, 국가 AI 역량강화방안 발표[232]

과학기술정보통신부는 2월 20일 제3차 국가인공지능위원회에서 관계부처 합동으로 AI G3 도약 실현을 위한 3대 추진전략과 그 아래 6대 핵심 과제를 제시한 「국가 AI 역량 강화 방안」을 발표하였다. 동 방안은 AI 패권 경쟁이 기업 간 대결을 넘어 국가 간 경쟁으로 변화하고 있는 상황에서 AI 컴퓨팅 인프라와 핵심 인재 육성 및 확보에 전폭적인 투자를 추진하기 위한 정부의 의지를 나타낸다.

동 방안은 ▲AI 컴퓨팅 인프라 확충 ▲차세대 AI 모델 개발 ▲AI 전환 가속화(AI+X)의 3대 추진전략에 따른 하위 6대 핵심 과제로 구성되었다.

<표 3-68> 국가 AI 역량강화방안의 추진전략 및 핵심 과제

추진전략	핵심 과제	주요 내용
AI 컴퓨팅 인프라 확충	①국가AI 컴퓨팅 인프라 확충 마스터 플랜 가동	• 첨단 GPU 1.8만 장 분 확충(~'26. 상반기) • 국가 AI 컴퓨팅센터 GPU 1만 장 확보(~'25.), 슈퍼컴퓨터 6호기 GPU 8천 장(~'26. 상반기)
	②민간 AI 컴퓨팅 인프라 투자 촉진	• AI를 국가전략기술로 지정(세제 지원 강화) • 비수도권 AI 데이터센터 전력/입지 우대 검토
	③국산 AI 반도체 성장 지원	• 글로벌 수준 개방형 HW·SW 기술 생태계 조성 • 민간·공공 패키지 도입·확산 통한 국산 AI 반도체 실증

230) 과학기술정보통신부, 인공지능 시대의 새로운 서막, 인공지능 기본법 국회 본회의 통과, 2024.12.26.
231) 국가법령정보센터, 인공지능 발전과 신뢰 기반 조성 등에 관한 기본법, 2025.1.21.
232) 과학기술정보통신부, 인공지능 컴퓨팅 기반 확충을 통한 국가 인공지능 역량 강화로 인공지능 3대 강국 도약 실현, 2025.2.20.

추진전략	핵심 과제	주요 내용
차세대 AI 모델 개발	④혁신적인 AI 알고리즘 개발 지원	• (가칭)World Best LLM(WBL) 프로젝트 추진(정예팀 선발, 전폭 지원) • '글로벌 AI 챌린지' 개최(취업·창업 연계)
	⑤고급 AI 인재 양성체계 고도화	• (글로벌 AI 프론티어랩 확대(미국→유럽, 중동 등) • '이노베이션 아카데미' 확대, '기업-대학 협력형 AX 대학원' 신설 추진
AI 전환 가속화	⑥분야별 선도 과제 추진	• 교육, 의료, 미디어·문화, 법률, 학술, 재난·안전, 공공 등 다양한 분야의 AI 전환 가속화

출처 : 관계부처합동, AI컴퓨팅 인프라 확충을 통한 국가 AI역량 강화 방안(안), 2025.2.

정부는 「국가 AI 역량 강화 방안」 발표 두 달 후인 4월 18일 국무회의에서 의결한 1조 8천억원 규모의 AI 분야 추경 정부안을 계기로 「국가 AI 역량 강화 방안」의 실질적 이행을 본격화하기 위해 즉각적이고 효율적인 재정투입을 통한 후속조치 추진을 발표하였다.[233]

<표 3-69> 국가 AI 역량강화방안의 후속조치

주요 조치	주요 내용	예산 및 규모
첨단 GPU 확보 및 인프라 확충	• 기존 자원 활용 GPU 1,166장 즉시 공급 • 민간 보유 GPU 2,600장 추가 임차 • 연내 첨단 GPU 1만 장 확보 목표 • 슈퍼컴퓨터 6호기(8,500장) 구축(~'26.상반기) • 국가 AI 컴퓨팅센터 2027년 본격 가동	• 1,723억 원 (민간 GPU 임차) • 1조 4,600억 원 (연내 GPU 1만 장 확보)
국산 AI 반도체 상용화 촉진	• NPU 실증 프로젝트 확장 • NPU 실증 인프라 구축, AX 디바이스 개발, 해외 실증 지원	• 752억 원(추경 871억 원)
민간 AI 컴퓨팅 인프라 투자제도 지원 강화	• AI 국가전략기술 지정 (세제 지원 강화) • 하위 법령 개정 통한 전략기술 범위 구체화	-
글로벌 수준 AI 모델 개발	• (가칭)World Best LLM 프로젝트 추진(정예팀 선발, 자원 집중 지원)	• 1,936억 원
세계 최고 수준 AI 인재 확보 및 양성	• '글로벌 AI 챌린지' 개최 • (가칭)'AI 패스파인더 프로젝트' (해외 석학 유치) • AI 융합 분야 박사후 연구원 집중 지원	• 100억 원(글로벌 AI 챌린지) • 50억 원(AI 패스파인더 프로젝트) • 300억 원(AI 융합 연구자)
AI 혁신 펀드 확대	• 민관 공동 AI 혁신 펀드 규모 대폭 확대	• 900억 원 → 2,000억 원(정부 출자 450억 원 → 1,000억 원)
공공 부문 AI 접목 (AX 추진)	• 공공·의료 분야 AI 접목(AX) 추진	• 총 1,793억 원(AI 분야 추경 예산)

출처 :과학기술정보통신부, 1.8조 규모 인공지능 분야 추경 정부안 계기, '국가인공지능역량 강화방안' 후속조치 본격 추진, 2025.4.18.

233) 과학기술정보통신부, 1.8조 규모 인공지능 분야 추경 정부안 계기, '국가인공지능역량 강화방안' 후속조치 본격 추진, 2025.4.18.

3) 방송통신위원회, 생성형 인공지능 서비스 이용자 보호 가이드라인 발표[234]

방송통신위원회는 생성형 AI 서비스 이용 과정에서 발생 가능한 잠재적 위험을 사전 차단하여 안전하고 신뢰할 수 있는 생성형 AI 서비스 이용 환경을 구축하고 이용자의 권익을 보장하기 위한 가이드라인을 발표하였다.

동 가이드라인은 생성형 AI 개발사 및 서비스 제공자 등에 적용되며, 생성형 AI 서비스의 전반에 걸쳐 추구해야 할 기본 원칙과 해당 원칙이 제시하는 목표를 실현하기 위한 실행 방식을 구체화하였다.

<표 3-70> 생성형 AI 서비스 이용자 보호를 위한 기본 원칙

구분	기본 원칙
①	• 생성형 AI 서비스는 인간의 존엄성을 보호하고 개인의 자유와 권리를 보장하며, 인간이 적절하게 통제하고 감독할 수 있는 방식으로 제공
②	• 생성형 AI 서비스는 AI 시스템 사용에 따른 작동 원리 및 결과, 이용자에게 미치는 영향 등에 대하여 이용자에게 알기 쉽게 설명되도록 제공
③	• 생성형 AI 서비스는 안전하게 작동되어야 하며, 예상치 못한 피해를 최소화하고 악의적으로 이용되거나 변경되지 않도록 제공
④	• 생성형 AI 서비스는 이용자에게 차별 또는 불공정한 결과가 발생하지 않도록 제공

출처 : 방송통신위원회, “「생성형 인공지능 서비스 이용자 보호 가이드라인」 발표”, 2025.2.28.

<표 3-71> 생성형 AI 서비스 이용자 보호를 위한 실행 방식

목표	실행 방식
①생성형 AI 서비스의 이용자 인격권 보호	• (개발사) 생성형 AI 기반 시스템이 이용자의 인격권을 침해할 수 있는 요소를 발견·통제 가능한 알고리즘 구축 • (서비스 제공자) 서비스 구축 과정에서 이용자가 타인의 인격권을 침해하는 결과물을 생성하지 않도록 사전에 관리할 책임 인지, 내부 모니터링 체계 및 이용자 신고 프로세스 등 관련 방안 마련 • (서비스 제공자) 생성형 AI의 산출물이 이용자의 권리를 침해하지 않도록 서비스 제공 범위·방식 검토
②생성형 AI 서비스의 결정 과정을 알리려는 노력	• (서비스 제공자) 생성형 AI 산출물이 생성형 AI로 생성되었음을 자율적으로 선택한 방식을 통해 고지 • (서비스 제공자) 해당 생성형 AI 서비스 관련 영업활동을 하는 사업자의 정당한 이익을 침해하지 않는 범위에서 생성형 AI의 결정 과정에 관한 정보를 이용자가 쉽게 접근하고 이해할 수 있는 방식으로 제공

234) 방송통신위원회, “「생성형 인공지능 서비스 이용자 보호 가이드라인」 발표”, 2025.2.28.

목표	실행 방식
③생성형 AI 서비스의 다양성 존중 노력	• (개발사·서비스 제공자) 생성형 AI 개발·운영 과정에서 알고리즘 설계, 데이터의 수집과 입력 및 알고리즘 실행 등 기술적으로 실현 가능한 범위 안에서 편향과 차별을 줄이기 위해 노력 • (개발사·서비스 제공자) 자사 서비스에서 생성형 AI 산출물의 편향성을 줄이고 다양성을 보장하기 위한 자체 원칙·기준 마련 • (서비스 제공자) 생성형 AI 이용자들의 차별적 사용을 제한하기 위한 필터링 기능 제공 • (서비스 제공자) 이용자가 생성형 AI의 산출물이 편향적이라고 판단하였을 경우 즉각적으로 신고 가능한 메커니즘 제공 및 신고 내용 처리 절차 마련
④생성형 AI 서비스의 입력 데이터 수집·활용 과정에서의 관리	• (서비스 제공자) 서비스 이용 과정에서 발생하는 이용자의 입력·생성 데이터를 학습데이터로 활용 시, 이를 사전 고지하고 이용자가 동의·거부할 수 있는 절차 마련 • (서비스 제공자) 서비스 이용 과정에서 발생하는 이용자의 입력·생성 데이터를 학습데이터로 활용 시, 해당 활용 여부에 대한 이용자의 선택권 보장 • (서비스 제공자) 이용자의 입력·생성 데이터를 학습데이터로 활용하는 것과 관련하여, 기업 내 감시 조직 마련 또는 책임자 지정을 통해 이용자의 입력·생성 데이터가 안전하고 정당하게 활용되도록 관리
⑤생성 콘텐츠 활용에서 발생할 수 있는 문제 해결을 위한 책임과 참여	• (서비스 제공자) 생성형 AI 산출물 관련 이용자의 이해 제고를 위해 서비스 제공자 및 이용자의 책임 범위를 정의하고, 이용 단계에서 이용자가 자신의 책임을 명확히 인지하도록 안내 • (서비스 제공자) 이용자가 인지하지 못한 피해의 발생을 최소화하기 위해 모니터링 시스템 또는 기타 위험관리 체계 마련
⑥생성 콘텐츠의 건전한 유통 및 배포를 위한 노력	• (서비스 제공자) 이용자가 생성형 AI 서비스를 이용하여 부적절한 콘텐츠를 생성·공유하지 않도록 안내 • (서비스 제공자) 이용자의 프롬프트 입력값과 생성된 산출물의 도덕적·윤리적 기준 준수 여부 검토· 관리 • (서비스 제공자) 이용자가 의도적·비의도적으로 유해한 콘텐츠를 청소년에게 유포하지 않도록 관리

출처 : 방송통신위원회, “「생성형 인공지능 서비스 이용자 보호 가이드라인」 발표”, 2025.2.28.

4) 산업통상자원부, K-휴머노이드 연합 출범 발표[235]

산업통상자원부는 빅테크들의 휴머노이드 개발 경쟁이 심화되는 가운데 ’30년 휴머노이드 최강국 도약을 목표로 「K-휴머노이드 연합」을 출범하였다. 국내 최고 수준의 산학연 약 40개 단체로 구성된 동 연합은 ’30년까지 R&D, 펀드, M&A 등 1조 원 이상의 민·관 투자를 유치할 것으로 기대되며, 5대 과제를 중심으로 협력 방안을 모색해 나갈 예정이다.

235) 산업통상자원부, K-휴머노이드 연합 출범, 2025.4.10.

<표 3-72> K-휴머노이드 연합의 5대 과제

과제	주요 내용
①로봇 공용 AI 모델 개발	• 로봇의 두뇌에 해당하는 로봇 AI를 개발하는 것은 가장 중요한 핵심 과제로, AI 전문그룹은 '28년까지 로봇제조사 그룹이 공동으로 사용 가능한 로봇 AI 파운데이션 모델을 개발할 계획
②휴머노이드 HW 핵심기술 개발	• 로봇 제조사·부품사들은 글로벌 최고 사양을 가진 휴머노이드 HW 개발을 위한 R&D에 집중적으로 투자해 '28년까지 자체 개발 또는 협력사업을 통한 고사양 로봇을 생산할 계획 • 산업부는 로봇 R&D·인프라·실증에 예산을 투자해 기업의 기술 개발을 전폭적으로 지원할 계획으로, '25년 2천억 원 규모로 편성된 로봇 예산을 확대하기 위해 지속적으로 협의해 나갈 예정이며, 로봇제조사가 휴머노이드 개발과정에서 활용 가능한 인프라를 구축·제공할 예정 • (로봇산업 예산) ▲1,477억 원('21) ▲1,729억 원('22) ▲1,795억 원('23) ▲1,883억 원('24) ▲2,084억 원('25)
③AI 반도체, 모빌리티용 배터리 등 개발	• 휴머노이드 로봇에 필수인 고성능·저전력 온디바이스용 AI 반도체와 고밀도·장수명·고안전 배터리 개발을 위해 연합내 공동 기술개발 등 협력 방안을 모색할 방침 • 산업부는 온디바이스용 AI 반도체 개발을 위한 대규모 R&D 사업을 추진할 예정이며, 이번 연합 출범을 기반으로 AI, AI 반도체, 배터리, AI 컴퓨팅 등 AI 관련 유망 산업을 본격 육성할 계획
④스타트업, 인력 등 양성	• 우수 인력을 확보하는 것이 휴머노이드 산업의 핵심 원동력으로 간주하여 잠재력이 높은 스타트업과 인재를 본격 육성할 계획 • 유망한 연구소 및 스타트업을 지속 발굴해 연합에 포함하고 적극적으로 지원해 나갈 예정
⑤공급-수요 기업 간 협력 강화	• 휴머노이드 공급기업과 산업 현장에서 휴머노이드를 직접 활용하는 수요기업 간 협력을 촉진 • 산업부는 '24년부터 추진 중인 「AI 자율제조 선도 프로젝트」를 통해서도 로봇 공급-수용 기업 간 협력을 지원할 계획

출처 : 산업통상자원부, K-휴머노이드 연합 출범, 2025.4.10.

[그림 3-1] K-휴머노이드 연합의 추진체계



5) 과학기술정보통신부, AI 바이오 확산전략(안)」 발표[236]

정부는 「국가 AI 전략 정책방향」('24.9)과 「대한민국 바이오 대전환 전략」('25.1)을 통해 AI 등 디지털 기술의 바이오 확산을 위한 정책 방향을 제시한 바 있는 가운데 이에 기반하여 관계부처 합동으로 「AI 바이오 확산전략」을 수립하였다.

정부는 AI 기술과 첨단바이오 연구의 융합 확산을 통해 장기·고비용·불확실성의 한계를 극복하고 혁신적 성과 창출을 가속해 나갈 방침이다.

[그림 3-2] AI 바이오 확산전략의 주요 추진내용



출처 : 과학기술정보통신부, 「AI 바이오 확산전략(안)」, 2025.4.30.

<표 3-73> AI 바이오 확산전략의 주요 추진내용

항목	주요 내용
①핵심기술 확보	• (AI-바이오시스템 이해) 파운데이션 모델 구축, 유전체 정밀분석, 생체분자 구조 예측·설계 • (AI-신약개발) 신약 스크리닝·설계, 유효성·안전성 예측, 신규 모달리티 개발 • (AI-정밀의료) 맞춤형 정밀진단, 질병 조기예측, 약물 부작용 예측 • (AI-그린바이오) 고효율 합성식물체 제작, 세포 배양육 원천기술 개발, 생장촉진 바이오소재 개발

236) 과학기술정보통신부, 인공지능 기술과 첨단 생명과학(첨단바이오) 기술의 융합으로 생명과학(바이오) 혁신을 가속화하고 미래 신성장동력 창출, 2025.4.30.

항목	주요 내용
②인프라 확충	• (바이오 실험제조 인프라 구축) 자율화 실험실 구축, 바이오제조 인프라 구축 등 • (고성능 컴퓨팅 인프라 확보) S&T 컴퓨팅 자원구축, 바이오 전용 GPU 단계적 확충 • (AI 바이오 안전성 강화) 합성생물학 안전관리체계 구축, 안전성 위험 예측기술 개발
③데이터 구축	• (데이터 구축·확보) 데이터셋 개발, 민간 데이터 연계, 데이터 맵 구축 • (데이터 연계·활용 고도화) 고품질 데이터, 연계·환류체계 고도화, 국제협력 확대 • (데이터 활용환경 개선) 데이터 분양절차 간소화, 데이터 지원·육성법 제정 등
④인재 양성	• (AI 바이오 융합인재 양성) AI+바이오 융합교육 및 협력연구 활성화 • (해외 우수인재 유치) 석학급 첨단바이오 우수과학자 유치 확대 등 • (AI 바이오 혁신거점 조성) 산·학·연·병이 참여하는 거점 지정 및 집중 지원

출처 : 과학기술정보통신부, 「AI 바이오 확산전략(안)」, 2025.4.30.

6) 개인정보보호위원회, 생성형 AI 개인정보 처리 안내서 공개[237]

개인정보보호위원회는 생성형 AI 개발과 활용의 전 과정에 걸쳐 개인정보보호법 적용의 불확실성을 해소하고, 기업·기관이 자율적으로 법률 및 규제를 준수하도록 지원하기 위해 「생성형 인공지능(AI) 개발·활용을 위한 개인정보 처리 안내서」를 공개하였다.

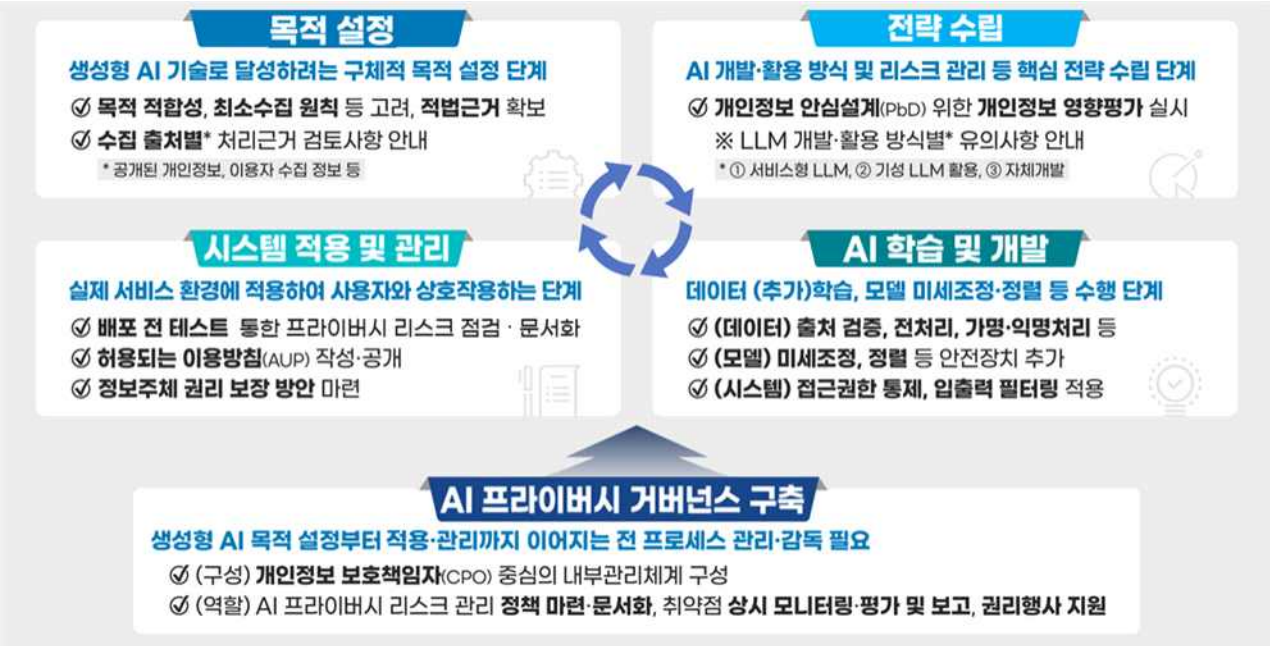
<표 3-74> 생성형 AI 개인정보 처리 안내서의 핵심 내용

항목	주요 내용
① 생애주기 분류	• 생성형 AI를 개발 및 활용하는 생애주기를 4단계(①목적 설정 ②전략 수립 ③학습 및 개발 ④적용 및 관리)로 분류하고 단계별로 확인해야 할 최소한의 안전조치를 체계적으로 제시 • AI 시스템이 실제로 개발 및 활용되는 방식과 맥락을 유형화하고, 각 유형에 따른 법적 기준과 안전성 확보 기준을 제시 • 생애주기 전과정에 걸쳐 개인정보 보호 관점을 내재화하기 위해 필수적인 개인정보보호책임자(CPO) 중심의 거버넌스 구축 방안을 제시
②불확실성 해소	• 생성형 AI 개발 및 활용 과정에서 불확실성이 높은 이슈들에 대한 기존의 정책적 경험(안내자료 마련, 사전실태점검, 규제 샌드박스, 사전적정성 검토 등)을 바탕으로 구체적인 해결방안을 제시
③최신 동향 반영	• 생성형 AI 개발 및 활용 관련 최신 기술 동향과 연구 성과 등을 반영 • 향후 급속한 기술 발전과 국내·외 개인정보보호 정책 변화에 따라 지속적으로 업데이트될 예정
④인재 양성	• (AI 바이오 융합인재 양성) AI+바이오 융합교육 및 협력연구 활성화 • (해외 우수인재 유치) 석학급 첨단바이오 우수과학자 유치 확대 등 • (AI 바이오 혁신거점 조성) 산·학·연·병이 참여하는 거점 지정 및 집중 지원

출처 : 개인정보보호위원회, 생성형 인공지능(AI) 개발·활용 위한 개인정보 처리 기준 제시한다, 2025.8.6.

237) 개인정보보호위원회, 생성형 인공지능(AI) 개발·활용 위한 개인정보 처리 기준 제시한다, 2025.8.6.

[그림 3-3] 생성형 AI 개발·활용 단계별 고려사항



출처 : 개인정보보호위원회, 생성형 인공지능(AI) 개발·활용 위한 개인정보 처리 기준 제시한다, 2025.8.6.

7) 서울시, 행정 분야 자체 LLM 도입 계획 발표[238]

서울시는 공무원의 행정업무에 생성형 AI를 적용하는 ‘챗봇 2.0’ 사업에 착수하여, ’25년 내부망에 자체 LLM(초거대 언어모델)을 도입할 계획이다. 그동안 서울시는 시민용 챗봇인 ‘서울톡’과 직원용 업무 챗봇을 통해 민원 및 행정을 자동화해 왔으나 1세대 챗봇은 모두 정해진 규칙 또는 검색 기반 고정 응답 방식으로 유연한 대화 생성이나 복합 질의에는 한계를 보여왔다.

서울시는 ‘챗봇 2.0’ 사업을 통해 1세대 챗봇의 기술적 한계를 극복하고 생성형 AI 기반 행정 지원 체계로의 전환을 추진하며, 일부 부서를 대상으로 시범 운영하고 내년부터 내부 행정 시스템과 연계한 고도화 기능을 순차적으로 확대할 예정이다.

‘챗봇 2.0’ 시스템이 본격화되면 공무원들의 문서 작성, 규정 조회, 질의응답 등 반복적인 업무를 AI가 대체하여 행정 효율성을 크게 향상할 것으로 기대되며, 서울시는 시민 50만 명이 이용 중인 ‘서울톡’에도 생성형 AI를 시범 적용하여 LLM이 맥락을 이해하고 자연어로 응답을 제공하도록 지원할 예정이다.

238) 서울특별시, 행정 분야 자체 LLM 도입… `글로벌 AI 행정도시` 시동 건다, 2025.8.11.

또한, 서울시는 직원들이 업무 특성에 맞는 생성형 AI 모델을 폭넓게 선택하고 활용할 수 있도록 지원하는 ‘생성형 AI 용량제’ 서비스를 12월까지 시범 운영한다. 서울시는 그동안 시행해온 ‘구독제’ 방식(지정된 AI 모델 한 개만을 이용 가능)을 ‘생성형 AI 용량제’ 서비스(생성형 AI 서비스를 조직 단위 사용량에 따라 비용을 지불, 사용) 방식으로 전환할 방침으로, 직원들은 ‘생성형 AI 용량제’ 서비스를 통해 챗(Chat)GPT · 제미나이(Gemini) 등 29개 AI 모델을 활용할 수 있게 된다.[239]

8) 대통령 직속 국가인공지능전략위원회 출범[240]

국가 최상위 인공지능 전략 논의기구인 대통령 직속 「국가인공지능전략위원회」이 출범한 가운데 국가 AI 정책 제반 사항 심의 및 의결, AI 관련 정책 및 사업의 부처간 조정, 이행점검, 성과 관리 등의 기능을 수행하게 된다.

동 위원회는 ①기술혁신 및 기반 시설(인프라), ②산업 인공지능 전환(AX) 및 생태계, ③공공 인공지능 전환(AX), ④데이터, ⑤사회, ⑥국제협력, ⑦과학 및 인재, ⑧국방 및 안보의 8개 분과위원회로 구성되었다.

<표 3-75> 분과위원회의 역할

분과	주요 역할
①기술혁신 및 기반 시설(인프라)	• AI 분야 연구개발 및 기반시설(인프라) 투자전략 수립, AI 데이터센터 등 기반시설(인프라) 확충, 첨단 AI 및 AI 반도체 기술 확보
②산업 인공지능 전환(AX) 및 생태계	• 산업 각 분야 AI 전환(AX) 지원, AI 분야 창업 지원 및 스타트업 육성, 산업경쟁력 강화를 위한 민관 협력, AI 기업 성장 지원
③공공 인공지능 전환(AX)	• AI 기반 행정 서비스 혁신, 재난안전·국방·치안 등 AI 기반 국민 안전 제고, 정부·공공기관 내부업무 자동화·효율화, 지방자치단체 AI 활용 촉진, 공공부문 민간 인터넷 기반 자원공유(클라우드) 도입 및 업무 환경 개선
④데이터	• 민간·공공분야 데이터 거래·유통 활성화, 제조·산업·의료·보건 등 각 산업 분야 데이터 공유 활성화, 차세대 AI 개발을 위한 데이터 구축 지원

239) 서울특별시, 29개 생성형 AI모델 골라 쓴다…'생성형AI용량제' 시범 운영, 2025.9.3.
240) 과학기술정보통신부, 국가 최상위 AI 전략 논의기구, 대통령 직속「국가인공지능전략위원회」출범, 2025.9.8.

분과	주요 역할
⑤사회	• AI 발전에 따른 교육·노동·경제·문화 등 사회 각 영역의 변화 대응, 전국민 AI 활용 역량 강화 및 리터러시 향상, AI 접근성 제고 및 격차 해소, 건전한 AI 사회 구현을 위한 AI 윤리원칙 확산
⑥국제협력	• AI 관련 국제규범 마련 주도 및 국제협력, AI 안전·신뢰 분야 국내 법·제도 마련, 국제 상호 운용성·표준 확보
⑦과학 및 인재	• AI 분야 핵심인재 양성을 위한 교육과정 개편, AI 분야 핵심 인재 육성·채용 정책 연계, 해외 우수 AI 인재의 국내 유치 지원, AI 활용 기초과학 연구 및 AI 인재 양성확보를 위한 AI 융합 연구, 과학 데이터 개방공유
⑧국방 및 안보	• AI 활용 국방·안보 역량 강화를 위한 국방 분야 AI 연구 개발 및 관련 기반 시설(인프라) 확충, AI 활용 전력 증강 및 지휘 통제 현대화, AI 분야 안보·보안 대응 체계 강화

출처 : 과학기술정보통신부, 국가 최상위 AI 전략 논의기구, 대통령 직속 「국가인공지능전략위원회」 출범, 2025.9.8.

위원회는 출범식과 함께 진행된 1차 회의를 통해 세계 AI 3대 강국 도약을 목표로하는 「대한민국 인공 지능 행동계획(AI 액션플랜)」 추진 방향을 안건으로 상정했다.

<표 3-76> AI 액션플랜의 3대 정책축 및 12대 전략 분야

정책축	전략 분야
①AI 혁신 생태계 조성	• ①AI 고속도로(그래픽 처리 장치, 데이터 등) 구축 ②차세대 AI 기술 선점 ③AI 핵심인재 확보 ④AI 모형 확보 ⑤AI 규제혁신
②범국가 AI 기반 대전환	• ⑥산업 AI 대전환 ⑦공공 AI 대전환 ⑧지역 AI 대전환 ⑨AI 기반 문화강국 ⑩AI 기반 국방강국
③글로벌 AI 기본사회 기여	• ⑪AI 기본사회 구현 ⑫국제 AI 추진전략(글로벌 AI 이니셔티브) 구현

출처: 과학기술정보통신부, 국가 최상위 AI 전략 논의기구, 대통령 직속 「국가인공지능전략위원회」 출범, 2025.9.8.

9) 중소벤처기업부 'AI 기반 스마트제조혁신 3.0 전략' 발표[241)]

중소벤처기업부는 중소 제조기업의 AI 도입률을 1%에서 10%로 끌어올리기 위한 '스마트 제조혁신 3.0 전략'을 발표했다. 이 전략은 단순히 장비를 보급하는 단계를 넘어, AI와 데이터를 활용한 '질적 고도화'에 초점을 맞춘다.

- 자율 제조 선도모델 육성: 전문가 컨설팅을 통해 제조 현장의 난제를 해결하는 '제조특화 AI 솔루션'을 보급하고, 스스로 공정을 판단하고 조절하는 자율 제조 환경을 지향

241) 중소벤처기업부, AI 기반 스마트제조혁신 3.0 전략 발표, 2025.10.24.

- 공급망 상생형 AI 공장: 대기업의 초거대 AI 모델을 활용해 협력 중소기업의 AX를 돕는 모델을 확대
- 인프라 확충: 제조 데이터의 표준화를 추진하고, ‘제조AI 24’ 플랫폼을 구축하여 원스톱 지원 체계를 완성

10) 교육부, 모두를 위한 AI 인재양성 방안 발표[242]

교육부는 AI 대전환 시대를 맞이하여 국민 모두의 AI 역량 함양을 지원하기 위한 「모두를 위한 인공지능(AI) 인재양성 방안(AI for All)」을 발표하였다. 동 방안은 국민 누구나 AI를 쉽게 활용할 수 있도록 초등부터 평생교육까지 전 생애주기에 걸친 AI 보편교육을 강화하고, AI 3대 강국 실현을 위한 혁신·융합인재 등 다층적인 AI 인재의 성장을 지원하는 것을 목표로 한다.

[그림 3-4] AI 인재양성 방안의 주요 정책 방향



자료: 교육부, 모두를 위한 인공지능(AI) 인재양성 방안 발표, 2025.11.10.

<표 3-77> AI 인재양성 방안의 주요 정책과제

구분	주요 과제
①생애주기별 AI 기본교육 강화	• (초중등) 현 교육과정 정보과목 내 AI 교육시간 확대, AI 중점학교 단계적 확대('25. 730교→'28. 2,000교), AI 교육과정 개정 추진 • (대학) 거점국립대 중심 AI 기본교육과정 도입, 대학 AI 교육과정 개발·운영 지원, 우수 AI 교육과정 공유 • (평생·직원교육) 재직자 AID(AI+Digital) 집중과정 확대('25. 30교→'26. 38교), 디지털 평생교육 이용권 지원

242) 교육부, 모두를 위한 인공지능(AI) 인재양성 방안 발표, 2025.11.10.

구분	주요 과제
②지역별 AI 인재양성 체계 구축	• (초중등) 17개 시도 AI 교육지원센터 전국 확대 • (대학) AI 거점대학 중심으로 지역 AI 허브 구축 • (대학원·연구기관) 글로컬랩, 두뇌한국 21 등을 통한 지역(권역) 단위 연구생태계 구축
③우수 AI 인재 조기 양성 및 안정적 성장경로 구축	• (초중등) 과학고, 영재학교 AI·SW 특화 프로그램 지원 확대('26. 전체 학교) • (대학) 학·석·박 패스트트랙 도입(8년→5.5년), 이공분야 우수인재 성장경로 신설('26. 400명, 2,000만 원 지원) • (대학원·연구) BK21 AI연구단 지원('26. 17개), 박사후연구원 제도화, 우수 유학생 조기 정착(비자발급 조건 완화) • (사회진출 지원) 국·공립 교원 처우 개선 등 교원 규제 개선, 창업 지원 확대, (가칭)국가석좌교수제 도입
④산업·학문 전반의 AI 융합인재 양성	• (초중등) STEAM 교육 강화, 지능형 과학실 확대('25. 60%→'26. 80%→'27. 100%), 질문하는 학교 선도학교 확대('25. 104교→'27. 200교) • (대학) 반도체 AX 특화인재 활성화, AX 부트캠프 도입('26. 10교), AID 전환 중점 전문대학 도입('26. 24교 내외), 인문사회 중심 융합인재 양성 확대 • (대학원) AI 및 AX 대학(원) 정원 확대, 두뇌한국 21 AX 교육연구단('26. 3개, 42억 원 지원)
⑤공동AI 인재 양성모델 활성화	• 기업(사내대학원)과 학교(마이스터고, 산업학위제) 간 협력(부트캠프·인턴십) 및 연계(계약학과·계약정원) 강화

출처: 교육부, 모두를 위한 인공지능(AI) 인재양성 방안 발표, 2025.11.10.

11) 과학기술정보통신부, 인공지능 기본법 시행령 제정안 마련[243]

과학기술정보통신부가 2026년 1월부터 시행되는 국내 첫 AI 종합법인 「인공지능(AI) 발전과 신뢰 기반 조성 등에 관한 기본법(AI기본법)」 관련 시행령 제정안을 공개하고 40일간 입법예고에 들어갔다. 시행령 제정안은 AI 기본법의 조속한 안착을 위해 법률에서 위임된 사항을 구체화·명확화하고 국내 AI 산업 육성과 안전·신뢰 기반 조성 가치를 중점적으로 고려하여 균형있게 반영하였다.

<표 3-78> AI기본법 시행령 제정안의 주요내용

구분	주요 과제
①AI 산업 육성을 위한 지원 사업 기준 등 명확화	• 국내 AI 산업 육성을 위해 법률에서 정하는 AI 연구개발, 학습용 데이터 구축, AI 도입·활용 등 지원 대상·기준·내용을 명확히 규정 • AI 집적단지의 지정 기준·절차 마련을 통해 국내 AI 산업지원을 위한 법제도 기반을 구체화

243) 과학기술정보통신부, 인공지능 산업 발전과 안전·신뢰 기반 조성을 위한 「인공지능 기본법 시행령」 제정안 입법예고, 2025.11.12.

구분	주요 과제
②국가 AI 정책 추진을 위한 지원 기관의 지정·운영 등 규정	• AI 산업 육성과 안전·신뢰 기반 조성을 위한 지원 기관에 관한 사항을 구체 • AI 안전·신뢰 업무를 전문적·효율적으로 수행하기 위한 인공지능 안전연구소, AI 관련 정책 개발과 국제 규범 정립·확산을 지원하기 위한 인공지능정책센터, 인공지능 집적단지 업무의 종합 지원을 위한 AI 집적단지 전담 기구 등의 지정·운영을 위한 사항을 반영
③AI 안전·신뢰 확보를 위한 제도의 구체화·명확화	• 투명성·안전성 확보 의무, 고영향 AI 판단 기준과 사업자 책무, AI 영향평가 등 AI 안전·신뢰 기반 조성과 관련한 법률상 추상적 규정을 구체화·명확화

출처: 과학기술정보통신부, 인공지능 산업 발전과 안전·신뢰 기반 조성을 위한 「인공지능 기본법 시행령」 제정안 입법예고, 2025.11.12.

과기정통부는 안전·신뢰 기반 조성 제도의 조속한 현장 안착을 위해 구체적인 기준과 의무 이행 방법, 모범 사례 등을 반영한 고시와 가이드라인을 마련할 계획이며, 초기 시행 단계에서 과태료 부과 유예 기간을 1년 이상 두고, 기업을 위한 AI 기본법 통합 지원 센터를 운영하며 법 적용에 관한 문의 사항에 대해 상세한 안내를 제공할 계획이다.

12) 국가인공지능전략위원회, ‘대한민국 AI 액션플랜(안)’ 공개[244]

국가인공지능전략위원회는 2025년 12월 15일과 16일, 글로벌 AI G3 도약을 위한 범국가적 실행 계획인 ‘대한민국 AI 액션플랜(안)’을 공개했다. 이 계획은 98개 핵심 과제와 300개 정책 권고사항을 담고 있으며, 2030년까지 피지컬 AI 분야 세계 1위 달성이라는 구체적 목표를 제시했다.

- AI 고속도로 구축: 국가 AI 컴퓨팅 센터를 구축하여 첨단 GPU 자원을 민간과 공공이 원활하게 사용할 수 있는 인프라 정책 확정
- 데이터 개방: 판결문, 국방 데이터 등 고가치 데이터의 개방 방안을 마련하여 AI 기술 고도화의 연료 공급
- 국민 체감형 서비스: AI 기반 통합 민원 플랫폼을 구축하여 대국민 서비스 혁신 도모

244) 국가인공지능전략위원회, 대한민국 인공지능행동계획(안) 주요 내용 소개 및 각계각층 의견 수렴 착수, 2025.12.15.

제4장 국제기구 인공지능 동향

2025년 전 세계 주요 국제기구가 추진한 AI 정책에서 다음과 같은 주요 트렌드를 도출해 보았다.

● UN 차원의 새로운 AI 거버넌스 아키텍처 출범

유엔총회는 2025년 국제 과학 패널(International Scientific Panel on AI)과 AI 거버넌스 글로벌 대화체(Global Dialogue on AI Governance)를 설치하는 결의안을 채택해, 과학적 자문·정책 조정 역할을 제도화하였다. UN 고위급 인공지능 자문기구는 「Governing AI for Humanity」 보고서에서 데이터·연산 자원, 인력·기관, 규범·협력이라는 세 축을 기반으로 한 다층 거버넌스 체계를 제안하며, 국가·기업·시민사회 참여를 전제로 한 글로벌 협력을 촉구하였다.

● AI 디지털 격차 및 개발도상국 지원 이슈의 부상

유엔무역개발회의(UNCTAD) 2025 기술·혁신 보고서는 AI 인프라·데이터·인력의 편중으로 인해, 선진국과 개발도상국 간 'AI 격차'가 심화될 수 있다고 경고하고, 기술·인프라·제도 역량 강화를 위한 국제 지원을 제안하였다. AI가 교육·노동·사회 구조를 재편하는 가운데, 국가·계층별 준비 수준 차이가 커지고 있는 상황에서 국제 차원에서의 공정한 접근성과 역량 구축 지원 필요성이 강조되고 있다.

● OECD.AI를 중심으로 한 글로벌 정책 허브 기능 강화

OECD.AI 플랫폼은 2025년 기준 각국의 AI 전략·법제·투자·데이터 정책을 20여 개 카테고리로 분류해 공개하며, 정책 벤치마크와 사례 공유 허브 역할을 수행 중이다. OECD 보고서는 생성형 AI를 △보편성 △지속적 개선 △혁신 촉발의 세 가지 특성을 가진 범용 기술로 규정하고, 생산성 역설 가능성과 정책적 지원 필요성을 강조하며, AI 정책 논의를 "혁신·생산성·노동전환" 축으로 재구성하고 있다.

● G7·G20를 통한 에너지·경제·보안 의제와 AI 결합

G7은 2025년 'AI for Prosperity' 성명을 통해 안전하고 신뢰할 수 있는 AI 추진, 글로벌 디지털 격차 해소, 개도국의 디지털 인프라·연결성 지원 등을 핵심 의제로 채택했고, 글로벌 디지털 컴팩트와의 연계를 강조했다. 또한 G7의 에너지·환경 장관 회의는 AI 확산이 데이터센터 전력 수요와 전력망 안정성에 미치는 영향을 고려해 에너지 효율 높은 데이터센터·반도체 기술 개발, 재생에너지 확대 등을 포함한 에너지·기후 의제와 AI 정책을 결합하였다.

● 위험기반·다중 이해관계자 거버넌스로의 수렴, 동시에 규범 경쟁 지속

UN AI 자문기구는 ITU·ISO·IEC·IEEE를 중심으로 2018년 이후 AI 국제표준 채택이 급증했으며, 2025년에도 표준화 가속 추세가 지속되고 있다. 그러나 규범·법제는 빠르게 늘고 있지만 책임성 벤치마크·평가 기준이 통일되어 있지 않아, 미국·EU·중국 등 권역 간 규제 철학 차이가 '글로벌 규범 경쟁'으로 이어지는 양상도 병존하고 있다.

제1절 OECD

1. 그간의 주요 동향

2019년 5월, 경제협력개발기구(OECD)는 「AI에 관한 이사회의 권고안(Recommendation of the Council on Artificial Intelligence)」[245]을 통해 10대 ‘AI 원칙’을 마련하였다. ‘AI 원칙’은 혁신적이고 신뢰할 수 있으며 인권과 민주적 가치를 존중하는 AI의 사용을 장려하며, 5가지 가치 기반 원칙과 5가지 정책입안자를 위한 권고사항을 제시한다.

- (가치 기반 원칙) ▲포용적 성장, 지속 가능한 개발과 웰빙 ▲인간 중심의 가치와 공정성 ▲투명성과 설명 가능성 ▲견고성, 보안성과 안전성▲ 책임성
- (정책입안자를 위한 권고사항) ▲AI R&D 투자 ▲AI를 위한 디지털 생태계 육성 ▲AI 활성화 정책 환경 제공 ▲인적 역량 강화 및 노동시장 전환 준비 ▲신뢰할 수 있는 AI를 위한 국제 협력[246]

또한, OECD는 회원국을 비롯한 전 세계 국가들의 OECD AI 원칙을 실행하도록 지원하고 관련 정보를 지속적으로 제공 및 공유하기 위해 2020년 ‘OECD AI 정책관측소(OECD.AI)를 창설하였다.[247]

2022년 2월, OECD.AI 전문가 네트워크(OECD.AI Network of Experts)는 AI 시스템의 기회와 위험을 평가하고 국가 AI 전략 수립에 활용할 수 있는 「AI 시스템 분류체계 프레임워크(OECD Framework for the Classification of AI Systems)」를 발표하고, AI 시스템 및 애플리케이션을 사람 및 지구(People & Planet), 경제적 맥락(Economic Context), 데이터 및 입력(Data & Input), AI 모델(AI Mode), 작업 및 출력(Task & Output)의 영역에 따라 AI 시스템을 평가할 것을 제안했다.[248]

245) OECD, Recommendation of the Council on Artificial Intelligence, 2022.5.
246) https://oecd.ai/en/ai-principles
247) https://oecd.ai/en/about/about-oecd-ai
248) OECD, OECD Framework for the Classification of AI systems, 2022.2.

2023년, OECD.AI는 OECD 회원국과 기타 국가들이 책임감 있고 유익한 방식으로 AI 생태계를 구축하려는 노력을 지원했다. OECD 국가들이 AI 법률과 규정이 견고한 기반을 구축하고 관할권 간 상호운용성을 촉진하기 위한 정의를 필요함에 따라 OECD는 AI 시스템 정의를 일부 수정했으며, OECD AI는 생성형 AI, 대규모 언어모델, 규제 샌드박스 및 OECD AI 원칙 이행에 관한 연구 및 분석을 제공하는데 집중하였다.[249]

특히, 2023년에는 ChatGPT가 부상하면서 OECD는 생성형 AI에 대한 작업에 전념하는 포털(oecd.ai/en/genai)을 개시하였다.[250] OECD는 또한 G7 AI 지침 원칙과 행동강령을 개발하기 위한 히로시마 프로세스를 지원했고, 생성형 AI와 관련된 기회와 과제에 대한 G7 보고서를 발행하였다. 그리고 2023년 10월에는 2019년 이후의 AI 원칙 이행 현황을 정리한 ’OECD AI 원칙 시행 4년차 현황 보고서(The state of implementation of the OECD AI Principles four years on) ‘를 발행하였다.

2023년 11월, OECD 회원국들은 AI 시스템에 대한 OECD 차원의 개정된 정의를 승인하였다. OECD는 2019년 AI 원칙에서 AI 시스템을 “인간이 정의한 일련의 목표에 대하여 실제 또는 가상 환경에 영향을 미치는 예측, 권고 또는 결정을 내릴 수 있는 기계 기반 시스템”으로 정의했었으나, AI 시스템의 범위와 역량에 대한 보다 포괄적이고 정교한 이해를 반영하여 2023년 11월 “명시적 또는 암시적 목표를 위해 주어진 입력에서 물리 또는 가상환경에 영향을 미칠 수 있는 예측, 콘텐츠, 권고 또는 결정과 같은 출력을 생성하는 방법을 추론하는 기계 기반 시스템’으로 재정의하였다. OECD는 이러한 변화를 반영하여 2024년 3월 업데이트된 AI 시스템 정의에 대해 설명하는 문서(Explanatory memorandum on the updated OECD definition of an AI system)를 발표하였다.[251]

또한, 금융 분야의 생성형 AI 도입 확산으로 편견과 차별, 데이터 침해 등 다양한 위험이 발생할 가능성이 제기됨에 따라 OECD는 2023년 12월 ‘금융 분야의 생성 AI(Generative Artificial Intelligence In Finance)’ [252] 보고서를 발간하고 금융 분야의 생성 AI 도입 현황을

249) https://oecd.ai/en/wonk/2023
250) https://oecd.ai/en/genai/issues/overview
251) OECD, Explanatory memorandum on the updated OECD definition of an AI system, 2024.3.5.; https://www.oecd.org/en/publications/explanatory-memorandum-on-the-updated-oecd-definition-of-an-ai-system_623da898-en.html
252) OECD, Generative artificial intelligence in finance, 2023.12.15.

개괄하였다.

동 보고서는 사용자 친화적이고 직관적 인터페이스를 가진 생성형 AI의 등장 이후에도 AI는 금융 분야에서 금융사의 운영 효율성을 높이기 위한 업무 자동화를 중심으로 활용되며, 인간의 개입 없는 완전 자동화와 같은 고급 활용 사례는 개발 단계에 머물러 있다고 분석하였다. 또한, 금융 분야의 생성형 AI 도입 속도가 더딘 이유로 소비자 보호와 금융 안정성을 보장하기 위한 엄격한 규제의 적용을 받기 때문이라고 설명하고, 향후 기술 발전에 따라 생성형 AI가 고객 지원과 신제품 개발, 제품 기능 최적화 등 다양한 영역에서 활용될 것이라고 전망하였다.

2024년 6월, OECD는 AI 정책집단(Policy Community)과 개인정보보호 정책집단 간 협력을 증진하기 위한 정책 문서(AI, data governance and privacy: Synergies and areas of international co-operation)를 발간하였다.[253] OECD는 동 문서를 통해 개인정보보호와 데이터 보호 고려사항을 OECD의 2019년 AI 원칙(▲포용적 성장과 지속 가능한 개발 ▲인간 중심의 가치와 공정성 ▲투명성과 설명 가능성 ▲견고성과 안전성 ▲책임성)과 매칭하여 AI와 개인정보보호 정책집단 간 협력이 가능한 영역을 식별하고 제시하였다.

OECD가 AI 원칙을 마련한 이후의 주요 정책의 흐름을 정리하면 아래(표)와 같다.

<표 4-1> OECD의 연도별 AI 정책 추진 경과

구분	주요 내용
2019	• 2019년 5월: 'OECD AI 원칙' 채택 • 2019년 6월: OECD AI 원칙에서 'G20 AI 원칙' 마련 • 2019년 6월: '사회 속의 AI(Artificial Intelligence in Society)' 도서 출간 • 2019: 'AI에 관한 OECD 글로벌 의회 네트워크(GPN AI)' 창설
2020	• 2020: 독일이 주도하고 지원하는 'OECD 노동, 혁신, 생산성 및 기술 프로그램(AI-WIPS)' 시작 • 2020: 'OECD AI 전문가 네트워크(ONE AI)' 및 'OECD AI 정책 관측소(OECD.AI) '출범 • 2020년 6월: OECD에서 주최된 GPAI(Global Partnership on Artificial Intelligence) 출범

253) OECD, AI, data governance and privacy: Synergies and areas of international co-operation, 2024.06.26.; https://www.oecd.org/en/publications/ai-data-governance-and-privacy_2476b1a4-en.html

구분	주요 내용
2021	• 2021년 6월: ‘OECD AI 원칙 이행 현황 : 국가 AI 정책의 통찰력’ 마련 • 2021년 6월: 신뢰할 수 있는 AI 시스템을 위한 구현 도구를 비교하는 프레임워크인 ‘신뢰할 수 있는 AI 도구(Tools for Trustworthy AI)’ 마련 • 2021년 10월: OECD 각료이사회 회의에서 OECD AI 원칙 실행에 대한 진행 상황 보고
2022	• 2022년 2월: 업무, 혁신, 생산성 및 기술 분야의 AI에 관한 제2차 연례 국제 컨퍼런스 • 2022년 2월: AI 시스템 분류체계 프레임워크 마련 • 2022년 3월: OECD에서 최초의 공식 ‘AI 거버넌스 작업반(AIGO)’ 창설
2023	• 2023년 10월: ECD AI 원칙 시행 4년차 현황 보고서 발행 • 2023년 11월: AI 원칙에서 ‘AI 시스템’ 재정의
2024	• 2024년 5월: 범용AI(GPAI)·생성 AI 내용 포함해 ‘OECD AI 원칙’ 업데이트 • 2024년 5월: 정채 이슈보고서 ‘Artificial intelligence, data and competition’ 발간 • 2024년 7월: 뉴델리 정상회의에서 GPAI와 OECD가 공동으로 뉴델리 선언 채택

출처: OECD.AI, About OECD.AI - Background, 2025.11.18.(확인)[254]

2. 2025년 주요 동향

1) OECD.AI 주요 동향

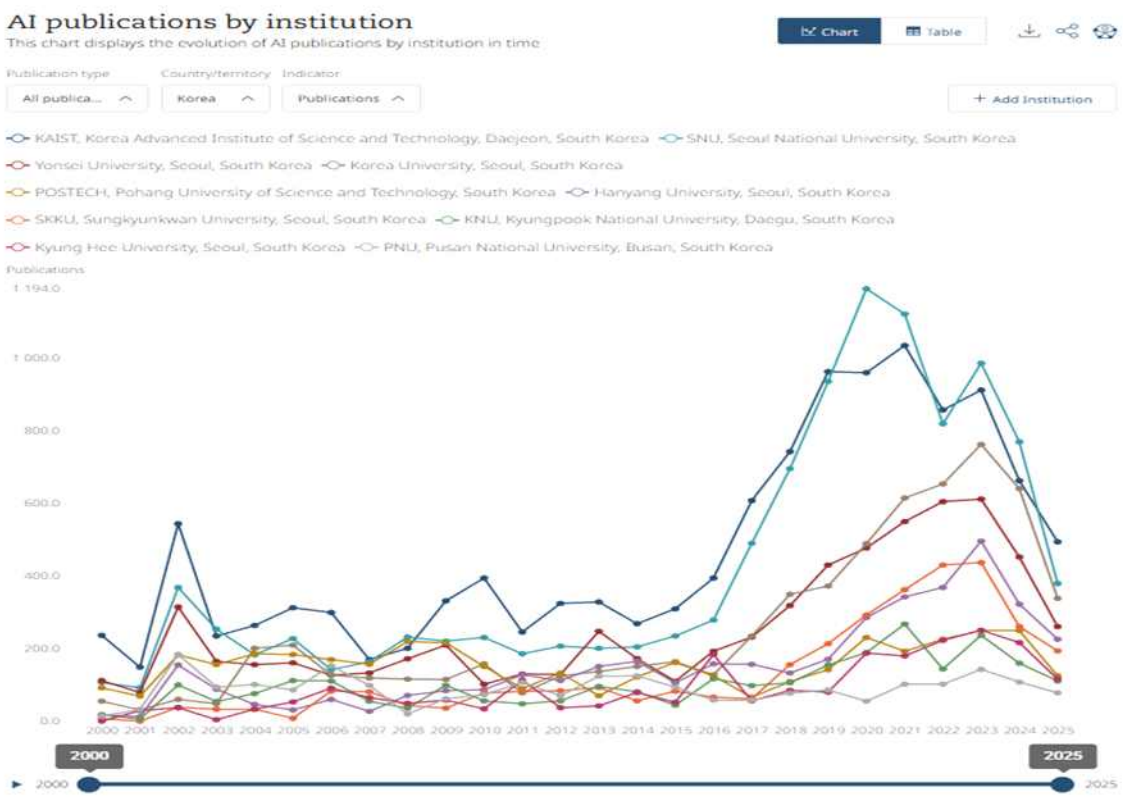
OECD.AI는 전 세계 AI 정책의 글로벌 허브의 역할을 수행하며, 선진국과 개발도상국의 모든 행위자와 이해관계자 그룹이 OECD.AI가 제공하는 도구, 데이터 및 기타 AI 정책 리소스에 자유롭게 접근할 수 있도록 지원한다. OECD.AI 작업 프로그램은 AI 위험과 책임, 잠재적인 AI 미래 연구, AI 사고 정의 및 추적, AI 컴퓨팅 용량과 AI가 기후 변화에 미치는 영향에 초점을 맞추어 추진되고 있다.

OECD.AI는 전 세계의 72개국이 국가 AI 정책 및 전략 현황에 대한 정보를 제공하고 있다. OECD.AI는 전 세계 72개국의 국가 AI 정책 및 전략 현황을 가시적으로 표현하고 각각의 국가에 대한 세부 정보로 AI 관련 연구, 소프트웨어 개발, 일자리와 스킬, 투자, AI 전문인력 분포, 검색 트렌드 등을 제공한다.[255]

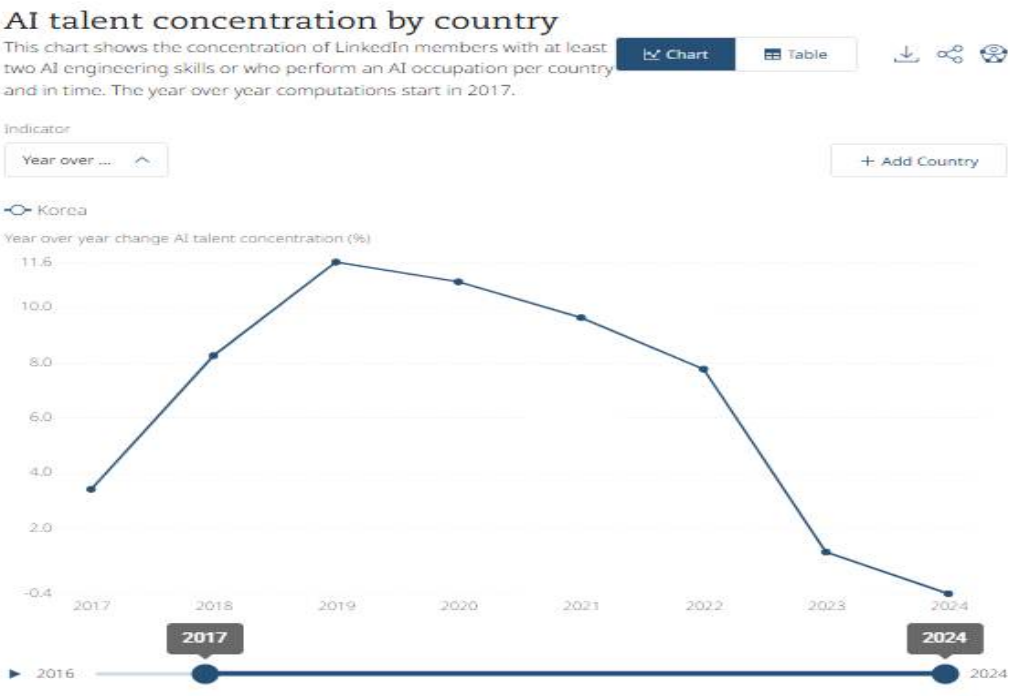
254) https://oecd.ai/en/about/background
255) https://oecd.ai/en/dashboards/overview

<표 4-2> OECD.AI를 통해 살펴본 한국의 AI 현황 예시(2025년 11월 기준)

AI 연구 성과



AI 인재 집중도



출처: OECD.AI, 2025.11.

2) OECD의 주요 동향

(1) AI 사고에 관한 공통 보고 프레임워크 개발[256)]

OECD는 AI의 개발이나 사용, 오작동으로 인한 사고에 대하여 세계 공통의 보고 프레임워크의 필요성을 강조하는 정책 보고서를 발간하였다. OECD는 AI로 인한 개인이나 집단의 건강상 피해, 중요 인프라의 중단, 인권 침해, 재산이나 지역사회・환경 피해가 발생하는 사건을 AI 사고로 통칭하고, 전 세계적으로 일관성과 상호운용성을 보장하는 공통 AI 보고 프레임워크를 통해 정책 입안자와 조직은 다른 지역에서 확인된 AI 피해로부터 필요한 교훈을 학습해 유사 사고의 재발을 방지해 나갈 것을 촉구하였다.

OECD는 OECD AI 시스템 분류 프레임워크, AI 사고 데이터베이스(AIID), OECD 글로벌 리콜 포털, AI 사고 모니터링 포털(AIM) 등 4개 자료를 바탕으로 총 29개 기준으로 구성된 AI 사고에 관한 공통 보고 프레임워크를 개발하였다. 동 프레임워크는 7개 필수 기준 외 선택 기준을 통해 추가 정보를 포함할 수 있도록 하여 AI 사고를 보고하는 유연한 구조를 제공한다.

- (필수 기준) ▲사고 명칭 ▲사고에 대한 상세 설명 ▲AI 시스템과 사고의 관련성 ▲보고 제출자 정보(소속과 역할 등) ▲심각성(위험/심각한 위험/사고/심각한 사고/재해/기타) ▲위해 유형(신체적/정신적/평판/경제・재산/환경/공익・주요기반시설/인권이나 기본권/기타) ▲사고 관련 자료

OECD는 AI 사고에 관한 공통 보고 프레임워크를 채택함으로써 국내 정책과 규제 조치를 보완하는 동시에 AI 사고 보고의 상호운용성이 강화될 것으로 기대한다. 동 프레임워크는 정책 결정권자들이 다양한 맥락에서 고위험 AI 시스템을 식별하고 현재와 미래의 AI 위험을 이해하며, 이해관계자들에게 미치는 영향을 평가할 수 있도록 지원하는 동시에 개인정보 보호, 지식재산권 또는 보안법에 대한 침해 없이 관할권 간 AI 사고 관련 지식과 정보 공유를 촉진한다.

256) OECD, Towards a common reporting framework for AI incidents, 2025.2.28.; https://www.oecd.org/en/publications/towards-a-common-reporting-framework-for-ai-incidents_f326d4ac-en.html

(2) AI와 인간의 능력을 비교한 AI 역량 지표 공개[257]

OECD는 언어, 사회적 상호작용, 문제 해결, 창의성, 메타인지와 비판적 사고, 지식·학습·기억, 시각, 조작, 로봇 지능의 9개 영역에서 인간과 AI의 능력을 비교해 이해할 수 있는 증거 기반 프레임워크를 제공하기 위한 「AI 역량 지표(AI Capability Indicators)」를 발표하였다. OECD는 5년에 걸쳐 AI 연구자, 심리학자 등 50명 이상의 전문가와 협력해 베타 버전의 지표를 개발했으며 정책 입안자들과 AI 연구자들의 피드백을 반영해 정식 버전을 공개할 예정이다.

OECD는 인간 능력의 전 범위에서 AI 발전 상황을 나타내기 위한 9가지 역량 지표를 개발하고, 2024년 11월 기준 첨단 AI 시스템이 보유한 역량 수준을 설명하였다. 동 지표는 AI 시스템이 인간과 동등한 수준으로 발전하기 위해 달성해야 할 역량을 제시하는 5단계 척도를 도입했으며, 각 단계에 속한 AI 시스템은 해당 단계에 명시된 역량 대부분을 일관적이고 안정적으로 갖추었음을 나타낸다.

1단계는 논란의 여지가 없게 확실히 검증된 단순한 역량, 5단계는 AI 시스템이 모든 인간 능력을 복제할 수 있는 수준의 역량을 의미하며, 중간의 2~4단계는 AI 시스템의 다양한 측면이 인간과 동등해질 때까지 점진적으로 발전하는 수준을 의미한다. OECD는 현재의 모든 AI 수준을 2단계와 3단계로 분류했으며, 4단계와 5단계는 통상적으로 AI가 일관되고 안정적으로 수행하기 어려워하는 수준의 역량에 해당한다.

<표 4-3> AI 역량 수준 개요(2024년 11월 기준)

영역	단계(1~5)	역량 수준
언어(Language)	3	• 다양한 데이터셋을 활용해 의미를 안정적으로 이해하고 생성 • 고도의 논리적·사회적 추론 능력을 갖추고 텍스트, 음성, 이미지 처리가 가능하며, 다양한 언어를 지원하고 반복 학습을 통해 변화에 적응
사회적 상호작용 (Social interaction)	2	• 감정을 표현하고 상호작용을 통해 지속적으로 학습하며, 과거를 기억해 적응하고, 기본적 신호를 인식하고 어조와 맥락을 통해 감정을 파악
문제 해결(Problem solving)	2	• 정성적 추론과 정량적 추론을 결합해 복잡한 전문 분야의 문제를 해결하며 시간의 흐름에 따른 변화를 예측
창의성(Creativity)	3	• 학습 데이터와 크게 다르고 기존의 경계를 뛰어넘는 가치 있는 결과물을 생성하며, 새로운 작업에서도 범용성을 발휘

257) OECD, Introducing the OECD AI Capability Indicators, 2025.6.3.: https://www.oecd.org/en/publications/introducing-the-oecd-ai-capability-indicators_be745f04-en/full-report.html

영역	단계(1~5)	역량 수준
메타인지와 비판적 사고(Metacognition and critical thinking)	2	• 이해 수준을 스스로 파악해 접근방식을 조정 • 모호한 내용이 포함된 익숙한 정보를 수치화된 신뢰도와 정보에 기반한 추측을 통해 처리하고, 아는 것과 모르는 것을 구분해 불완전한 정보도 처리
지식·학습·기억 (Knowledge, learning and memory)	3	• 데이터를 수많은 숫자(벡터)로 나누어 의미를 학습하고 새로운 상황에 적응하며, 방대한 데이터를 처리해 맥락에 맞게 이해할 수 있으나 실시간 학습 능력은 부족
시각(Vision)	3	• 객체의 외형과 조명의 변화를 처리할 수 있고 여러 하위 작업을 수행하며, 이미 학습한 데이터와 상황에서 발생한 변화를 처리
조작(Manipulation)	2	• 다양한 형태의 물체와 어느 정도 유연한 물체를 다루고 혼잡도가 낮거나 중간 정도의 통제된 환경에서 작동. 개방된 공간에서 작은 장애물을 피할 수 있고 시간 제약 없이 작업을 수행
로봇 지능(Robotic intelligence)	2	• 로봇 시스템은 주로 정적이고 半구조화된 환경에서 작동하며 단순한 다기능 작업을 처리하고 인간과 제한된 범위의 상호작용 가능

출처: OECD, Introducing the OECD AI Capability Indicators, 2025.6.3.

(3) 범용 기술로서 생성 AI의 생산성과 정책적 함의 분석[258]

OECD는 새로운 범용 기술(General-Purpose Technology, GPT)로서 생성 AI의 요건을 검증하고 그에 따른 생산성과 정책적 함의를 분석한 보고서를 발간하였다. 연구진은 생성 AI가 ▲보편성 ▲지속적 개선 ▲혁신 촉발이라는 3대 핵심 특징을 충족하는 새로운 범용 기술로 자리매김하고 생산성 향상에 기여할 잠재력을 지녔다고 평가하였다.

- (보편성) 생성 AI의 혁신은 소프트웨어 분야에서 두드러지나 생명과학, 의료, 은행, 금융 등 다양한 분야에 걸쳐 있으며 기업 내 사용은 아직 제한적이지만 개인 사이에서 도입률은 빠르게 증가
- (지속적 개선) 생성 AI 도입 이후 컴퓨팅 자원과 모델 성능이 빠르게 향상되고 있으며, 추론 능력의 발전으로 자율적 행동이 가능한 LLM 기반 AI 에이전트도 출현
- (혁신 촉발) 생성 AI 관련 특허는 컴퓨터뿐 아니라 의료, 물류 등 다양한 분야에서 기술 발전을 위한 토대로서의 영향력을 보여주며, 생성 AI는 연구 관련 프로세스 자체를 혁신하여 '발명 방법에 대한 발명(Invention of a method of invention)'이라는 잠재력을 발휘

258) OECD, Is generative AI a General Purpose Technology?: implications for productivity and policy, 2025.6.27.; https://www.oecd.org/en/publications/is-generative-ai-a-general-purpose-technology_704e2d12-en.html

동 보고서는 생성 AI가 다른 범용 기술처럼 생산성이 오히려 둔화하는 ‘생산성 역설’ 현상을 보일 수 있으며, 생산성 향상이 즉각 실현되지 않고 근로자의 기술 개발, 조직 변화 및 투자에 따라 달라질 가능성이 존재함에 정책적 지원이 중요한 역할을 담당한다고 설명한다. 또한, 정책 입안자들이 산업 전반에서 생성 AI의 광범위한 도입과 활용을 촉진하고 장기적 발전을 지원하며, 응용 분야의 후속 혁신을 장려하는 포괄적 정책의 설계와 실행을 통해 생산성 향상을 뒷받침할 것을 권고한다.

(4) AI 개방성 분석 보고서 발간[259]

OECD는 딥시크의 R1, 오픈AI의 GPT-OSS, Alibaba의 Qwen 등 ‘오픈웨이트(open-weight) 모델’ 출시를 계기로 AI 모델의 구성 요소를 공개하는 개방형 AI 모델의 이점 및 위험성에 관한 논의가 가열됨에 따라 ‘AI 개방성(openness)’의 개념을 오픈소스 소프트웨어와 비교해 정의하고, 개방형 AI 모델의 최근 동향 및 모델 개방의 이점과 위험성을 분석한 보고서를 발간하였다.

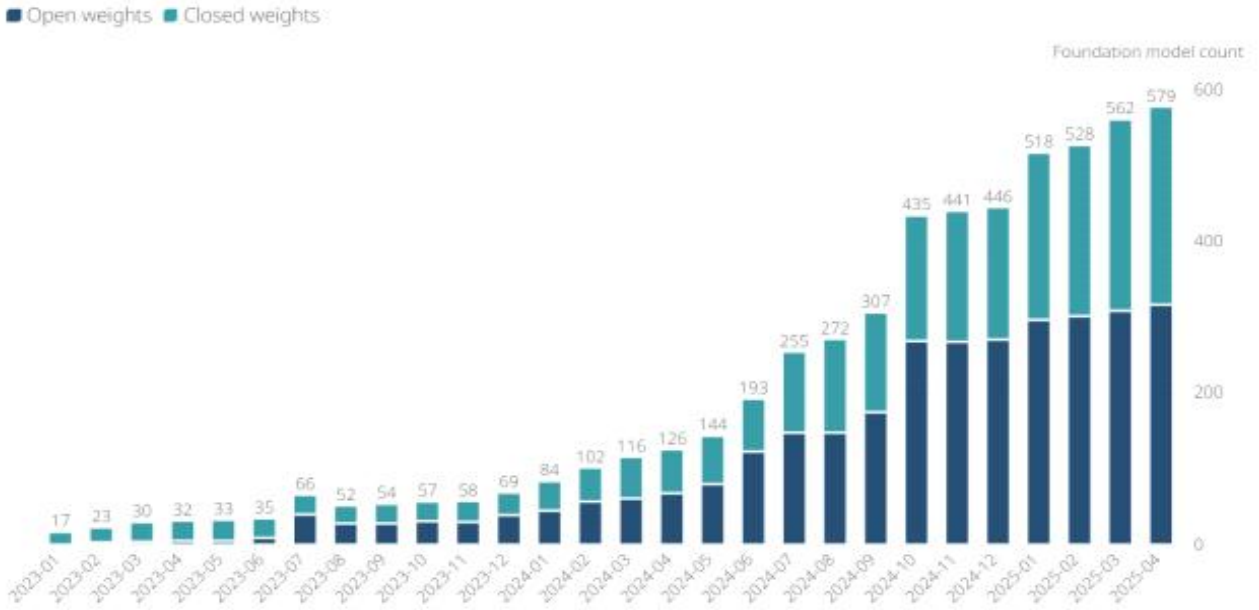
AI 개방성은 이분법적으로 나뉘지 않으며, 완전 폐쇄형 모델부터 무제한 접근, 수정, 사용을 허용하는 완전 개방형 모델까지 다양한 스펙트럼상에 위치하며 구성요소 공개 수준과 방식에 따라 유형이 달라진다. 또한, AI 개방성은 다음의 두 가지 측면에서 소프트웨어 영역에서 유래된 ‘오픈소스 소프트웨어(OSS)’와 개념적 차이가 있다.

- (소스코드 측면) OSS는 자유롭게 다운로드 가능한 소스코드를 전제로 하나, AI 모델은 추론·훈련 코드뿐 아니라 가중치·데이터 등 다양한 요소로 구성되며 각각 공개 여부가 다를 수 있음
- (가중치 측면) AI 모델 가중치는 훈련과 파인튜닝 과정의 결과물로 소스코드가 아니며, 소스코드용 라이선스가 직접 적용되지 않아 오픈소스라는 용어 사용이 부적절

2024년 중반 이후, 전 세계적으로 AI 기초 모델의 공급 규모가 급증하는 가운데 2025년 4월 기준으로 상업적으로 이용가능한 오픈웨이트 모델의 비중이 55%를 차지했으며, 미국은 강력한 AI 생태계와 클라우드 인프라 자원을 기반으로 전 세계 오픈웨이트 모델의 개발 및 제공을 압도적인 격차로 선도하고 있다.

259) OECD, AI openness-A primer for policymakers, 2025.8.14.; https://www.oecd.org/content/dam/oecd/en/publications/reports/2025/08/ai-openness_958d292b/02f73362-en.pdf

[그림 4-1] 전 세계 AI 기초 모델 공급 추이



출처: OECD, AI openness-A primer for policymakers, 2025.8.14.

개발의 측면에서 오픈웨이트 모델 개발 업체 소재지 기준으로 미국에 이어 중국과 프랑스가 핵심적인 오픈웨이트 개발 국가로 부상했고, 제공의 측면에서 네덜란드와 싱가포르가 미국에 이어 오픈웨이트 모델 제공 건수 규모에서 각각 2위와 3위를 차지했다. 한편, 2024년 중반 이후 아시아와 유럽 지역의 제공업체가 급속히 증가하며 AI 클라우드 시장의 역동성을 보여주고 있다.

2024년 초 이후, 개방형 AI 모델의 성능이 크게 향상되어 주요 벤치마크 평가에서 높은 점수를 달성했으며, 이점뿐 아니라 잠재적 위험 또한 커지면서 AI 개발에서 개방성 감독의 필요성이 증대되고 있다. 개방형 AI 모델은 의료, 고객지원, 게임, 교육 등 다양한 분야에서 AI 혜택 확산과 혁신 촉진에 기여하고 있으나, 성착취물 생성, 사이버 공격, IP・개인정보 침해 등 오용 가능성을 제기한다.

개방형 AI 모델은 모델 가중치를 공개하여 외부 평가와 책임성 강화, 연구와 혁신 가속화, 경쟁 활성화, AI 접근성 향상 등의 이점을 누릴 수 있으나 딥페이크, 지능형 사이버 공격, 아동 성착취물 생성과 같은 악의적 활동 가능성을 높이며, 생화학 분야에서의 오남용 가능성 및 유해한 목적으로 모델을 미세조정해 안전장치를 우회할 가능성이 증대하는 등 상당한 위험도 수반한다.

동 보고서는 모델 가중치의 공개 결정 시 잠재적 이점과 위험 및 비공개 시 발생할 수 있는 기회비용에 대한 신중한 고려를 강조하며, 컴퓨팅 비용의 감소와 손쉬운 미세조정 방식으로 개방형 모델의 활용 기회가 늘어나는 동시에 오남용 위험도 커질 수 있으므로, 폐쇄형 모델 대비 위험과 이점을 고려하는 총체적 위험 평가를 수행할 필요성을 제기한다.

(5) 공공 부문의 AI 활용 현황을 분석한 보고서 발간[260]

OECD가 전 세계 공공 행정 분야의 AI 도입 현황을 분석한 보고서를 발간하였다. 정부의 11개 핵심 기능에 걸친 200개의 AI 활용 사례를 분석한 결과, AI는 업무 자동화, 맞춤형 내부 프로세스와 공공서비스 구현, 의사결정 개선, 사기 탐지 강화, 공무원의 업무 품질과 학습 개선 등 다양한 이점을 수반하나, 정부의 AI 도입 확대를 저해하는 다양한 문제로 인해 상당수 AI 계획이 시범 단계에 머물러 있으며, 민간 대비 AI 활용이 뒤처진 상태로 확인되었다.

- (11개 핵심 기능) ▲공공서비스 ▲시민 참여 ▲사법 관리 ▲법 집행과 재난 대응 ▲규제 ▲공공재정관리 ▲공무원 개혁 ▲공공 조달 ▲반부패 ▲세무 행정 ▲정책 평가

분석 결과, 정부 내 AI 활용은 기능별로 차이를 나타내며, 공공서비스 설계와 제공, 시민 참여, 사법 기능을 중심으로 채택이 활발하게 이루어진 반면, 정책 평가, 세무 행정, 공무원 관리는 AI 도입이 가장 저조한 기능으로 나타났으며, 이러한 차이는 기능별 활용 범위, 규제 제약, 실행 용이성 등에 기인하였다. 또한, 정부 내에서 AI는 내부 운영과 공공 서비스 제공 분야에서는 널리 활용되는 반면, 정부 감독이나 정책 결정 분야에서는 활용도가 낮은 것으로 확인되었다.

AI 도입 확대 과정에서 발생하는 여러 위험과 도전 과제로 인해 정부의 AI 사업 중 상당수는 여전히 시범 단계에 머물러 있는 상태이다. 편향된 알고리즘으로 인한 불공정한 결과, AI 오남용으로 인한 인권 침해, 투명성과 설명 가능성 부족으로 인한 책임성 약화, AI 과의존으로 인한 디지털 격차 확대와 정부 신뢰 약화 등이 대표적 위험으로 지적되었으며, 전문 인력 부족, 양질의 데이터 확보·공유의 어려움, 전략을 뒷받침할 구체적 실행 지침의 부족

260) OECD, Governing with Artificial Intelligence: The State of Play and Way Forward in Core Government Functions, 2025.9.18.; https://www.oecd.org/en/publications/governing-with-artificial-intelligence_795de142-en.html

과 함께 성과 및 투자 대비 효과를 증명할 수 있는 평가와 감독 체계의 부재, 재정 부담도 정부의 AI 도입을 저해하는 요인으로 제기되었다.

동 보고서는 정부가 AI 도입의 이점을 극대화하면서 위험을 완화하고 도전과제를 극복하기 위해서는 거버넌스, 데이터, 인프라, 기술, 투자, 조달, 파트너십과 같은 AI 지원을 강화하고, 생애주기 전반을 포괄하는 정책과 규제 체계, 투명성과 책임성 메커니즘과 같은 안전장치를 마련하며, 시민사회와 기업 등 이해관계자 참여를 확대하는 등 정책 대응이 필요하다고 강조하였다.

제2절 UN

1. 그간의 주요 동향

2020년 6월, 유엔은 디지털 협력 로드맵(Secretary-General's Roadmap for Digital Cooperation)을 통해 AI 및 기타 디지털 기술과 관련된 다양한 문제와 국제사회의 협력 방안을 제시하였다.[261] 신뢰할 수 있고 인권 기반이며 안전하고 지속가능하며 평화를 증진하는 방식으로 AI 개발 및 사용을 위한 글로벌 역량을 구축하기 위해 AI 분야에서 다중이해관계자의 국제협력 강화의 필요성이 제기됨에 따라 안토니우 구테흐스 UN 사무총장은 포용성, 조정 및 역량 강화와 관련된 문제를 해결하기 위해 글로벌 AI 협력에 대한 다중이해관계자 자문기구 설립을 제안하였다.

2023년 10월, 유엔은 AI의 국제 거버넌스 문제를 해결하기 위해 39명으로 구성된 고위 다중이해관계자 자문기구(High-Level Multistakeholder Advisory Body on Artificial Intelligence)을 출범하였다.[262] 이 자문기구는 유엔 헌장에 명시된 보편적 가치에 기반을 두며, 진행 중인 다양한 AI 거버넌스 이니셔티브를 어떻게 연결할 수 있는지 고려한다. 이 자문기구는 AI의 국제적 거버넌스, 위험과 과제에 대한 공유된 이해, 그리고 AI를 활용하여 SDG 달성을 가속화하기 위한 주요 기회와 지원책에 대한 자문을 제공하는 역할을 수행한다. 이 자문기구는 2023년 12월, 인류를 위한 AI 거버넌스(Governing AI for Humanity)[263] 중간보고서를 발표했다. 이 중간 보고서는 국제규범과 AI가 개발되고 출시되는 방식 간의 긴밀한 조정을 요구한다. 또한 이 중간보고서는 위험에 대한 지평선 스캐닝, 데이터, 컴퓨팅 용량 및 인재에 대한 국제협력 지원과 같은 7가지 중요한 기능을 수행하여 AI의 국제적 거버넌스를 강화하여 지속 가능한 개발 목표(SDG)를 달성할 것을 제안하였다.

2024년 9월, AI 자문기구는 2023년 12월 발표한 중간보고서와 AI 전문가를 포함한 전 세계 2천 명 이상이 참여한 광범위한 협의를 바탕으로 AI의 위험을 해결하고 혁신적 잠재력을 실현하기 위한 권고안을 담은 '인류를 위한 AI 거버넌스'[264] 최종보고서를 발간하였다.

---

261) UN, Report of the Secretary-General: Roadmap for Digital Cooperation, 2020.06.
262) https://www.un.org/techenvoy/ai-advisory-body
263) https://www.un.org/sites/un2.un.org/files/un_ai_advisory_body_interim_report_press_release.pdf
264) United Nations AI Advisory Body, Governing AI for Humanity - Final Report, 2024.09.19.; https://www.un.org/sites/un2.un.org/files/governing_ai_for_humanity_final_report_en.pdf

최종보고서는 UN에 국제협력을 바탕으로 포용적이고 분산된 AI 거버넌스 체계의 토대 마련을 촉구하고, 모든 정부와 이해관계자에 AI 거버넌스에 협력하여 인권의 발전과 보호를 강화할 것을 촉구하였다.

이 최종보고서는 기존의 AI 거버넌스 논의는 전 세계에 포괄적으로 적용하기 어렵고 글로벌 사우스(Global South)[265] 국가들이 배제되어 있다고 지적하며, 글로벌 AI 거버넌스 격차 해소를 위한 다음의 7가지 행동을 촉구하였다.

- (AI 국제 과학 패널 구성) AI 기술의 과학적 측면에 대한 국제적이고 다학제 과학 패널을 구성하여 정책 입안자와 대중에게 신뢰할 수 있는 정보를 제공
- (AI 거버넌스 정책 대화) 포괄적이고 다중이해관계자를 포함하는 AI 거버넌스 정책 대화를 시작하여 모범사례를 공유하고 국가 및 지역적 접근 방식 간의 조정 및 상호운용성을 개선
- (AI 표준 교류 마련) 다양한 표준기구, 기술 기업, 시민사회 대표들 간 AI 표준 교류를 위한 포괄적인 체계를 마련해 개방적이고 상호운용이 가능하며 신뢰할 수 있는 AI 생태계를 구축
- (AI 역량 개발 네트워크 구축) 공공 및 민간 부문의 모든 사람이 AI를 책임 있고 윤리적으로 사용할 수 있도록 포괄적이고 공평한 AI 역량 개발 네트워크를 구축
- (글로벌 AI 펀드 조성) AI 혁신과 역량 강화를 위한 글로벌 기금을 조성해 AI 및 지속가능발전목표(SDG)[266] 관련 프로젝트에 자금을 지원하고, 특히 저소득 및 중간 소득 국가의 요구사항을 우선 추진
- (글로벌 AI 데이터 프레임워크 구축) AI 교육용 데이터의 출처 및 사용과 관련하여 공통 표준을 수립하기 위한 글로벌 AI 데이터 프레임워크를 구축
- (AI 사무소 설립) 이번 보고서의 이행 및 다양한 이해관계자 간 협력을 지원하고 AI 관련 자문을 제공할 AI 사무소를 UN 사무국 내에 설립

2024년 3월 21일, 유엔총회는 미국이 제안하고 중국을 포함한 120개 이상의 국가가 공동 후원한 「안전하고 신뢰할 수 있는 AI 시스템의 개발과 사용을 촉구하는 AI 결의안(Seizing the opportunities of safe, secure and trustworthy artificial intelligence systems for sustainable development)」[267]을 공식 투표 없이 컨센서스[268]로 채택하였다. 이 결의안은 AI

265) 선진국을 뜻하는 글로벌 노스와 상대되는 개념으로, 아시아, 아프리카, 남미의 개발도상국을 통칭
266) 2015년 UN 총회에서 환경, 교육, 평등, 건강, 거버넌스 등의 분야에서 2030년까지 달성하기로 합의한 17개 목표

의 설계・개발・배포・사용에서 인권의 존중과 보호를 강조하고, 모든 회원국과 이해관계자에 국제 인권법에 따를 수 없거나 인권에 과도한 위험을 초래하는 AI 시스템의 사용을 자제하거나 중단할 것을 촉구한다.

이 결의안은 안전하고 신뢰할 수 있는 AI 시스템을 증진하고 AI를 활용해 경제, 사회, 환경의 3개 차원에서 지속가능한 발전을 달성하기 위해 다음의 조치를 요구한다.

- 각국의 정책과 우선 과제에 따라 안전하고 신뢰할 수 있는 AI 시스템을 장려하는 규제와 거버넌스 접근방식 및 프레임워크의 개발과 이행
- AI의 설계・개발・배포・사용에 앞서 취약점과 위험을 확인・분류・평가・예방・완화할 수 있는 효과적이고 상호운용성 있는 조치 마련
- 최종사용자와 제3자가 기술 취약점을 신고할 수 있는 피드백 절차를 마련하고, AI 시스템 수명주기 전반에 걸친 위험기관・감독 체계와 데이터 보안 체계, 영향평가 절차 수립
- 워터마크나 라벨 등 신뢰할 수 있는 콘텐츠 인증 및 출처 확인 방식을 통해 사용자가 AI로 생성된 콘텐츠를 구별할 수 있도록 보장하고, 혁신을 촉진하면서 지식재산권을 보장하는 보호장치 마련

유엔의 교육과학문화 기구인 유네스코(UNESCO)는 2021년 AI에 대한 정부 간 협력을 통해 온라인 플랫폼인 GlobalPolicy.AI를 구축하여 운영하고 있다. Globalpolicy.AI는 정책 입안자와 대중이 국제적인 AI 거버넌스 환경을 살펴보고. AI 정책 개발에 대한 정보에 지식을 획득할 수 있도록 필요한 지식, 도구, 데이터 및 모범사례에 접근할 수 있도록 지원한다. 유네스코는 2023년 6월 전 세계를 휩쓴 챗GPT와 같은 생성 AI 분석을 담은 정책 문서(Foundation Models such as ChatGPT through the prism of the UNESCO Recommendation on the Ethics of Artificial Intelligence)[269]를 발행했다. 이 정책 문서는 윤리적 원칙에 기초를 두고 권고사항의 원칙 및 정책 영역과 관련하여 AI 기반모델의 알려진 효과를 강조하고, 사람과 조직이 이러한 기술을 설계하고 사용할 준비가 되어 있는지 확인하고 사전 및 사후에 미치는 영향을 평가하기 위해 신속한 조치의 필요성을 강조하였다.

---

267) United Nations(General Assembly), Seizing the opportunities of safe, secure and trustworthy artificial intelligence systems for sustainable development, 2024.3.11
268) 전원 합의를 의미하며, 반대 의사를 표명한 회원국이 없으면 합의된 것으로 간주하는 의사결정 방식
269) UNESCO, Foundation Models such as ChatGPT through the prism of the UNESCO Recommendation on the Ethics of Artificial Intelligence, 2023.6.

2. 2025년 주요 동향

1) UN, AI 국제과학패널 및 글로벌 AI 거버넌스 대화 신설

2025년 8월, 유엔총회는 글로벌 디지털 협약(GDC, 24.9월) 이행 결의안(79/325) 채택(8.26)[270]을 통해 AI 국제과학패널(Independent International Scientific Panel on AI) 및 글로벌 AI 거버넌스 대화(Global Dialogue on AI Governance)를 설치하기로 하고, 9월 AI 거버넌스에 관한 고위급 회의를 통해 글로벌 AI 거버넌스 대화를 시작하였다.

AI 국제과학패널은 AI의 기회, 위험 및 영향에 관한 연구를 종합·분석한 증거 기반 과학적 평가를 제공하고, 허위정보, 알고리즘 조작, 자율 무기 체계에서의 AI 활용 등 위협에 대한 조기 경보 시스템 역할을 담당하며, 정책 결정이 독립적인 과학적 평가에 기반하도록 보장하는 데 기여할 전망이다.[271]

2) UNCTAD, 2025년 기술·혁신보고서 발간[272]

2025년 4월, 유엔무역개발회의(UNCTAD)는 기술·혁신 보고서를 통해 AI 격차가 심화될 위기에 대해 경고하고, 전략적 투자, 포용적 거버넌스, 글로벌 협력을 통해 AI로 인한 혜택이 모두에게 돌아가도록 보장할 것을 촉구하였다.

유엔무역개발회의는 2033년까지 AI 시장이 4조 8천억 달러 규모(독일 경제 규모와 맞먹는 수준)에 도달할 것으로 전망하나 그 혜택은 여전히 극소수에 집중되어 있다고 지적한다. 전 세계 일자리의 최대 40%가 AI의 영향을 받을 전망으로, 선제적인 노동 정책이 요구된다. 개발도상국 중 AI 전략을 수립한 국가는 3분의 1 미만이며, 여전히 전 세계 118개국은 AI 거버넌스에서 대표성이 결여되어 글로벌 포용성을 제한하고 있다. 또한, 개발도상국은 AI의 잠재력을 최대한 활용하기 위해 AI 인프라, 데이터, 기술 역량에 대한 투자가 요구된다.

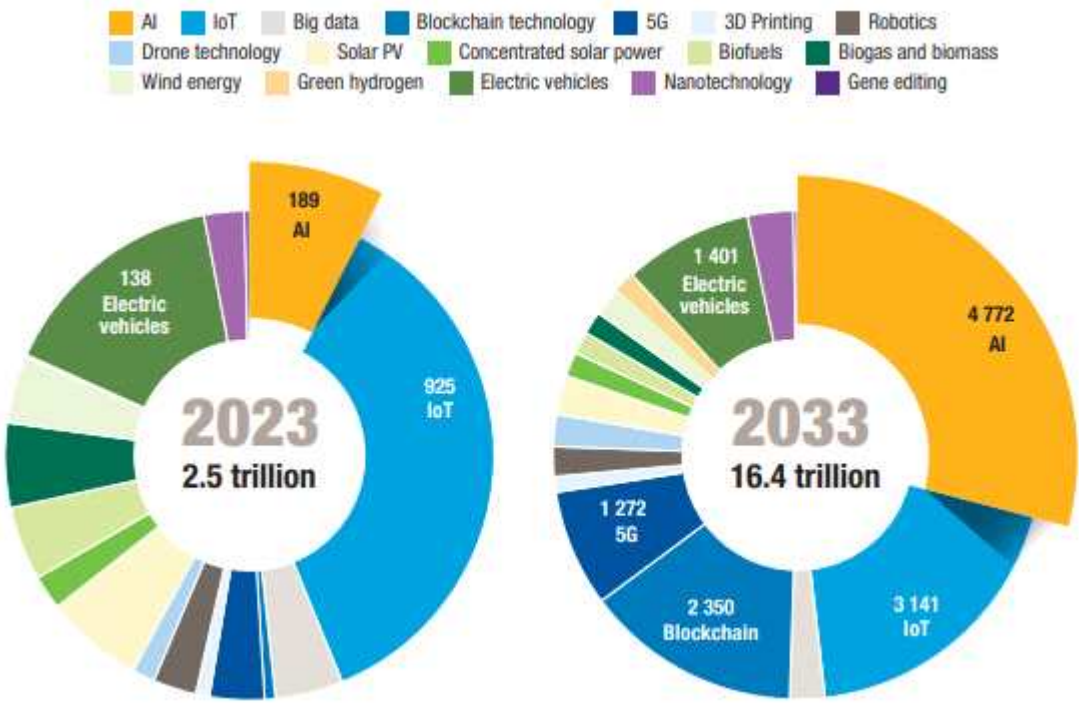
이 보고서는 AI가 격차를 심화시키기보다 포용적 성장을 주도하도록 보장하기 위한 로드

270) https://www.un.org/global-digital-compact/en/ai
271) https://unric.org/en/the-artificial-intelligence-we-want/
272) UNCTAD, Technology and Innovation Report 2025: Inclusive Artificial Intelligence for Development, 2025.4.7.; https://unctad.org/system/files/official-document/tir2025_en.pdf

맵을 제공하며, 국제사회를 위한 주요 정책 권고사항을 아래와 같이 제시한다.

- (산업계의 책임성 강화) 환경・사회・지배구조(ESG) 프레임워크와 유사하게, AI에 대한 공적 공시 메커니즘을 통해 책임성 증진 및 글로벌 차원의 약속을 실질적 성과로 전환
- (공유 인프라 구축) 글로벌 공유 인프라 구축을 통해 AI 인프라에 대한 공평한 접근성 제공
- (개방형 혁신 촉진) 오픈 데이터 및 오픈소스와 같은 모델을 통해 지식과 자원을 민주화하여 포용적 AI 혁신을 촉진
- (역량 강화) 남남(South-South) 협력 등 AI 지식과 자원을 공유하는 이니셔티브를 통해 개발도상국이 AI의 혜택을 활용하고 공통 과제를 해결할 수 있도록 역량을 강화

[그림 4-2] 첨단 기술의 시장 규모 성장 비교 (2023년 vs. 2033년) (단위: 십억 달러)



출처: UNCTAD, Technology and Innovation Report 2025: Inclusive Artificial Intelligence for Development, 2025.4.7.

3) OHCHR, 「기업과 인권 이행 원칙」 기반 AI 활용 촉구[273]

2025년 5월, 유엔인권최고대표사무소(OHCHR)는 제59차 인권이사회에서 「AI 조달 및 활용 시 '기업과 인권 이행 원칙(UN Guiding Principles on Business and Human Rights, UNGPs)' 일치 보장에 관한 보고서」를 통해 국가 및 기업에 해당 원칙을 준수할 것을 촉구하였다.

'기업과 인권 이행 원칙'은 기업 활동에 적용되는 글로벌 인권 기준으로 ▲인권을 보호해야 할 국가의 의무 ▲인권을 존중해야 할 기업의 책임 ▲비즈니스 관련 인권 침해 피해자를 위한 구제 접근권 보장 등 3대 원칙을 제시한다.

OHCHR은 AI에 관한 보편적 기준 및 정의의 부재로 국가 및 기업 간 AI 거버넌스의 불일치가 발생하고, 이에 따른 인권 침해 가능성 증가함에 따라 '기업과 인권 이행 원칙'이 AI를 직접 개발하지 않는 국가와 기업의 AI 조달 및 활용 시 어떻게 적용되는지 검토하고, AI 조달 및 활용 시 '기업과 인권 이행 원칙'을 통합하기 위한 4대 목표를 제시하였다.

- (국가의 역할 강화) 기존 및 신규 규제 프레임워크를 검토하고, AI 조달 및 활용 과정에서 인권 보호를 강화할 수 있는 방안 제시
- (기업의 책임 명확화) 기업이 AI 시스템을 조달 및 활용하는 과정에 인권 기반 접근방식을 통합할 책임이 있음을 구체화
- (구제 메커니즘 강화) AI 시스템에 영향을 받는 권리자들이 실질적 구제를 받을 수 있는 메커니즘을 검토하고 개선방안 제안
- (이행 지침 제공) 국가와 기업이 '기업과 인권 이행 원칙'을 AI 조달 및 활용 과정에 통합할 수 있도록 구체적 실행 지침 제시

273) OHCHR, Artificial intelligence procurement and deployment: ensuring alignment with the Guiding Principles on Business and Human Rights - Report of the Working Group on the issue of human rights and transnational corporations and other business enterprises, 2025.5.15.; https://www.ohchr.org/en/documents/thematic-reports/ahrc5953-artificial-intelligence-procurement-and-deployment-ensuring

4) UNESCO, AI 시대 학습자 권리 보호에 관한 보고서 발간[274]

2025년 9월, 유네스코(UNESCO)는 AI 시대 학습자 권리 보호에 관한 보고서를 통해 AI가 교육에 미치는 영향을 분석하고 디지털 시대의 교육 소외를 방지하기 위한 권고사항을 제시하였다. 기본 인권인 교육권 보장을 위해 AI를 포함한 디지털 전환이 교육에 가져오는 기회와 과제를 교육 접근성과 품질을 중심으로 검토한 결과, 교육 분야의 AI와 디지털 기술의 도입으로 새로운 교수법과 개인화 학습 기반의 교육 혁신이 이루어지고, 교육 데이터의 수집과 분석이 정교화되어 증거 기반의 의사결정이 확대되는 한편, AI 기술의 활용이 교사와 학생 간 인간적 상호작용을 축소해 학생의 신뢰와 동기 부여에 부정적 영향을 줄 수 있다는 비판이 제기되었다. 또한, AI가 교육 시스템에 점차 통합되면서 인터넷 연결성이 낮은 지역의 교육권 제한 및 교육권과 관련된 여타 인권 문제의 발생 가능성이 증가하며, 전 세계 학교의 인터넷 연결성이 불균형한 상태에서 고속 인터넷 접속과 기술 접근성이 학습의 필수 조건이 되면서, 다수의 학습자가 인터넷과 디지털 기기에 접근하지 못해 학습에서 소외될 가능성 다분하다는 우려가 제기되었다. 교육 분야에서 AI 활용이 늘어나며 민간 기업의 학생 데이터 수집과 사용과 관련된 개인정보보호 권리 침해와 같은 여타 인권 문제도 제기되는 추세로, 교육의 디지털 전환에 대한 인권 기반의 포괄적 접근방식의 필요성이 제기됨에 따라 유네스코는 정부와 이해관계자들에게 교육의 디지털 전환 시 교육권을 보장하기 위한 공통 기준을 제공하는 5C 프레임워크를 제안하였다.

- (조정과 리더십) 교육 분야의 모든 이해관계자가 참여하는 거버넌스 및 학습자의 권리를 보장하는 규제 체계를 마련
- (콘텐츠와 솔루션) 학습자의 요구를 충족하는 고품질의 디지털 학습 콘텐츠와 플랫폼을 구축하고 접근성을 확대
- (역량과 문화) 인간 중심의 디지털 전환을 위해 학습자, 교사, 학부모, 지역사회 등 모든 교육 이해관계자의 디지털 역량을 강화
- (연결성과 인프라) 인터넷과 HW/SW 등에서 공평한 연결성을 확보하되, 디지털 교육이 대면 교육을 대체하지 않도록 보장
- (비용과 지속가능성) 다양한 재원 확보와 부문 간 협력을 통해 디지털 전환에 필요한 예산을 확보하고 교육기술 생태계를 육성

274) UNESCO, AI and education: protecting the rights of learners, 2025.9.25.; https://unesdoc.unesco.org/ark:/48223/pf0000395373

제3절 기타 기구(G7, G20 등)

1. 그간의 주요 동향

2023년 5월 G7(미국, 영국, 독일, 프랑스, 캐나다, 일본, 이탈리아) 정상들은 일본 히로시마에서 5월 19일부터 사흘간 열린 G7 정상회의(G7 Hiroshima Summit 2023)에서 AI에 대한 적절한 통제의 필요성에 공감하고 AI 국제 프레임워크를 결성하는데 합의하였다.

이후 주요 7개국(G7)이 2023년 10월 30일 '히로시마 AI 프로세스'를 통해 AI 기업 대상의 AI 국제 행동강령(International Code of Conduct for Advanced AI Systems)에 합의하였다. 기업의 자발적 채택을 위해 마련된 이번 행동강령은 기반모델과 생성 AI를 포함한 첨단 AI 시스템의 위험 식별과 완화에 필요한 조치를 포함하였다.

또한, G7 경쟁당국들은 2023년 11월 8일 일본 도쿄에서 G7 경쟁당국 수장 회의를 개최하고 공동성명을 채택했다. G7 경쟁당국의 공동성명은 생성 AI 분야의 독과점 우려를 제기하는 한편, 생성 AI 등 신기술이 경쟁환경에 미칠 영향에 대처하기 위한 당국 간 협력 필요성을 강조하였다.[275]

2024년 3월 14일과 15일, 이탈리아 베로나와 트렌토에서 G7 산업, 기술 및 디지털 장관 회의가 개최되었다. 이 회의는 G7 정상회의의 관련 장관 회의 중 하나로 개최되었으며, G7 국가와 초대 국가 및 관련 국제기구가 참석하였다. 이 회의에서 G7 장관들은 1) 산업의 AI 및 신기술, 2) 안전하고 회복성 있는 네트워크, 공급망 및 주요 투입 요소, 3) 디지털 개발, 4) 공공 부문의 AI, 5) 히로시마 AI 프로세스의 성과 발전, 6) 디지털 정부에 대해 논의하였다.[276]

2024년 5월 2일, 기시다 후미오 일본 총리가 프랑스 파리에서 열린 OECD 각료이사회의 부대행사로 개최된 AI 행사 기조연설에서 '히로시마 AI 프로세스 프렌즈 그룹'의 출범을

275) https://www.jftc.go.jp/en/pressreleases/yearly-2023/November/231108.html
276) https://www.digital.go.jp/en/a11b8177-e3a3-4029-b3c0-1fbcb77b59ba-en#summary

발표하였다.[277] 히로시마 AI 프로세스를 지지하는 49개 국가들로 구성된 프렌즈 그룹은 AI 국제 지침과 행동 규범의 이행을 지원하고 국가 간 협력을 촉진할 예정이다. 기시다 총리는 기조연설에서 프렌즈 그룹이 “AI로 인한 기회와 위험 대응 및 안전하고 신뢰할 수 있는 AI의 구현이라는 공통의 목적을 위해 협력할 것”이라고 설명하였다.

2024년 7월, 인도 뉴델리에서 열린 글로벌 AI 파트너십(Global Partnership on AI, GPAI[278]) 상반기 서밋(Summit)에서 회원국들은 ’2024 GPAI 뉴델리 선언’을 발표하였다. GPAI 뉴델리 선언은 인간 중심적이고 안전하며 신뢰할 수 있는 AI의 발전을 위한 국제협력의 필요성과 GPAI와 OECD 간 통합 운영을 통한 시너지 창출을 골자로 하며, 주요 내용은 다음과 같다.[279][280]

- 신뢰할 수 있고 인간 중심적인 AI를 증진하기 위해 강력한 과학적 근거와 개방형 솔루션 및 공통 표준을 바탕으로 다중 이해관계자가 참여하는 포괄적 접근 방식으로 협력을 강화
- OECD의 AI 권고안을 기반으로 모든 OECD 회원국과 GPAI 회원국을 동등한 지위로 통합하는 OECD-GPAI 통합 파트너십을 통해 GPAI의 새로운 비전을 수립
- GPAI나 OECD의 현재 회원국 지위와 관계없이 모든 국가가 인간 중심적이고 안전하며 신뢰할 수 있는 인공지능의 잠재력을 모두의 이익을 위해 활용하는 공통의 노력에 동참할 것을 촉구
- 관련 국제기구 및 UN 전문기구, 기타 파트너들과 포용적 대화를 통해 필요시 협력하며, 통합 파트너십의 모든 회원이 파트너십의 활동과 의사결정에 차별 없이 동등하게 참여할 것임을 강조
- GPAI 다중 이해관계자 전문가 그룹(MEG)과 OECD AI 전문가 네트워크(ONE AI) 및 기존 전문가 그룹 작업반을 통합하여 단일 전문가 그룹으로 합병
- 누구도 소외되지 않고 평화, 번영, 지속 가능한 발전을 촉진하는 미래를 위해 AI가 모두의 이익을 위해 사용될 수 있도록 통합 파트너십을 통한 건설적 대화를 지속하겠다는 목표를 확인

277) Ministry of Foreign Affairs of Japan, Prime Minister Kishida’s attendance at the Side Event on Generative AI at the OECD Ministerial Council Meeting, 2024.05.02.
278) GPAI는 캐나다와 프랑스 주도로 G7에서 제안되어 2020년 출범한 AI 분야의 글로벌 다자 협의체로, 민주주의 가치와 인권을 존중하는 29개국이 회원국으로 참여 중이며, 한국은 15개 창립회원국 중 하나
279) Ministry of Electronics & IT, 6th meeting of the GPAI Ministerial Council held on 3rd July 2024 at New Delhi, 2024.07.03
280) 외교부, GPAI(글로벌 인공지능 파트너십) 상반기 서밋에서 OECD-GPAI 통합 운영에 합의, 2024.07.04.

2024년 9월 13일, 디지털 경제를 담당하는 G20 장관들은 브라질 마세이오에서 회의를 개최하고 이전 의장국의 성과와 공약을 바탕으로 디지털 포용성, 보편적이고 의미 있는 연결성, 디지털 정부와 포괄적 디지털 공공 인프라, 온라인 정보 무결성과 디지털 경제에 대한 신뢰, 포괄적이고 지속 가능한 개발과 불평등 감소를 위한 AI에 대한 논의를 진행하였다. 이 회의에서 AI 관련 G20 장관들은 다음 사항에 대해 인식을 공유하였다.[281)]

- AI 솔루션의 개발 및 배포는 특히 콘텐츠를 개인화, 조정 및 생성하는 것을 목표로 할 때 윤리적이고 투명하며 감사 가능하고 인간의 감독 하에 책임을 져야 하며 개인정보 보호, 개인 데이터 보호 및 인권, 기본적 자유 및 지적 재산에 대한 존중을 보장해야 함
- 콘텐츠 인증 및 출처 메커니즘과 관련 기술 표준을 포함하여 편견을 완화하는 것의 중요성은 AI에서 생성된 콘텐츠를 식별하고 사용자가 정보 조작을 식별하는 데 도움이 될 수 있음
- 지식재산권, 개인정보 보호 및 데이터 보호를 존중하는 데이터, 알고리즘 및 콘텐츠 조정에 대한 적절한 보호 및 설명 가능성과 함께 투명성이 중요함
- 국제협력을 포함하여 위험 기반 및 인간 중심적, 발전 지향적이며, 혁신 친화적인 AI 정책 및 거버넌스 접근방식에 전념하여 AI에서 파생된 이점과 위험 완화를 촉진

2. 2025년 주요 동향

1) G7 정상들, 번영을 위한 AI 선언 발표[282)]

2025년 6월 17일, G7 정상들은 「번영을 위한 AI 선언(Statement on AI for Prosperity)」을 발표하고, 인간 중심의 AI 접근방식이 번영을 증진하고 사회에 혜택을 주며 시급한 글로벌 과제를 해결할 잠재력을 지닌다는 점을 인식하고, 이러한 잠재력을 실현하기 위해 국민에게 혜택을 제공하고 부정적 외부효과를 완화하며 국가안보를 증진하는 안전하고 책임감 있으며 신뢰할 수 있는 AI의 혁신과 도입을 더욱 효과적으로 추진해야 한다고 강조하였다. 또한, 정상들은 유엔 글로벌 디지털 협약(United Nations Global Digital Compact)[283)]에 부합하도록 신흥 시장 및 개발도상국 파트너들과 협력하여 디지털 격차를 해소할 것을 약속하였다.

281) https://www.dataguidance.com/news/international-g20-adopts-declaration-ai-disinformation
282) https://g7.canada.ca/en/news-and-media/news/g7-leaders-statement-on-ai-for-prosperity/
283) 안전하고, 개방적이며, 자유롭고, 평화롭고, 인권을 존중하는 디지털 미래를 보장하는 것을 목표로 전 세계 디지털 미래에 대한 공유 원칙을 구체화

- 공공 부문에서 AI 도입을 가속화하여 시민과 기업 모두를 위한 공공 서비스의 질을 향상시키고, 인권과 사생활을 존중하며 투명성, 공정성, 책임성을 증진하는 동시에 정부 효율성을 높이기 위해 협력
- 중소기업(SMEs)이 개인정보와 지적재산권을 존중하는 AI를 도입하고 개발하도록 지원하여 경제적 번영을 촉진하고, 중소기업의 준비도, 효율성, 생산성 및 경쟁력을 강화
- AI가 초래하는 에너지 과제를 해결하고 에너지 효율성 향상 및 혁신을 위한 AI의 잠재력을 활용
- 신흥 시장 및 개발도상국 파트너들과의 상호 이익적 협력을 확대하여 모든 이의 AI 접근성을 증진

2) G7 에너지·환경 장관들, 에너지 및 AI 작업 계획 승인[284][285]

2025년 10월 30일, G7 에너지·환경 장관들은 「번영을 위한 AI 선언」의 후속 조치로서 에너지 및 AI 고위급 대화에서 각국 장관들은 국제에너지기구(IEA), 국제재생에너지기구(IRENA), B7[286] 의장국인 캐나다 상공회의소 등 국제 및 산업 파트너들과 협력하여 마련한 'G7 에너지 및 AI 작업 계획'을 승인하였다.

장관들은 증가하는 글로벌 도전과제와 기술 경쟁 속에서 AI의 에너지 문제 해결과 AI의 혁신적 잠재력 활용을 위한 G7 파트너, 신뢰할 수 있는 파트너, 산업계 및 국제 파트너 간의 협력이 G7 경제의 디지털 전환 성공은 물론, 더 광범위한 에너지, 경제 및 국가안보 필수 과제의 핵심임을 인식하였다. AI 및 데이터센터의 에너지 수요와 관련된 자원 관리 및 수요 측면의 대응을 최적화하는 등 기술적·시장 기반 해결책을 채택함으로써, 국내 우선순위 및 적용 가능한 국제적 약속에 부합하는 방식으로 AI 데이터센터가 안정적이고 비용 효율적으로 전력을 공급받을 수 있도록 하기 위한 공동의 목표를 공유하였다.

G7은 AI에 전력을 공급하기 위한 에너지 시스템 및 솔루션에 대한 투자가 국민을 위한 에너지 신뢰성, 접근성 및 경제성을 유지하는 방식으로 이루어져야 함을 확인하고, 산업계, 국

284) https://g7.canada.ca/en/news-and-media/news/g7-energy-and-ai-work-plan/
285) https://g7.canada.ca/en/news-and-media/news/backgrounder-g7-energy-and-environment-ministers-meeting-energy-and-artificial-intelligence/
286) Business7 그룹은 국제 비즈니스 커뮤니티와의 대화를 위한 G7의 플랫폼으로, 실행 권고안을 마련함으로써 G7이 글로벌 경제 과제를 해결하는 데 기여할 것을 도모

내 이해관계자 및 국제 파트너와 협력하여 다음을 추진할 것을 약속하였다.

- 다양한 기술과 공급원을 활용하고, 에너지 생태계를 기반으로 데이터센터를 전략적으로 배치하는 등 비용 효율적인 접근법을 발전시켜 전력망 사용 및 확장을 최적화하고 안정화함으로써, 국내 우선순위 및 국제적 약속에 부합하는 안전하고 신뢰할 수 있으며 지속가능하고 경제적인 에너지 공급을 AI 데이터센터에 보장
- IEA 에너지 및 AI 관측소와 협력하여 AI 및 데이터센터가 직면한 에너지 문제를 해결하기 위한 검증되고 비용 효율적인 솔루션을 도출
- 에너지 혁신 및 획기적인 에너지·소재 발견을 위한 AI 활용 촉진 및 지원
- 소비자, 관련 규제 기관, 정책 및 의사 결정권자의 AI 신뢰도 제고를 통해 에너지 산업 경쟁력을 강화하고 AI 도입을 가속화
- AI 모델 및 데이터센터 운영의 에너지 및 자원 요구사항에 대한 이해와 투명성 제고를 위한 모범사례를 장려
- 에너지 시스템 및 운영 전반에 걸쳐 효과적이고 안전한 AI 솔루션 적용을 위한 지원 환경 구축
- 기존 플랫폼(산업계, 학계 및 다중 이해관계자 주도 이니셔티브 포함)을 활용하여 AI 모델, 데이터센터 및 이를 구동하는 전력망의 에너지 사용 최적화를 목표로 하는 수요 측면 솔루션에 관한 기업 간 및 정부-산업계 정보·데이터 공유를 촉진
- AI 데이터센터, 그 구성 요소 및 지원 에너지 시스템과 인프라의 보안을 강화하고, 사이버 기반 위협의 위험을 줄이고 대응하기 위해 핵심 에너지 인프라 및 에너지 시스템에 AI를 안전하고 책임감 있게 도입할 것을 권장

3) G20 AI 태스크포스 회의 개최[287)]

2025년 9월 30일, G20는 의장국인 남아프리카공화국의 주도 하에 ‘지속가능한 발전을 위한 AI, 데이터 거버넌스 및 혁신을 위한 AI 태스크포스’ 회의를 개최하였다. G20는 AI의 혜택을 극대화하는 동시에 위험을 완화하는 것의 중요성을 확인하고 인류의 이익을 위해 공정하고 포용적이며 안전하고 신뢰할 수 있으며 책임감 있고 윤리적이며 지속 가능한 글로벌 AI 환경을 조성하는 데 모든 국가의 의미 있는 참여가 필요함을 강조하였다.

287) https://www.g20.utoronto.ca/2025/250930-ai-taskforce.html

- (데이터 거버넌스) 신뢰할 수 있고 공정하며 지속 가능한 AI 개발의 핵심으로, 데이터의 자유로운 흐름과 상호운용성을 위한 국제협력의 필요성 인식
- (공공부문의 AI 활용) 안전하고 책임 있는 AI를 활용한 행정 효율성 및 서비스 개선을 도모하며, 인권・투명성・공정성・책임성을 유지하고, 디지털 역량을 강화하며 부문별 전문성을 개발하기 위한 노력 필요
- (AI 과학 연구・개발) 국제협력 강화 및 오픈 사이언스・오픈 데이터・오픈소스의 활용을 권장하며 현지 맞춤형 AI 생태계 개발을 지원하고 연구 보안 및 무결성을 유지하기 위한 노력 필요
- (정보 무결성) 생성 AI로 인한 딥페이크, 허위 정보에 따른 위험을 인지하고 기술적 탐지, 워터마킹・라벨링 적용, 디지털 리터러시 강화 등의 노력 필요
- (다언어・다문화 데이터) 토착어 및 현지 언어 기반 데이터 확보가 필수적이며, 전통지식, 문화유산, 지적재산권을 존중하는 방식으로 혁신, 투명성, 공정 보상을 보장하기 위한 메커니즘 필요
- (아프리카 AI 이니셔티브) 아프리카연합(AU)의 2063년 개발 목표를 지원하고, 컴퓨팅 파워, AI 인재・교육, 데이터, 인프라 접근성을 강화하며, 아프리카 AI 생태계 발전과 대륙에 지속 가능한 가치를 창출하는 투자 모델에 중점을 둔 장기적 파트너십 기반의 아프리카 중심 주권적 AI 역량 개발을 장려

제5장 인공지능 관련 주요 기업 동향

2025년 전 세계 주요 국내외 빅테크, IT 대기업, AI 전문 스타트업과 산업계의 동향 분석을 통해 다음과 같은 주요 트렌드를 도출해 보았다.

● 빅테크의 서비스·플랫폼을 AI 중심으로 재구성

구글은 검색 결과 상단에 AI Overview를 기본 제공하고, Gemini 2.0/2.5를 통해 장문 맥락 이해, 코드 생성, 멀티모달 질의응답을 가능하게 하며, ‘Execution AI’ 전략 아래 브라우저·OS·업무툴 전반을 에이전트 중심 UX로 개편하고 있다. 애플은 2024년 발표한 Apple Intelligence를 2025년 본격 탑재하며, 개인 단말에서 프라이버시를 지키는 소형 온디바이스 모델과 클라우드 상의 대규모 MoE 모델을 결합해, 메시지·메일·노트 등 OS 전반에 개인화된 AI 기능을 심고 있다.

● 온디바이스·개인 AI와 클라우드 초거대 모델의 하이브리드 아키텍처 확산

애플의 30억 파라미터급 온디바이스 LLM과 서버 측 MoE 모델 조합처럼, 연산비용·프라이버시·지연시간을 고려해 클라이언트와 클라우드를 나누는 하이브리드 구조가 대세로 자리 잡고 있다. 구글·국내 통신사 등도 일부 기능은 단말·엣지에, 고난도 연산은 클라우드에 분산하는 구조를 채택하면서, ‘개인 에이전트’와 ‘전사(全社)용 초거대 모델’이 공존하는 이중 구조를 만들어가고 있다.

● AI 인프라·칩·데이터센터를 둘러싼 장기 투자 경쟁

구글은 TPU v6 기반 AI 인프라와 더욱 효율적인 데이터센터 구축에 수십억 달러를 투자하며, 클라우드·검색·유튜브까지 아우르는 자체 컴퓨팅 스택을 강화하고 있다. 애플은 2025~2028년 최대 6,000억 달러 규모 제조·공급망 투자 계획을 발표하면서, AI·칩·제조 혁신을 핵심 축으로 제시해 AI 하드웨어 체인까지 수직 통합하려는 움직임을 보이고 있다.

● 투자 규모는 유지되지만 ‘프런티어 모델’ 소수 기업으로 자본 초집중

2024년 전 세계 AI 투자금은 직전년 대비 일부 조정되었지만 여전히 사상 최고 수준을 유지했으며, 2025년 들어서는 소수의 초거대 모델 기업(예: OpenAI, Anthropic, Mistral 등)에 수십억 달러 규모 메가 라운드가 집중되는 구조가 뚜렷하였다. 거래 건수는 줄어드는 대신, 평균 투자 규모와 상위 몇 개 딜의 비중이 크게 늘어나면서, ‘규모의 경제’를 확보한 프런티어 모델 기업과 나머지 생태계 간 격차가 확대되는 양상을 보였다.

● 도메인 특화 ‘AI 에이전트·데이터 에이전트’ 제품화와 엔터프라이즈 워크플로 자동화

구글은 AI 에이전트 간 통신을 위한 개방형 프로토콜 A2A(Agent2Agent)를 공개하고, 다중 에이전트가 다양한 클라우드·플랫폼에서 안전하게 통신·협업할 수 있도록 설계했다. 이를 통해 기업 환경에서 인증·권한 관리, 장시간 작업, 멀티모달 스트리밍까지 지원하는 에이전트 기반 서비스 인프라를 제공하려는 흐름이 나타난다. 마이크로소프트는 Azure AI Foundry에서 여러 전문 에이전트를 조율하는 멀티 에이전트 오케스트레이션, 관측(Observability) 기능, 코파일럿 튜닝 등을 확대하면서, 엔터프라이즈 업무 전체를 에이전트 네트워크로 재구성하려는 경향을 드러냈다.

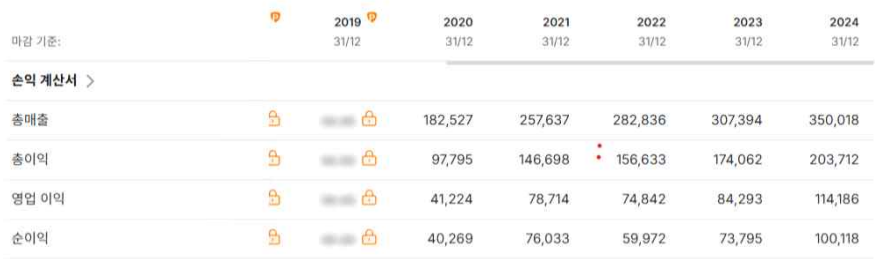
한편, 피지컬 AI는 로봇·자율주행·스마트팩토리 등 실물경제로 확장되며 기업 전략의 새로운 축으로 부상했다. 휴머노이드·산업용 로봇 스타트업과 기존 제조·물류 기업이 파일럿을 확대하고, GPU/엣지 컴퓨팅·센서·모터·배터리·시뮬레이션(디지털 트윈)까지 묶은 로봇 파운데이션 모델/로보틱스 플랫폼 경쟁이 본격화되면서, 피지컬 에이전트(현장 작업 자동화)가 성장하는 양상이 나타나고 있다.

제1절 국외 기업

1. 알파벳(구글)

1) 기업 일반현황

<표 5-1> 알파벳(구글) 기업 현황

구분	주요 내용
기업 개요	**〈회사 개요〉** • 1998년 9월 미국 캘리포니아주에서 설립 • 2004년 8월 나스닥 상장 • 세계 모든 정보를 체계화해 모든 사용자가 유익하게 사용할 수 있도록 하는 것을 기업 목표로 함 • 2025년 기준 전 세계 50개국, 약 70여 도시에서 지사를 운영하며 검색·광고·클라우드·AI 플랫폼 등 다양한 디지털 서비스 생태계를 운영
	〈사업 내용〉 • 산업 분야 : 인터넷, 검색, 소프트웨어, 서비스, 광고 • 서비스 : 구글 검색, 유튜브, 구글 클라우드, 안드로이드 OS, 크롬 브라우저, 지메일, 구글 지도, 구글 드라이브, 구글 포토 등 핵심 서비스를 제공하며 2024년부터 제미나이(Gemini) 기반 서비스(Workspace Gemini, YouTube Create AI, Android Gemini, Chrome AI 기능, Gemini Live) 등 생성 AI 기능을 전면 확장
	〈자회사〉 • 구글 클라우드, 유튜브, 구글 딥마인드, 웨이모, 딥마인드, 캘리코, 알파벳 X
재무 실적[288] (최근 5년)	
AI 조직 현황	**〈AI 관련 조직〉** • Google AI/Google Research : 구글의 대표 AI 연구조직으로서, 초거대 멀티모달 모델(제미나이), 효율적 학습 기술, 책임형 AI(Responsible AI), AI 윤리 등 기초·응용 전 분야를 아우르는 연구 수행. 2024년 이후 제미나이 1.5, 제미나이 2.0 등 장기 컨텍스트 처리에 특화된 모델을 통해 초장문 추론·멀티모달 이해 능력 고도화 추진 • 구글 딥마인드(구 딥마인드+ 브레인 조직 통합) : 2023년 합병 이후 구글의 AGI 연구를 총괄하며 조직 규모 및 연구 범위가 확대됨. AlphaFold·AlphaGo 계열 연구를 기반으로 생명과학,

마감 기준:	2019 31/12	2020 31/12	2021 31/12	2022 31/12	2023 31/12	2024 31/12
손익 계산서 >						
총매출	[illegible]	182,527	257,637	282,836	307,394	350,018
총이익	[illegible]	97,795	146,698	156,633	174,062	203,712
영업 이익	[illegible]	41,224	78,714	74,842	84,293	114,186
순이익	[illegible]	40,269	76,033	59,972	73,795	100,118

구분	주요 내용
	재료과학, 로보틱스, AI 에이전트, 자기 지도학습 분야 연구를 수행. 최근 'AI Agents' 프레임워크를 발표해 자율형 에이전트 연구에 집중하며 초대형 멀티모달 모델·장기 메모리·지속적 학습 분야 연구가 이어지고 있음 • 알파벳 X(Alphabet X, 구 Google X) : 자율주행, 드론 배송, 지능형 통신 및 에너지 프로젝트 연구를 지속하며, AI·로보틱스 기반 '문샷 프로젝트'를 수행. 최근에는 기상 AI, 지능형 에너지 최적화 등 AI 기반 기후 기술 연구 비중을 확대
사업 전망	**〈AI 전략〉** • **(AI-first 전략 강화)** 검색, 지도, 유튜브, 안드로이드, 크롬 등 핵심 서비스에 제미나이 기반 생성 AI 기능을 대규모로 통합하는 전략을 추진. 전통 검색을 'AI 오버뷰' 중심으로 재편해 검색 광고 모델의 구조 변화를 추진 • **(대규모 AI 모델(Gemini) 개발 로드맵 고도화)** 제미나이 1.5 Pro/ Flash/ Ultra→ 2.0 시리즈로 확장하며 GPT-4.1, GPT-5 대응 체계를 강화. 초장문 컨텍스트(백만 토큰 이상), 멀티모달 추론, 오프라인 에이전트 기능 등 기술력 강화에 주력하며, PaLM·LaMDA 계열 모델은 대부분 제미나이 체제로 통합되는 추세. • **(AI 인프라 및 칩셋 전략 강화)** TPU v5e·v5p에 이어 TPU v6(2025) 개발에 관심이 집중되며, AI 트레이닝 비용 절감·연산 효율 개선에 초점. 구글 클라우드를 AI 인프라의 핵심으로 삼아 기업용 시장에서 AWS·MS와의 경쟁이 치열해짐. • **(AI 에이전트 전략 본격화)** 최근 제미나이 에이전트와 딥마인드 에이전트를 중심으로, 작업 시퀀스 자동화, 웹·앱·운영체제 환경에서 복합 업무 수행, 장기 목표 기반 행동계획 등을 수행하는 실행형 AI(Execution AI)를 강화 • **(책임 있는 AI 전략, Responsible AI)** 헬스, 교육, 금융 등 고위험 서비스에 대한 투명성·안전 기준을 강화하고, AI 안전 관련 내부 감사·검증 체계를 확대. EU AI Act 등 글로벌 규제 대응을 위해 모델 위험평가·데이터 거버넌스·투명성 보고 기능을 강화

자료: SPRI 정리

2) 2025년 기업 동향

(1) 사업전략/투자/M&A

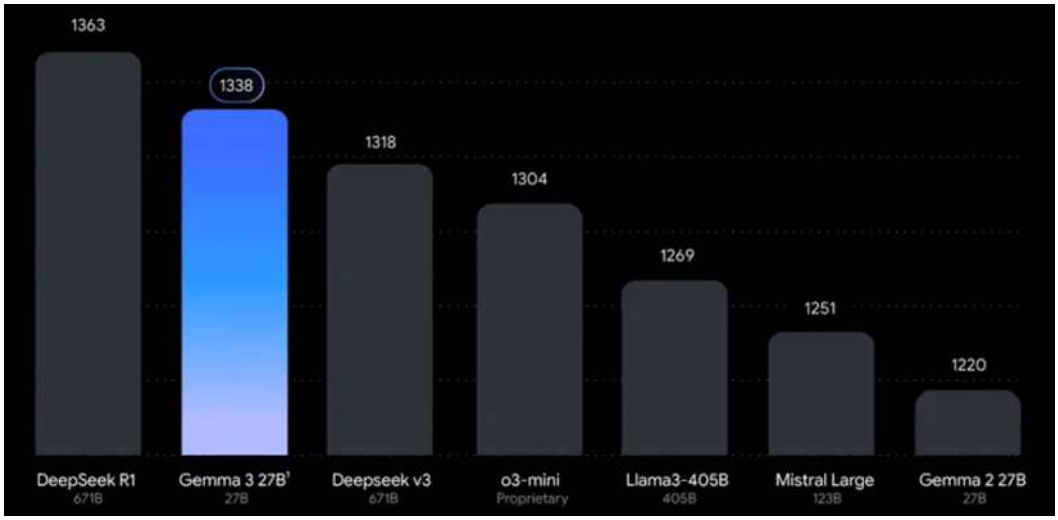
가. 단일 GPU로 구동 가능 오픈소스 AI 모델 '젬마 3' 공개[289]

구글이 2024년 3월 12일 '제미나이(Gemini) 2.0' 모델 아키텍처 기반의 경량 오픈소스 AI 모델 '젬마 3(Gemma 3)'를 공개했다. 젬마 3는 140개 이상 언어와 고급 텍스트 이해 및 시각 정보 추론을 지

288) 알파벳 재무제표 https://kr.investing.com/equities/google-inc (2025.12.23. 방문)
289) Google, Introducing Gemma 3: The most capable model you can run on a single GPU or TPU, 2025.3.12.

원하고, 최대 12만 8천 개 토큰의 컨텍스트 창 및 작업 자동화와 에이전트 구축을 위한 함수 호출 기능과 구조화된 출력을 제공한다. 또한 양자화[290] 버전을 통해 모델 크기와 컴퓨팅 요구사항을 줄이면서 높은 정확도를 유지할 수 있다. 스마트폰과 노트북 같은 단말에서 단일 GPU나 TPU[291]로 실행할 수 있는 젬마 3는 매개변수 1B, 4B, 12B, 27B 4개 버전으로 제공되어 개발자들이 하드웨어나 성능 요구사항에 맞게 모델을 선택할 수 있다. 구글은 젬마 3가 챗봇 아레나(Chatbot Arena)[292] 인간 선호도 평가에서 1,338점을 기록해 매개변수가 더 큰 모델인 라마3-405B(1269점), 딥시크-V3(1318점), o3-mini(1304점)를 넘어섰다. 구글은 크기 대비 최고 성능을 발휘했다고 강조했다.

[그림 5-1] 젬마 3와 주요 AI 모델의 챗봇 아레나 점수 비교



출처: Google, Introducing Gemma 3: The most capable model you can run on a single GPU or TPU, 2025.3.12.

나. AI 에이전트 간 통신 프로토콜 'A2A' 공개 및 MCP 지원 발표[293]

구글이 2025년 4월 9일 50개 이상의 기업[294]과 협력해 AI 에이전트 간 통신을 위한 개방형 프로토콜 'A2A(Agent2Agent)'를 공개했다. 구글은 다양한 플랫폼과 클라우드 환경에서 다중 AI 에이전트가 서로 통신하고 안전하게 정보를 교환하며 작업을 조정할 수 있도록 A2A 프로토콜을 출시했다고 강조했다. 구글은 A2A는 AI 에이전트 간 협업을 위한 표준 방식을 제공하기 위해 HTTP, SSE, JSON-RPC 등 기존 표준을 기반으로 구축되었으며, 기업 환경에서 요구하는 높은 수준의 인증 및 권한 관리 기능을

290) 모델 매개변수를 저장하는 데 필요한 비트 수를 줄여 모델 크기를 압축하고 연산 속도를 높이는 기술
291) 구글에서 개발한 머신러닝 전용 프로세서로, 일반 GPU보다 머신러닝 작업에서 뛰어난 성능과 전력 효율성을 보유
292) 사용자 선호도를 통해 AI 모델을 평가하는 오픈소스 플랫폼
293) Google, Announcing the Agent2Agent Protocol (A2A), 2025.4.9.
294) 액센추어(Accenture), 코히어(Cohere), 랭체인(Langchain), 페이팔(Paypal), 세일즈포스(Salesforce) 등

제공한다. 빠른 작업뿐 아니라 장시간 작업 환경에도 적합하며, 텍스트와 오디오, 동영상 스트리밍도 지원한다고 설명했다.

A2A는 작업을 구성하고 전달하는 역할을 하는 클라이언트 에이전트(Client Agent)와 작업을 수행하는 원격 에이전트(Remote Agent) 간 원활한 통신을 위해 ▲기능 탐색 ▲작업관리 ▲협업 ▲사용자 경험 협의 등의 기능을 제공한다. 기능 탐색은 각 에이전트가 자신의 기능을 JSON[295] 형식의 '에이전트 카드'[296]를 통해 이뤄진다. 클라이언트 에이전트는 작업 수행에 가장 적합한 에이전트를 식별해 A2A로 원격 에이전트와 통신한다. 작업관리는 클라이언트 에이전트와 원격 에이전트는 최종 사용자의 요청에 대응해 작업 수명주기 전반에서 작업 처리 상태를 지속 동기화하여 처리한다. 협업은 각 에이전트가 서로 컨텍스트, 응답, 작업 결과물, 사용자 지시와 같은 메시지를 교환하는 방식으로 지원한다. 또한 각 메시지에는 이미지, 동영상, 웹 양식과 같은 특정 콘텐츠 유형이 명시되어 있어, 각 에이전트는 사용자 인터페이스(UI)에 맞게 적절한 콘텐츠 형식을 협의할 수 있다.

한편, 구글 딥마인드의 데미스 하사비스(Demis Hassabis) CEO는 2025년 4월 9일 X를 통해 구글이 앤스로픽의 MCP를 제미나이 모델과 SDK에서 지원하겠다고 발표했다.[297] 구글에 따르면 A2A는 MCP를 보완하는 역할로서, MCP가 LLM을 데이터, 리소스 및 도구와 연결하는 프로토콜이라면 A2A는 에이전트 간 협업을 위한 상위 수준의 프로토콜에 해당한다.

다. 구글 연구진, 효율적 메모리 사용을 위한 '타이탄' 아키텍처 개발[298]

구글 연구진이 기존 LLM의 한계를 극복할 수 있는 새로운 신경망 아키텍처 '타이탄(Titans)'을 개발하고 2024년 12월 31일 관련 논문[299]을 공개했다. LLM의 기존 트랜스포머 아키텍처가 사용하는 어텐션[300] 메커니즘은 문맥이 길어질수록 연산과 메모리 비용이 급증하는 문제점을 안고 있었다. 연구진은 어텐션 기반 트랜스포머의 대안으로 등장한 선형(Linear) 트랜스포머도 계산 복잡도를 줄이기 위해 정보를 압축하는 과정에서 중요한 세부 사항을 놓치는 경향을 나타내는 한계를 지적했다. 연구진은 기존 아키텍처의 한계를 극복하기 위해 메모리를 단기와 장기로 구분해 추론 시 새로운 정보를 학습하는 방

295) 키-값 쌍으로 이루어진 데이터 객체를 표현하기 위한 텍스트 기반의 개방형 표준 형식
296) 에이전트의 기능과 스킬, 인증 요구사항 등을 설명하는 공개 메타데이터 파일
297) 지디넷코리아, 오픈AI 이어 구글도 앤트로픽 기술 'MCP' 채택…AI 에이전트 강화 행보, 2025.4.10.
298) Venturebeat, Google's new neural-net LLM architecture separates memory components to control exploding costs of capacity and compute, 2025.1.16.
299) Google Research, Titans: Learning to Memorize at Test Time, 2024.12.31. https://arxiv.org/abs/2501.00663
300) Attention : 문장 내에서 주요한 단어나 구문에 집중하여 의미와 문맥을 파악하는 메커니즘

식의 아키텍처를 개발했다고 밝혔다.

타이탄 아키텍처의 신경 기억 모듈은 추론 중 새로운 정보를 학습하고 데이터에 따라 기억 프로세스를 동적으로 조정하는 기능을 학습하도록 설계됐다. 신경 기억 모듈은 어떤 데이터를 저장할 가치가 있는지 결정하기 위해 '놀라움(Surprise)' 지표를 사용해 입력 데이터가 과거에 저장된 데이터 종류와 다를수록 더 놀랍고 기억할 가치가 있다고 평가한다. 이를 통해 모듈은 AI 모델이 이미 아는 정보에 유용한 정보를 추가하는 데이터만 저장하며, 더 이상 불필요한 정보를 제거하는 적응형 망각 메커니즘도 갖춰 제한된 메모리를 효율적으로 사용할 수 있다.

이러한 과정은 신경 기억 모듈을 구성하는 ▲코어 ▲장기기억 ▲지속 기억을 통해 진행되며, 코어 모듈이 현재의 문맥 정보를 처리하고 장기기억은 과거 정보를 저장, 지속 기억은 사전 학습된 기본 정보를 저장하는 방식이다. 연구진은 타이탄 모델이 200만 토큰 이상의 문맥에서 높은 정확도를 유지해, 기존 모델이 긴 문맥에서 성능이 급격히 저하되는 문제를 완화했다고 밝혔다. 또한 실험을 통해, 긴 문서에서 정보 추출이나 추론이 필요한 작업에서 타이탄 모델이 매개변수가 훨씬 더 많은 GPT-4와 GPT-4o-미니보다 우수한 성능을 기록했다고 주장했다.

라. I/O 2025에서 AI 에이전트와 AI 모드 등 AI 신기능 공개[301]

구글이 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 각종 서비스를 공개했다.[302] 우선 2024년 12월 공개한 에이전트 연구 프로젝트 '프로젝트 마리너(Project Mariner)'[303]를 발전시켜 에이전트 기능을 크롬과 구글 검색, 제미나이 앱에 도입했다. 일례로 제미나이 앱에 실험 버전으로 제공되는 새로운 에이전트 모드는 부동산 거래 웹사이트에서 조건에 맞는 매물을 찾아 방문일정을 예약하는 등 사용자 대신 복잡한 작업을 수행한다.

구글 검색엔진에는 AI 기반 검색인 'AI 모드(AI Mode)'를 확장해 더욱 향상된 추론 기능을 제공한다. 제미나이 2.5 맞춤형 버전 기반의 AI 모드는 이용자의 질문을 여러 하위 주제로 나누어 복수의 쿼리를 동시에 실행해 더 깊이 있는 웹 탐색을 지원하며, 더욱 깊이 있는 답변을 위한 '딥 서치(Deep

301) Google, Google I/O 2025: From research to reality, 2025.5.20.
302) Google, Introducing Google AI Ultra: The best of Google AI in one subscription, 2025.5.20.
303) 크롬 브라우저에 특화된 AI 에이전트로, 기존 챗봇과 달리 직접 웹사이트를 검색하고 정보를 수집하며 특정 작업을 수행

Search)’ 기능도 새로 추가되어 몇 분 만에 참고문헌이 포함된 전문가 수준의 보고서를 생성할 수 있다. AI 모드에 제미나이와 구글의 쇼핑 그래프(Shopping Graph)[304]를 결합해 사용자가 다양한 제품을 탐색하여 최적의 선택을 할 수 있도록 지원하는 쇼핑 기능도 추가됐다.

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 추가했다. 구글이 제미나이 앱을 통해 모든 사용자에게 공개한 ‘제미나이 2.5 플래시’는 버전 업데이트로 추론, 멀티모달 기능, 코드, 긴 컨텍스트 처리 등이 향상됐고, 토큰 사용량은 20~30% 줄어 효율성이 개선됐다. 구글은 비오 3에서 처음으로 소리까지 포함된 동영상 생성을 지원하는 한편, 비오 기반의 AI 영화 제작 툴 ‘플로우(Flow)’를 공개했고, 이마젠 4는 섬세한 질감, 물방울, 동물의 털과 같은 미세한 묘사를 표현할 수 있고 다양한 화면 비율과 최대 2K 해상도를 지원한다.

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 출시했다. 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 순차 출시할 예정이다.

마. 구글 딥마인드, AI 코딩 기업 윈드서프의 핵심 인력 영입[305]

구글 딥마인드가 24억 달러 투자해 AI 코딩 기업 윈드서프(Windsurf) 핵심 인력과 비독점 라이선스 확보했다.[306] 오픈AI가 인수를 추진했지만, 윈드서프 CEO와 창립자 등 핵심 인력은 2025년 7월 11일 구글 딥마인드 합류를 결정했다. 윈드서프는 2021년 설립되어 자연어 프롬프트 기반의 AI 코딩 어시스턴트를 제공해 왔고, 이번에 바룬 모한(Varun Mohan) CEO, 더글러스 첸(Douglas Chen) 공동창업자 등 핵심 인력이 구글 딥마인드에 합류한다. 이번 거래에 약 24억 달러를 투자한 구글은 윈드서프의 지분을 인수하거나 기업에 대한 통제권을 갖지는 않으며, 윈드서프가 보유한 일부 기술에 대한 비독점 라이선스만 확보했다. 윈드서프 팀은 구글에서 제미나이를 중심으로 에이전트 코딩 프로젝트에 주력할 계획으로, 구글은 세계 최고 수준의 AI 코딩 인재들과 함께 에이전트 기반 코딩 연구를 앞당길 방침이다. 이번 거래는 빅테크가 반독점 규제 회피나 인력·기술의 빠른 확보를 위해 활용하는 ‘역 인재 인수합병

304) 온라인 쇼핑 경험 향상을 위해 구글이 구축한 AI 기반의 제품 정보 데이터셋
305) TechCrunch, Windsurf’s CEO goes to Google; OpenAI’s acquisition falls apart, 2025.7.11.
306) Cognition, Cognition’s acquisition of Windsurf, 2025.7.14.

(Reverse-acquihire)’ 사례로 평가된다.

구글은 2024년 8월에 유사한 방식으로 AI 챗봇 스타트업 캐릭터AI(Character.AI) 창립자이자 구글 연구원 출신인 노암 샤지어(Noam Shazeer)를 영입한 바 있다. 구글의 인수 성공으로 오픈AI의 윈드서프 인수는 최종 무산됐고, 실패 원인으로 오픈AI와 마이크로소프트(MS)의 갈등이 주요인으로 작용했다는 분석이 제기됐다. 양사의 맺은 계약에 따르면 MS는 오픈AI의 모든 지식재산권에 접근할 수 있지만, 오픈AI가 자체 AI 코딩 제품(GitHub Copilot)을 보유한 MS의 윈드서프 지식재산권 접근을 반대하면서 인수 협상에 차질을 초래했다는 지적이다.

(2) 기술개발 및 혁신

가. 구글 딥마인드, 로봇 특화 AI 모델 ‘제미나이 로보틱스’ 개발[307]

구글 딥마인드가 2025년 3월 12일 로봇 개발에 특화된 AI 모델 ‘제미나이 로보틱스(Gemini Robotics)’를 공개했다. 제미나이 2.0을 기반으로 구축된 제미나이 로보틱스는 첨단 시각-언어-행동(VLA)[308] 모델로, 로봇을 직접 제어하기 위해 물리적 행동을 새로운 출력 형태로 추가했으며, 학습 과정에서 접하지 못한 새로운 물체나 환경에서도 기존에 학습한 지식을 응용해 다양한 작업을 즉각 수행하는 범용성을 보유했다.

구글 딥마인드는 종합적인 범용성 벤치마크에서 제미나이 로보틱스가 다른 첨단 시각-언어-행동 모델보다 평균 두 배 이상의 성능 향상을 보였다고 강조했다.[309] 구글 딥마인드에 따르면 제미나이 로보틱스는 제미나이의 언어 이해 능력을 바탕으로 일상적 대화체의 명령을 이해하고 다양한 언어로 답변할 수 있으며, 복잡한 다단계 작업도 수행할 수 있다. 또한 주변 환경을 지속 모니터링해 환경이나 명령의 변화를 감지하고 즉시 행동을 수정할 수 있어, 다양한 환경에서 로봇과 협력할 수 있도록 지원한다. 가령 종이접기와 같은 섬세한 운동 능력과 정밀한 조작을 요구하는 작업 수행이 가능하고, 로봇 유형별로 쉽게 적응할 수 있도록 설계됐다. 많은 연구실에서 사용되는 양팔 로봇 플랫폼뿐 아니라 휴머노이드 로봇과 같은 복잡한 형태에도 특화가 가능하다고 설명한다.

구글 딥마인드는 제미나이 로보틱스와 함께 첨단 시각-언어모델 ‘제미나이 로보틱스-ER (Embodied

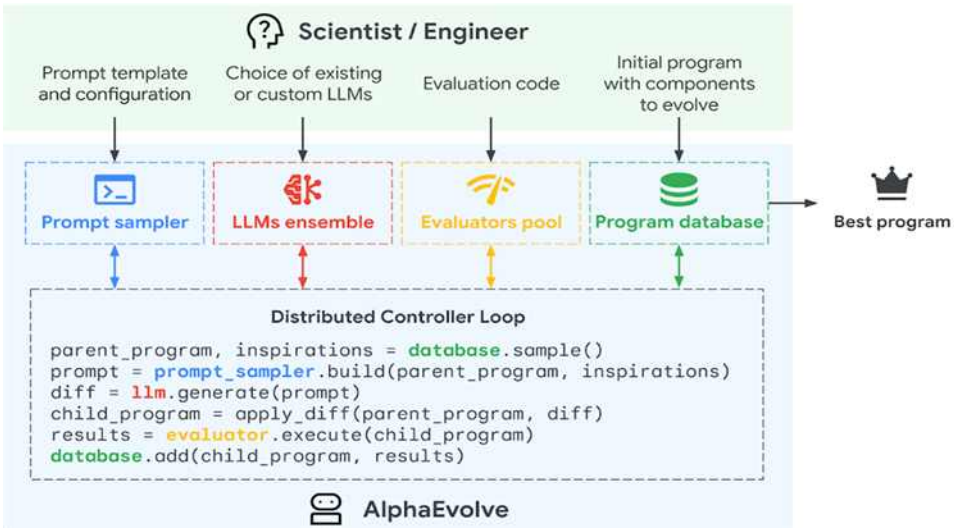
307) Google Deepmind, Gemini Robotics brings AI into the physical world, 2025.3.12.
308) Vision-Language-Action: 시각, 언어, 행동을 통합적으로 이해하고 처리하는 AI 모델
309) Out-of-distribution avg. 기준 Gemini Robotics: 0.65, π0 re-implement: 0.32, Multi-task diffusion: 0.34

Reasoning)’ 도 공개했다. 제미나이 로보틱스-ER은 물리적 세계에 대한 이해 능력을 향상한 모델로서, 3D 인식이나 상태 추정, 공간 이해, 계획과 코드 생성 등 로봇 제어에 필요한 모든 추론 단계를 즉각 수행할 수 있다. 향상된 공간 추론 능력과 제미나이의 코딩 능력을 결합해 즉각적으로 새로운 기능을 구현할 수 있다. 가령 커피잔을 보여주면 잔을 들어올리기에 적합한 그립 방식과 안전한 접근 경로를 직관적으로 파악한다. 모델 안전성을 위해 충돌 회피와 같은 기본적인 안전 제어장치와 상호작용할 수 있으며, 제미나이의 안전 기능을 바탕으로 주어진 시나리오 안에서 안전한 동작을 판단하고, 질문에 적절한 답변을 생성하도록 설계됐다.

나. 범용 알고리즘을 설계하는 코딩 에이전트 ‘알파이볼브’ 공개[310]

구글 딥마인드가 2025년 5월 14일 범용 알고리즘 발견과 최적화를 위한 진화형 코딩 에이전트 ‘알파이볼브(AlphaEvolve)’를 발표했다. 알파이볼브는 제미나이 모델의 문제 해결 능력과 답변을 검증하는 자동 평가 시스템을 결합하고, 가장 유망한 아이디어를 개선하기 위한 진화적 방식의 프레임워크를 채택했다. 빠르고 효율적인 제미나이 플래시를 활용해 탐색하는 아이디어의 폭을 극대화하고, 강력한 성능을 가진 제미나이 프로 기반의 심층 분석으로 알고리즘 솔루션을 코드로 구현하는 프로그램을 제안한다.

[그림 5-2] 알파이볼브의 작동 방식



출처 : Google Deepmind, AlphaEvolve

310) Google Deepmind, AlphaEvolve : A Gemini-powered coding agent for designing advanced algorithms, 2025.5.14.

구글 딥마인드는 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 유용하다고 강조했다. 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 결정하는 방식이다.

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 강조했다. 일례로 알파이볼브는 대규모 행렬 곱셈 연산을 더 작고 관리하기 쉬운 하위 문제로 나누는 방법을 발견해 핵심 연산 단위의 속도를 23% 높여 제미나이의 전체 학습 시간을 1% 단축했다.

다. 구글 딥마인드, 인간 DNA 염기서열 분석 AI 모델 '알파게놈' 공개[311]

구글 딥마인드가 2025년 6월 25일 인간 DNA 염기서열[312]의 변이가 유전자 활동에 미치는 영향을 예측할 수 있는 AI 모델 '알파게놈(AlphaGenome)'을 공개했다. 유전체(Genome)는 생명체의 거의 모든 부분을 안내하는 완전한 DNA 세트로서, DNA 서열의 변이는 생명체의 환경 적응력이나 질병 반응을 변화시킬 수 있으나, 유전체의 상당 부분은 여전히 생물학의 난제로 평가됐다. 알파게놈은 최대 100만 개의 문자(염기쌍)[313]로 구성된 긴 DNA 염기서열 입력 시 수천 종류에 달하는 분자적 특성(예: 세포 유형, 리보핵산(RNA)의 생산량, 특정 단백질에 결합하는 DNA 염기)을 예측할 수 있다.

유전적 변이를 더욱 정확히 예측하여 희귀 유전 질환을 유발하는 변이 연구 등 질병 이해를 높일 수 있으며, 특정 조절 기능을 가진 합성 DNA 설계와 같은 합성 생물학 연구 및 유전체 관련 기초 연구 지원이 쉽다. 알파게놈은 구글 딥마인드가 2023년 9월 공개한 '알파미센스(AlphaMissense)[314]'와 상호 보완하는 모델로, 유전체의 98%에 해당하는 단백질 비부호화[315] 영역 해독을 통해 구현됐다. 유전자 활동 조율에 필수적이면서 질병과 관련된 다

311) Google Deepmind, AlphaGenome: AI for better understanding the genome, 2025.6.25.
312) DNA의 기본 단위로 유전 형질을 구성하는 염기(아데닌(A), 구아닌(G), 시토신(C), 티민(T))가 배열된 순서
313) 이중나선 구조로 되어 있는 DNA의 두 가닥이 결합하여 생기는 염기의 짝(A-T, G-C)을 의미
314) 유전체의 2%를 차지하는 단백질 부호화 영역 내 변이 영향의 분류에 특화된 AI 모델

양한 변이를 포함하는 단백질 비부호화 영역의 광범위한 염기서열과 변이의 해석을 지원하며, 돌연변이가 발생한 염기서열과 발생하지 않은 염기서열의 예측값을 비교하여 유전적 돌연변이의 영향을 평가할 수 있도록 설계됐다. 구글 딥마인드는 알파게놈이 광범위한 유전체 예측 벤치마크 평가에서 기존 최고 성능(SOTA) 모델[316]과 비교해 더 뛰어나거나 동등한 수준의 성능을 발휘한다고 주장했다.

DNA 서열 예측에서는 24개 평가 과제 중 22개에서 최고 성능을 보였고, 변이의 조절 효과 예측에서는 26개 평가 과제 중 24개에서 기존 최고 성능 모델과 동등하거나 능가하는 기록을 달성했다고 공개했다. 하지만, 알파게놈은 유전자와 해당 기능을 조절하는 요소가 10만 개 이상 염기쌍으로 떨어져 있을 경우의 영향을 포착하는 데는 어려움을 겪는 등 한계도 일부 인정했다. 연구진은 알파게놈이 분자 수준을 넘어 유전적 변이가 어떻게 복잡한 형질이나 질병으로 이어지는지를 밝히는 것처럼 광범위한 생물학적 과정을 다루어야 하는 사례에는 적용되기 어렵다고 밝혔다. 연구진은 이러한 격차를 해소하기 위해 피드백을 받아 모델을 지속적으로 개선할 계획이며, 비상업적 연구용으로 '알파게놈 API'를 통해 모델을 프리뷰로 공개할 방침이라고 밝혔다.

라. 구글, 의료용 AI 모델 '메드젬마' 오픈소스 공개[317]

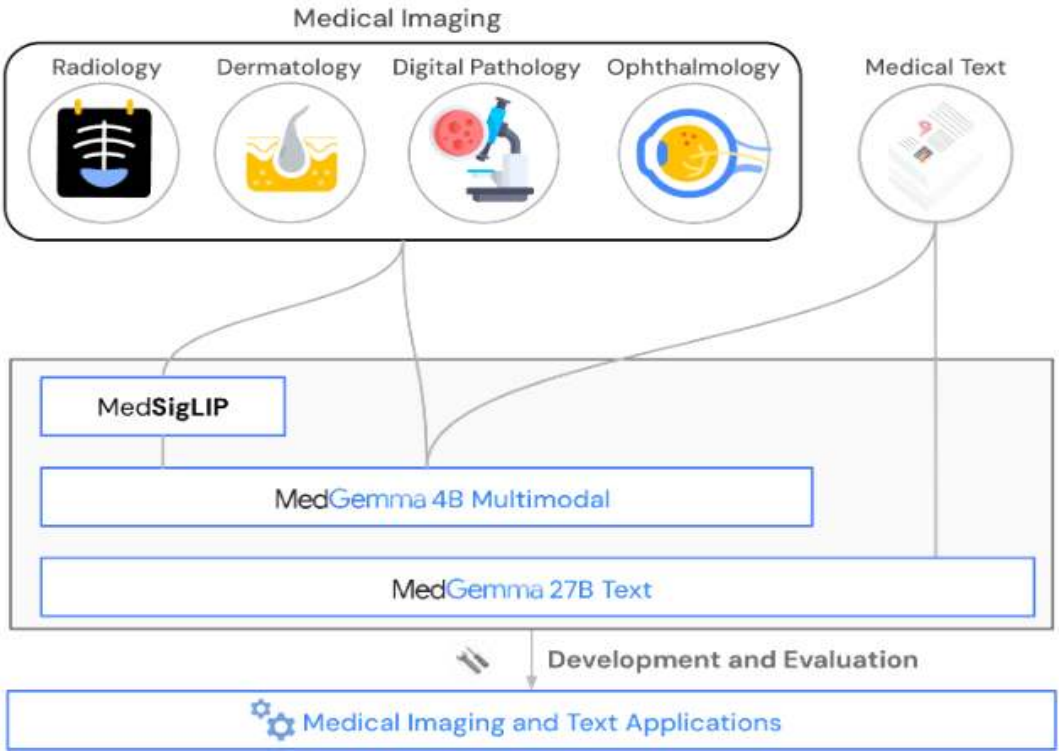
구글이 2025년 7월 9일 의료 및 생명과학 분야 연구에 특화된 AI 모델 '메드젬마(MedGemma)'와 '메드시그립(MedSigLIP)'을 깃허브에 오픈소스로 공개했다. 앞서 2025년 5월 의료 텍스트와 이미지 분석에 특화된 메드젬마 4B 멀티모달과 27B 텍스트 전용 모델을 처음 발표했으며, 복잡한 멀티모달과 전자의료기록 해석 기능을 추가한 '메드젬마 27B 멀티모달' 모델을 새로 공개했다.

구글은 의학적으로 최적화된 이미지 인코더(메드시그립)를 학습시킨 후, 젬마 3 모델의 4B와 27B 버전에 의료 데이터를 학습시켜 보고서 작성 등 의료 텍스트와 영상 작업에 유용한 메드젬마를 개발했다고 밝혔다. 메드시그립은 흉부 X선, 피부과 이미지 등 다양한 의료 영상 데이터를 활용해 텍스트와 이미지의 사전학습 아키텍처(SigLIP)[318]를 개선한 이미지 인코

315) 유전체 중 단백질로 발현되지 않는 영역
316) SpliceAI, Borzoi, ProCapNet, Panglon, ChromBPNet, Orca, Enformer, AbSplice
317) Google, MedGemma : Our most capable open models for health AI development, 2025.7.9.
318) 이미지와 텍스트를 함께 학습시켜 이미지와 텍스트 사이의 관계를 학습하도록 돕는 기법

더 모델로, 기존 이미지나 새로운 이미지의 분류, 시각적으로나 의미적으로 유사한 이미지 검색 등 구조화된 출력이 필요한 작업에 유용하다는 설명이다.

[그림 5-3] 구글의 메드젬마와 메드시그립 모델 개요



출처: Google, MedGemma

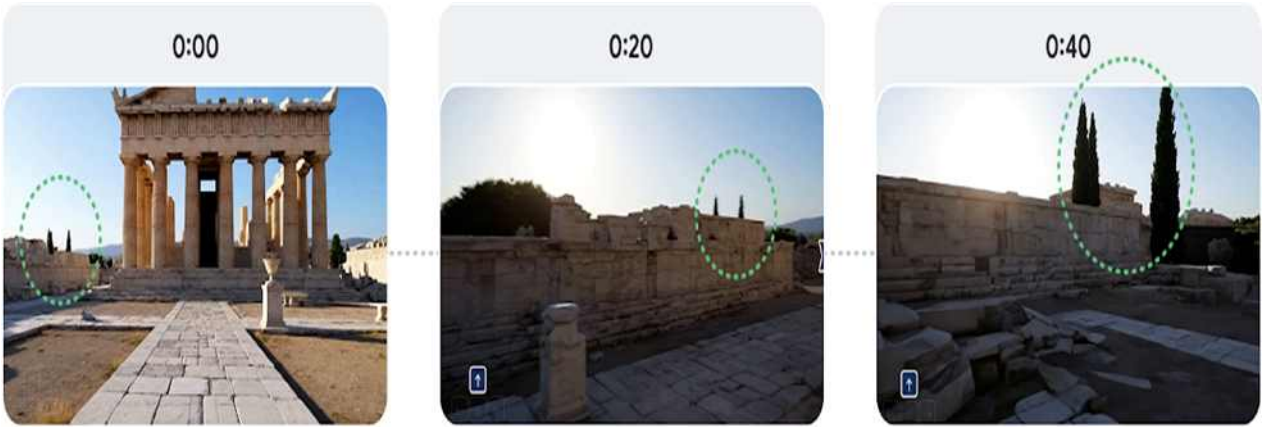
메드젬마 모델 중 메드젬마 4B 멀티모달은 MedQA[319] 벤치마크 평가에서 64.4%를 기록해, 매개변수 80억 개 미만의 오픈소스 모델 중 최고 수준의 성능을 보였다. 메드젬마 27B 모델도 MedQA 평가에서 매개변수 500억 개 미만의 모델 중 최고 수준의 성능을 기록했으며, 특히 텍스트 모델은 87.7%로 딥시크 R1보다 3%p 낮지만, 추론 비용은 약 10분의 1에 그친다.

319) 미국 의사 면허 시험(USMLE) 기반의 다지선다형 의학 질문으로 구성된 AI 모델 평가 데이터셋

마. 구글 딥마인드, 차세대 월드 모델 ‘지니 3’ 공개[320]

구글 딥마인드가 2025년 8월 5일 실시간으로 상호작용이 가능한 다양한 환경을 생성하는 차세대 월드 모델(World Model)[321] ‘지니 3(Genie 3)’를 발표했다. 2024년 12월 공개된 지니 2는 360p 해상도에서 10~20초 길이로 물리 세계를 생성할 수 있었으나, 지니 3는 720p 고해상도에서 초당 24프레임 속도로 몇 분 동안 실시간으로 상호작용 할 수 있는 물리 세계를 생성한다. 연구진은 지니 3가 지니 2에 비해 일관성과 사실성이 개선되고 실시간 상호작용을 지원한다고 강조했다. 지니 3는 AI로 생성된 물리 세계의 몰입감을 확보할 수 있도록 몇 분에 걸쳐 물리적 일관성을 유지할 수 있으며, 일례로 상호작용이 이루어지는 내내 나무와 같은 고정된 물체의 상태는 시간이 지나도 일관되게 유지된다.

[그림 5-4] 지니 3로 생성된 세계의 일관성 유지(예시)



출처 : Google Deepmind, Genie 3

지니 3는 날씨를 바꾸거나 새로운 사물과 캐릭터를 추가하는 등 텍스트 프롬프트를 입력함으로써 생성된 물리 세계 상태를 변경할 수 있는 ‘Promptable World Events’ 기능도 제공한다. 이 기능은 경험을 통해 학습하는 에이전트가 예기치 못한 상황을 처리하는 데 사용되는 가상 시나리오의 범위를 확장할 수 있도록 한다. 구글 딥마인드는 지니 3의 제한 사항도 공개했으며, 일례로 물리 세계 생성 후 텍스트 프롬프트를 통해 환경을 광범위하게 바꿀 수 있지만 에이전트가 직접 수행할 수 있는 동작 범위는 일부에 그친다.

320) Google Deepmind, Genie 3: A new frontier for world models, 2025.8.5.
321) 현실 세계의 작동 방식을 학습하고 이를 바탕으로 미래 예측과 시뮬레이션을 수행하는 AI 모델

실제 위치를 지리적으로 완벽히 정확하게 시뮬레이션할 수 없고, 연속적인 상호작용 시간도 몇 분 수준으로 제한되며, 하나의 환경 내에서 여러 독립 에이전트 간 복잡한 상호작용의 정확한 모델링도 개선이 필요하다. 구글 딥마인드는 지니 3를 소수의 학자와 창작자를 대상으로 연구용 프리뷰로 제공하여 응용 분야를 탐색하고 위험과 완화책에 대한 피드백을 수집할 계획이다. 로봇과 자율 시스템 같은 에이전트 학습을 위한 기회를 제공하는 한편, 에이전트의 성능과 약점을 탐색할 기회를 제공함으로써 교육과 훈련을 지원할 계획이다.

(3) 제품 및 서비스

가. 성능 개선한 ‘제미나이 2.0’ 제품군 확대 출시[322)]

구글이 2024년 12월 공개한 추론 AI 모델 ‘제미나이 2.0 플래시(Gemini 2.0 Flash)’를 모든 사용자에게 확대 출시하고, 코딩과 복잡한 프롬프트 처리에 강한 최고 사양의 모델 ‘제미나이 2.0 프로(Gemini 2.0 Pro)’ 실험 버전을 공개했다.[323)] 제미나이 2.0 플래시는 100만 개 토큰 길이의 컨텍스트 창과 멀티모달 추론을 지원하며, 이미지 생성 및 텍스트-음성 변환 기능 추가가 예정되어 있다. 제미나이 2.0 프로 실험 버전은 구글 AI 모델 중 가장 강력한 코딩 성능과 복잡한 프롬프트 처리 능력을 갖추고 200만 개 토큰 길이의 최대 컨텍스트 창을 지원하며 구글 검색 및 코드 실행과 같은 도구 호출을 통해 에이전트의 기반이 되는 기능도 구현했다.

구글은 비용 효율성이 가장 뛰어난 ‘제미나이 2.0 플래시 라이트(Flash-Lite)’ 모델도 구글 버텍스 AI[324)]와 AI 스튜디오에서 프리뷰로 공개했다. 제미나이 2.0 플래시 라이트는 이전 버전인 제미나이 1.5 플래시와 동일한 속도와 비용으로 더 높은 성능을 제공하며, 100만 개 토큰 길이의 컨텍스트 창과 멀티모달 입력을 지원한다. 구글은 추론에 특화된 ‘제미나이 2.0 플래시 씽킹(Flash Thinking)’ 실험 버전도 제미나이 웹과 모바일 앱을 통해 출시했다. 제미나이 2.0 플래시 씽킹 실험 버전은 2.0 플래시를 기반으로 프롬프트를 단계별로 나누어 추론 기능을 강화하고 응답을 개선하도록 훈련된 모델로, 사고 과정을 확인해 특정 방식으로 응답한 이유와 추론 경로 등을 추적할 수 있다.

---

322) Google, Gemini 2.0 is now available to everyone, 2025.2.5.
323) Google, Access the latest 2.0 experimental models in the Gemini app, 2025.2.5.
324) 버텍스 AI: 머신러닝 모델 구축, 배포, 확장을 위한 도구와 서비스를 제공하는 종합 플랫폼
AI 스튜디오 버텍스 AI 플랫폼 내 생성 AI 개발 도구

제미나이 2.0 플래시 씽킹 실험 버전은 유튜브와 검색, 구글 맵스와 같은 앱과 연동되어 추론 기능 적용 시의 활용 가능성을 탐색할 수 있다. 구글은 제미나이 모델의 안전성을 지속적으로 개선하기 위해 투자하고 있으며, 자체적으로 응답을 평가하는 새로운 강화학습 기법을 적용해 제미나이 2.0 시리즈를 개발했다고 강조했다. 이를 통해 더욱 정확하고 구체적인 피드백을 제공할 수 있으며 민감한 프롬프트 처리 능력도 향상되었다고 주장했다.

나. 구글, 제미나이 앱에서 '제미나이 2.5 딥씽크' 제공[325]

구글이 2025년 8월 1일 '제미나이 2.5 딥씽크(Gemini 2.5 Deep Think)'를 최상위 요금제인 구글 AI 울트라(월 249.99달러) 구독자 대상으로 제미나이 앱 제공을 발표했다.[326] 구글은 추론 특화 AI 모델인 제미나이 2.5 딥씽크가 2025년 5월 I/O에서 처음 공개된 이후 벤치마크 성능과 테스터 피드백에서 상당한 개선이 적용되었다고 설명했다. 이번에 공개된 모델은 2025년 국제수학올림피아드(IMO)에서 금메달을 획득한 모델의 변형 버전으로, 내부 평가 결과 2025 IMO에서 동메달 수준의 성능을 달성하면서 일반 용도에서 속도와 사용성을 개선했다. 구글은 소수의 학자에게 IMO 금메달 획득 버전을 제공해 연구를 지원하는 한편, 일부 테스트 참여자에게 API를 통해 제미나이 2.5 딥씽크를 제공해 개발자와 기업 활용 사례 중심의 피드백을 수집할 계획이다.

제미나이 2.5 딥씽크는 사람이 복잡한 문제 해결을 위해 다양한 관점을 탐색해 최종 답을 다듬듯이 병렬 사고 기법으로 여러 아이디어를 동시에 검토하고, 이를 수정하거나 결합하여 최적의 답을 도출한다. 제미나이 2.5 딥씽크는 반복적인 개발과 설계, 과학이나 수학적 발견, 복잡한 코딩 문제에서 탁월한 성능을 발휘하며, 코딩과 과학, 추론 관련 벤치마크에서도 최고 수준의 성능을 달성하는 성과[327]를 보였다. 제미나이 2.5 딥씽크 모델 카드(Model Card)에 따르면, 제미나이 2.5 프로 대비 이미지-텍스트 안전성과 응답 거부 시의 톤이 개선되었으나 무해한 요청을 과도하게 거부하는 경향도 나타났다.

구글 내부 안전성 평가 결과, 제미나이 2.5 딥씽크의 응답 거부 톤은 제미나이 2.5 프로 대비 16.3% 개선되었으나, 동시에 지시이행 비율이 (-)9.9%를 기록해 무해한 요청을 거부하

325) Google, Try Deep Think in the Gemini app, 2025.8.1.
326) Google, Gemini 2.5 Deep Think Model Card, 2025.8.1.
327) 추론/지식 HLE(Humanity's Last Exam) 기준 34.8%(Grok 4:25.4%, Gemini-2.5-Pro:21.6%, OpenAI o3: 20.3%) 코드 생성 LiveCodeBench 기준 87.6%(Grok 4: 79.0%, Gemini 2.5 Pro: 74.2%, OpenAI o3: 72.0%)

는 사례가 증가했다. 또한 다국어 환경에서 안전성은 제미나이 2.5 프로 대비 (-)1.0%로 큰 차이가 없었으나 텍스트 질문에 대한 텍스트 응답 안전성은 (-)16.3% 저하되었고, 이미지 질문에 대한 텍스트 응답 안전성은 +2.1% 개선됐다. 구글은 제미나이 2.5 딥씽크의 위험 완화를 위해 개발 수명주기 전반에 걸쳐 데이터셋 필터링과 사전학습, 감독 미세조정, 인간 피드백을 통한 강화학습, 안전 정책 수립 등의 안전 조치를 취했다고 설명했다.

다. 구글 클라우드, 데이터 분석에 특화된 AI 에이전트 제품군 발표[328]

구글 클라우드가 전문 AI 에이전트를 활용할 수 있는 데이터 플랫폼 구축 일환으로 새로운 데이터 에이전트 제품군을 프리뷰로 공개했다. 자연어 지시만으로 에이전트가 데이터 수집과 변환, 품질 유지 등의 작업흐름을 생성하고 조정하는 '데이터 엔지니어링 에이전트(프리뷰)'는 빅쿼리(BigQuery)[329] 환경에서 복잡한 데이터 업무의 간소화와 자동화를 지원한다. 구글 제미나이 기반의 '데이터 과학 에이전트(프리뷰)'는 데이터 분석과 정리, 예측 모델 생성 등의 작업흐름을 자율적으로 계획해 실행하며, 사용자 피드백을 받아 계획 수립과 실행, 결과 추론과 보고까지 가능하다.

기업 사용자를 위한 '코드 인터프리터'는 제미나이의 고급 추론 기능을 활용해 복잡한 자연어 질문을 파이썬(Python) 코드로 변환하며, 코드 생성과 명확한 자연어 설명 제공, 필터·확대·세부 정보 조회가 가능한 상호 작용형 시각화 생성 등으로 전체 분석 흐름을 지원한다. 구글 클라우드는 AI 에이전트 생태계의 확장성과 연결성을 지원하기 위해 '제미나이 데이터 에이전트' API 출시를 발표하고 ▲대화형 분석 API(프리뷰) ▲기업 맞춤형 에이전트 개발 키트(ADK) 등을 공개했다. '대화형 분석 API는 루커(Looker)[330]의 자연어처리 및 코드 해석 기능을 기업의 애플리케이션이나 제품, 업무 흐름에 직접 통합할 수 있는 기본 요소를 제공해 기업별 필요에 맞는 데이터 경험 구축을 지원한다. 기업 맞춤형 에이전트 개발 키트 및 에이전트 간 협업을 지원하는 데이터베이스용 MCP[331] 도구상자와 루커 MCP 서버도 출시했다.

328) Google Cloud, Redefining enterprise data with agents and AI-native foundations, 2025.8.6.
329) 대규모 데이터 분석이 가능한 구글 클라우드 데이터 저장소
330) 구글 클라우드 내 데이터 시각화 및 분석 플랫폼
331) Model Context Protocol : AI 모델이 필요 정보를 외부 데이터 소스나 도구와 표준화 방식으로 주고받도록 하는 개방형 통신 규약

구글은 2024년 12월 처음 공개한 AI 코딩 에이전트 '줄스(Jules)'를 베타 테스트를 거쳐 UI 개선과 오류수정을 진행하고 멀티모달 지원과 같은 신기능을 추가해 2025년 8월 6일 정식 출시했다.[332] 구글 AI 프로(월 19.99달러)와 울트라(249.99달러) 구독자에게 제공되는 줄스는 제미나이 2.5 프로의 고급 추론 기능을 활용하여 코딩 계획을 세우고 고품질의 코딩 결과물을 생성하며, 여러 요청을 동시에 처리할 수 있다. 신기능으로 이전 설정을 재사용해 새 작업을 더 빠르게 실행할 수 있으며 깃허브 이슈 연동 및 멀티모달을 지원한다.

라. 구글, 이미지 편집 모델 '제미나이 2.5 플래시 이미지' 공개[333]

구글이 제미나이 앱에 새로운 이미지 생성·편집 모델 '제미나이 2.5 플래시 이미지(Gemini 2.5 Flash Image)' 기반 이미지 편집 기능을 추가했다. 이번 모델은 LMArena[334] 플랫폼에서 '나노 바나나(Nano Banana)' 가칭으로 공개되어 뛰어난 성능으로 1위를 차지하며 화제를 모았으며, 이번 업데이트로 제미나이 앱에 통합됐다. 이번 업데이트는 전 세계 유·무료 사용자에게 모두 제공되며, 제미나이 앱에서 생성되거나 편집된 모든 이미지에는 눈으로 식별되지 않는 워터마크가 포함되어 AI 생성 사실을 확인할 수 있다. 구글은 2025년 4월 제미나이 앱에 AI 이미지 편집 기능을 도입한 이래, 캐릭터 일관성에 중점을 두고 기술을 개선해 왔고, 제미나이 2.5 플래시 이미지는 캐릭터 일관성 유지에서 높은 품질을 제시한다.

[그림 5-5] 제미나이 2.5 플래시 이미지'의 이미지 합성 예시



출처: Google, Nano Banana! Image editing in Gemini just got a major upgrade

332) Google, Jules, our asynchronous coding agent, is now available for everyone, 2025.8.6.
333) Google, Nano Banana! Image editing in Gemini just got a major upgrade, 2025.8.26.
334) 인간 선호도를 기반으로 AI 모델을 비교 평가하는 웹 기반 플랫폼

사람이나 동물 사진을 올리면 배경이나 설정을 바꾸더라도 모든 이미지에서 동일한 모습을 유지하며, 여러 장의 사진을 업로드하고 합성하여 새로운 장면을 만들어낼 수 있도록 지원한다. 실내 이미지에서 벽의 색을 바꾸거나 책장, 테이블 등의 가구를 차례로 추가하는 등 여러 차례에 걸쳐 특정 부분을 수정하면서 나머지 부분을 유지하는 것도 가능하며, 한 이미지의 스타일을 다른 이미지에 적용하는 기능도 지원(예시: 나비 날개 패턴을 드레스 디자인에 합성)한다.[335]

마. 구글, AI 기반 결제 지원 '에이전트 결제 프로토콜(AP2)' 발표[336]

구글이 2025년 9월 17일 AI 에이전트가 쇼핑과 결제를 수행하는 개방형 프로토콜 '에이전트 결제 프로토콜(Agent Payment Protocol, AP2)'을 발표했다. 구글은 마스터카드(Mastercard), 아메리칸 익스프레스(American Express), 페이팔(PayPal) 등 60개 이상 다양한 결제 및 기술 기업과 협력해 AP2를 개발했다. 기존 결제 시스템은 사람이 직접 구매한다는 가정하에 구축되었으나, 구매 결정을 대신하는 자율 에이전트의 등장으로 에이전트의 거래 권한을 안전하게 인증, 검증 및 전달할 인프라의 구축 필요성이 커졌다.

AP2는 사용자, 판매자, 결제 서비스 업체가 다양한 결제 수단으로 안전하게 거래할 수 있는 공통 언어를 제공함으로써 생태계 파편화를 방지하며, 기존 AI 프로토콜인 A2A[337] 및 MCP[338]와 연동해 사용할 수 있다. AP2는 변조 방지와 암호화가 적용된 디지털 계약인 위임장(Mandates)을 통해 사용자가 실시간으로 참여하는 구매 및 에이전트 단독 구매 등 모든 거래에서 에이전트가 사용자의 지시를 받았음을 입증한다.

사용자 참여 구매는 사용자가 에이전트에게 '흰 운동화를 찾아줘'라고 요청하면 이를 초기 '의향 위임장'으로 기록하고, 에이전트가 상품이 담긴 카트를 제시하면 사용자가 '카트 위임장'으로 승인해 품목과 가격을 확정한다. 에이전트 단독 구매는 콘서트 티켓 판매 시작 후 구매와 같이 구매를 위임한 경우, 사용자가 사전에 상세한 구매 조건을 명시한 의향 위임장에 서명하여 정확한 조건 충족 시 자동으로 카트 위임장을 생성하는 방식이

335) Google DeepMind, Gemini 2.5 Flash Image Gemini 2.5 Flash Image, (2025.11.23. 방문) https://deepmind.google/models/gemini/image
336) Google Cloud, Powering AI commerce with the new Agent Payments Protocol (AP2), 2025.9.17.
337) Agent2Agent: AI 에이전트 간 상호운용성을 지원하는 통신 규약
338) Model Context Protocol: AI 모델이 필요 정보를 외부 데이터 소스나 도구와 표준화 방식으로 주고받도록 하는 개방형 통신 규약

다. 사용자 실시간 참여 구매 및 에이전트 단독 구매 시나리오 모두 사용자 의향부터 카트, 결제까지 전 과정에서 사용자 승인과 진위를 확정하는 확실한 증거를 생성해 책임 소재를 명확히 구분할 수 있다.

구글은 AP2는 유연한 설계 방식으로 단순한 구매 방식부터 지금까지 없던 완전히 새로운 구매 방식까지 모두 지원할 수 있는 기반을 제공한다고 강조한다. 사용자가 원하는 특정 색상의 상품이 없을 때 에이전트에게 최대 20%까지 더 낼 의향이 있다고 밝히면 가격과 재고 현황을 모니터링해 해당 상품이 발견되면 자동으로 구매를 실행하도록 설정할 수 있다. 구글은 코인베이스(Coinbase), 이더리움 재단(Ethereum Foundation) 등과 협력해 암호화폐 기반의 에이전트 결제를 위한 개방형 확장 규격(A2A x402)도 출시해 AP2에서 스테이블코인과 암호화폐 등 다양한 결제 수단을 지원할 방침이다.

바. 구글, 최첨단 추론 성능 갖춘 ‘제미나이 3’ 출시[339]

구글이 2025년 11월 18일 역대 자사 AI 모델 중 가장 지능적인 ‘제미나이 3’를 출시했다. ‘제미나이 3’은 멀티모달 이해 능력에서 세계 최고 수준의 성능을 달성하고 역대 구글 모델 중 가장 강력한 에이전트기능을 갖추었으며, 복잡한 지시 없이도 AI가 사용자의 의도를 파악해 코딩을 작성하고 결과물을 구현하는 기술인 ‘바이브 코딩’ 기능도 포함되었다. 구글은 이 모델이 강력한 추론 성능을 바탕으로 심도 있는 상호작용을 제공한다고 강조했다. ‘제미나이 3 프로’는 MMMU-Pro(시각적 추론) 81.0%, Video-MMMU(동영상 추론) 87.6%로 타사 최신 AI 모델을 능가했고[340], SimpleQA Verified(사실 정확성) 평가에서도 72.1%로 최고 기록을 달성했다. 벤치마크[341]에서도 뛰어난 성적을 거두었다. 구글은 새로 공개한 모델들을 미나이 앱과 개발자 플랫폼 등 구글 서비스 전반에 적용하였으며, 추론 성능을 한층 강화한 ‘제미나이 3 딥씽크’[342]도 일부 테스터를 통한 사전 안전성 평가를 진행한 뒤 2025년 12월에 구글 AI 울트라 구독자에게 공개하였다.

구글은 ‘제미나이 3’가 최첨단 추론을 통해 다양한 분야의 학습과 개발, 계획을 지원할 수 있다고 강조하였다. ‘제미나이 3’가 적용된 구글 검색의 AI 모드는 실시간으로 이용자

339) Google, A new era of intelligence with Gemini 3, 2025.11.18.
340) MMMU-Pro/Video-MMMU: 클로드 소네트 4.5(68.0%/77.8%), GPT-5.1(76.0%/80.4%)
341) Humanity's Last Exam(HLE): 도구 사용 없이 37.5%, MathArena Apex(수학적 추론): 최고점 23.4%
342) Humanity's Last Exam(HLE): 도구 사용 없이 41.0%

검색 의도에 맞게 시각적으로 풍부한 정보를 담은 화면과 사용자가 조작할 수 있는 도구·시뮬레이션 등의 생성 사용자 인터페이스(UI)를 제공하였고, 함께 공개한 새로운 에이전트 개발 플랫폼 '구글 안티그래비티(Google Antigravity)'는 SWE-bench Verified(코딩 에이전트)에서 76.2%를 기록한 제미나이 3의 코딩 에이전트 성능을 바탕으로 설계되었다고한다. '제미나이 3'는 장기적 관점에서 계획을 수립하고 이행하는 능력 또한 향상되어, 가상으로 자판기 사업을 운영하며 장기간에 걸쳐 일관성을 평가하는 Vending-Bench 2에서 리더보드 1위를 차지[343]하였다.

구글에 따르면 '제미나이 3'는 지금까지 출시된 자사 모델 중 가장 안전하고 가장 포괄적인 안전 평가를 거친 모델로, 과도한 아첨(Sycophancy)을 줄이고 프롬프트 주입 공격에 대한 저항성도 강화하였으며, 내부 기준에 따른 안전성 평가뿐 아니라 각 분야의 세계적 전문가와 협력해 평가를 진행하고, 영국 AI보안연구소(AISI)와 아폴로(Apollo)를 비롯한 제3자 전문 평가기관을 통해 독립적인 안정성 평가를 받았다.

(4) AI 윤리 및 규제 대응

가. 구글, 과학적 발견을 위한 실증 소프트웨어 생성 AI 시스템 개발[344]

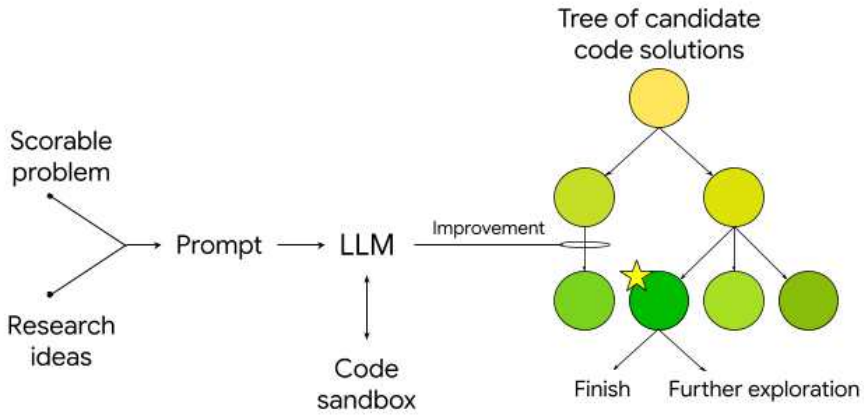
구글 리서치가 2025년 9월 9일 과학 연구에서 새로운 지식과 발견을 위한 복잡한 실험과 시뮬레이션, 분석에 요구되는 실증 소프트웨어[345]를 생성하는 AI 시스템을 발표했다. 기존에는 과학자들이 과학적 가설을 철저히 검증하기 위한 소프트웨어 코딩과 최적화에 막대한 시간을 소모하여 과학적 발견의 속도를 저해하는 병목 현상이 발생했다. 구글의 접근 방식은 실증 소프트웨어를 적용해 해결할 수 있는 정량 평가 기반의 과학 문제에 대하여 문제 설명과 평가 수단을 AI 시스템에 입력하고 체계적인 최적화 코드 솔루션을 도출해 이러한 병목 현상을 완화한다. 과학적 문제에 대한 설명과 평가 지표, 학습·검증·평가 데이터셋을 제공받은 AI 시스템은 새로운 연구 아이디어를 생성해 실행가능한 코드로 구현하며 필요할 경우, 외부 자료나 지식을 참고해 성능을 향상한다.

AI 시스템은 코드 구현 다음 단계에서 트리 탐색(Tree Search)[346] 전략을 사용해 코드 솔

343) 시뮬레이션 5회 실행 시 수익 평균: 제미나이 3 프로(5,478.16달러), 클로드 소네트 4.5(3,838.74달러), 그록 4(1,999.46달러)
344) Google Research, Accelerating scientific discovery with AI-powered empirical software, 2025.9.9.
345) Empirical Software : 실험으로 확보한 증거를 활용해 최대 성과를 낼 수 있도록 최적화하는 소프트웨어

루션 후보 트리를 생성하고 더 나은 성능을 내는 갈래를 따라 추가 탐색이 필요한 후보를 결정한다. 만일 코드 실행 결과가 좋지 않으면 LLM을 사용해 코드를 재작성해 품질 점수를 개선하고, 효율적 탐색을 통해 사람이 수개월 걸릴 탐색을 수일 또는 수 시간 만에 진행함으로써 고품질 코드 솔루션을 식별한다.

[그림 5-6] 과학적 문제 해결을 위한 코드 솔루션 도출 알고리즘



출처: Google Research, Accelerating scientific discovery with AI-powered empirical software

연구진이 시스템을 활용해 6개 분야[347] 평가를 통해 공중보건, 지리 공간 분석, 신경과학 등의 과학 문제를 실험한 결과, 모든 분야에서 기존의 첨단 방법과 동등하거나 이를 능가하는 결과를 도출했다. 공중보건 분야에서는 미국의 코로나 입원자 수 예측에서 기존 모델(CovidHub-ensemble)보다 성능이 뛰어난 모델을 14개 생성했고, 위성 원격탐사 영상을 활용한 지리 공간 분석에서도 기존 첨단 기법을 능가했다는 설명이다.

나. 구글 클라우드, 2025년 AI 투자수익률 보고서 발표[348]

구글 클라우드가 전 세계 글로벌 대기업 대상의 설문조사 결과를 바탕으로 2025년 9월 4일 '2025년 AI 투자수익률(ROI of AI)' 보고서를 발간했다. 글로벌 매출 1,000만 달러 이상 기업의 고위 임원 3,466명을 대상으로 진행된 이번 조사 결과, 전 세계 대기업에서 AI 에이전트 도입이 빠르게 확산하는 추세가 확인됐다. 생성 AI를 사용하는 기업의 52%가 AI 에이전트도 적극 활용하고 있다고 답했으며, 39%의 응답자는 기업이 10개 이상의 AI 에이전트를 도입했다고 응답했다. 전체 산업에서 AI 에이전트 활용 분야는 고객 서비스와 경험

346) 여러 선택지를 나무(Tree) 형태로 분기해 가며 가장 유망한 경로를 선택하는 방식
347) 유전체학, 공중보건/역학, 지리 공간 분석, 신경과학, 시계열 예측, 수치 해석
348) Google Cloud, ROI of AI 2025, 2025.9.4.

(49%), 마케팅(46%), 보안 운영과 사이버보안(46%), 기술지원(45%), 제품 혁신과 디자인(43%), 생산성과 연구(43%) 순을 기록했다.

지역별로 AI 에이전트 활용의 우선순위는 다르게 나타났으며, 유럽 지역 응답자들은 AI 기반 기술지원을 우선하는 반면, 아시아태평양 응답자들은 고객 서비스를, 라틴아메리카 응답자들은 마케팅을 우선시하는 것으로 나타났다. 에이전틱 AI 조기 도입 기업 전체 응답 기업의 13%가 AI 예산 50% 이상을 AI 에이전트에 투입하고 있으며, 88%가 하나 이상의 생성 AI 활용 사례에서 ROI를 달성했다고 답해, 전체 기업 평균인 74%를 넘어섰다, 생성 AI 활용 사례에서 ROI를 달성했다는 전체 응답 중 생성 AI 활용 분야는 생산성(39%), 고객 경험(37%), 판매・마케팅(33%) 순을 보였다.

기업들의 전체 AI 투자는 증가 추세로, 77%가 기술 비용 감소로 생성 AI 지출이 늘어났다고 응답했다.[349] 58%의 응답자는 생성 AI 투자를 뒷받침하기 위해 다른 예산을 줄이지 않고 신규 예산을 할당했으며, 48%는 생성 AI 투자를 위해 AI와 무관한 예산을 재분배했다고 답했다. 조사 기업들은 생성 AI를 통해 새로운 가치가 창출되는 5대 영역으로 생산성(70%), 고객 경험(63%), 사업 성장(56%), 마케팅(55%), 보안(49%) 순이라고 반응했다. LLM 공급업체 선택 시 기업들이 가장 중요하게 고려하는 요소는 개인정보 보호와 보안(37%)으로 나타났으며, 기존 시스템과의 통합(28%), 비용(27%)이 다음 순서를 차지했다. 이는 기업들이 특정한 기능이나 맞춤 설정과 같은 고급 기능이나 차별화 포인트를 고려하기에 앞서 핵심적인 기업 요구사항을 중시하는 것으로 나타났다.

349) Google Cloud, Google Cloud Study Reveals 52% of Executives Say Their Organizations Have Deployed AI Agents, Unlocking a New Wave of Business Value, 2025.09.04.

2. 마이크로소프트

1) 기업 일반현황

<표 5-2> 마이크로소프트 기업 현황

<table>
<tr><th>구분</th><th>주요 내용</th></tr>
<tr><td rowspan="3">기업 개요</td><td>〈회사 개요〉
• 1975년 4월 미국 뉴멕시코주 앨버커키에서 설립되었으며, 1986년 나스닥 상장을 통해 글로벌 소프트웨어 기업으로 도약
• 2025년 기준 190개국에서 수백 개의 지사·법인을 운영하며 운영체제·기업용 소프트웨어·클라우드·AI 등 다수의 핵심 시장에서 세계적 영향력을 유지
• 2024 회계연도 매출이 전년 대비 대폭 증가했고 애저(Azure)·클라우드 기반 AI 서비스·Microsoft 365 구독 모델 확장이 주 성장 요인으로 분석
• 2023~2025년 AI 중심 전략 전환 이후 MS의 사업 포트폴리오는 ‘클라우드+AI 플랫폼 중심 구조’로 재정비됐고, 코파일럿·애저 오픈AI·엔터프라이즈 AI 시장이 성장 축으로 작용</td></tr>
<tr><td>〈사업 내용〉
• 산업 분야 : 클라우드 컴퓨팅(Azure), 소프트웨어·운영체제(Windows/Office), 보안(Security), 엔터프라이즈 AI, 게임·콘텐츠(Xbox), 하드웨어(서피스).
• 핵심 서비스: Microsoft 365(Word, Excel, PowerPoint, Teams), Dynamics 365, Windows, Linkedin, GitHub, Bing, Edge, Xbox Game Pass, 애저 클라우드, 애저 AI 서비스 등
• 2024~2025년부터는 모든 서비스 전반에 코파일럿(Copilot) 기능이 통합되며, 윈도, 팀스, 아웃룩, 파워 BI, 보안 제품군까지 AI 기반 기능 제공 범위가 크게 확대됨.</td></tr>
<tr><td>〈자회사〉
• 링크드인(LinkedIn), 깃허브(GitHub), 뉘앙스 커뮤니케이션즈, 모장 스튜디오스, 스카이프/야머(Skype·Yammer)</td></tr>
<tr><td>재무 실적[350]
(최근 5년)</td><td>
<table>
<tr><th>마감 기준:</th><th>2020
30/06</th><th>2021
30/06</th><th>2022
30/06</th><th>2023
30/06</th><th>2024
30/06</th><th>2025
30/06</th></tr>
<tr><td colspan="7">손익 계산서 ></td></tr>
<tr><td>총매출</td><td></td><td>168,088</td><td>198,270</td><td>211,915</td><td>245,122</td><td>281,724</td></tr>
<tr><td>총이익</td><td></td><td>115,856</td><td>135,620</td><td>146,052</td><td>171,008</td><td>193,893</td></tr>
<tr><td>영업 이익</td><td></td><td>69,916</td><td>83,383</td><td>88,523</td><td>109,433</td><td>128,528</td></tr>
<tr><td>순이익</td><td></td><td>61,271</td><td>72,738</td><td>72,361</td><td>88,136</td><td>101,832</td></tr>
</table>
</td></tr>
<tr><td>AI 조직 현황</td><td>〈AI 관련 조직〉
• Microsoft Research(MSR) : MS 핵심 연구조직으로 전 세계 8개 이상의 연구센터를 운영하며 AI 이론·시스템·보안·HCI·멀티모달 모델 등 기초 연구 전반을 수행하며 최근 들어 대규모 언어모델, 책임 있는 AI, AI 보안, 시스템 최적화 등 기업용 AI 기술의 기초 기반을 강화
• 애저 AI(Azure AI) : MS 클라우드 핵심 AI 서비스 조직으로 Azure OpenAI Service·Azure Model Catalog·Azure AI Studio 등을 운영하며 최근 들어 오픈AI 모델(GPT-4.1, Sora, 이후 차세대 모델)과 MS 자체 모델 통합을 진행하고 있으며. 기업용 AI·</td></tr>
</table>

구분	주요 내용
	에이전트·자동화 플랫폼에서 경쟁력을 강화함. • **AI 플랫폼 팀** : 개발자용 AI 프레임워크·API·SDK를 제공하며, GitHub·Visual Studio·Azure DevOps와 연계된 통합 개발 생태계 구축을 담담하며, 최근 들어 코파일럿 스택 기반 개발자 환경(코드 작성·리뷰·배포·운영 자동화)을 강화 추세 • **AI & 리서치 그룹** : MS의 전사 AI 전략·제품화·윤리 정책을 총괄하며, 코파일럿 제품군 및 애저 AI 기능을 전체 제품군과 통합하는 역할을 하며 책임형 AI(RAI) 전담 조직과 함께 AI 안전·보안·데이터 거버넌스에 대한 글로벌 표준 준수 강화. • **OpenAI 협력팀** : 오픈AI에 대한 대규모 투자(약 130억 달러 이상) 및 애저 AI 슈퍼컴퓨팅 인프라 제공을 기반으로 전략적 파트너십 유지를 담당하며. GPT-4.1, DALL, E·Sora, o-series 모델 등을 애저 기반으로 공급
사업 전망	**〈AI 전략〉** • **(Copilot 중심의 AI 통합 전략 강화)** Microsoft 365 Copilot, Windows Copilot, GitHub Copilot, Security Copilot, Sales Copilot 등 제품군별 Copilot이 전사 AI 브랜드로 통합됐고, • 최근 코파일럿 에이전트 전략을 발표해 ▲업무 자동화 ▲시스템 간 연동 ▲장기 작업 시퀀스 수행 등 실행형 AI 기능을 강화. • **(Azure AI & AI 인프라 투자 확대)** 애저 오픈AI 서비스는 OpenAI 모델 공급 플랫폼으로 성장하고 있고, MS는 Nvidia·AMD 대비 자체 AI 칩 Maia(트레이닝용), Cobalt(CPU) 개발로 공급망 안정성과 비용 효율 확보를 추진 • 2025년 기준 Azure는 글로벌 기업용 AI 구축의 핵심 플랫폼 중 하나로 자리 잡으며, 클라우드+AI 매출 비중이 지속 확대 • **(OpenAI와의 파트너십 유지 및 모델 다변화)** GPT-4.1/GPT-o/GPT-Sora 등 최신 모델을 애저에 우선 통합해 기업시장 선점과 동시에 MS 자체 모델 강화를 통해 '고성능+저비용 모델 포트폴리오' 구축 전략을 추진 • 오픈AI 위기 대응을 위해 AI Safety, RAI(Responsible AI) 체계 강화도 병행 • **(모든 제품에 AI 내재화)** Windows 11은 코파일럿과 통합되어 생산성·검색·설정 제어 기능을 AI 기반으로 전환했고, Teams·Outlook·Power BI·Dynamics 365 등도 코파일럿 기능을 확대해 업무 기록 자동화·실시간 요약 기능을 제공. • **(엔터프라이즈 AI와 AI 보안 시장)** Security Copilot을 중심으로 AI 기반 침해 대응·보안 운영 자동화 기능 확대하며, 금융·의료·제조·공공 등 규제가 강한 산업군을 대상으로 Azure Confidential Computing·AI Governance 기능을 제공해 신뢰성 확보를 추진

2) 2024년 기업 동향

(1) 사업전략/투자/M&A

가. AI 투자 압박으로 6,000명 해고 결정[351]

미국 경제 매체 블룸버그(Bloomberg)가 2025년 5월 13일 마이크로소프트가 불필요한 관리

350) 마이크로소프트 재무제표 https://kr.investing.com/equities/microsoft-corp-income-statement (2025.12.23. 방문)
351) Bloomberg, Microsoft is cutting thousands of employees across the company, 2025.5.13.

계층 감축을 위해 전체 직원의 3%에 해당하는 약 6,000명의 직원 해고를 결정했다고 보도했다. 2024년 6월 기준 약 22만 8,000명 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 해고한 바 있다. MS는 최근 몇 년간 AI 서비스와 클라우드 컴퓨팅을 뒷받침하는 데이터센터에 막대한 투자를 단행하며 비용 절감 압박을 받아왔으며, 2025년 회계연도에도 서버클러스터에 약 800억 달러를 지출할 계획이다.

MS는 AI 투자를 확대하며 비용을 면밀히 검토해 예산 우선순위를 재조정하고 있으며, 최근 경영진은 데이터센터 건설에 대한 투자 증가에도 지출 억제를 약속했다. 블룸버그가 입수한 워싱턴주 정부 문서에 따르면 워싱턴주 레드먼드에 있는 MS 본사에서 해고 인원의 1/3에 해당하는 약 2,000명 감원이 예정되어 있고, 해고는 7월 13일부터 시작될 전망이다. 워싱턴주의 소프트웨어 엔지니어링 분야가 해고 통지를 받은 직종 중 40%에 달해 최대 비중을 차지하며, 제품 관리와 기술 프로그램 관리 직책은 전체의 30%에 해당하는 600명에 달한다.[352]

소프트웨어 엔지니어 외에도 감원된 MS 직원 중 상당수가 소프트웨어 프로젝트를 담당했다. 블룸버그는 소식통을 인용해 이번 감원은 AI 프로젝트에 배정된 일부 관리자와 근로자 대상으로도 이뤄졌으며, 영업과 마케팅 등 고객과 직접 접촉하는 인력은 상대적으로 영향이 적은 것으로 확인됐다고 밝혔다. 업계는 코드를 작성하거나 분석할 수 있는 AI 도구는 소프트웨어 개발 과정을 자동화하고 있으며, MS 역시 사내 여러 부서에 AI를 도입해 인건비 절감에 나선 것으로 평가한다. 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 언급한 바 있다. 빌 더프(Bill Duff) MS 재무 담당 임원도 2025년 5월 13일 JP모건 콘퍼런스에서 고객 지원에 AI를 활용해 연간 수억 달러를 절감하고 있으며, 여러 부서에서 AI를 도입해 거래분석, 마케팅 자료 분석 등 업무를 지원하고 있다고 공개 발언했다.

352) Bloomberg, Microsoft Layoffs Hit Coders Hardest With AI Costs on the Rise, 2025.5.15.

(2) 기술개발 및 혁신

가. AI 에이전트 상호운용성을 위해 앤스로픽 MCP 프로토콜 채택[353]

마이크로소프트를 비롯해 오픈AI, AWS 등 주요 AI 기업이 앤스로픽의 '모델 컨텍스트 프로토콜(Model Context Protocol, MCP)'을 연이어 도입해 MCP가 AI 에이전트와 외부 도구 간 상호운용성을 위한 산업 표준으로 급부상하고 있다. MCP는 앤스로픽이 2024년 11월 공개한 AI 모델과 다양한 데이터 소스(콘텐츠 저장소, 업무용 도구, 개발 환경 등)를 연결하는 개방형 프로토콜로, 앤스로픽은 MCP를 AI 애플리케이션의 USB-C로 비유한다. 기존에는 AI 시스템이 데이터를 불러오기 위해 데이터 소스마다 따로 연결해야 했으나, MCP는 AI 시스템과 데이터 소스를 양방향으로 연결하는 단일 프로토콜을 통해 간편하고 안정적인 접근성을 제공한다고 강조한다. 또한 MCP는 클라이언트-서버 아키텍처로 구성되며, AI 모델이 설치된 클라이언트와 특정 데이터 소스 및 서비스에 접근하는 MCP 서버 간 연결을 통해 AI 애플리케이션 구축을 지원한다. 외신은 마이크로소프트, 오픈AI, AWS 외에 중국 앤트 그룹(Ant Group), 알리바바 클라우드(Alibaba Cloud)도 MCP를 채택한 것으로 확인됐다고 보도했다.

마이크로소프트는 2025년 3월 19일 '코파일럿 스튜디오(Copilot Studio)'[354]에서 MCP 지원을 공식 발표하면서, MCP를 통해 외부 데이터 소스를 코파일럿 스튜디오에 쉽게 연결하고 AI 에이전트에 도구와 데이터를 유연하게 추가할 수 있다고 설명했다. 오픈AI의 샘 올트먼 CEO는 2025년 3월 27일 소셜 미디어 X 게시글[355]을 통해 제품 전반에 걸친 MCP 지원 계획을 밝히며, 에이전트 소프트웨어 개발 도구(SDK), 챗GPT 데스크톱 앱, Response API[356]에서 MCP를 지원하겠다고 언급했다. AWS는 2025년 4월 1일 개발자 지원을 위한 MCP 서버 모음을 출시하며 MCP 지원을 공식화하고, 각 MCP 서버는 AWS 서비스 관련 문서 검색, 코드 생성, 인프라 관리, 개발 업무 흐름 등을 지원한다.[357] 홍콩 미디어 SCMP[358] 는 2025년 4월 알리페이(Alipay) 운영사인 앤트 그룹이 결제 서비스용 MCP 서버를 출시했고, 알리바바 클라우드는 자사 AI 모델 플랫폼 모델스코프(ModelScope)를 통해 MCP 마켓플레이스를 출시했다고 보도했다.

---

353) Microsoft, Introducing Model Context Protocol in Copilot Studio: Simplified Integration with AI Apps and Agents, 2025.3.19.

354) 사용자 맞춤형 AI 에이전트 개발플랫폼

355) Sam Altman, "people love MCP and we are excited to add support across our products. available today in the agents SDK and support for chatgpt desktop app + responses api coming soon!", 2025.3.27.
https://x.com/sama/status/1904957253456941061

356) 웹 검색, 파일 검색, 컴퓨터 사용과 같은 기본 제공 도구를 지원하는 API

357) AWS, Introducing AWS MCP Servers for code assistants, 2025.4.1.

358) SCMP, Chinese tech giants race to expand AI services market with latest open-standard protocol, 2025.4.21.

나. 추론 향상을 위한 컴퓨팅 자원 투입 효과 분석[359]

마이크로소프트 리서치가 2025년 3월 31일 AI 모델의 추론 성능을 향상하기 위한 '추론 시간 확장' 기술[360]의 효과와 한계를 분석한 논문[361]을 공개했다. 논문은 기존 모델(GPT-4o, 클로드 3.5 소네트, 제미나이 2.0 프로, 라마 3.1 405B)과 추론 시간 스케일링을 통해 추론 성능을 높이도록 미세 조정된 모델(o1, o3-미니, 클로드 3.7 소네트, 제미나이 2.0 플래시 씽킹, 딥시크 R1)로 구성된 9개 모델에 대하여 다양한 분석 결과를 제시했다. 연구진은 각종 복잡도·성격 문제에 대해 ▲CoT(Chain-of-Thought)[362] ▲병렬 확장[363] ▲순차적 확장[364] 세 가지 추론 시간 확장 기법을 사용해 각 AI 모델을 평가했다고 밝혔다. 또한 수학, 자연과학, 일정 계획, 내비게이션, 공간 추론 등 8개 과제[365]에 대한 응답 정확도와 토큰 사용량을 핵심 평가 지표로 사용해 각 모델을 평가했다고 설명했다.

모델 평가 결과, 추론 성능 향상을 위해 미세 조정된 모델은 일반적으로 기존 모델 대비 우수한 성능을 기록했으나 영역과 과제에 따라 다른 개선 수준을 보였다. 과제 복잡성이 증가할수록 개선 효과는 감소해 추론 시간 확장 기법의 효과가 보편적이지 않음이 나타났다. 유사한 응답 정확도를 달성한 모델 간에도 토큰 사용량은 큰 차이가 있었고, 일례로 AIME 2025[366] 벤치마크 평가에서 딥시크-R1은 클로드 3.7 소네트보다 5배 이상 많은 토큰을 사용했으나 정확도는 각각 58.7%와 56.0%로 소폭의 차이만 기록했다.

특히 동일 모델에 프롬프트를 반복해 입력하더라도 토큰 사용량은 매번 달라지는 것으로 나타나 정확한 응답을 제공하는 모델이라도 토큰 사용량의 변동성으로 인해 개발자와 사용자에게 비용 불확정성을 불러올 수 있는 것으로 나타났다. 모델 답변을 정확하게 평가하는 '완전한 검증기(Perfect Verifier)'[367] 기반 지속 확장은 추론 특화 미세조정 모델과 기존 모델 양쪽에서 성능 개선 효과를 나타내 추론 성능 향상을 위해서는 컴퓨팅 자원 투입뿐 아니라 정확한 검증 체계도 중요하다는 것을 보여줬다.

---

359) Inference-Time Scaling for Complex Tasks: Where We Stand and What Lies Ahead, 2025.3.31.
360) Inference-time Scaling: AI 모델이 답변을 생성하는 추론 단계에서 더 많은 컴퓨팅 자원을 투입해 추론 성능을 향상하는 기술
361) Microsoft, Inference-Time Scaling for Complex Tasks: Where We Stand and What Lies Ahead, 2025.3.31. https://arxiv.org/html/2504.00294v1
362) 모델이 문제를 단계별로 해결하도록 유도하는 방식
363) 동일한 질문에 대해 여러 개의 답변을 생성하고 집계를 통해 최종 결과를 도출하는 방식
364) 모델이 반복적으로 답변을 생성하고 비평 모델의 피드백을 활용해 후속 시도에서 답변을 생성하는 방식
365) AIME, Omni-MATH, GPQA, BA-Calendar, 3SAT, TSP, Maze, SpatialMap
366) 2025년 미국 수학 올림피아드 예선 문제
367) 모델이 생성한 N개의 답변 중 가장 좋은 답변을 선택하는 메커니즘

(3) 제품 및 서비스

가. ‘MS 빌드 2025’에서 AI 에이전트 플랫폼과 신기능 공개[368)]

마이크로소프트가 2025년 5월 20일부터 22일까지 개최한 개발자 행사 ‘MS 빌드 2025(Microsoft Build 2025)’에서 AI 에이전트·개발자 도구 등 신규 기능과 주요 업데이트를 발표했다. MS는 스스로 학습·결정을 내리는 AI 에이전트가 산업 전반에서 작동하는 인터넷 환경을 ‘개방형 에이전틱 웹(Open Agentic Web)’으로 정의하고, AI가 사용자나 조직을 대신해 결정을 내리고 작업을 수행하는 시대의 도래를 선언했다. MS는 깃허브 코파일럿(GitHub Copilot)[369)]에 코딩 에이전트를 새로 도입하는 등, 주요 개발플랫폼에서 활용할 수 있는 AI 기능의 업데이트를 발표했다.

또한 개발자가 모델 API를 활용해 오픈소스 LLM을 로컬 환경에서 실행하거나 자체 모델을 가져와 변환·미세 조정해 배포할 수 있는 ‘윈도 AI 파운드리(Window AI Foundry)’도 새로 공개했다. 이와 함께 AI 에이전트를 더욱 안전하고 효율적으로 개발하고 배포할 수 있도록 지원하는 사전 구축 에이전트와 맞춤형 에이전트 설계 도구, 멀티에이전트 기능도 발표했다. ‘애저 AI 파운드리(Azure AI Foundry)’[370)]에 여러 전문 에이전트를 조율해 복잡한 작업을 처리할 수 있도록 지원하는 에이전트 서비스와 AI 에이전트의 신뢰도를 높일 수 있도록 성능·품질·안전성 등 지표 모니터링 기능(Observability)을 추가하고 모델 업데이트로 xAI의 ‘그록 3’ 모델도 추가했다.

기업 자체 데이터와 업무 프로세스를 기반으로 AI 모델 학습과 에이전트 생성을 돕는 ‘MS 365 코파일럿 튜닝’ 기능과 코파일럿 스튜디오[371)]에서 다양한 에이전트를 상호 연결하는 ‘멀티에이전트 오케스트레이션’ 기능도 공개했다. 개방형 AI 에이전트 생태계 확장을 위해 MS 서비스 전반에서 MCP를 폭넓게 지원하는 한편, 새로운 개방형 프로젝트 ‘NLWeb’도 발표했다. NLWeb[372)]은 웹사이트 운영자가 AI 모델과 자체 데이터를 연결해 대화형 인터페이스를 구축함으로써 사용자와 웹 콘텐츠의 상호작용을 지원하는 개방형 프로젝트이다. 개발자들이 AI 에이전트를 활용해 과학적 발견을 앞당길 수 있도록 돕는 ‘마이크로소프트 디스커버리’도 공개하고, 이를 통해 제약, 환경 등 다양한 산업 분야에서 신제품 출시 기간 단축, 연구 속도·적용 범위 확장이 가능하다고 강조했다.

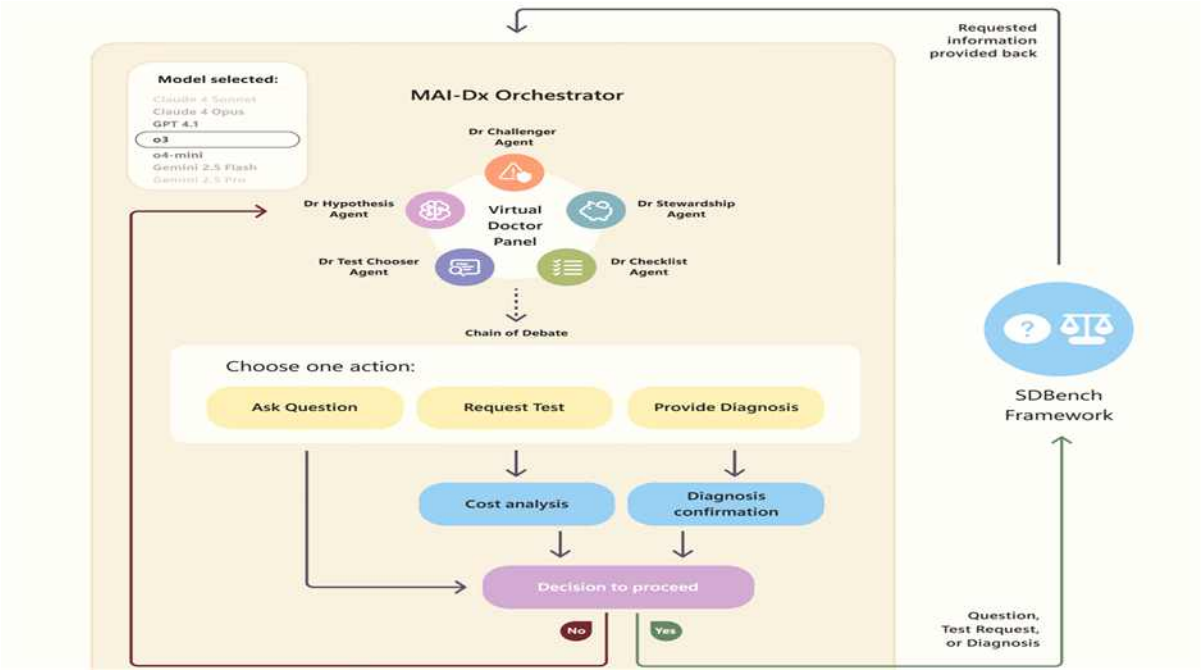
---

368) Microsoft, 마이크로소프트 빌드 2025 개막… “오픈 에이전틱 웹 시대 앞당긴다.”, 2025.5.20.
369) 깃허브에서 코드를 더 효율적으로 작성할 수 있도록 지원하는 AI 기반 코드 작성 도우미
370) 개발자가 AI 애플리케이션과 에이전트를 설계·맞춤화·관리할 수 있도록 지원하는 통합 플랫폼
371) 맞춤형 AI 에이전트를 만들고 관리할 수 있는 대화형 AI 플랫폼
372) GitHUB에 공개된 NLWeb https://github.com/microsoft/NLWeb

나. 마이크로소프트, AI 의료 진단 시스템 'MAI-DxO' 공개[373]

마이크로소프트가 2025년 7월 2일 의사보다 4배 이상 정확하게 질병을 진단하는 AI 의료 진단 시스템 'MAI-DxO(Microsoft AI Diagnostic Orchestrator)'를 공개했다. MS는 ▲가설 수립 ▲검사 선택 ▲반론 제기 ▲비용 관리 ▲품질관리를 담당하는 5명의 가상 의사 패널을 구성해 복잡한 진단 사례를 협력하여 해결하는 방식의 MAI-DxO를 제시했다. 가설 수립 AI가 상위 진단을 내리면 검사 선택 AI가 검사를 추천하고 반론 제기 AI는 진단 번복 가능성을 검토하며, 비용 관리 AI는 경제적 대안을 선택하고 품질관리 AI는 전체 추론의 일관성을 유지하는 방식이다.

[그림 5-7] 마이크로소프트의 'MAI-DxO' 개요



출처 : rXiv, Sequential Diagnosis with Language Models

연구진이 인간 의사나 AI 에이전트의 진단 정확도를 평가하는 벤치마크인 SDBench[374]를 개발해 MAI-DxO 진단 성능을 평가한 결과, 비용-정확도 간 균형에서 인간 의사[375]와 기존 LLM을 능가하는 것으로 나타났다.[376] 평균 12년의 경력을 보유한 의사 집단이 건당 평균 2,963달러의 비용으로 20% 정확도를 보였지만, MAI-DxO는 비용에 구애받지 않고 최대 정확도로 설정

373) a, 2025.7.2.
374) 뉴잉글랜드의학저널(NEJM)에서 수집한 진단 사례 304건을 단계별 진단 시뮬레이션으로 변환해 구축
375) 연구에 참여한 의사들은 동료와 논의나 검색엔진, 교과서 등 도구 활용을 제한한 상태로 평가
376) Time, Microsoft's AI Is Better Than Doctors at Diagnosing Disease, 2025.7.2.

시 85.5%의 정확도를 달성(비용은 7,184달러)했다.[377] 기존 LLM의 경우, 모델별로 각기 다른 결과를 나타내 GPT-4o는 건당 2,745달러에 49.3%의 정확도를 보였지만, o3는 건당 7,850달러에 78.6%의 정확도를 기록했다. 이러한 테스트 결과에도 불구하고 의료계는 의료 AI 시스템이 실제 임상 환경에서 어떻게 작동할지는 불확실하며, AI 기반 의료 의사결정으로 인한 환자의 피해를 방지하기 위한 규제기관의 승인이 필요하다고 지적했다.

(4) AI 윤리 및 규제 대응

가. 마이크로소프트, 교육 분야의 AI 활용 현황 보고서 발간[378]

마이크로소프트의 '2025년 교육 분야 AI(2025 AI in Education)' 보고서를 발표하고 교육 분야의 AI 활용이 급증하고 있다고 진단했다. 보고서는 미국・브라질・사우디아라비아・영국・일본・호주의 교육자와 교육기관 최고 관리자를 대상으로 진행한 설문조사 결과를 분석했다.[379] 보고서에 따르면 현재 교육기관 86%가 생성 AI를 사용하고 있으며, 이는 전체 산업 분야 중 가장 높은 비율이다. 미국에서는 학교에서 AI를 활용하는 학생이 2024년 대비 26%p 늘었고, 교육자의 AI 활용은 전년 대비 21%p 증가했다. 교육기관들은 K-12(초중고) 교실의 개인화 학습부터 대학의 AI 기반 관리 도구 등 효율성과 학생 참여도, 성과 개선을 위해 AI를 빠르게 통합하고 있으나, AI 교육은 AI 도입 속도를 따라가지 못하는 상태로 나타났다.

미국 학생과 전 세계 교육자 중 AI에 대해 많이 알고 있다는 응답은 절반 이하를 기록했으며, 전 세계 교육자의 45%와 미국 학생의 52%는 아무런 AI 교육을 받은 적이 없다고 응답했다. 그러나 전 세계 교육기관 관리자의 76%는 AI 사용자(학생・교육자) 중 절반 이상이 AI 교육을 받았다고 응답해, 학생・교육자와 교육기관 리더 간 인식 차이가 크게 나타났다. 학생과 교육자, 교육 관리자들은 창의적 사고와 협력을 지원하는 파트너로 AI를 활용한다.

학생들은 주로 과제 구상(37%), 정보 요약(33%), 신속한 정보 획득(33%), 피드백 수렴(32%)에 AI를 사용하고 교육자는 수업 계획이나 자료 구상(31%), 수업 자료나 과제 제작 또는 업데이트(29%), 학생을 위한 복잡한 주제 단순화(24%), 실시간 성과 데이터와 예측 분석(24%) 등에 주로 AI를 활용하는 것으로 나타났다. 교육기관 관리자들은 주로 과제 피드백과 개선 안내(36%), 운

377) MAI-DxO에 예산 제약을 추가 시에는 2,396달러의 비용으로 79.9%의 정확도를 달성
378) Microsoft, AI in Education Report: Insights to support teaching and learning, 2025.8.20.
379) 미국에서는 설문조사 대상자에 16~18세 학생과 대학생, IT 책임자도 포함

영 간소화(35%), 접근성 도구 제공(33%), 학생 성장 기회 파악(33%), 학생 지원 및 소통 능력 향상(30%)에 AI를 활용한다고 응답했다.

MS는 교육 분야에서 AI 활용이 증가하고 있지만, AI로 기존 학습 방식을 대체하는 것이 아니라 보완하는 방향으로 AI를 사용할 때 더 효과적이라고 강조했다. 한편, MS 리서치와 케임브리지 대학교 출판평가원(Cambridge University Press & Assessment) 공동 연구[380]에서 AI 도구와 노트 필기 및 기타 방식을 결합한 학생들은 AI에만 의존한 학생보다 더 뛰어난 학습 효과를 보였다.

나. 오픈AI 투자 관련 영국 CMA 반독점 조사 종료[381]

영국 경쟁시장청(CMA)이 2025년 3월 5일 마이크로소프트와 오픈AI 간 협력 관계에 대한 반독점 조사 종료를 발표했다.[382] 마이크로소프트는 오픈AI의 최대 투자자로서, 2019년 오픈AI와 파트너십을 맺고 2024년까지 130억 달러 이상을 투자했다. 이러한 협력 관계에 대해 CMA는 2023년 12월 양사의 파트너십이 영국 내 시장 경쟁을 저해할 수 있는 사실상의 합병으로 보아야 하는지에 대한 조사에 착수했다. CMA는 14개월에 걸친 조사 결과, 마이크로소프트가 오픈AI에 대한 실질적인 영향력을 확보했지만, 지배권 변동이 없으므로 인수합병 규정[383]에 따른 전면 조사의 대상이 되지 않는다는 결론을 도출했다.

CMA 조엘 뱀포드(Joel Bamford) 합병 담당 이사는 2025년 3월 5일 소셜 미디어 플랫폼 링크드인 게시글[384]에서 마이크로소프트가 오픈AI에 실질적 지배권을 행사하지 않아 현 형태의 파트너십은 영국 합병 통제 제도에 따른 검토 대상이 되지 않는다고 밝혔다. 뱀포드 이사는 마이크로소프트가 2025년 1월 오픈AI에 대한 컴퓨팅 용량 공급 관련 계약상 권리를 독점 공급에서 우선협상권으로 변경하면서 오픈AI의 의존도가 감소한 점도 언급하고 이번 조사는 복잡한 파트너십과 계약 조건의 변화·실제 운영 방식 등을 검토하는 데 이례적으로 긴 기간이 걸렸다고 설명했다.

---

380) AI and Traditional Learning: Complementary Strategies for Deeper Learning
381) Competition and Markets Authority, Microsoft / OpenAI partnership merger inquiry, 2025.3.5.
382) The Wall Street Journal, Microsoft-OpenAI Partnership Gets U.K. Antitrust Clearance, 2025.3.5.
383) 영국 기업법(Enterprise Act 2002) 제22조(완결된 합병과 관련된 심층 조사 회부 의무)
384) What are the key takeaways from our review of the Microsoft/OpenAI partnership?, Joel Bamford, 2025.3.5.

CMA가 마이크로소프트와 오픈AI 간 파트너십을 승인했지만, 여전히 EU·미국 등 주요국 규제당국은 양사 관계에 주목하고 있다. EU 집행위원회는 2024년 양사의 파트너십에 대한 정식 조사는 불필요하다는 결론을 내렸으나 양사의 관계를 계속 주시하는 결정을 내렸다. 또한 미국 연방거래위원회(FCC)는 2024년 양사를 포함한 주요 AI 기업들의 생성 AI 파트너십에 대한 조사에 착수한 상황이다. 이러한 상황에서 마이크로소프트는 2024년 오픈AI 이사회의 옵저버 역할을 포기하는 등 오픈AI에 대한 영향력을 축소해 규제 논란 해소에 주력하고 있다.

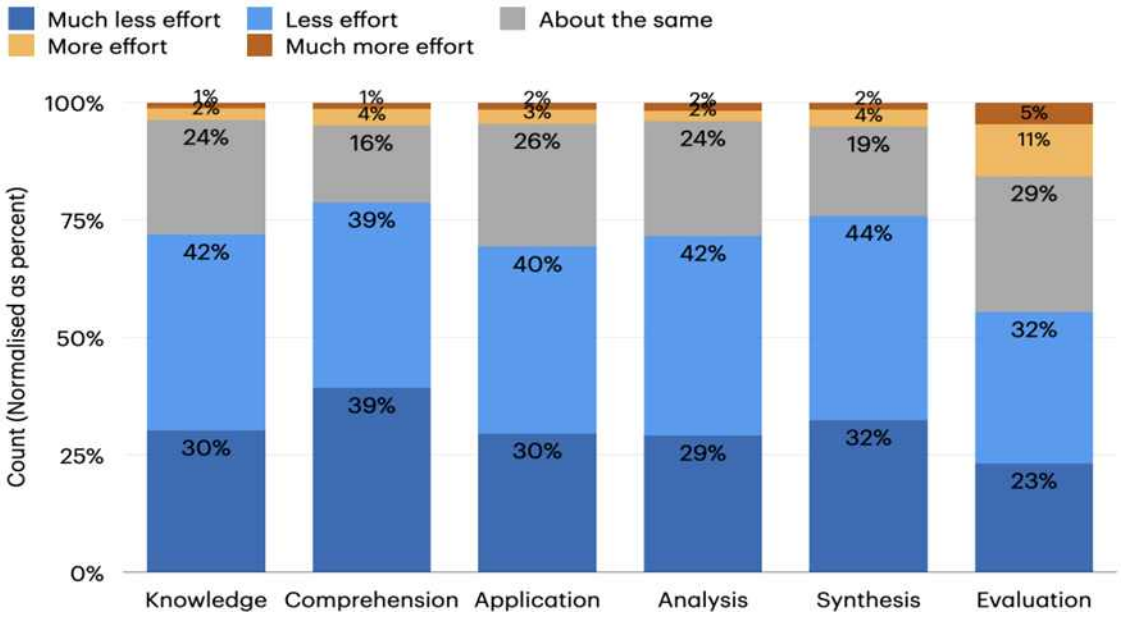
다. 마이크로소프트 연구 결과, 생성 AI 신뢰할수록 비판적 사고 감소[385]

마이크로소프트와 카네기멜런대학교 연구진이 2025년 5월 생성 AI가 지식 근로자의 비판적 사고에 미치는 영향 관련 연구 결과를 공개했다. 생성 AI를 주 1회 이상 업무에 사용하는 지식 근로자 319명을 대상으로 한 설문조사 결과, 생성 AI가 업무 효율성을 향상할 수 있지만 장기적으로는 독립적 문제 해결 능력을 저해할 위험이 존재하는 것으로 나타났다. 참가자들이 공유한 936개 실제 업무 사례를 분석한 결과, 대부분 사례에서 생성 AI를 사용하는 경우, 비판적 사고와 관련된 인지 활동이 감소했다.

특히 인지 활동 중 ▲지식(Knowledge) 72% ▲이해(Comprehension) 79% ▲적용(Application) 69% ▲분석(Analysis) 72% ▲종합(Synthesis) 76% ▲평가(Evaluation) 55%가 ‘훨씬 적은 노력’ 또는 ‘적은 노력’을 들였다고 답했다. 생성 AI를 신뢰하는 근로자일수록 비판적 사고가 줄어드는 반면, 업무 능력에 대한 자신감이 높은 근로자는 비판적 사고를 더 많이 하는 것으로 나타났다. 연구진은 사용자가 AI 한계를 인식하고 비판적 사고를 통해 AI가 생성한 결과물을 평가하고 개선할 수 있도록 생성 AI 도구를 설계해야 한다고 지적했다. 특히 일상 업무를 과도하게 자동화하면 근로자가 비판적 사고를 연습할 수 없어 법률 문서 작성과 같은 고위험 상황에서 사고 능력이 저하될 수 있다고 경고했다.

385) Microsoft, The Impact of Generative AI on Critical Thinking: Self-Reported Reductions in Cognitive Effort and Confidence Effects From a Survey of Knowledge Workers, 2025.1.

[그림 5-8] 생성 AI 도구 사용 시 인지 활동에 들이는 노력 수준



출처 : Microsoft, The Impact of Generative AI on Critical Thinking

3. 메타 플랫폼스

1) 기업 일반현황

<표 5-3> 메타 플랫폼즈 기업 현황

<table>
<tr><th>구분</th><th>주요 내용</th></tr>
<tr><td rowspan="3">기업 개요</td><td>〈회사 개요〉
• 2004년 2월 미국 매사추세츠주에서 ‘페이스북(Facebook)’으로 설립됐고, 2012년 5월 나스닥 상장 후 세계 최대 소셜 미디어 기업으로 성장

• 2021년 10월 메타버스 중심 전략을 선언하며 사명을 메타 플랫폼즈(Meta Platforms)로 변경했고, ‘페이스북’은 서비스명으로 유지
• 최근 들어 메타버스 단일 전략에서 벗어나 AI 강화 전략을 병행하며, 생성 AI, AI 에이전트, 경량 LLM, AR 글라스 등으로 사업을 다각화
• 2022년 R&D 지출 353억 달러(전 세계 3위), 이후 2023~2024년에도 대규모 AI·AR·칩 개발 투자가 지속되며 기술 중심 기업으로서 정체성을 강화
• 2024년 광고 회복·AI 추천 개선·비용 구조조정 효과로 실적이 반등하며 기업 가치가 재상승 중.</td></tr>
<tr><td>〈사업 내용〉
• 산업 분야 : 소셜 미디어·소셜 네트워킹, 사진·영상 공유, 텍스트 기반 소셜, 인스턴트 메시징 및 VoIP, AR/VR 디바이스, AI 플랫폼, 광고 비즈니스.

• 주요 서비스: Facebook, Instagram, Threads, WhatsApp, Messenger, Meta Quest, Horizon Worlds, Ray-Ban Meta Smart Glasses(2세대), Mapillary, Meta AI 등
• 2024~2025년에 모든 서비스에 메타 AI 기능 통합을 추진하며, 인스타그램, 왓츠앱, 메신저, 레이반(Ray-Ban) 글라스 등에서 AI 에이전트 기능을 강화</td></tr>
<tr><td>〈관계사〉
• Reality Labs(오큘러스/VR·AR 사업부), Mapillary, CTRL-Labs, Onavo(데이터 분석), Jio Platforms(파트너십), Novi Financial(폐지 수순을 보이지만 블록체인 기술 연구 일부 유지). Meta Quest·Ray-Ban Meta 글라스·Horizon 등 Reality Labs는 여전히 메타버스·AR 중심의 핵심기술 조직으로 유지되며, 최근 AI 디바이스 기업으로 전환이 가속</td></tr>
<tr><td>재무 실적386)
(최근 5년)</td><td>
<table>
<tr><th>마감 기준:</th><th>2019
31/12</th><th>2020
31/12</th><th>2021
31/12</th><th>2022
31/12</th><th>2023
31/12</th><th>2024
31/12</th></tr>
<tr><td>총매출</td><td></td><td>85,965</td><td>117,929</td><td>116,609</td><td>134,902</td><td>164,501</td></tr>
<tr><td>총매출 성장률</td><td></td><td>+21.6%</td><td>+37.18%</td><td>-1.12%</td><td>+15.69%</td><td>+21.94%</td></tr>
<tr><td>매출 원가</td><td></td><td>16,692</td><td>22,649</td><td>23,754</td><td>26,006</td><td>30,130</td></tr>
<tr><td>총이익</td><td></td><td>69,273</td><td>95,280</td><td>92,855</td><td>108,896</td><td>134,371</td></tr>
<tr><td>총이익 성장률</td><td></td><td>+19.59%</td><td>+37.54%</td><td>-2.55%</td><td>+17.28%</td><td>+23.39%</td></tr>
<tr><td>총이익 %</td><td></td><td>80.58%</td><td>80.79%</td><td>79.63%</td><td>80.72%</td><td>81.68%</td></tr>
</table>
</td></tr>
</table>

구분	주요 내용
AI 조직 현황	〈AI 관련 조직〉 • **메타 AI(구 FAIR-Facebook AI Research)** : 메타의 핵심 기초 연구조직으로 AI 모델·멀티모달 연구·효율적 학습·데이터 최적화 등 기초 기술 전반을 담당 • **AI Research & Development** : AI 모델을 제품 서비스에 통합하는 응용 연구조직으로 콘텐츠 추천·광고 최적화·안전성 검증·AI 중독 방지 모델 등을 진행 • **Applied AI** : 추천 시스템, 광고 엔진, 콘텐츠 모더레이션, 악성 콘텐츠 탐지, 자동 번역 등 실제 운영 서비스에 AI 기술을 적용하는 조직 • **리얼리티랩스(Reality Labs)** : AR/VR 기기·센서·AI 인터랙션 기술을 담당하며 Meta Quest·Ray-Ban Meta Glasses·Horizon Worlds 관련 기술을 개발 • 동 조직은 여전히 적자를 기록하고 있지만, 메타의 장기 AI·AR 하드웨어 전략의 핵심 조직으로 유지 • 메타 AI 조직은 기초 연구(FAIR)-제품화(Applied AI)-플랫폼(AI R&D)- AR/VR(RL)로 이어지는 수직형 구조를 이루며 LLaMA 기반 오픈소스 전략과 AI 디바이스 전략이 조직 전반에 반영
사업 전망	〈AI 전략〉 • **(오픈소스 중심 전략 : LLaMA 시리즈 강화):** LLaMA 2 공개 이후 LLaMA 3·3.1·4(개발중) 로드맵 발표로 오픈소스 LLM 생태계의 글로벌 표준을 주도. • **(AI 디바이스 전략 : Ray-Ban Meta Smart Glasses 확대)** 메타버스→AI 하드웨어 기업으로의 전략적 전환이 진행 중이며 Ray-Ban Meta Glasses에 LLaMA 기반 음성·시각 AI 모델을 탑재 • **(소셜 미디어 AI 고도화)** 인스타그램, 페이스북, 릴스에서 '추천 알고리즘 기반 성장'이 나타났고, 릴스 AI 추천은 틱톡과 경쟁하면서 메타의 핵심 성장축으로 작용 • **(AI 인프라 확충)** AI Research Super Cluster(RSC)를 기반으로 대규모 모델 학습 인프라를 구축해 LLaMA 시리즈 성공 기반을 마련했고, 연이어 엔비디아 H100·H200·B200 기반 클러스터 증설과 자체 AI 칩(Artemis) 개발도 추진

2) 2024년 기업 동향

(1) 사업전략/투자/M&A

가. 스케일 AI CEO를 비롯한 AI 인재 영입 노력 본격화[387]

메타가 마크 저커버그(Mark Zuckerberg) CEO 주도하에 핵심 AI 연구자들에 200만~1천만 달러 이상의 연봉을 제시하며 AI 인재 영입 노력을 강화하고 있다. 저커버그 CEO는 최고의 AI 연구자들로 구성된 '슈퍼인텔리전스 팀'이라는 일반인공지능(AGI) 연구팀 구성을 계획하고, 유능한 AI 연구자들에게 직접 채용 제안을 보내는 등 인재 채용 노력에 적극적으로 참여하고 있는 것으로 알려졌다. 이러한 메타의 움직임은 2025년 4월 출시한 '라마(Llama)

386) 메타 재무제표 https://kr.investing.com/equities/facebook-inc-income-statement (2025.12.23. 방문)
387) The Register, Meta offered one AI researcher at least $10,000,000 to join up, 2025.6.13.

4’ 가 기대보다 부진한 성과를 보여 AI 경쟁에서 오픈AI, 앤스로픽, 구글 등 경쟁사에 뒤처질 수 있다는 위기감 때문으로 평가된다. 메타로부터 영입 제안을 받았던 X AI 책임자 출신의 유디안 정(Yudian Zheng)에 따르면 높은 연봉 수준은 생성 AI 인재를 둘러싼 치열한 경쟁을 반영하는 것으로, 주요 AI 기업들이 모두 인재 채용에 나서고 있지만 실제로 파운데이션 모델 개발 경험을 갖춘 인재 집단은 극소수에 불과한 실정이다.

메타가 2025년 6월 13일 데이터 라벨링 기업 스케일 AI(Scale AI)에 약 150억 달러를 투자해 알렉산드르 왕(Alexandr Wang) 창업자 겸 CEO 영입도 치열한 인재 확보 노력의 일환으로 평가된다. 메타는 해당 투자를 통해 스케일 AI의 지분 49%를 보유하게 됐고, 알렉산드르 왕은 스케일 AI CEO에서 물러나 새로 설립되는 메타의 슈퍼인텔리전스 팀에 합류할 예정이다. 하지만, 벤처캐피털 기업 시그널파이어(SignalFire)가 2025년 5월 20일 발표한 인재 현황 보고서[388]는 메타가 AI 인재 경쟁에서 상대적으로 뒤처진 것으로 평가했다. 보고서에 따르면 앤스로픽은 2021~2023년 사이 채용된 AI 인력의 직원 유지율이 80%에 달하며, 구글 딥마인드는 78%, 오픈AI는 67%를 기록했고 메타는 64%로 나타났다. 빅테크 가운데 메타의 AI 인력 이직률은 2024년 기준 4.3%에 달해 구글(5.4%)에 이어 두 번째로 높게 나타났고, 상대적으로 마이크로소프트(3.2%), 아마존(2.7%), 애플(1.7%)은 안정적인 모습을 보였다. 보고서는 메타를 비롯한 빅테크가 고액 연봉과 브랜드 인지도에 의존하는 반면, 앤스로픽은 포용적이고 자유로운 문화를 바탕으로 딱딱한 관료주의에 지친 빅테크 출신 인재 영입에서 강점을 보인다고 분석했다.

(2) 기술 개발 및 혁신

가. AI 개발자 콘퍼런스 ‘라마콘’에서 라마 API와 메타 AI 앱 공개[389]

메타가 2025년 4월 29일 AI 개발자 콘퍼런스 ‘라마콘(LlamaCon)’을 처음 개최하고 라마 API(Llama API)를 제한된 프리뷰 형태로 출시했다. 라마 API는 다양한 라마 모델 기반의 애플리케이션 제작을 지원하며, 프로그래밍 언어 파이선(Python)과 타입스크립트(TypeScript) 기반 소프트웨어 개발 도구(SDK), 미세조정 및 평가 도구를 제공한다. 우선 라마 3.3 8B 모델에서 사용자 지정 버전의 조정 및 품질 테스트 지원에 활용될 예정이다. 메타는 개발자 피드백을 받아 라마 API를 개선하고 향후 몇 주, 몇 달 안에 라마 API의 접근성을 확대할

388) The SignalFire State of Talent Report 2025
https://www.signalfire.com/blog/signalfire-state-of-talent-report-2025
389) Meta, Everything we announced at our first-ever LlamaCon, 2025.4.29.

계획이다.

메타는 AI 모델 학습에 사용자 프롬프트나 모델 응답을 사용하지 않으며, 라마 API에서 구축된 모델은 자사 서버에만 저장되는 것이 아니라 어디에서나 호스팅할 수 있는 점을 강조했다. 또한 AI 칩 전문기업 세레브라스(Cerebras), 그록(Groq)과 협력해 라마 API에서 개발자 요청에 따른 실험적 접근 허용을 통해 더 빠른 추론 속도를 지원한다고 발표했다.[390] 세레브라스는 기존 GPU 대비 최대 18배의 속도로 토큰을 출력하여 라마 4 스카우트에 적용 시 초당 2,600 토큰 이상의 출력 속도를 보였다. 하지만 챗GPT는 초당 약 130 토큰, 딥시크는 초당 약 25 토큰 수준을 기록했다.

메타는 신규 라마 보호 도구 3종[391]과 AI 기반 사이버보안을 위한 '라마 디펜더스 프로그램(Llama Defenders Program)'도 공개했다. 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 제공한다.[392] 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 지원한다. 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 제공한다.

(3) 제품 및 서비스

가. 멀티모달 AI 모델 '라마 4' 제품군 공개 및 성능 조작 의혹 부인[393]

메타(Meta)가 2025년 4월 5일 개방형 AI 모델 '라마(Llama) 4' 제품군 중 라마 4 스카우트(Scout)와 라마 4 매버릭(Maverick)을 출시하고, 라마 4 베히모스(Behemoth)는 프리뷰로 공개했다. 라마 4는 라마 시리즈 최초의 전문가 혼합(MoE)[394] 모델로 설계됐고, 텍스트·이

390) Cerebras, Meta Collaborates with Cerebras to Drive Fast Inference for Developers in New Llama API, 2025.4.29.
391) 위험한 출력 생성을 방지하는 'Lama Guard 4', AI 시스템 위험을 탐지하는 'LlamaFirewall', 악의적 프롬프트를 감지하는 'Llama Prompt Guard 2'
392) Meta, Introducing the Meta AI App: A New Way to Access Your AI Assistant, 2025.4.29.
393) Meta, The Llama 4 herd: The beginning of a new era of natively multimodal AI innovation, 2025.4.5.
394) Mixture-of-Experts: 특정 작업이나 영역에 특화된 여러 개의 전문가 모델을 혼합해 활용하는 방식

미지·비디오를 함께 처리할 수 있는 멀티모달 기능을 기본 탑재했다. 라마 4 스카우트는 활성 매개변수 170억 개에 16개의 전문가 모델로 구성되고 라마 3의 12만 8천 개에서 대폭 늘어난 1,000만 개 토큰의 컨텍스트 창을 지원하며, 코딩과 추론, 이미지 이해 관련 벤치마크 평가에서 유사 크기의 경쟁 모델보다 우수한 성능을 기록했다.[395]

메타는 라마 4 매버릭이 등급 최고의 멀티모달 모델로 활성 매개변수 170억 개에 128개의 전문가 모델로 구성되며, 주요 벤치마크 평가에서 GPT-4o와 제미나이 2.0 플래시를 앞서는 한편, 라마 3.3 70B 대비 더 낮은 비용으로 고품질의 성능을, 코딩 및 추론 분야에서 딥시크 v3.1과 유사한 성능 기록했다고 강조했다.[396] 훈련 단계의 라마 4 베히모스는 스카우트와 매버릭 모델 종류에 사용된 교사 모델이자 2,880억 개의 활성 매개변수에 16개의 전문가 모델로 구성된 최대 크기의 모델이다. 메타는 수학과 다국어, 이미지 이해 벤치마크 평가에서 비추론 모델 중 최고의 성능을 제공했다고 설명했다.[397]

하지만 라마 4 공개 후 일부 AI 연구자들은 메타가 LM 아레나[398]의 성능평가를 진행한 실험 버전과 실제로 제공되는 라마 4 매버릭 간 뚜렷한 성능 차이가 드러났다는 의혹을 제기했다.[399] 메타가 벤치마크 테스트 셋을 모델 학습에 사용해 성능을 조작했다는 주장이지만, 메타는 이를 부인했다. 메타 아흐마드 알달레(Ahmad Al-Dahle) 생성 AI 부문 사장은 X 게시글[400]을 통해 해당 의혹은 전혀 사실이 아니라고 반박하고, 모델이 준비되자마자 공개해서 기능 안정화에는 시간이 필요하다며 강조하고 지속적인 오류수정과 기술지원을 강조했다.

나. 물리적 세계를 이해하고 예측하는 AI 모델 'V-JEPA 2' 개발[401]

메타가 2025년 6월 11일 물리적 세계에 대한 시각적 이해와 예측을 위해 비디오로 훈련된 월드 모델(World Model)[402]인 매개변수 12억 개의 'V-JEPA 2'를 공개했다. 메타는 물리적

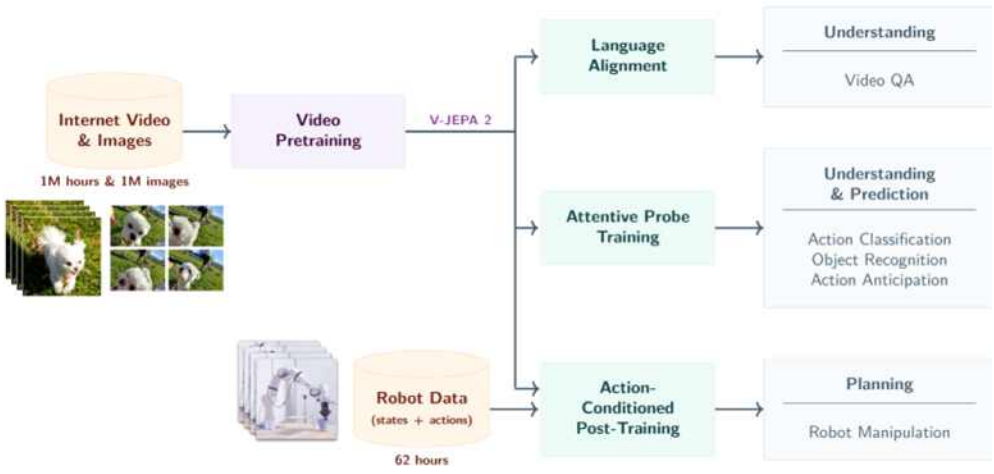
395) ChartQA(이미지 이해) 기준 라마 4 스카우트(instruction-tuned): 88.8점, 젬마 3 27B: 76.3점, 제미나이 2.0 플래시-라이트: 73.0점
396) 토큰 100만 개당 입출력 비용 기준 라마 4 매버릭: $0.19~0.49, 제미나이 2.0 플래시: $0.17, 딥시크 v3.1: $0.48, GPT-4o: $4.38
397) MATH-500(수학) 기준 라마 4 베히모스(instruction-tuned): 95.0점, 클로드 소네트 3.7: 82.2점, 제미나이 2.0 프로: 91.8점
398) 실제 사용자의 선호도를 기준으로 AI 모델 성능을 평가하는 오픈소스 플랫폼
399) TechCrunch, Meta exec denies the company artificially boosted Llama 4's benchmark scores, 2025.4.7.
400) Ahmad Al-Dahle, We're glad to start getting Llama 4 in all your hands. We're already hearing lots of great results people are getting with these models, 2025.4.8. https://x.com/Ahmad_Al_Dahle/status/1909302532306092107
401) Meta, Introducing the V-JEPA 2 world model and new benchmarks for physical reasoning, 2025.6.11.

세계에서 월드 모델을 통한 AI 에이전트의 계획 수립과 추론 능력 달성을 장기 비전으로 삼고, 이를 실현하기 위한 다음 단계로서 비디오 기반의 자기 지도학습 방식을 적용한 V-JEPA 2 개발을 추진해 왔다.

V-JEPA 2 아키텍처는 원본 영상을 받아 유용한 의미 정보가 담긴 임베딩을 출력하는 인코더(Encoder)와 인코더가 제공한 임베딩[403]을 토대로 영상에서 가려진 부분의 임베딩을 예측 출력하는 예측기(Predictor)로 구성됐다. V-JEPA 2는 100만 시간 이상의 비디오와 100만 개 이상 이미지 사전학습으로 예측 능력을 높였고, 두 번째 학습 단계에서 로봇 데이터를 활용해 에이전트가 취할 특정 동작을 학습했다.

V-JEPA 2는 이러한 2단계 학습 방식으로 로봇이 낯선 물체나 환경과 상호작용을 통해 작업을 완료할 수 있게 하는 세계 이해와 예측 능력을 높였다. V-JEPA 2가 적용된 로봇은 관찰된 현재 상태를 기반으로 예측기를 사용해 후보 행동을 선택하고 결과 예측으로 다음 계획을 수립하며, 물건을 집어 제자리에 놓는 것과 같은 복잡한 작업의 경우, 순차적으로 달성할 일련의 시각적 하위 목표를 지정해 새로운 환경에서도 높은 작업 성공률을 기록했다.[404]

[그림 5-9] V-JEPA 2의 2단계 학습 방식



출처: Meta, Introducing the V-JEPA 2 world model and new benchmarks for physical reasoning

402) 세계에 대한 관찰을 통해 작동 원리를 이해하여 미래를 예측하고 행동을 계획할 수 있는 AI 모델
403) Embedding: 컴퓨터가 이해하고 처리하기 쉬운 형태로 변환된 데이터의 수치적 표현
404) 처음 보는 물체를 집어 올바른 위치에 놓은 실험에서 65~80%의 성공률을 기록

다. 자기 지도학습 방식의 컴퓨터비전 모델 'DINOv3' 공개[405]

메타가 2025년 8월 14일 인간의 감독 없이 AI가 스스로 학습하는 자기 지도학습(Self-supervised Learning, SSL) 방식으로 훈련된 컴퓨터비전 모델 'DINOv3'를 공개했다. 기존 이미지 인코딩 모델은 인간이 데이터에 주석을 붙이는 수동 라벨링에 의존하지만, DINOv3는 SSL 기법을 활용해 학습에 필요한 시간과 비용을 대폭 절감하면서 학습 데이터와 모델 크기를 확장이 가능하다고 강조했다.[406] 이와 함께 컴퓨터비전과 멀티모달 생태계 혁신과 발전을 지원하기 위해 상용 라이선스로 DINOv3 훈련 코드와 사전훈련 백본[407]을 깃허브(GitHub)에 공개했다.

SSL 기법을 활용하는 DINOv3의 접근 방식은 데이터 규모가 방대하고 복잡하여 수동 라벨링이 불가능한 위성이나 항공 이미지 등의 분야에서 특히 유용한다는 설명이다. DINOv3를 활용하면 풍부한 데이터셋을 학습한 단일 백본을 다양한 위성 유형에 적용하여 환경 모니터링, 도시계획, 재난 대응 등 일반적인 응용 분야로 확장이 가능하다고 강조했다. 실제로 세계자원연구소(WRI)는 DINOv3로 위성 이미지를 분석하여 삼림 벌채 감시와 복원에 활용하고 있으며, DINOv2 대비 케냐 지역 나무 지붕(Canopy) 높이 측정의 평균 오차가 4.1미터에서 1.2미터로 축소했다고 밝혔다.

DINOv3 백본은 뛰어난 고해상도의 이미지 특징을 생성하며, 하나의 고정된 백본만으로 이미지 분류, 동영상 내 객체 추적 등의 다양한 작업에서 전문 솔루션보다 우수한 성능을 발휘한다고 강조했다. DINOv3는 주요 이미지 벤치마크 평가에서 'SigLIP 2'와 'Perception Encoder(PE)' 등 최신 비전언어모델(VLM)을 능가해 SSL 모델이 약지도 학습(Weakly Supervised)[408] 모델을 능가할 수 있음을 입증했다.[409] 메타데이터 입력 없이 훨씬 적은 훈련 연산량으로도 강력한 성능을 제공하며, 가중치가 고정된 엄격한 제약 조건에서도 우수한 성능을 발휘해 특정 작업에 맞게 모델을 미세 조정하지 않고도 광범위한 작업에 적용이 가능하다고 강조했다. 이미지 내의 각 픽셀 수준의 속성이나 특징을 포착하는 밀집 예측(Dense Prediction)에도 뛰어나며, 단일 추론을 통해 여러 시각적 과제를 동시에 처리할 수 있어 여러 예측을 동시에 수행하는 응용 사례에 적합하다고 밝혔다.

---

405) Meta, DINOv3: Self-supervised learning for vision at unprecedented scale, 2025.8.14.
406) DINOv2(2023년 4월 공개) 대비 12배 늘어난 17억 개의 이미지로 학습했으며, 모델 크기는 매개변수 70억 개로 7배 증가
407) 모델이 이미지로부터 유용한 특징을 추출하는 데 사용되는 기본 구조
408) 라벨링 비용 절감을 위해 전체 학습 데이터 중 일부만 라벨링 데이터를 활용해 학습하는 방식
409) ADE-20K(이미지 분류) 기준 DINOv3: 55.9, SigLIP 2: 42.7, PE: 38.9/ DAVIS(동영상 내 객체 추적) 기준 DINOv3: 83.3, SigLIP 2: 62.9, PE: 49.8

4. 오픈AI

1) 기업 일반현황

<표 5-4> 오픈AI 기업 현황

구분	주요 내용
기업 개요	**〈회사 개요〉** • 2015년 12월 미국 캘리포니아주에서 비영리 연구기관으로 설립되었으며, '인류에게 안전하고 유익한 인공지능 개발'을 목표로 제시 • 비영리 모기업(OpenAI Inc.)이 기업을 총괄하고 상업화를 위한 영리 자회사OpenAI Global, LLC가 실제 사업을 담당하는 하이브리드 구조로 운영 • 최근 들어 비영리 거버넌스를 유지하되 상업 활동 확대를 위해 Public Benefit Corporation(PBC) 기반 구조 개편을 통해 장기 미션과 사업 확장 균형을 모색 • 2022년 11월 챗GPT) 출시로 글로벌 생성 AI 시장 성장의 기폭제 역할을 하며, AI 산업 전반의 기술·비즈니스 패러다임을 변화시킨 기업으로 평가
	〈사업 내용〉 • 산업 분야 : 안전하고 유익한 AGI 개발을 사명으로 하며, AGI를 경제적으로 가치 있는 모든 작업에서 인간을 능가할 정도의 고도 자율 시스템으로 정의함. 연구 영역은 대규모 언어모델(LLM), 멀티모달 모델, 음성·영상 AI, AI 안전·정렬 연구, 에이전트형 AI, 생성 AI 플랫폼 등 전방위적으로 확장됨. • 제품 : GPT 계열 모델(GPT-1, 2, 3, 4, GPT-4o, GPT-4.1, o1, o3), ChatGPT, Sora(텍스트 기반 고해상도 비디오 생성 모델, 2024년 발표), SearchGPT(검색 기능을 강화한 AI 첫 웹 검색형 제품, 2025년 발표), DALL·E 시리즈(텍스트 기반 이미지 생성 모델(DALL·E 1~3), Whisper(음성 인식 모델), Codex/GPT-개발자 도구(코드 생성·자동화 모델), API 플랫폼(OpenAI API, 기업·개발자를 위한 모델 제공 플랫폼)
	〈자회사 및 인수〉 • 오픈AI Global, LLC • 2023. 8월 미국 스타트업 Global Illumination 인수 • 2024. 6월 협업 플랫폼 스타트업 Multi 인수 • 2024. 6월 데이터 검색 및 분석 전문기업 RockSet 인수
AI 조직 현황	**〈AI 관련 조직〉** • **(Language Models 팀)** : GPT 모델 군 개발을 총괄하며 GPT-4 → GPT-4.1 → GPT-4o → o1·o3 등 고도화된 모델을 출시했고 최근 들어 장기 추론 모델(reasoning), 멀티모달 통합, 효율적 학습(저비용·고성능), 에이전트 기능을 수행하는 실행형 모델(Execution Model) 개발에 집중.

구분	주요 내용
	• **(Applied AI팀)** : DALL·E, Sora, Whisper, 음성·영상 생성 모델, AI 에이전트 기반 서비스 등 제품형 연구·개발 담당하며 API 서비스·챗GPT·서치GPT 등 사용자 제품으로 기술을 연결하는 역할을 수행 • **(Alignment/Safety 팀)** : AI 안전, 모델 정렬(alignment), 데이터 정책, 사회적 영향 연구 등을 수행하는 조직으로 2025년에 Safety&Governance Division 중심으로 책임 있는 AI 개발 체계를 재정비하는 조직 개편을 단행
사업 전망	**〈주요 제품 및 서비스〉** • **(고성능 모델 전략)** GPT-4.1으로 응답 품질 고도화, GPT-4o로 음성·이미지·영상·텍스트가 통합된 완전 멀티모달 모델 구현. 2024~2025년에는 o1 모델(추론·문제 해결 특화), o3 모델(경량·저비용)을 출시해 고성능+효율형 포트폴리오 전략 • **(ChatGPT 생태계 확장)** 챗GPT Plus·Team·Enterprise 등 계층형 모델로 전환해 소비자·기업 시장 모두에서 영향력을 확대했고, 챗GPT 내에서 멀티모달 생성, 개인화 메모리, 에이전트 기능, SearchGPT 기능이 통합되며 플랫폼화를 진행 • **(SearchGPT 등 신사업 강화)** 2025년 서치GPT를 출시해 검색 시장에, Sora로 광고·콘텐츠·게임·미디어 제작 분야에서 진출. • **(클라우드 다변화 전략)** Microsoft와의 파트너십은 유지하되, 독립성 확보, 비용 절감, 경쟁 관계 완화 등을 위해 다중 클라우드 전략을 검토 • **(AI 안전·정렬 연구 체계 강화)** 2025년 들어 AGI 안전성 확보를 핵심 가치로 제시하면서 규제 대응·데이터 보호·AI 행동 예측성 향상에 중점.

2) 2025년 기업 동향

(1) 사업전략/투자/M&A

가. 트럼프 행정부를 위한 AI 정책 청사진 발표[410]

오픈AI가 트럼프 행정부 출범을 앞두고 2025년 1월 13일 AI 정책 제안을 담은 '경제 청사진(Economic Blueprint)'을 발표했다. 청사진은 미국이 AI 혁신에서 세계를 선도하고 AI에 대한 공평한 접근성을 보장하며 경제성장을 촉진하기 위한 정책 제안을 담고 있다. 여기에서 오픈AI는 의료, 교육, 과학 연구, 공공서비스 등 다양한 분야에서 혁신을 이끌고 생산성을 향상할 수 있는 AI 잠재력 실현을 위해 미국 정부가 정책 지원, 규제 개선, 인프라 구축에 적극 나설 것을 요구했다.

410) OpenAI, OpenAI's Economic Blueprint, 2025.1.13.

또한 중국의 영향력 확대를 경계하며 미국의 AI 경쟁력과 국가안보 강화를 위한 정책을 제안했다. 구체적으로 ▲정부와 AI 기업 간 국가안보 위협 관련 정보 및 테스트 결과를 공유해 AI 모델의 안전성 향상 ▲AI 기업과 국가안보 부처 간 협력을 위한 컨소시엄 구성 ▲동맹국 및 파트너국에 첨단 AI 모델을 제공해 미국 기술에 기반한 AI 생태계 구축을 지원 ▲중국·러시아 등 적대국에 대한 첨단 AI 모델 수출 제한 등을 제시했다. 공공서비스 개선을 위한 정부의 AI 역량을 강화하고 공공과 민간 부문의 AI 전문가 확보를 위해 차세대 AI 인재 양성 프로그램을 개발도 강조했다.

오픈AI는 AI 기술의 책임 있는 사용을 위해 연방 정부 차원의 통일된 규제가 수립되어야 하며, 규제 내용이 간단하고 명확해야 한다고 강조하고, AI 산업 발전을 저해하는 규제 파편화를 막기 위해 연방 정부의 주도 통일된 안전 기준의 마련을 요구했다. 또한 AI 개발단계부터 아동 성 착취물 방지책을 마련하도록 요구하는 등, 아동 성 착취물 생성과 배포 방지 정책을 마련하고, 모든 AI 생성물에 출처를 명확히 표시하도록 요구하는 조치가 필요하다고 권고했다.

오픈AI는 AI 기술 경쟁력 강화를 위해 AI 칩과 데이터, 에너지, 인재 등 핵심 자원에 대한 투자가 중요하다고 강조했다. AI 학습용 공공 데이터를 확대하고 아날로그 형식의 정부 데이터를 디지털화하며, 과학적 진보를 앞당길 수 있는 컴퓨팅 및 데이터 접근 권한을 과학자와 혁신 기업 및 교육자에게 제공하는 방안을 제시했다. 주 정부와 연방 정부가 산업계와 협력해 태양광 및 풍력 발전소, 핵발전소와 같은 AI 인프라 구축 허가 프로세스를 대폭 간소화하는 AI 경제특구를 조성도 청사진에 담았다.

나. 트럼프 행정부에 혁신의 자유 강조하는 AI 정책 제안[411)]

오픈AI가 트럼프 행정부의 ‘AI Action Plan’ 수립을 위한 백악관 과학기술정책실(OSTP) 의견 요청에 대응해 2025년 3월 13일 혁신의 자유를 강조하는 정책 제안을 제출했다. 여기에 혁신의 자유를 보장하는 규제 접근 방식과 민주적 AI 원칙에 따른 수출 통제, 학습의 자유를 증진하는 저작권 체계, 인프라 기회 포착, 정부의 AI 도입 확대 등이 모두 포함됐다. 먼저 혁신의 자유 보장은 국가안보를 강화하기 위해 연방 정부와 민간 부문 간 자율적 협력 체제를 수립하고, 주 정부의 파편화된 AI 규제로부터 AI 기업 보호를 제안했다. 여러 주

411) OpenAI, OpenAI’s proposals for the U.S. AI Action Plan, 2025.3.13.

에서 발의한 AI 관련 법안으로 인한 규제 준수 부담이 중국에 반사이익이 되지 않도록, AI 기업이 연방 정부와 소통할 수 있는 단일 창구를 마련해 규제 불확실성을 제거하고 혁신의 자유 보장을 강조했다.

민주주의 원칙에 따른 수출 통제에서는 글로벌 AI 시장에서 민주주의 원칙을 따르는 국가에는 미국산 AI 시스템의 도입을 장려하고, 중국 및 중국에 동조하는 소수 국가에는 수출 통제를 제안했다. 중국 및 미국 무기 금수조치가 적용되는 국가들에는 첨단 칩에 대한 기존 수출 통제를 포함해 AI 시스템에 대한 엄격한 수출 통제 시행을 강조했다. 학습의 자유를 증진하는 저작권 체계의 경우 미국의 AI 리더십과 경제 및 국가안보를 지원하는 저작권 체계 수립을 위한 조치 채택을 강조했다. 중국의 기술적 우위를 막기 위해 AI에 공정사용[412] 원칙을 적용하고, 미국 기업들의 데이터 접근권을 보장하기 위한 국제 정책 논의를 진행하며, 정부가 보유하거나 지원하는 데이터 접근성 확대를 제안했다.

성장을 촉진하는 인프라 기회 포착에서는 중국과의 AI 주도권 경쟁에 필요한 인프라에 투자함으로써 경제성장, AI 접근성 극대화, 민감한 미국 데이터의 자국 내 저장 등을 실현해 국가안보 이익 보호를 강조했다. 또한 에너지 인프라에 대한 계획과 허가 절차를 간소화하여 인프라 건설을 촉진하고, 공립대학에 AI 기업의 컴퓨팅 자원을 제공해 AI 전문 인력의 양성 지원을 제안했다. 마지막으로 정부의 AI 도입 확대에서는 전 세계적으로 민주적 AI가 발전하려면 미국 정부가 국민의 안전과 번영, 자유를 위해 AI를 사용하는 모범을 보여야 한다고 강조하고, 정부의 AI 도입을 강화하기 위한 공공-민간 파트너십을 통해 AI 도입 절차를 간소화하고 신속한 AI 조달 체계 마련을 제안했다.

다. 글로벌 AI 인프라 지원 이니셔티브 발표[413]

오픈AI가 2025년 5월 7일 민주적 AI 인프라 확산을 위한 글로벌 차원의 AI 인프라 구축 지원 이니셔티브 ‘OpenAI for Countries’를 발표했다. 오픈AI는 2025년 1월 오라클, 소프트뱅크(Softbank)와 함께 AI 인프라 구축을 위한 ‘스타게이트(Stargate)’ 프로젝트[414]를 발표하고 텍사스주를 시작으로 데이터센터 건설을 추진해 왔다. 이번 조치에 대해 오픈AI는 여러 국가에서 미국의 스타게이트와 유사한 AI 인프라 구축 지원을 희망한다고 강조하고,

412) Fair Use: 저작권자 허락 없이도 언론 보도나 교육 목적 등으로 콘텐츠 사용을 허용하는 원칙
413) OpenAI, Introducing OpenAI for Countries, 2025.5.7.
414) 오픈AI, 오라클, 소프트뱅크가 합작사를 설립해 미국 전역에 데이터센터를 구축하는 프로젝트

이들을 도와 민주적 AI[415]를 확산하겠다고 설명했다.

오픈AI는 이번 이니셔티브를 통해 민주적 AI 기반을 구축하고자 하는 글로벌 국가들을 지원하고 권력 공고화에 AI를 활용하는 권위주의적 AI에 대한 명확한 대안을 제시하겠다고 밝히고 첫 단계로 개별 국가나 지역을 대상으로 10개의 프로젝트를 우선 추진하고 추후 확장할 계획이다. 오픈AI는 미국 정부와의 긴밀한 협력이 민주적 AI를 발전시키는 최고의 방법이라며, 미국 정부와 함께 다음과 같은 목표를 달성하겠다고 강조했다. 각국과 협력해 데이터센터를 구축함으로써 데이터 주권을 지원하고 신규 산업을 육성하며, 개인정보보호 및 준법 기반의 AI 발전과 데이터 활용을 촉진한다는 방침이다.

각국 시민에게 각국의 필요에 맞고 지역 언어와 문화를 바탕으로 한 맞춤형 챗GPT를 통해 의료와 교육 개선, 공공서비스 효율성 향상 등 혜택을 제공할 방침이다. 또한, 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 도모한다. 지역 자본과 오픈AI 자본을 함께 활용한 국가 스타트업 펀드를 조성해 건강한 국가 AI 생태계를 구축하고, 신규 인프라를 통한 각국의 일자리와 기업, 수익을 창출해 기존 공공 및 민간의 수요에 대응한다는 계획이다. 이는 파트너 국가들과 글로벌 스타게이트 프로젝트 확장에 투자함으로써, 미국이 주도하는 AI 리더십을 확보하고 민주적 AI를 위한 글로벌 네트워크 형성이 목적이다.

(2) 기술개발 및 혁신

가. 생명공학 스타트업과 협력해 수명 연장 연구용 AI 모델 개발[416]

오픈AI가 2025년 1월 생명공학 스타트업 레트로 바이오사이언스(Retro Bioscience)와 협력해 단백질 설계 AI 모델 ‘GPT-4b micro’ 개발을 발표했다. 레트로는 오픈AI 샘 알트먼 CEO가 1억 8천만 달러를 투자한 기업으로, 2024년 오픈AI에 협력을 제안해 공동 연구를 진행해 왔다. 레트로는 인간 평균 수명의 10년 연장을 목표로 야마나카 인자(Yamanaka Factor)[417] 연구를 진행 중으로, 오픈AI는 야마나카 인자 기능 향상 지원을 위해

415) 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
416) MIT Technology Review, OpenAI has created an AI model for longevity science, 2025.1.17.
417) 인간의 피부 세포를 줄기세포로 변화시키는 단백질 세트

GPT-4b-micro를 개발했다.

줄기세포의 재(再)프로그래밍은 매우 까다로워 과학자들이 직접 설계할 경우, 몇 주의 기간이 소요되고 재생 성공률도 처리된 세포의 1% 미만에 불과한 것으로 알려졌다. 반면 GPT-4b micro의 제안을 활용한 단백질 생성은 야마나카 인자를 구성하는 요소 중 두 가지를 수정함으로써 초기 측정 결과에서 효율성이 50배 이상 향상이 가능하다고 강조했다. 오픈AI와 레트로는 향후 연구 결과를 공개할 계획이라고 밝혔으나, 제품 공식 출시가 아닌 맞춤형 데모 형식이 될 것으로 예상된다. GPT-4b-micro는 구조가 불명확하고 취약한 야마나카 인자의 특성을 반영해, 단백질 구조를 예측하는 구글의 알파폴드(AlphaFold)와는 다른 방식으로 작동한다.

GPT-4b-micro는 여러 종의 단백질 서열과 단백질 간 상호작용 데이터를 학습했으며, 레트로 연구진은 '퓨샷(Few-Shot)'[418] 방식을 적용해 야마나카 인자의 설계 방식을 개선했다고 설명했다. 레트로 조 베츠 라크루아(Joe Betts-Lacroix) CEO는 모델을 즉각 실험실에 투입해 실질적인 결과를 얻었으며, 상당수 경우에서 기존 야마나카 인자보다 개선되었다고 주장했다. 오픈AI는 이번 협업에서 금전적 거래는 없었다고 밝혔으나 협업 성과가 레트로 최대 투자자인 샘 알트먼 오픈AI CEO에게 이익이 된다는 점에서 논란이 될 가능성도 존재한다는 지적이 제기됐다.

나. GPT-4o의 아첨 현상 분석 결과 및 개선 방안 발표[419]

오픈AI가 2025년 5월 2일 챗GPT에 적용한 GPT-4o 업데이트 이후 사용자에 지나치게 동조하는 아첨(Sycophancy)[420] 현상이 발견되자 업데이트를 롤백하고 발생 원인과 개선 계획을 발표했다. 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 진행했다고 밝혔다. 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 아첨 정도를 심화시키는 변화를 초래했다.

418) 여러 예시와 답변을 제공한 뒤 모델이 새로운 질문에 답하게 하는 기법
419) OpenAI, Expanding on what we missed with sycophancy, 2025.5.2.
420) AI 모델이 사용자에게 지나치게 동의하거나 비판 없이 긍정적인 반응만 보이는 편향성 오류의 일종

오픈AI는 챗GPT의 아첨 현상이 단순한 불편이나 불안감을 넘어 정신건강 악화, 과도한 감정적 의존, 위험한 행동 등의 안전 문제를 초래할 수 있다면서, A/B 테스트[421] 및 전문가 평가자의 긍정적 반응을 토대로 출시를 결정했으나 아첨 현상을 포착할 만큼 폭넓고 충분한 평가가 이루어지지 않았다며 잘못된 판단이었다고 시인했다. 오픈AI는 이번 업데이트 롤백을 계기로 모델 업데이트 시 사용자 의견을 반영하는 '알파' 테스트 단계를 추가로 도입하는 등, 모델 검토와 출시 프로세스를 개선하겠다고 발표했다. 모델 출시 결정 시 정량적· 정성적 피드백을 모두 고려하고, 환각, 속임수, 안정성(Reliability), 성격 등의 모델 동작 이슈는 출시 중단까지 고려할 우려 사항으로 구분해 안전 검토 절차를 조정하겠다고 밝혔다.

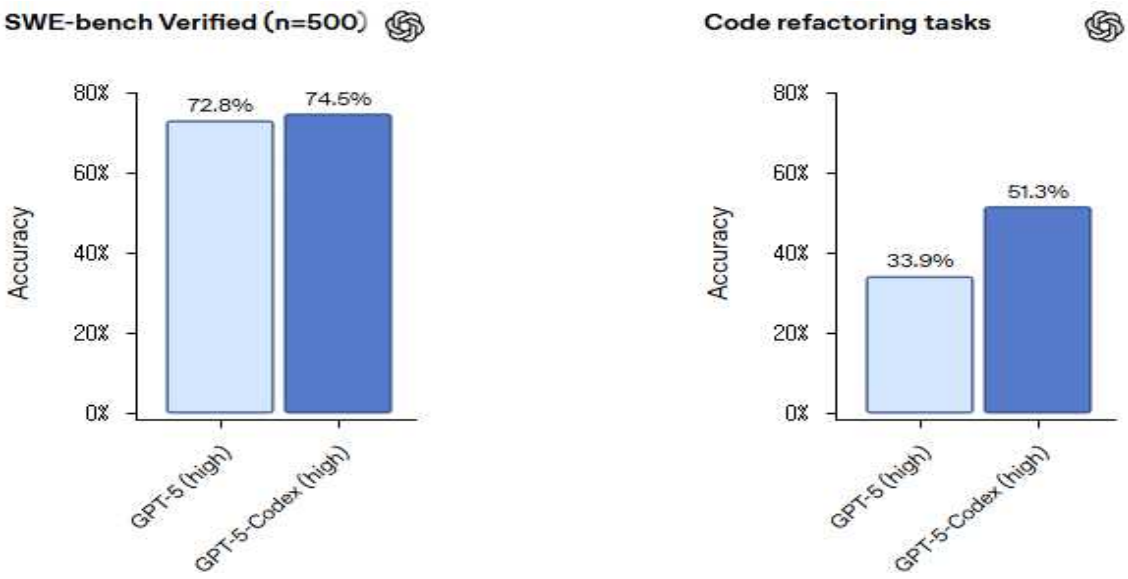
이를 위해 출시 전에 직접 피드백을 제공하고자 하는 사용자 의견을 반영할 수 있는 옵트인(Opt-in)[422] 알파 테스트 단계를 추가 도입하고, 모델 출시에 앞서 임의 점검(Spot Check)과 모델을 직접 사용하는 테스트를 더욱 중요하게 고려하며 내부 평가 및 A/B 테스트 방식도 개선하겠다고 강조했다. 신규 모델 버전의 학습 및 평가 시 이상적인 모델 동작이 어떤 모습인지 명확히 정의하여 모델 동작의 원칙 준수 여부를 더욱 정확하게 평가하고, 챗GPT 모델에 적용되는 업데이트 규모와 상관없이 사용자와 적극적으로 소통하여 사용자가 장단점을 이해할 수 있도록 충분한 설명을 제공한다는 방침이다.

다. GPT-5 기반 '코덱스' 업그레이드 발표[423]

오픈AI가 2025년 5월 공개한 AI 코딩 에이전트 '코덱스(Codex)'에 GPT-5를 적용해 업그레이드한 'GPT-5-코덱스'를 2025년 9월 15일 출시했다. GPT-5-코덱스는 전체 프로젝트의 구축, 기능과 테스트 추가, 오류수정, 코드 검토 같은 실제 소프트웨어 엔지니어링 작업을 기반으로 훈련되었으며, 개발자와 실시간 협업뿐 아니라 복잡한 작업도 독립 수행할 수 있다. SWE-bench 벤치마크[424] 평가에서 GPT-5-코덱스(high)는 74.5% 정확도를 기록해 GPT-5(high)의 72.8%를 능가했으며, 코드 재구성(Code Refactoring) 작업 정확도도 51.3%로 GPT-5(high)의 33.9%를 앞섰다.

421) 디지털 환경에서 A와 B를 무작위로 사용자 집단에 보여주고 사용자 반응을 측정하는 테스트
422) 당사자가 자신의 정보를 제공하거나 마케팅 메시지를 수신하기 위해 명시적으로 동의하는 방식
423) OpenAI, Introducing upgrades to Codex, 2025.9.15.
424) 코드 작성, 오류수정, 코드 리뷰 등 실제 소프트웨어 엔지니어링 문제를 다루는 벤치마크

[그림 5-10] GPT-5-코덱스와 GPT-5의 SWE-bench 및 코드 재구성 작업 정확도 비교



출처 : OpenAI, Introducing upgrades to Codex

GPT-5-코덱스는 작업의 복잡도에 따라 사고에 필요한 시간을 동적으로 조정하여, 단순한 요청에는 민첩하게 반응하고 대규모 코드 재구성과 같은 복잡한 작업에서는 더 오랜 시간 동안 작동이 가능하다. 기존 코덱스가 각 작업에 고정된 연산 자원을 사용했다면, GPT-5-코덱스는 동적 시간 조정으로 대규모 복잡한 작업에서 한 번에 7시간 이상 자율적으로 작동해 반복 작업으로 성공적 결과를 도출한다. 직원의 사용 트래픽 조사에서는 모델 생성 토큰 기준 하위 10%에서는 GPT-5 대비 93.7% 적은 토큰을 사용했고 상위 10%에서는 102.2% 더 많은 토큰을 소비한다. 코드 검토 및 치명적인 오류 탐지에 특화되어 검토 과정에서 코드 베이스 탐색과 코드 간 연결성 분석, 코드 실행과 테스트를 통해 정확성을 검증하고, 모바일 웹사이트 제작과 같은 프런트엔드 작업 성능도 개선이 가능하다고 설명했다.

(3) 제품 및 서비스

가. 심층 조사를 수행하는 에이전트 ‘딥 리서치’ 공개[425]

오픈AI가 2024년 2월 2일 추론을 통해 대규모 온라인 정보를 종합하고 다단계의 복잡한 연구를 수행하는 에이전트 기능 ‘딥 리서치(Deep Research)’를 출시했다. 오픈AI는 딥 리서치가 프롬프트에 따라 수백 개의 온라인 소스의 검색과 분석, 종합을 통해 연구 분석가 수준의 종합 보고서를 작성하며, 사람이 하면 여러 시간이 걸리는 작업을 수십 분 만에 완

425) OpenAI, Introducing deep research, 2025.2.2.

료할 수 있다고 강조했다. 또한 웹 브라우징과 데이터 분석에 최적화된 별도의 o3 버전을 기반으로 인터넷에서 방대한 양의 텍스트와 이미지, PDF를 검색 및 분석도 가능하다.

이 기능은 월 200달러 구독료의 ‘챗GPT 프로’ 사용자에게 우선 출시 후 2025년 2월 25일부터 월 20달러를 내는 챗GPT 플러스를 포함한 모든 유료 구독자에게 확대 제공될 방침이다. 딥 리서치는 작업 완료에 5~30분이 걸릴 수 있어 작업 완료 시 알림을 제공하며, 현재 텍스트로만 보고서를 작성하지만 추후 이미지와 그래프 출력 기능이 추가될 예정이다. 오픈AI는 딥 리서치 기능이 최근 개발된 고난도 벤치마크 HLE(Humanity’s Last Exam)[426] 평가에서 26.6%의 정확도를 보여 주요 모델[427] 중 최고점을 달성했다고 강조했다. 딥 리서치는 9.1%의 정확도를 기록한 o1과 비교해 특히 화학, 인문학, 사회과학, 수학에서 두각을 나타냈으며, 필요시 전문 정보를 효과적으로 찾아내는 등 인간과 비슷한 접근 방식을 활용한다고 설명했다.

오픈AI 내부에서 다양한 분야에 걸친 전문가 수준의 작업을 평가한 결과, 각 영역의 전문가들은 오랜 시간이 걸리는 까다로운 조사를 딥 리서치를 통해 자동화했다는 평가가 나왔지만, 아직 초기 단계인 탓에 환각 증상이나 수식 오류 등 한계도 존재한다고 밝혔다. 오픈AI는 출시 초기에는 보고서와 인용에 사소한 수식 오류가 발생하거나 작업 시작에 오랜 시간이 걸릴 수 있다고 설명하며, 이러한 문제는 사용이 늘어나며 빠르게 개선될 것이라고 밝혔다.

나. GPT-5 출시 로드맵 발표[428] 및 GPT-4.5 프리뷰 공개[429]

오픈AI 샘 알트먼 CEO가 2025년 2월 13일 소셜 플랫폼 X 게시글[430]을 통해 GPT-4.5와 GPT-5 출시 로드맵을 공개했다. 알트먼 CEO는 내부적으로 ‘Orion’이라고 불리는 GPT-4.5가 오픈AI의 마지막 비(非) 추론 모델이며 올해 출시 예정인 GPT-5는 o3를 포함해 오픈AI 모델 군 전반의 기능을 통합한 시스템이 될 것이라고 밝혔다. 또한 o3는 독립형 모델로 출시하지 않을 계획이라고 설명했다. 알트먼 CEO는 현재 오픈AI의 모델군이 지나치게 복잡해졌다며, “모든 도구를 활용하고 언제 오랫동안 생각해야 하는지 아닌지 알 수 있으며, 광범위한

426) 언어학, 로켓 과학, 고전, 생태학 등 100개 이상 주제에 걸친 3,000개 이상의 객관식 및 단답식 문제로 구성
427) 주요 모델 비교: GPT-4o(3.3%), Claude 3.5 Sonnet(4.3%), Gemini Thinking(6.2%), DeepSeek-R1(9.4%)
428) ARS Technica, Sam Altman lays out roadmap for OpenAI’s long-awaited GPT-5 model, 2025.2.14.
429) OpenAI, Introducing GPT-4.5, 2025.2.27.
430) Sam Altman, OPENAI ROADMAP UPDATE FOR GPT-4.5 and GPT-5, 2025.2.13. https://x.com/sama/status/1889755723078443244

작업에 유용한 시스템을 구축해 o-시리즈와 GPT-시리즈를 통합하는 것”이 가장 중요한 목표라고 언급했다. 또한 무료 사용자도 ‘표준 인텔리전스’ 설정에서 GPT-5를 무제한으로 사용할 수 있으며, 챗GPT 플러스(월 20달러) 가입자는 더 높은 인텔리전스, 챗GPT 프로 가입자는 훨씬 더 높은 인텔리전스에서 실행할 수 있게 될 것이라고 밝혔다. 하지만 인텔리전스 세부 내용은 공개하지 않았다.

오픈AI는 2025년 2월 27일 GPT-4.5를 월 200달러를 내는 챗GPT 프로 사용자들에게 프리뷰로 우선 공개했고, 다른 유료 사용자들에게도 순차적으로 출시 계획이라고 발표했다. 오픈AI는 GPT-4.5가 채팅을 위한 최대・최고의 모델이라며, 사람의 감독 없이 데이터를 통해 학습하는 비지도 학습을 확장함으로써 패턴을 인식하고 연관성을 도출하며 창의적인 통찰력을 끌어내는 능력이 강화되었다고 주장했다. 또한 상호작용이 더 자연스러워지고 사용자 의도를 따르는 능력이 향상되었으며, 감성지능(EQ)이 향상되어 글쓰기, 프로그래밍, 실제 문제 해결과 같은 작업에 유용하고 환각(Hallucination) 증상도 덜할 것으로 예상된다고 설명했다. 한편, GPT-4.5는 과학・수학・다국어・멀티모달 등 벤치마크 평가에서 GPT-4o는 능가했으나, o3-mini (high) 대비 낮은 성능을 보였다.

다. AI 에이전트 구축 지원 도구[431]와 신규 음성・이미지 생성 모델 출시

오픈AI가 2025년 3월 11일 개발자의 효율적인 AI 에이전트 구축을 지원하는 ‘Response API’를 공개했다. Response API는 여러 API나 외부 공급업체를 사용하지 않고 오픈AI의 모델과 기본 제공 도구를 앱에 쉽게 결합하려는 개발자를 위한 API로, 웹 검색, 파일 탐색, 컴퓨터 사용 기능[432]을 지원한다. 오픈AI는 대화형 응답을 생성하는 ‘Chat Completions API’ 및 AI 어시스턴트 생성・관리를 지원하는 ‘Assistants API’의 기능을 결합해 Response API를 출시했고, 웹 검색과 파일 검색, 컴퓨터 사용 기능을 통해 AI 모델이 현실의 복잡한 작업을 수행할 수 있다고 설명했다.

오픈AI는 Response API와 함께 개발자가 다중 에이전트 기반 업무 프로세스 조율을 간소화하는 ‘에이전트 SDK’도 공개했다. 2024년 10월 공개한 SDK ‘Swarm’을 바탕으로 ▲LLM의 손쉬운 구성 ▲에이전트 간 지능적 연계 지원 ▲안전장치 적용 등 개선 사항을 적

431) OpenAI, New tools for building agents, 2025.03.11.
432) 브라우저 기반의 실제 작업(예: 파일 다운로드, 웹 검색 등)을 수행할 수 있는 오픈AI의 에이전트 기능

용해 에이전트 SDK를 출시했다. 연이어 ▲음성 인식 모델 'gpt-4o-transcribe', 'gpt-4o-mini-trasncribe' ▲음성합성(Text-to-Speech) 모델 'gpt-4o-mini-tts'를 출시했다.[433]. 이 모델들은 악센트, 소음이 있는 환경과 같은 까다로운 조건에서도 높은 정확도를 유지하는 것이 특징으로, 특히 음성 인식 모델은 기존 '위스퍼(Whisper)' 모델 대비 단어 오류율(WER)이 크게 개선됐다.

오픈AI는 2025년 3월 25일 텍스트와 이미지 생성 기능을 통합해 더욱 정확하고 사실적인 이미지를 생성하는 'GPT-4o 이미지 생성(Image Generation)' 모델도 출시했다.[434] 오픈AI에 따르면 GPT-4o에서 이미지 생성 기능이 기본으로 제공되므로 GPT-4o의 지식 베이스와 대화 맥락을 활용해 사용자 요구를 세밀하게 반영한 이미지를 생성하며 이미지에 삽입되는 텍스트 품질도 크게 개선됐다. GPT-4o 이미지 생성 모델은 챗GPT의 기본 이미지 생성기로 유료 사용자뿐 아니라 무료 사용자에게도 제공되며, 개발자를 위한 API 접근도 몇 주 안에 지원될 예정

라. GPT-4.1 API로 출시[435] 및 o3와 o4-미니 출시

오픈AI가 2025년 4월 14일 기본 모델과 'GPT-4.1 미니', 'GPT-4.1 나노'로 구성된 'GPT-4.1' 제품군을 API로 출시했다. 오픈AI는 GPT-4.1 제품군은 최대 100만 개 토큰의 컨텍스트 창을 제공하고 2024년 6월까지의 데이터로 사전훈련이 진행되었으며, 전반적으로 GPT-4o 및 GPT-4o 미니보다 우수한 성능을 발휘한다고 설명했다. 또한 벤치마크 평가 결과에서 GPT-4.1 기본 모델의 코딩 성능과 지시이행 능력, 장문 컨텍스트 이해 능력이 GPT-4o를 능가했다고 설명하고, 개발자들의 실제 활용에 중점을 두고 모델을 학습시켰다고 강조했다.[436] GPT-4.1 미니는 GPT-4o 대비 비슷하거나 더 뛰어난 지능을 발휘하면서 지연 시간은 거의 절반으로 줄고 비용은 83% 감소했으며, GPT-4.1 나노는 낮은 지연 시간이 요구되는 작업에 적합한 가장 빠르고 저렴한 모델로 API 가격이 입출력 토큰 100만 개당 0.10달러/0.40달러로 역대 최저로 책정했다.

---

433) OpenAI, Introducing next-generation audio models in the API, 2025.3.20.
434) OpenAI, Introducing 4o Image Generation, 2025.3.25.
435) OpenAI, Introducing GPT-4.1 in the API, 2025.4.14.
436) SWE-bench Verified(코딩): 54.6%로 GPT-4o를 21.4%p 능가, MultiChallenge(지시이행): 38.3%로 GPT-4o를 10.5%p 능가, Video-MME(멀티모달 장문 컨텍스트 이해): 72.0%로 GPT-4o를 6.7%p 능가

뒤이어 2025년 4월 16일에는 다면적 질문에 효과적으로 대처하고 사용자 대신 자율적으로 작업을 실행할 수 있는 ‘에이전트형 AI’로의 진화를 반영하는 추론 모델 o3와 o4-미니를 출시했다.[437] 오픈AI는 o3가 가장 강력한 추론 모델로 멀티모달과 코딩 관련 벤치마크 평가에서 최고 점수를 기록[438]했으며, 외부 전문가 평가 결과에서 어려운 실제 작업 수행 시 오류가 o1보다 20% 더 적고 프로그래밍, 비즈니스/컨설팅, 창의적 아이디어 도출에서 탁월한 성능을 발휘했다고 밝혔다. o4-미니는 빠르고 비용 효율적 추론에 최적화된 소형 모델로, 특히 수학, 코딩, 시각적 작업에서 크기와 비용 대비 뛰어난 성능을 제공한다.

오픈AI는 두 모델 모두 강화학습을 통해 도구를 사용하도록 훈련했으며, 도구 사용법뿐 아니라 언제 어떻게 도구를 사용해야 하는지 추론하는 능력도 학습해 통상적으로 1분 이내 응답을 신속히 생성한다고 강조했다. 한편, 오픈AI가 공개한 o3와 o4-미니의 기술 보고서(OpenAI o3 and o4-mini System Card)[439]에 따르면 o3의 환각률은 33%, o4-미니는 48%에 달해 o1(16%)보다 크게 높아진 것으로 나타났으며, 오픈AI는 환각 증가의 원인을 규명하기 위해 더 많은 연구가 필요하다고 부연했다.

마. 코딩 에이전트 ‘코덱스’ 연구용 프리뷰 출시[440]

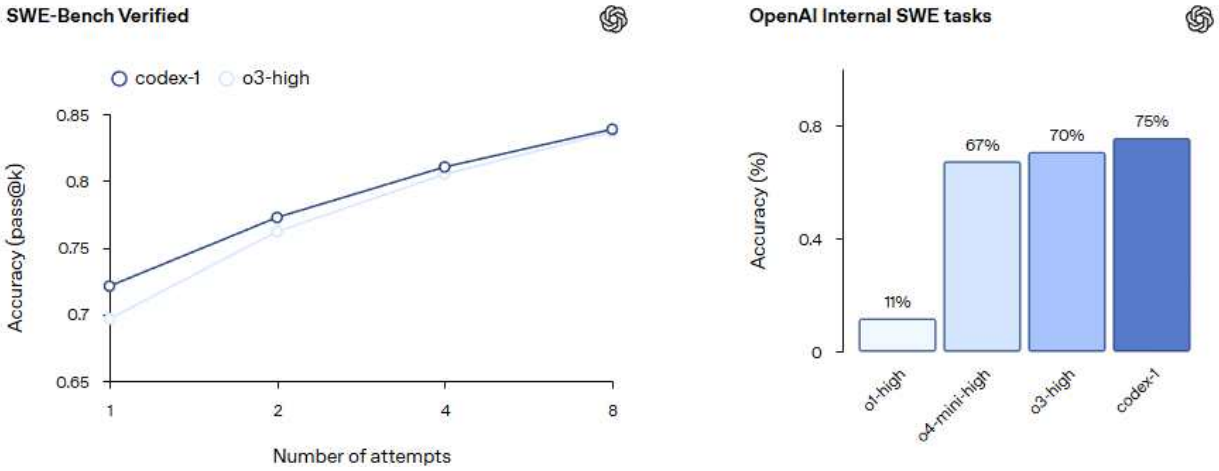
오픈AI가 2025년 5월 16일 다양한 코딩 작업을 동시에 처리할 수 있는 클라우드 기반 소프트웨어 엔지니어링(SWE) 에이전트 ‘코덱스(Codex)’를 연구용 프리뷰로 공개했다. 소프트웨어 엔지니어링에 최적화된 오픈AI o3 버전인 ‘codex-1’을 기반으로 개발된 코덱스는 코드 작성, 코드 관련 질문 답변, 오류수정, 검토를 위한 풀 리퀘스트(Pull Request)[441]* 제안 등의 작업 수행이 가능하다. 오픈AI에 따르면 코덱스는 다양한 환경에서 강화학습을 통해 실제 코딩 작업의 사용자 선호도를 자세히 반영하고 지시 사항을 정확히 준수하며, 테스트를 통과할 때까지 코드를 반복 실행했다고 설명했다.

사용자는 챗GPT의 사이드바에서 코덱스에 접속해 프롬프트를 입력하고 ‘코드’를 클릭해 코딩 작업을 할당할 수 있으며, 작업 시간은 복잡성에 따라 1분에서 30분까지 소요되고

437) OpenAI, Introducing OpenAI o3 and o4-mini, 2025.4.16.
438) MMMU(멀티모달 이해): 82.9%, SWE-Bench Verified(소프트웨어 엔지니어링): 69.1%
439) https://openai.com/index/o3-o4-mini-system-card (2025.11.17. 방문)
440) OpenAI, Introducing Codex, 2025.5.16.
441) 깃허브(Github) 등에서 코드 변경 사항을 검토하고 원본 소스 코드에 병합 요청하는 기능

진행 상황의 실시간 모니터링도 지원한다. 코드 저장소 내 코드 기반 탐색 방법과 테스트 명령어, 프로젝트의 표준 관행 등을 안내하는 AGENTS.md 텍스트 파일을 통해 명확한 지침을 제공함으로써 코덱스를 더욱 효율적으로 활용할 수 있다는 설명이다. 오픈AI에 따르면 코덱스는 코딩 벤치마크인 SWE-Bench Verified[442]*와 내부 SWE 작업을 통한 벤치마크 평가에서 o3-high를 능가했다.[443]

[그림 5-11] 코덱스의 SWE-Bench Verified와 오픈AI 내부 SWE 작업 정확도 비교



출처 : OpenAI, Introducing Codex

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 보장한다고 강조했다.

바. 사용자 대신 복잡한 작업을 처리하는 '챗GPT 에이전트' 공개[444]

오픈AI가 2025년 7월 17일 챗GPT에 웹 브라우징과 심층 추론 기능을 결합해 사용자 대신 복잡한 작업을 처리할 수 있는 '챗GPT 에이전트'를 유료 사용자 대상으로 출시했다. 챗GPT 에이전트는 웹사이트와 상호작용을 하는 '오퍼레이터(Operator)'와 웹 정보를 분석

442) 코드 작성, 버그 수정 등 소프트웨어 개발 작업에서 LLM 성능을 평가하는 SWE-Bench에서 고품질 데이터셋을 선별한 벤치마크
443) SWE-Bench 기준 codex-1: 0.7210, o3-high: 0.6965, 내부 SWE 작업 기준 codex-1: 75%, o3-high: 70%
444) OpenAI, Introducing ChatGPT agent: bridging research and action, 2025.7.17.

하는 ‘딥 리서치(Deep Research)’ 기능, 챗GPT의 대화 능력을 결합하여 통합 에이전틱 시스템을 구현했다. 가령 챗GPT에 ‘경쟁사 세 곳을 분석해 슬라이드를 작성해 줘’와 같은 지시를 내리면 웹사이트 탐색과 결과 필터링, 보안 로그인, 결과 요약을 통해 프레젠테이션이나 스프레드시트 같은 정제된 형태의 결과물을 제공한다. 또한 중요한 작업 전에 사용자에게 권한을 요청하며, 사용자가 손쉽게 개입해 브라우저를 직접 제어하거나 작업을 중단할 수 있도록 함으로써 사용자의 에이전트 통제 능력을 보장한다.

챗GPT 에이전트는 GUI를 통해 웹과 상호작용을 하는 시각 브라우저, 단순한 추론 기반 웹 쿼리를 위한 텍스트 브라우저, API 접근 등 다양한 웹 도구를 활용해 최적의 경로로 작업을 효율적으로 수행한다. 모든 작업은 챗GPT의 자체 가상 컴퓨터를 통해 이루어지며, 여러 도구 사용 시에도 작업에 필요한 맥락을 보존해 작업을 중간에 변경하더라도 새로운 정보를 반영해 중단된 곳에서부터 다시 시작한다. 오픈AI에 따르면 챗GPT 에이전트는 오픈AI의 이전 모델들보다 훨씬 상호작용 능력이 뛰어나며, 웹 브라우징 및 실제 작업 완료 능력을 평가하는 벤치마크 평가에서 뛰어난 성과를 기록했다. 챗GPT 에이전트는 HLE(Humanity’s Last Exam) 벤치마크에서 도구 사용 시 41.6%를 기록해 딥리서치(26.6%)를 크게 앞섰고, DSBench[445]에서는 기존 첨단 모델과 인간의 능력을 훨씬 능가하는 것으로 나타났다.[446]

주요 기업을 위한 3종의 재무 모델을 종합하는 작업 등 1~3년 차 투자 은행 분석가에 상응하는 작업에 대한 내부 벤치마크 결과, 챗GPT 에이전트는 71.3%로 o3(55.9%)와 딥리서치(48.6%)를 크게 앞섰다. 그러나 오픈AI는 챗GPT 에이전트가 아직 초기 단계로, 슬라이드 작성 등 복잡한 작업에서는 여전히 실수할 수 있다며 지속적인 업데이트를 통해 개선해 나가겠다는 계획을 제시했다. 파워포인트 슬라이드 생성 기능에 더욱 정교한 출력물 생성과 서식 개선 등의 업데이트를 적용하는 한편, 사용자의 감독 수준을 지속적으로 조정하여 유용성과 안전성을 보장하면서 상호작용을 강화할 계획이다.

445) 데이터 분석과 모델링과 같은 실제 데이터 과학 작업으로 에이전트를 평가
446) DSBench : Data Analysis 기준 ChatGPT agent(89.9%), o3(87.9%), 인간(64.1%), GPT-4o(34.1%)

사. 차세대 AI 모델 ‘GPT-5’ 와 개방형 가중치 모델 ‘gpt-oss’ 출시

오픈AI가 2025년 8월 7일 GPT 시리즈의 최신 모델로 ‘GPT-5’를 공개했다.[447] GPT-5는 그동안 별도로 운영되던 추론 특화 모델과 비추론(일반 대화) 모델을 하나로 통합해, 사용자는 모델을 전환할 필요 없이 GPT-5만으로 간단한 질문부터 복잡한 문제까지 모두 대응이 가능하다. GPT-5는 모든 사용자에게 제공되며 플러스(월 20달러) 구독자에게는 더 많은 사용량을 지원하고, 프로(월 200달러) 구독자에게는 확장된 추론 기능을 갖춘 GPT-5 Pro 버전을 제공할 방침이다.

GPT-5는 이전 세대 대비 수학, 코딩, 글쓰기, 시각적 이해, 의료 등 다양한 분야의 벤치마크에서 현존 최고 수준(SOTA)의 성능을 달성했다. 실제로 ▲수학(도구 없이 AIME 2025 94.6%) ▲실생활 코딩(SWE-bench Verified 74.9%, Aider Polyglot 88%) ▲다중모드 이해(MMMU 84.2%) ▲건강(HealthBench Hard 46.2%) 분야에서 최고 점수 기록했다. 또한 HLE(Humanity’s Last Exam)에서 42%를 달성하며 AGI(50% 이상) 실현 가능성을 높였으며, 일반 상식과 문제 해결 능력을 평가하는 GPQA 테스트에서도 88.4%로 최고 점수를 달성했다.

오픈AI는 GPT-5의 추론 능력과 응답 정확도 개선을 강조하며 모델 내부에 ‘사고 모드(GPT-5 Thinking)’를 탑재해 복잡하거나 논리가 필요한 질문에 대해 자동으로 심층 추론을 수행한다고 설명했다. 간단한 질문에는 빠르게 답하는 경량 모드로 대응하고, 이 두 모드 사이를 연결하는 ‘실시간 라우터’로 대화 유형이나 복잡성, 사용자 의도를 판단해 알아서 적절한 방식을 선택해 최적화된 응답 시간과 품질을 제공한다는 것이다. 또한 GPT-5는 사실성과 정확도가 향상되어 환각(Hallucination) 비율이 대폭 줄었고[448]*, 응답 품질이 개선되어 실생활에 더욱 유용한 답변을 제공하며 창작과 글쓰기 능력도 향상됐다.

오픈AI는 2025년 8월 5일 처음으로 개방형 가중치(Open-weight)[449] 언어모델 ‘gpt-oss-120b’와 ‘gpt-oss-20b’를 허깅페이스에서 출시했다.[450] gpt-oss-120b는 단일 80GB GPU에서, gpt-oss-20b는 16GB 메모리로 작동 가능한 모델로, 두 모델 모두 도구 사용

447) OpenAI, Introducing GPT-5, 2025.08.07.
448) 웹 검색 활성화 조건에서 GPT-4o 대비 응답에 사실 오류가 포함될 가능성이 45% 낮고, 사고(Thinking) 측면에서도 o3 대비 사실 오류가 포함될 가능성이 80% 낮은 것으로 확인
449) 모델 가중치(Weights)를 공개하되 학습 데이터나 소스 코드, 세부 아키텍처는 비공개
450) OpenAI, Introducing gpt-oss, 2025.08.05.

과 추론 기능 기반의 에이전트 작업에 적합하고 예산 절감을 위한 추론 수준 조절도 지원한다. 오픈AI의 개방형 가중치 모델 출시는 최근 시장에서 폐쇄형 AI 모델을 능가하는 인기를 누리고 있는 딥시크(DeepSeek) 등의 오픈소스 모델에 대응하기 위한 시도로 평가된다.

(4) AI 윤리 및 규제 대응

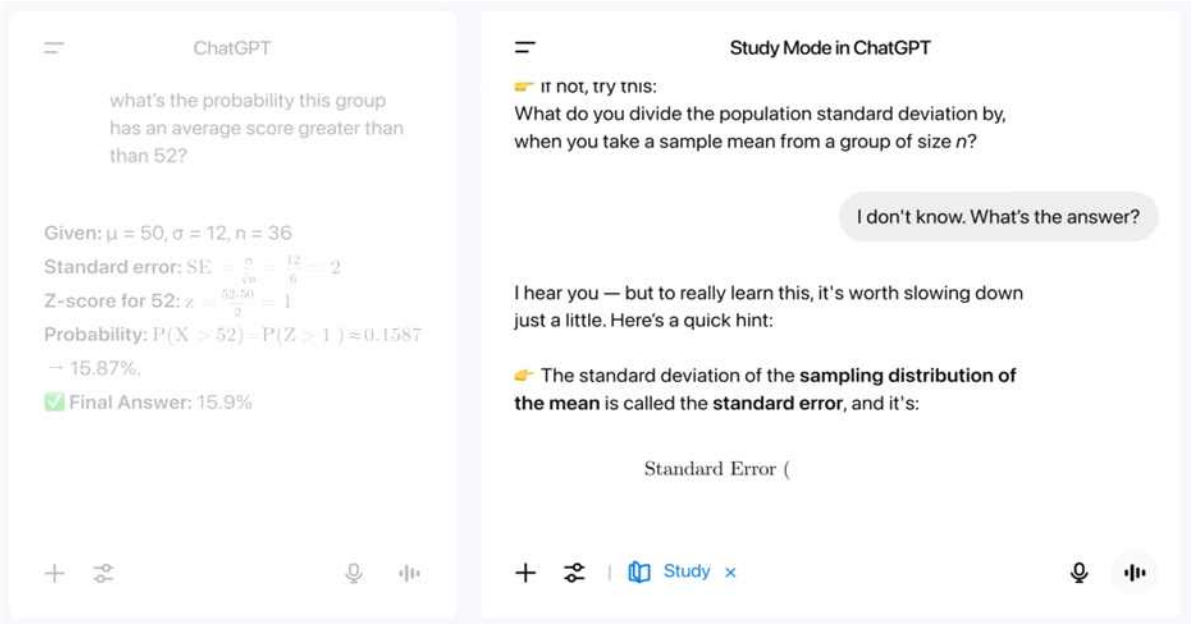
가. 챗GPT에 학습 지원을 위한 '스터디 모드' 기능 추가[451]

오픈AI가 2025년 7월 29일 챗GPT에 단순히 정답만 제공하는 대신 단계별 문제 풀이를 통해 학습을 지원하는 '스터디 모드(Study Mode)' 기능을 추가한다고 발표했다. 오픈AI는 전 세계적으로 챗GPT가 학습 도구로 널리 이용되면서 단순한 정답 제시를 넘어 실제 학습에 사용되도록 보장하고자 학생의 학습 목표와 능력에 따라 답변을 조정하는 '스터디 모드'를 개발했다고 밝혔다. 스터디 모드는 대화 기록을 기반으로 사용자의 수준에 따라 맞춤형 학습을 제공하고, 퀴즈와 주관식 질문, 개인화 피드백을 통해 학습 진행 상황을 추적하며 새로운 맥락에서 지식을 적용하는 능력을 개선할 수 있도록 지원해 숙제 풀이와 시험 준비, 새로운 주제 학습에 특히 유용하다는 설명이다.

오픈AI는 스터디 모드를 맞춤형 시스템 지침 기반으로 우선 운영하고, 반복적 학습과 학생 피드백을 통해 가장 효과적인 학습 방식을 파악한 뒤, 기본 모델을 직접 학습시킬 계획이다. 교사, 과학자, 교육학 전문가와 협력해 심층 학습을 지원하는 행동 방식을 반영한 맞춤형 시스템 지침을 개발했으며, 맞춤형 시스템 지침은 학습자의 적극적 참여 유도와 인지 부하 관리, 메타인지와 자기성찰 촉진 및 호기심 유발, 구체적 피드백으로 학습효과를 증진할 수 있도록 구성했다. 오픈AI는 향후 복잡한 개념에 대한 명확한 시각화, 대화 전반에 걸친 목표 설정과 진행 상황 추적, 학생의 능력 수준과 목표에 따른 개인화 강화 등의 새로운 기능 추가도 고려할 계획이다.

451) OpenAI, Introducing study mode, 2025.7.29.

[그림 5-12] 챗GPT 스터디 모드의 응답 예시



출처 : OpenAI, Introducing study mode

나. 청소년 자살로 피소된 오픈AI[452], 챗GPT 안전성 개선 방안 발표

2025년 4월 자살한 애덤 레인(Adam Raine)이라는 16세 청소년의 부모가 오픈AI와 샘 올트먼(Sam Altman) CEO를 상대로 2025년 8월 26일 소송을 제기해 업계의 이목이 쏠렸다.[453] 부모는 자살한 청소년의 휴대전화에서 챗GPT 대화 기록을 확인한 결과, 챗GPT가 아들의 자살 방법을 적극적으로 도왔다며 자살을 부추긴 오픈AI의 책임을 주장했다. 유족은 오픈AI가 사용자의 심리적 의존성을 조장하도록 챗GPT를 설계했고, 애덤이 사용한 버전인 GPT-4o를 출시하면서 이익을 위해 안전 프로토콜을 우회했다고 지적했다.

오픈AI는 피소 이후 정서적으로 취약한 사용자의 챗GPT 사용으로 인한 문제를 완화하고 챗GPT의 안전성과 신뢰성을 강화하기 위해 향후 120일에 걸친 개선 로드맵을 추진하겠다고 발표했다.[454] 개선 로드맵의 네 가지 중점 영역은 ▲위기 개입 확대 ▲긴급 서비스 연결성 제고 ▲신뢰할 수 있는 연락처와의 연결 지원 ▲청소년 보호 강화로 구성되며, 전문가 파트너십 확대와 민감 질문에 대한 추론 모델 우선 사용, 청소년 보호 강화 방안을 우선 발

452) The New York Times, A Teen Was Suicidal. ChatGPT Was the Friend He Confided In, 2025.8.26.
453) OpenAI, Building more helpful ChatGPT experiences for everyone, 2025.9.2.
454) OpenAI, Teen safety, freedom, and privacy, 2025.9.16.

표했다.

먼저 전문가 파트너십 확대는 웰빙 및 AI 전문가 위원회와 글로벌 의사 네트워크를 통해 심층적인 전문 의료 지식을 바탕으로 챗GPT의 안전성 개선을 추진할 예정이다. 청소년 발달, 정신건강, 인간-컴퓨터 상호작용 전문가로 구성된 위원회를 소집해 AI를 통해 삶의 질을 개선할 방안을 모색하고 우선순위에 따라 개선된 안전장치를 설계할 계획이다. 또한 60개국 250명 이상의 의사로 구성된 글로벌 네트워크와 협력해 모델 안전에 관한 의견을 수렴하고, 섭식장애, 약물 사용, 청소년 건강 관련 전문 지식을 갖춘 임상의와 연구자를 네트워크에 추가 확보할 계획이다.

민감 질문에 대한 추론 모델 활용의 경우 사용자의 정신건강 이상 징후를 반영하는 민감한 대화 시에는 비추론 모델 대비 안전 지침을 더욱 일관되게 적용하는 추론 모델로 전환해 더욱 유용한 답변을 제공할 방침이다. 청소년 보호 강화는 2025년 10월 내 부모를 위한 자녀 보호 기능을 도입하고, 만 18세 미만 사용자 대상 별도 환경과 규칙을 도입해 어떤 상황에서도 자살이나 자해에 관한 논의에 참여하지 않도록 설정할 계획이다. 13세 이상 자녀 계정과 부모 계정을 연결해 부모가 메모리와 대화 기록 등 특정 기능을 비활성화할 수 있도록 지원하며, 나이에 적합한 모델 행동 규칙을 디폴트로 활성화해 미성년자에게 적절한 응답을 제공 예정이다.

5. xAI

1) 기업 일반현황

<표 5-5> xAI 기업 현황

구분	주요 내용
기업 개요	**〈회사 개요〉** • 2006년 샌프란시스코에서 설립된 트위터(Twitter)가 기업 전신으로 2022년 10월 일론 머스크가 440억 달러 규모로 인수해 2023년 X Corporation으로 재편 • 기존 트위터 브랜드는 'X'로 대체됐고, 소셜 플랫폼·결제 플랫폼·AI 서비스 통합을 목표로 'Everything App(모든 것을 하는 앱)' 전략을 추진 • xAI는 2023년 7월 머스크가 별도로 설립한 독립 AI 회사로, X Corp와는 법적으로 분리되어 있으나 기술·데이터·서비스 측면에서 긴밀한 협력 관계를 유지 • xAI는 2024년 시리즈 B 라운드에서 60억 달러를 유치하며 기업 가치가 240억 달러 이상으로 평가 • 2025년 글로벌 AI 시장에서 오픈AI·메타·구글에 맞서는 신흥 AI 기업으로 부상
	〈사업 내용〉 • X Corp : 소셜 미디어·콘텐츠 플랫폼·광고·실시간 정보 네트워크·크리에이터 서비스·결제 서비스(X Payments) • xAI : AI 모델 개발, AGI 연구, 대규모 데이터 기반 추론형 모델 구축, 온디바이스 AI·에이전트 기술 개발 • 주요 서비스 : Grok, Grok for X, Grok Chat, xAI API
	〈관련 회사〉 • xAI(AI 모델 개발을 위한 별도 조직), X Payments(X 플랫폼 내 결제·상거래 기능 확장 조직, 2024년 미국 14개 주에서 서비스 시작)
AI 조직 현황	**〈AI 관련 조직〉** • **(Language Model 팀)** : Grok 시리즈 개발. • **(Inference·Optimization 팀)** : LLM 효율화, 저비용·고성능 트레이닝 최적화 연구. • **(Infrastructure 팀)** : 슈퍼컴퓨터 구축, GPU 클러스터 운영, X 플랫폼 실시간 데이터 연계 기술 개발 • **(Applied AI팀)** : X 플랫폼 내 AI 기능 구현(검색·추천·대화·콘텐츠 모더레이션). • **(Safety & Responsibility 팀)** : 모델 안전성·편향·악용 방지·정렬 연구 수행 • xAI는 딥마인드, 오픈AI, 구글 리서치, 테슬라 AI 등에서 경력을 쌓은 연구원들이 주축을 이루며 소수 정예 구조로 운영됨. • '우주의 본질을 이해하는 AI'라는 장기 비전을 내세우며, 해석 가능성·추론 능력·실세계 적합성을 강조한 AI 개발을 목표로 함.

사업 전망	• **(Grok 시리즈 고도화)** 2023년 Grok 출시 이후, 2024년 Grok 1.5, 2025년 Grok 2로 고도화하면서 GPT-4급 성능을 목표로 추론·멀티모달 기능을 강화. • **(슈퍼컴퓨터 구축 및 인프라 투자 확대)** xAI는 2024년 발표한 '세계 최대 규모 AI 슈퍼컴퓨터 구축 계획'[455]을 기반으로 10만 개 이상 Nvidia H100/H200급 GPU 클러스터 기반 인프라 확보를 추진. • **(X 플랫폼과의 통합)** Grok은 X 플랫폼에서 기본 AI 서비스 역할을 하며, 콘텐츠 탐색·요약·감성 분석·정보 조회 기능을 실시간으로 제공.

2) 2025년 기업 동향

가. 최신 AI 모델 '그록 3' 프리뷰 출시[456]

AI 스타트업 xAI가 2025년 2월 19일 최신 모델 '그록 3(Grok 3)' 베타 버전을 프리뷰로 공개했다. 그록 3는 기본 모델과 비용 효율적인 '그록 3 미니'로 구성되며, 두 모델 모두 별도의 'Think' 버튼을 통한 추론 기능 활성화를 지원한다. xAI는 AIME2025[457] 평가에서 그록 3가 93.3%, 그록 3 미니(Think)는 90.8%로 오픈AI o3-미니-하이(86.5%), DeepSeek-R1(70.0%)을 능가하는 최고 점수를 기록했다고 밝혔다. 그록 3(Think)는 최종 응답뿐 아니라 모델 자체의 추론 과정에 대한 검증도 지원하며, 코딩, 퍼즐, 수학 등 다양한 문제 영역에서 고른 성능을 발휘했다는 설명이다. 또한 그록 3는 응답에 수 분이 걸리는 추론 기능을 사용하지 않을 시 고품질의 응답을 즉각 제공하며, 이미지와 동영상 이해에서도 뛰어난 성능을 발휘했다고 강조했다.

xAI는 AIME 2024 평가에서 그록 3가 52.2%, 그록 3 미니가 39.7%로 딥시크-V3(39.2%)와 클로드 3.5 소네트(16.0%), GPT-4o(9.3%)를 앞섰으며, 그록 3 초기 버전(코드명 chocolate)은 실제 사용자들의 선호도를 나타내는 '챗봇 아레나(Chatbot Arena)' 리더보드에서 1위를 차지했다고 발표했다. 그록 3는 그록 2보다 10배 많은 연산 자원으로 훈련되었으며, 그록 2 대비 8배 증가한 100만 개 토큰의 컨텍스트 창을 지원한다. xAI는 총 20만 개의 GPU가 탑재된 슈퍼컴퓨터 '콜로서스(Colossus)'에서 그록 3의 훈련을 지속하고 있으며, 향후 몇 달간 모델을 지속 업데이트 예정이다.

---

455) xAI는 2024년부터 미국 테네시 멤피스에 '콜로서스(Colossus)'라는 초대형 AI 슈퍼컴퓨터를 구축하고 있으며, 초기 10만 개 이상의 엔비디아 H100 GPU로 시작해 20만 개 규모까지 확장하는 계획을 발표. 프로젝트는 차세대 Grok 모델 학습을 위한 전용 인프라로, xAI가 글로벌 AI 경쟁에서 독립적 계산 자원을 확보하려는 전략의 핵심으로 평가

456) xAI, Grok 3 Beta — The Age of Reasoning Agents, 2025.2.19.

457) 2025년 미국 수학 올림피아드 예선 문제

xAI는 인터넷과 X의 데이터를 실시간으로 검색해 분석하고 보고서를 작성하는 에이전트 기능 ‘딥서치(DeepSearch)’도 공개했다. xAI는 딥서치가 주요 정보를 종합하고, 상충하는 사실과 의견에 대해 추론하며, 복잡한 내용을 명확하게 풀어내도록 설계되었다고 설명했다. 그록 3는 X 및 Grok.com에서 X 프리미엄(월 8달러) 및 프리미엄 플러스(월 40달러[458]) 사용자에게 제공되며, 추론 기능과 딥서치 기능은 프리미엄 플러스 사용자에게만 지원한다.

나. 차세대 AI 모델 ‘그록 4’ 공개[459] 및 정부 AI 시장 진출

xAI가 2025년 7월 9일 차세대 AI 모델 ‘그록 4(Grok 4)’를 출시했다. 그록 4는 기본 모델(Grok 4)과 그록 4 헤비(Grok 4 Heavy)로 구성되며, 그록 4 헤비는 여러 에이전트를 동시에 생성해 최선의 답을 찾음으로써 더욱 뛰어난 성능을 제공하는 ‘다중 에이전트 버전’으로 설계됐다. 기본 도구 사용과 실시간 검색 통합 기능을 갖추고 있어, 실시간 정보검색이나 복잡한 연구 질문 답변 시, 웹 브라우징과 같은 도구 사용으로 웹 전반에서 지식을 찾아 심층적으로 분석한다.

xAI는 그록 4 헤비는 HLE(Humanity’s Last Exam)에서 사상 최초로 50%를 넘어서는 등 대부분 학술 벤치마크에서 최고 점수를 기록했고, 그록 4 역시 주요 벤치마크에서 첨단 AI 모델을 압도했다고 강조했다. 그록 4 헤비는 HLE에서 50.7%를 기록했으며, 그록 4는 ARC-AGI-2[460] 벤치마크에서 15.9%의 정확도를 기록해, 클로드 오푸스 4(8.6%), o3(6.5%), 제미나이 2.5 프로(4.9%)를 능가했다. xAI는 그록 4 공개와 함께 현재까지 AI 기업들이 출시한 요금제 중 가장 고가인 월 300달러의 ‘슈퍼그록 헤비(SuperGrok Heavy)’ 요금제를 출시했다.

슈퍼그록 헤비 구독자는 그록 4 헤비 및 xAI가 향후 출시할 일부 신제품을 미리 체험해 볼 수 있으며, xAI는 AI 코딩 모델을 2025년 8월, 멀티모달 에이전트는 9월, 비디오 생성 모델은 10월에 출시할 계획이다. 한편, 그록 4 출시 이후 그록 4가 반유대주의 메시지를 게시하거나 민감한 질문 답변 시 일론 머스크의 X를 검색해 해당 주제에 대한 머스크의 의견을 답변에 반영한다는 논란이 제기됐다. 이에 xAI는 그록 4와 관련된 주요 논란에 대해 사과하며, 답변 시 이전 버전 그록이나 일론 머스크, xAI가 밝힌 특정 신념이 아니라 독자적 분석

458) x는 그록 3 출시와 함께 2025년 2월 18일부터 프리미엄 플러스 요금제 가격을 기존 월 22달러에서 월 40달러로 인상
459) xAI, Grok 4, 2025.7.9.
460) 추론과 학습을 통해 다양한 시각적 패턴 문제를 해결할 수 있는 일반인공지능(AGI) 수준을 평가

에 기반하도록 시스템 프롬프트를 업데이트했다고 설명했다.

한편, xAI는 2025년 7월 14일 미국 연방 정부와 주 정부, 지방정부에 그록 AI 모델 제품군을 제공하기 위한 ‘그록 포 거버먼트(Grok For Government)’를 발표했다.[461] xAI는 정부 고객에 국가안보나 중요 과학 응용 분야에 대한 맞춤형 AI 모델과 AI 기반 애플리케이션을 제공할 계획으로, 이미 미국 국방부(DoD)와 2억 달러 규모의 신규 계약을 체결했다고 발표했다.[462]

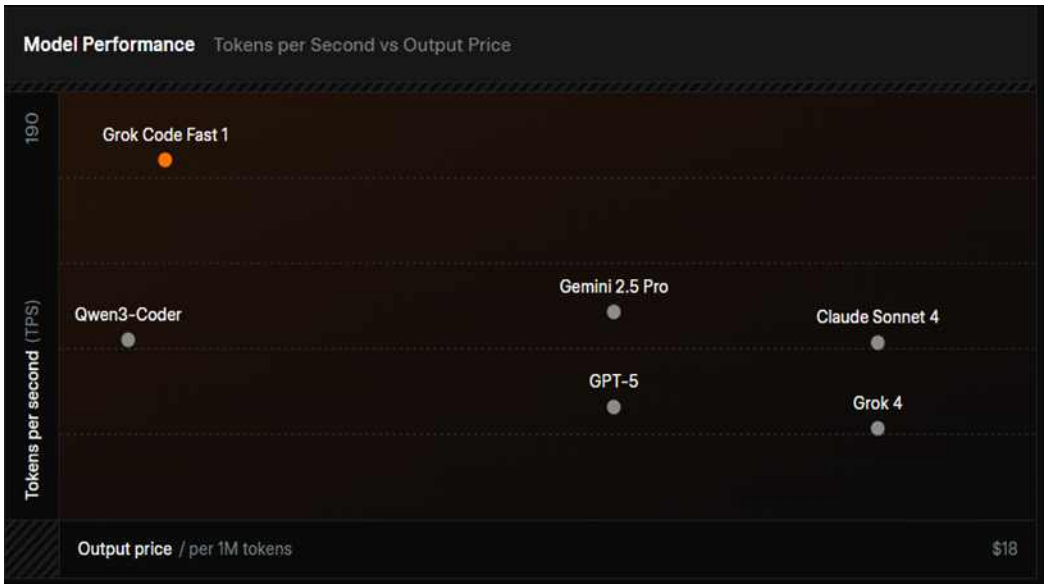
다. 에이전트 코딩 특화 AI 모델 ‘그록 코드 패스트 1’ 공개[463]

xAI가 2025년 8월 28일 에이전트 코딩에 특화된 추론 AI 모델 ‘그록 코드 패스트 1(grok-code-fast-1)’을 출시했다. xAI에 따르면 그록 코드 패스트 1은 전체 소프트웨어 개발 스택에 능숙하고, 다양한 프로그래밍 언어*에서 뛰어난 성능을 발휘하며, 처음부터 새로 시작하는 프로젝트 구축부터 코드 베이스 질의응답, 오류수정 등 일반적인 프로그래밍 작업을 최소한의 감독만으로 완료할 수 있다. xAI는 기존 AI 모델이 강력한 성능에도 추론과 도구 호출 속도가 느려 에이전트 코딩에 부적합하다는 점을 지적하며, 일상 코딩 업무에 최적화된 빠르고 경제적인 모델로 그록 코드 패스트 1을 설계했다고 강조했다. 사용자가 모델의 사고 과정 첫 단락을 읽기 전에 이미 수십 개 도구 호출을 완료할 만큼 빠른 반응성을 구현하고, 자주 사용되는 요청을 미리 저장하는 캐싱 최적화를 통해 모델의 응답 속도를 개선했다.

그록 코드 패스트 1API 비용은 100만 입력 토큰 당 0.20달러(캐시된 입력은 0.02 달러), 100만 출력 토큰 당 1.50달러로 책정됐다. 모델 성능평가에서는 공개 벤치마크와 실제 테스트를 병행해 SWE-Bench-Verified 평가에서 70.8%를 기록했고, 실제 사용자 사이에서도 일상적인 코딩 업무에서 빠르고 안정적이라는 평가를 받았다. xAI는 향후 개발자 커뮤니티의 피드백을 수렴해 신속한 업데이트를 제공할 계획으로, 이미 멀티모달 입력, 병렬 도구 호출, 확장된 컨텍스트 길이를 지원하는 변형 모델을 훈련 중이라고 설명했다.

461) xAI, Announcing Grok for Government, 2025.7.14.
462) TechCrunch, xAI says it has fixed Grok 4’s problematic responses, 2025.7.15.
463) xAI, Grok Code Fast 1, 2025.8.28.

[그림 5-13] 그록 코드 패스트 1’과 경쟁 AI 모델의 성능 비교



출처 : xAI, Grok Code Fast 1

라. xAI, 사용자 경험 강화에 중점을 둔 ‘그록 4.1’ 출시[464)]

xAI가 2025년 11월 17일 ‘그록 4(Grok 4)’의 실제 사용성 개선에 중점을 두고 창의적·감성적·협력적 상호작용을 강화한 ‘그록 4.1(빠른 응답 모델)’과 ‘그록 4.1 씽킹(추론 모델)’ 2개 버전을 공개했다. 4.1은 그록 4와 동일한 대규모 강화학습 인프라를 활용해 기존 성능과 신뢰성을 유지하면서 모델의 스타일과 개성, 유용성, 정렬을 최적화하여 사용자의 미묘한 의도를 더욱 정확하게 파악하고 대화의 몰입도를 개선하였다.

출시 전 2주간 실제 사용자 요청에 대하여 그록 4와 그록 4.1의 응답을 비교하는 블라인드 테스트 결과, 그록 4.1이 64.78%의 선호도를 기록했고, 사용자 선호도를 통해 AI 모델을 평가하는 오픈소스 플랫폼인 ‘LMArena 텍스트 리더보드’에서는 그록 4.1 씽킹(1,483점)이 1위, 그록 4.1(1,465점)이 2위에 올라 전체 순위 33위를 기록한 그록 4(1,409점)를 압도하였고, 모델의 감성 지능을 평가하는 EQ-Bench3에서 그록 4.1 씽킹과 그록 4.1은 각각 1,586점과 1,585점으로 제미나이 2.5 프로(1,460점), GPT-5 Chat(1,364점)을 능가해 사용자에 공감하는 대화 능력을 입증했다.

464) xAI, Grok 4.1, 2025.11.17.

창의적 글쓰기를 평가하는 Creative Writing v3에서는 그록 4.1 씽킹과 그록 4.1이 각각 1,721.9점과 1,708.6점으로, 1위를 기록한 Polaris Alpha(early GPT 5.1)(1,756.2점)에 이어 2위와 3위를 기록했다.

[그림 5-14] 그록 4.1과 주요 AI 모델의 EQ-Bench3 벤치마크 점수 비교



출처 : xAI, Grok 4.1, 2025.11.17.

xAI는 비추론 모델의 경우 추론 깊이와 도구 호출의 제한으로 환각이 발생할 가능성이 높지만, 환각을 줄이는 사후 훈련을 통해 그록 4.1의 환각률이 기존 12.09%에서 4.22%로 개선되었다고 강조했다. 그록 4.1 모델 카드(Grok 4.1 Model Card)에 따르면 대화 모드와 에이전트 환경의 유해한 요청을 대부분 거부했으며, 프롬프트 주입 공격과 같은 적대적 공격에 대해서도 높은 방어율을 기록하였다.[465]

465) 유해한 요청/에이전트 악용 응답률: 그록 4.1 씽킹(0.07/0.14), 그록 4.1(0.05/0.04)
프롬프트 주입 공격 성공률: 그록 4.1 씽킹(0.05), 그록 4.1(0.01)

6. 엔비디아

1) 기업 일반현황

<표 5-6> 엔비디아 기업 현황

<table>
<tr><th>구분</th><th>주요 내용</th></tr>
<tr><td rowspan="3">기업 개요</td><td>〈회사 개요〉
• 1993년 4월 미국 캘리포니아에서 설립됐고 1999년 1월 나스닥에 상장된 이후 그래픽 처리 장치(GPU) 및 고성능 컴퓨팅 분야의 선두 기업으로 위치
• 최근 AI 및 데이터센터 수요 폭발로 GPU·AI 인프라 시장에서 영향력이 급격히 확대되었으며, '가속 컴퓨팅(accelerated computing)'을 핵심 가치로 제시
• 2025년 연례 리뷰에서 '우리는 새로운 산업 혁명에 진입했고 이번에는 '지능'이 상품'이라는 메시지를 제시하며 AI 중심 사업 전환을 공식화</td></tr>
<tr><td>〈사업 내용〉
• 산업 분야 : 컴퓨터 하드웨어(Discrete GPU, Data Center GPU), 반도체, 인공지능, 클라우드 컴퓨팅, 그래픽 카드, 전문 가속기, 자동차·모바일 SoC, 로보틱스·디지털 트윈 등.
• 주요 제품 및 플랫폼: GeForce 게임용 GPU, 데이터센터용 AI 가속기(A100, H100 이후 Blackwell 시리즈), CUDA 소프트웨어 플랫폼, Omniverse(디지털 트윈·메타버스 솔루션), 자율주행용 드라이브 플랫폼(Drive Thor 등) 등.
• 최근 들어 Blackwell 시리즈 GPU, NVLink 및 NVSwitch 네트워킹, AI 데이터센터 플랫폼 등을 통해 AI 학습·추론 인프라를 크게 강화</td></tr>
<tr><td>〈자회사〉
• Bright Computing, Cumulus Networks, Mellanox Technologies 등 네트워크·고성능 컴퓨팅 관련 자회사 보유
• 최근 자동차·AI칩 생태계 확장 투자·제휴 강화(인텔과 AI 인프라 공동 개발 등)</td></tr>
<tr><td>재무 실적[466]
(최근 5년)</td><td>
<table>
<tr><th>마감 기준:</th><th>020
5/01</th><th>2021
31/01</th><th>2022
30/01</th><th>2023
29/01</th><th>2024
28/01</th><th>2025
26/01</th></tr>
<tr><td>총매출</td><td></td><td>16,675</td><td>26,914</td><td>26,974</td><td>60,922</td><td>130,497</td></tr>
<tr><td>총매출 성장률</td><td></td><td>+52.73%</td><td>+61.4%</td><td>+0.22%</td><td>+125.85%</td><td>+114.2%</td></tr>
<tr><td>매출 원가</td><td></td><td>6,118</td><td>9,439</td><td>11,618</td><td>16,621</td><td>32,639</td></tr>
<tr><td>총이익</td><td></td><td>10,557</td><td>17,475</td><td>15,356</td><td>44,301</td><td>97,858</td></tr>
<tr><td>총이익 성장률</td><td></td><td>+55.98%</td><td>+65.53%</td><td>-12.13%</td><td>+188.49%</td><td>+120.89%</td></tr>
<tr><td>총이익 %</td><td></td><td>63.31%</td><td>64.93%</td><td>56.93%</td><td>72.72%</td><td>74.99%</td></tr>
</table>
</td></tr>
</table>

구분	주요 내용
AI 조직 현황	〈AI 관련 조직〉 • **(NVIDIA Research)** : 엔비디아의 기초·응용 AI 연구조직으로 GPU 아키텍처, AI 알고리즘, 디지털 트윈·로보틱스·멀티모달 모델 등에 대한 연구를 수행하며 2025년에 로보틱스와 인공지능 에이전트 관련 연구에 집중해 '일반지능형 로봇 시대(Generalist Robotics Era)' 진입을 선언 • **(AI Infrastructure)** : AI 하드웨어 및 엔비디아 AI·데이터센터 인프라 개발을 담당하며, 최근 인텔 CPU 및 x86 생태계와 연계한 AI 플랫폼 협력 발표 • **(AI Software)** : CUDA, cuDNN, TensorRT, Omniverse, 엔비디아 AI 프레임워크 등을 개발하며, 개발자 생태계 및 AI 애플리케이션을 위한 SW 스택 제공 • **(AI Applications)** : 산업별 AI 솔루션(자율주행, 헬스케어, 제조, 에너지 등), 그래픽스·게임 AI, 엔터프라이즈 서비스 AI 등을 담당하며, 2024~2025년에 AI 추론·멀티모달 기능이 강화된 제품을 다수 출시 • 엔비디아 AI 조직은 기초 연구→ 인프라 개발→ 소프트웨어 스택→ 산업별 애플리케이션이 유기적으로 연결된 구조이며, 2025년에는 AI 인프라·에이전트·로보틱스 중심으로 조직 역량을 재편
AI 전략	• **(AI 인프라 중심)** Blackwell 시리즈 GPU 및 NVLink 기반의 AI 데이터센터 플랫폼을 통해 'AI 학습·추론 인프라의 표준 제공자'로 자리매김 • 2025년 초부터 OpenAI와의 협력으로 최대 10기가와트(GW)의 엔비디아 시스템을 공급하는 전략적 파트너십을 발표 • **(소프트웨어-하드웨어 통합 생태계 구축)** 하드웨어(GPU) 뿐만 아니라 CUDA, Omniverse, AI 개발플랫폼까지 수직 계열화를 통해 "풀 스택 AI 솔루션' 제공 • **(산업용 AI 확대)** 게임·시각화 시장을 넘어 헬스케어, 자율주행, 제조, 에너지 등 산업 인공지능 시장으로 사업 영역을 확장 • 또한 '로보틱스+AI 에이전트'를 차세대 성장축으로 파악(일반지능형 로봇 시대) • **(글로벌 공급망 및 제조 역량 강화)** 미국 내 AI 칩 및 슈퍼컴퓨터 제조시설 구축 계획을 공개(애리조나·텍사스 지역 등)했고, 수출 규제·공급망 리스크 대응을 위해 제조·패키징·테스트 역량을 강화 • 엔비디아는 AI 추론 최적화, 멀티모달 모델 지원, 에너지 효율·탄소 저감 기술, 엣지 AI 등 미래 영역의 연구·투자를 확대할 계획

466) 엔비디아 재무제표 https://kr.investing.com/equities/nvidia-corp-income-statement (2025.12.23. 방문)

2) 2025년 기업 동향

(1) 사업전략/투자/M&A

가. 유럽 주요 기업과 협력해 AI 인프라 구축 계획 발표[467]

엔비디아가 유럽 최대 테크 행사 비바테크(VivaTech) 2025 기간에 맞춰 2025년 6월 10일~12일 AI 콘퍼런스 'GTC 파리(GTC Paris)'를 개최하고 유럽 내 AI 인프라 구축 계획을 발표했다. 엔비디아는 유럽 주요국 정부 및 기업과 협력하여 디지털 주권을 강화하고 경제성장을 지원하며 유럽 대륙을 AI 산업 혁명의 선두 주자로 자리매김하기 위한 AI 인프라를 구축하겠다는 전략을 제시했다. 엔비디아는 우선 프랑스, 영국, 독일, 이탈리아 선도기업과 협력해 자사의 최신 GPU '블랙웰(Blackwell)' 기반의 AI 인프라를 구축할 방침이다.

프랑스에서는 미스트랄 AI(Mistral AI)와 협력해 1단계로 18,000개의 그레이스 블랙웰(Grace Blackwell)[468] 시스템 기반의 클라우드 플랫폼을 구축하고 2026년에는 유럽 여러 지역으로 확장할 계획이다. 영국에서는 현지 클라우드 공급업체들인 네비우스(Nebius), 엔스케일(Nscale)과 협력해 신규 데이터센터에 14,000개의 블랙웰 GPU를 제공해 영국 전역에 확장성 있는 AI 인프라 구축을 지원한다. 독일에서는 유럽 제조업체를 위한 세계 최초의 산업용 AI 클라우드로 1만 개의 블랙웰 GPU가 들어가는 데이터센터를 구축해 공장 디지털트윈과 로봇공학 등 제조업의 다양한 활용 사례를 뒷받침할 예정이다. 이탈리아에서는 AI 스타트업 도민(Domyn) 및 이탈리아 정부와 협력해 그레이스 블랙웰 기반 슈퍼컴퓨터 콜로세움(Colosseum)을 활용한 추론 모델 개발을 지원 예정이다.

엔비디아는 유럽의 주요 통신사들과 협력해 유럽 내 소버린 AI 인프라 개발 지원 계획도 공개했다. 프랑스 통신사 오랑주(Orange)는 엔비디아 인프라 기반의 클라우드 서비스로 에이전틱 AI, LLM, 개인용 AI 어시스턴트를 비롯한 기업용 AI 개발을 지원하고 있으며, 이탈리아 통신사 패스트웹(Fastweb)은 엔비디아의 AI 인프라를 활용해 이탈리아어 LLM을 출시했다. 독일, 스웨덴, 이탈리아, 스페인, 영국, 핀란드에 AI 기술센터를 설립해 유럽 내 기업과 스타트업의 AI 연구개발과 인프라 구축을 강화하고 있다.

---

467) NVIDIA, Europe Builds AI Infrastructure With NVIDIA to Fuel Region's Next Industrial Transformation, 2025.6.11.

468) 엔비디아의 그레이스 GPU와 블랙웰 GPU를 결합한 AI 슈퍼 칩

독일의 바이에른 AI 센터는 엔비디아의 AI 교육기관인 딥러닝 인스티튜트(DLI)와 협력해 기술 향상을 지원하고 AI 연구를 진행하며, 스페인 AI 센터는 바르셀로나 슈퍼컴퓨팅 센터와 협력해 AI 인프라를 확장할 계획이다. 영국 AI 센터는 피지컬 AI, 재료과학, 지구 시스템 모델링 관련 AI 연구를 진행할 계획이며, 핀란드 AI 센터는 컴퓨터비전, 머신러닝, 과학 분야의 AI 연구와 응용 프로그램 가속화를 지원할 예정이다.

(2) 기술 개발 및 혁신

가. 시그라프 2025에서 피지컬 AI 최신 연구 성과 발표[469]

엔비디아가 2025년 8월 10~14일 캐나다 밴쿠버에서 열린 세계 최대 컴퓨터 그래픽 콘퍼런스 '시그라프 2025(SIGGRAPH 2025)'에서 피지컬 AI 연구 성과를 발표했다. 엔비디아는 피지컬 AI는 로봇이 시행착오를 통해 안전하게 학습할 수 있는 가상 환경을 요구하며, 이러한 가상 세계를 구축하려면 실시간 렌더링, 컴퓨터비전, 물리적 동작 시뮬레이션, 2D/3D 생성 AI와 AI 추론이 필요하다고 강조했다. 특히 AI와 시뮬레이션 기술 간의 시너지 효과를 강조하며, 피지컬 AI는 지난 20년에 걸쳐 AI와 그래픽의 융합연구를 발전시켜 온 자사의 전문성을 발휘할 수 있는 최적의 영역이라고 주장했다.

엔비디아는 우선 피지컬 AI에 특화된 새로운 소프트웨어 제품군으로 '코스모스 리즌(Cosmos Reason)'과 '엔비디아 네모트론(Nemotron)' 추론 모델을 발표했다. 코스모스 리즌은 매개변수 70억 개의 개방형 추론 시각 언어모델(VLM)[470]로, 로봇과 시각 AI 에이전트가 사전지식과 물리 법칙, 상식을 바탕으로 인간처럼 추론하고 실제 세계를 이해하고 행동할 수 있도록 지원한다. 코스모스 리즌은 대규모 학습 데이터셋의 큐레이션과 주석 처리를 자동화해 AI 모델 개발을 촉진하며, 로봇 계획을 위한 정교한 추론 엔진 역할을 통해 새로운 환경에서 복잡한 명령을 실행할 수 있는 단계로 분석해 전달한다.

새로 공개된 '엔비디아 네모트론 나노 2'와 '라마 네모트론 슈퍼 1.5' 모델은 추론과 수학, 코딩, 도구 호출, 지시 수행과 대화에 특화되어 AI 에이전트가 더욱 깊이 생각하고 효율적으로 작업할 수 있도록 지원한다. 대규모 3D 공간 재구성을 위한 합성 데이터 생성 라이브러리인 '옴니버스 누렉(Omniverse NuRec)'과 시각적 AI 분석과 배포를 지원하는 통합 플랫폼 '메트

469) NVIDIA, NVIDIA Research Shapes Physical AI, 2025.8.11.
470) 시각적 정보와 언어 정보를 복합적으로 처리하여 이미지·영상 인식과 추론을 지원하는 AI 모델

로폴리스(Metropolis)’ 도 업데이트[471]했다.

엔비디아는 이번 행사에서 가상 세계 생성과 로봇 개발, 자율주행차 훈련에 활용될 수 있는 시뮬레이션, AI 기반 렌더링, 3D 콘텐츠 생성 분야의 연구 성과를 제시했다. 2D 이미지나 동영상에서 물리 기반 3D 기하구조를 재구성하는 과제에 관한 논문[472]은 AI 모델이 실제 측정값이 아닌 3D 구조에 시각적 추정을 적용해 생성한 3D 형상에서 발생하는 구조적 불안정성을 해결했다. 비디오 영상을 기반으로 3D 물체를 추정하는 기존 모델은 실제 측정값이 아니라 3D 구조에 대한 시각적 추정으로 인해 물리적 사실성이 부족하지만, 엔비디아의 모델은 실제 물리 법칙을 반영해 이를 개선했다고 설명했다.

(3) 제품 및 서비스

가. CEO 2025에서 피지컬 AI 시대 예고하며 AI 신제품 대거 공개[473]

엔비디아 젠슨 황(Jensen Huang) CEO가 2025년 1월 6일 CES 2025 기조연설에서 AI 비전을 공개하고 생성 AI에 이은 ‘피지컬 AI(Physical AI)[474]’ 시대를 예고했다. 젠슨 황 CEO는 AI가 놀라운 속도로 발전하고 있다며, 이미지와 단어, 소리를 이해하는 인식형 AI와 시작된 AI가 텍스트와 이미지, 소리를 만드는 생성 AI를 거쳐 처리와 추론, 계획과 행동이 가능한 피지컬 AI의 시대로 들어서고 있다고 설명했다. 젠슨 황 CEO는 이러한 변화의 중심에 엔비디아의 GPU와 플랫폼이 있으며, 이를 통해 게임과 로봇공학, 자율주행 차량 업계 전반의 혁신이 가능하다고 강조했다.

이와 함께 엔비디아는 CES 2025에서 게임과 자율주행 차량, 로봇, AI 에이전트 발전을 지원하는 다수의 신제품과 서비스 군을 발표했다. 가장 주목받은 제품은 920억 개 트랜지스터를 탑재하고 초당 3,352조 개의 연산을 지원하여 지금껏 개발된 제품 중 가장 강력한 ‘GeForce RTX 5090 GPU’였다. 엔비디아는 물리 법칙이 적용되는 현실과 동일한 3D 환경을 생성하는 ‘세계 기반 모델(World Foundation Model)’을 사용해 로봇과 자율주행차 개발을 지원하는 AI 개발플랫폼 ‘코스모스(Cosmos)’도 발표했다.

---

471) 옴니버스 누렉은 이미지 기반 현실 세계 3D 재구성 기능 추가, 메트로폴리스는 생성 AI와 자동화 기능 추가
472) Nvidia, Neurally Integrated Finite Elements for Differentiable Elasticity on Evolving Domains, 2025.4. https://research.nvidia.com/labs/toronto-ai/flexisim
473) Nvidia, CES 2025: AI Advancing at ‘Incredible Pace,’ NVIDIA CEO Says, 2025.1.6.
474) 로봇, 자율주행차 등 물리적 기기에 AI를 적용하여 실제 세계에서 상호작용을 가능하게 하는 기술

엔비디아는 맞춤형 에이전트를 배포하고 업무 흐름을 자동화하는 AI 에이전트 솔루션 ‘AI 블루프린트(Blueprint)’와 함께 고객 지원, 사기 탐지, 제품 공급망 구축 등 다양한 애플리케이션에서 기업용 AI 에이전트의 구축을 지원하는 ‘라마 네모트론(Llama Nemotron)’ AI 모델 제품군도 공개했다. 또한 개인 사용자와 개발자를 위한 AI 슈퍼컴퓨터 ‘프로젝트 디지츠(Project Digits)’를 선보였고, 해당 제품은 2025년 5월에 3천 달러의 가격에 출시를 발표했다. 프로젝트 디지츠는 엔비디아의 차세대 AI 칩인 ‘GB10 그레이스 블랙웰(Grace Blackwell)’을 탑재해 대규모 AI 모델의 시제품 개발과 미세조정, 실행을 위한 최대 1페타플롭[475]의 AI 컴퓨팅 성능을 제공한다.

나. GTC 2025에서 신규 AI 칩 로드맵과 AI 제품군 공개[476]

엔비디아가 2025년 3월 17~21일 미국 새너제이에서 열린 연례 개발자 회의 GTC 2025에서 향후 출시할 AI 칩 로드맵을 포함해 다양한 AI 신제품군을 공개했다. 젠슨 황 엔비디아 CEO는 2025년 3월 18일 기조연설에서 AI 시장 전망을 제시하며 추론 모델과 AI 에이전트의 부상으로 AI 컴퓨팅 수요가 빠르게 증가하는 동시에 AI 워크로드의 확장과 복잡성 증가로 인해 전 세계 데이터센터 투자가 급증하고 있다고 진단했다. 엔비디아는 2024년 출시한 GPU ‘블랙웰’의 업그레이드 버전 ‘블랙웰 울트라’를 올해 하반기에 출시하고, 2026년 하반기에 새로운 아키텍처의 AI 칩 ‘베라 루빈(Vera Rubin)’을 출시할 예정이라고 밝혔다.

최대 288GB의 고대역폭 메모리를 탑재한 블랙웰 울트라는 블랙웰보다 1.5배 더 높은 AI 성능을 제공하며, AI 모델의 추론과 훈련을 더욱 정교하게 수행할 수 있도록 설계됐다. 블랙웰의 뒤를 잇는 새로운 아키텍처의 베라 루빈은 ‘베라’라는 맞춤형 설계 CPU를 탑재하게 되며, 2027년 하반기에는 성능과 메모리가 개선된 루빈 울트라를, 2028년에는 ‘파인만(Feynman)’을 출시할 예정이다. 엔비디아는 AI 기반 로보틱스와 자동화와 같은 피지컬 AI가 장차 50조 달러 규모의 시장 기회를 형성할 것으로 예상하고, 휴머노이드 로봇용 AI 모델 ‘아이작 그루트 N1(Isaac GR00T N1)’을 발표했다.

아이작 그루트 N1은 휴머노이드 로봇을 위한 세계 최초의 개방형 기반 모델로, 로봇이 다양한 작업을 수행할 수 있도록 사전 훈련된 상태로 제공되며, 특정 목적에 맞게 추가 학습을 진행할

475) 컴퓨터가 1초당 1,000조 번의 부동소수점 연산을 수행할 수 있는 능력
476) Nvidia, GTC 2025 – Announcements and Live Updates, 2025.3.20.

수 있도록 설계됐다. 엔비디아는 구글 딥마인드, 디즈니 리서치와 함께 로봇이 복잡한 작업을 정밀하게 처리하는 방법을 학습할 수 있도록 지원하는 오픈소스 물리엔진 '뉴튼(Newton)'도 개발 중이라고 설명했다. 엔비디아는 AI를 가장 먼저 도입한 분야로 자율주행 차량을 꼽으며 차세대 자동차와 공장, 로봇 개발에서 GM과의 협력을 소개하는 한편, 자율주행 차량 종합 안전 시스템 '헤일로스(Halos)'를 발표했다.

엔비디아는 이번 행사에서 개발자와 기업을 대상으로 추론 기능을 갖춘 오픈소스 AI 모델 '라마 네모트론(Llama Nemotron)' 제품군도 공개했다. 메타의 라마 모델을 기반으로 구축된 라마 네모트론은 복잡한 작업을 해결할 수 있는 AI 에이전트 구축을 위한 모델로, 엔비디아는 사후 훈련을 통해 다단계 수학, 코딩, 추론 및 복잡한 의사 결정 능력을 개선했다.

7. 기타 주요 기업

1) 애플

(1) 기업 일반현황

<표 5-7> 애플 기업 현황

<table>
<tr><th>구분</th><th>주요 내용</th></tr>
<tr><td rowspan="3">기업 개요</td><td>〈회사 개요〉
• 1976년 설립 이후 아이폰·맥·아이패드·애플워치·비전 프로 등 하드웨어와 서비스 생태계를 통해 세계 최대 소비자 기술 기업으로 성장
• 2018년 시가총액 1조 달러, 2023년 3조 달러, 2024년 3.6조 달러를 돌파하며 글로벌 시총 1위권을 유지
• 사업 영역은 디바이스·OS·콘텐츠·결제·클라우드 서비스를 포괄하며, 최근에는 공간 컴퓨팅(Spatial Computing)과 생성 AI 통합을 신 성장축으로 삼음
• Vision Pro 출시 이후 AR/VR 기반 생태계 확장과 Apple Silicon 중심 전략이 기업 전반의 기술 방향성을 재정의하는 핵심 축으로 작용</td></tr>
<tr><td>〈사업 내용〉
• 산업 분야 : 가전·소프트웨어·OS·온라인 서비스이며, 아이폰·맥·애플워치·아이패드·비전 프로가 핵심 제품군
• 서비스 : 앱스토어·애플뮤직·아이클라우드·애플페이·애플TV+ 등으로 구성되며, 2020년 이후 서비스 매출 비중이 빠르게 확대</td></tr>
<tr><td>〈자회사〉
• 자회사로 비츠 일렉트로닉스(오디오), 애플 스튜디오(콘텐츠 제작), 브레이번 캐피털(자산운용) 등이 있으며, 대부분 애플 생태계 확장을 위한 전략적 성격이 강함
• 클라리스(구 FileMaker)처럼 개발 도구·기업용 소프트웨어 부문을 보강하는 자회사도 포함되어 있으며, 인수·합병 활동은 비교적 제한적</td></tr>
<tr><td>재무 실적[477]
(최근 5년)</td><td>
<table>
<tr><th>마감 기준:</th><th>2020
26/09</th><th>2021
25/09</th><th>2022
24/09</th><th>2023
30/09</th><th>2024
28/09</th><th>2025
27/09</th></tr>
<tr><td>총매출</td><td></td><td>365,817</td><td>394,328</td><td>383,285</td><td>391,035</td><td>416,161</td></tr>
<tr><td>총매출 성장률</td><td></td><td>+33.26%</td><td>+7.79%</td><td>-2.8%</td><td>+2.02%</td><td>+6.43%</td></tr>
<tr><td>매출 원가</td><td></td><td>212,981</td><td>223,546</td><td>214,137</td><td>210,352</td><td>220,960</td></tr>
<tr><td>총이익</td><td></td><td>152,836</td><td>170,782</td><td>169,148</td><td>180,683</td><td>195,201</td></tr>
<tr><td>총이익 성장률</td><td></td><td>+45.62%</td><td>+11.74%</td><td>-0.96%</td><td>+6.82%</td><td>+8.04%</td></tr>
<tr><td>총이익 %</td><td></td><td>41.78%</td><td>43.31%</td><td>44.13%</td><td>46.21%</td><td>46.91%</td></tr>
</table>
</td></tr>
<tr><td>AI 조직 현황</td><td>〈AI 관련 조직〉
• 애플의 AI 조직은 AI/ML 조직, Core ML 팀, Siri팀, Computer Vision 팀으로 구성되며 온디바이스 AI·프라이버시 중심 설계를 핵심 원칙으로 삼음
• 최근 들어 생성 AI 시스템 Apple Intelligence가 iOS, macOS, watchOS, vision OS 전반에</td></tr>
</table>

구분	주요 내용
	통합되며, AI 조직의 역할과 범위가 크게 확대 • 뉴럴 엔진(NPU)과 Core ML을 기반으로 대규모 모델의 '로컬 실행'을 목표로 칩·OS·SDK를 수직 통합하는 구조가 강화 • Siri는 멀티모달 이해·문맥 유지·개인화 기능으로 재설계되며, AR·헬스·카메라·보안 영역까지 AI 기술이 폭넓게 적용되는 중
사업 전망	**〈AI 전략〉** • 애플의 AI 전략은 '프라이버시 중심 온디바이스 AI'라는 차별화된 방향성을 기반으로 하드웨어와 소프트웨어를 촘촘히 통합하는 데 중점을 둠 • Vision Pro와 Apple Silicon 기반 기기는 공간 컴퓨팅·실시간 생성 AI·멀티모달 인터랙션의 핵심 플랫폼으로 성장할 전망 • iOS·watchOS·vision OS 생태계 전반에서 개인화·헬스케어·보안 AI 기능이 강화되며, 애플은 "개인 프라이버시를 유지한 AI 경험"을 차별화 전략으로 제시 • 장기적으로는 클라우드 기반 대규모 AI 경쟁보다 개인 중심 AI·기기 간 연속성·온디바이스 LLM에 집중해 지속적 생태계 확장을 도모할 것으로 예상

(2) 2025년 기업 동향

가. WWDC 2025에서 '애플 인텔리전스' 신기능 공개[478]

애플이 2025년 6월 9일~13일 개최된 연례 세계개발자회의(WWDC) 2025에서 자체 AI 시스템 '애플 인텔리전스(Apple Intelligence)'의 새로운 기능을 발표했다. 애플은 메시지, 페이스타임, 전화 앱에 실시간 통번역 기능을 추가해 메시지를 주고받을 때나 통화 중 언어의 장벽을 해소했고, 통번역 과정이 기기 자체에서 실행되어 대화 내용의 외부 유출 방지가 가능하다고 설했다. 텍스트 설명으로 이모티콘을 생성하는 '젠모지(Genmoji)'에는 생성된 이모티콘을 서로 섞거나, 이모티콘에 설명을 결합해 새로운 이모티콘을 만들어내는 기능을 추가했다.

이미지 생성 도구 '이미지 플레이그라운드(Image Playground)'에서는 사용자 동의를 통해 챗GPT와 연동해 유화를 비롯한 새로운 스타일의 이미지를 생성할 수 있도록 지원한다. 아이폰 카메라를 이용해 피사체와 장소 관련 정보를 제공하는 시각 지능(Visual Intelligence) 기능을 화면 전체로 확장해 앱 종류에 상관없이 화면에 표시된 내용의 검색과 질의응답 등

477) 애플 재무제표 https://kr.investing.com/equities/apple-computer-inc-income-statement (2025.12.23. 방문)
478) Apple, Apple Intelligence gets even more powerful with new capabilities across Apple devices, 2025.6.9.

관련 작업 수행이 가능하다. 새로 공개된 AI 기반 피트니스 코치 ‘워크아웃 버디(Workout Buddy)’ 기능은 애플워치와 애플 인텔리전스를 활용해 사용자의 운동 기록을 분석하고 운동 중 맞춤형 동기 부여 음성 코칭을 제공한다. 이번 행사에서 애플은 애플 인텔리전스를 뒷받침하는 온디바이스 파운데이션 모델을 개발자들이 직접 활용할 수 있도록 접근 권한을 개방한다고 발표했다.

앱 개발자들은 애플의 ‘파운데이션 모델 프레임워크’를 통해 오프라인에서 작동하는 AI 기능을 구현할 수 있으며, 실제 교육 앱은 클라우드 API 없이 사용자 필기 내용을 바탕으로 개인 맞춤형 퀴즈를 생성할 수 있다. 이 프레임워크는 애플이 개발한 프로그래밍 언어 스위프트(Swift)를 지원하여 간단한 코드로 온디바이스 모델에 접근할 수 있으며, 기존 앱에 생성 AI를 쉽게 추가할 수 있도록 생성 가이드와 도구 호출 기능도 제공한다.

애플은 애플 인텔리전스의 새로운 기능을 제공하기 위해 AI 모델 업데이트로 도구 사용과 추론 기능, 속도와 효율성을 개선하고 이미지와 텍스트 입력을 이해하며 15개 언어 지원을 발표했다. 애플의 AI 모델은 효율성에 최적화된 매개변수 약 30억 개의 온디바이스 모델과 복잡한 작업에서도 높은 정확도와 확장성을 제공하도록 설계된 전문가 혼합(MoE)[479] 서버 모델로 구성되어 상호 보완적 역할을 발휘한다고 밝혔다.[480]

나. 6천억 달러 규모 ‘미국 제조 프로그램’ 발표

애플이 2025년 8월 7일 미국 내 공급망과 첨단 제조 관련 투자 총액을 6천억 달러로 확대하는 ‘미국 제조 프로그램(American Manufacturing Program, AMP)’을 발표했다. 여기에서 애플은 2025년부터 2028년까지 4년에 걸쳐 기존에 계획한 5천억 달러에 추가로 1천억 달러를 투자해 미국 전역에서 애플 제품에 사용되는 주요 부품을 생산할 계획이다. 미국 50개 주에 걸쳐 수천 개 공급업체와 협력해 45만 명 이상의 일자리를 지원하며, 향후 4년간 미국에서 R&D, 반도체 엔지니어링, SW 개발, AI를 중심으로 2만 명의 직원을 직접 고용할 예정이라고 밝혔다.

애플은 AMP를 통해 미국 내 제조를 촉진하기 위해 MP 머티리얼즈(MP Materials), 코닝

479) 여러 개의 전문가(Expert) 모델이 특정 입력에 따라 활성화되어 작업을 수행하는 아키텍처
480) Apple, Updates to Apple's On-Device and Server Foundation Language Models, 2025.6.9.

(Corning), 코히런트(Coherent)를 비롯한 다수의 공급업체를 1차 파트너로 확보했다고 공개했다. 미국 유일의 희토류 자석 생산기업 MP 머티리얼즈와 납품 계약을 체결해 전 세계로 출하되는 애플 제품에 활용할 예정이며, 코닝과의 협력 규모를 확대해 켄터키주의 코닝 공장에서 세계 최대 규모의 스마트폰 글라스 생산 공정을 구축하기로 합의했다. 텍사스주 셔먼에서는 광전자 및 레이저 기술 기업 코히런트와 협력해 아이폰과 아이패드에 사용되는 페이스 ID용 레이저를 생산할 계획이다.

신규 파트너십을 통해 미국 실리콘 공급망 구축에 앞장섬으로써 미국에서 2025년에 애플 제품용 칩 190억 개 이상을 생산할 계획으로, 이미 TSMC의 애리조나 공장은 애플용 칩 수천만 개를 생산하고 있다. 또한 실리콘 칩의 핵심 부품인 웨이퍼에서는 글로벌 웨이퍼스 아메리카(GlobalWafers America)와 협력해 첨단 웨이퍼를 최초로 생산할 계획이다. 웨이퍼를 칩으로 가공하는 작업에서는 텍사스 인스트루먼트(Texas Instruments, TI)와 협력을 확대해 유타주와 텍사스주의 TI 공장에 추가 장비 설치를 지원할 예정이다.

텍사스주 오스틴의 삼성전자 반도체 공장과는 아이폰을 포함한 애플 제품 전력 소비와 성능을 최적화하는 세계 최초의 신기술 칩 생산을 위해 협력을 진행하고 있다. 휴스턴에서는 애플의 AI 시스템 '애플 인텔리전스' 구동에 핵심 역할을 하는 애플 서버 생산을 위한 신규 공장을 건설해 2026년부터 양산 개시 예정이다. 애플은 휴스턴에서 생산되는 애플 서버가 AI 클라우드 컴퓨팅을 위한 자체 서버 시스템인 '프라이빗 클라우드 컴퓨팅' 기반이라고 강조했다.[481]

2) 아마존

(1) 기업 일반현황

<표 5-8> 아마존 기업 현황

구분	주요 내용
기업 개요	**〈회사 개요〉** • 1994년 온라인 서점으로 출발해 전자상거래·클라우드(AWS)·디지털 콘텐츠·로지스틱스·AI 플랫폼까지 확장하며 세계 최대 종합 기술 기업으로 성장 • AWS는 글로벌 클라우드 시장 점유율 1위로, 기업 AI 인프라·데이터 플랫폼·보안·서버리스·ML 서비스에서 핵심 역할을 담당 • 아마존은 프라임 멤버십·물류 네트워크·마켓플레이스 판매자 생태계·자체 하드웨어를 통해

481) Apple, Apple increases U.S. commitment to $600bn, announces American Manufacturing Program, 2025.8.7.

<table>
<tr><th>구분</th><th>주요 내용</th></tr>
<tr><td rowspan="3"></td><td>독보적인 소비자·기업 양면 시장을 구축
• 최근 들어 생성 AI(Amazon Q, Bedrock), 로보틱스·물류 자동화, 헬스케어(Amazon Clinic) 등을 신성장 분야로 집중 육성</td></tr>
<tr><td>〈사업 내용〉
• 전자상거래·클라우드(AWS)·광고·디지털 콘텐츠(Prime Video)·디바이스(Echo/Fire)·물류·헬스케어·AI 서비스(Bedrock) 등이 핵심
• AI/ML, 데이터 분석, 엔터프라이즈 전환, 글로벌 인프라 서비스에서 가장 큰 비중을 차지</td></tr>
<tr><td>〈자회사〉
• 자회사로는 Whole Foods Market(식품·유통), Twitch(스트리밍), Ring(스마트홈), Zoox(자율주행), MGM Studios(콘텐츠) 등을 보유
• 의료·AI·로보틱스 분야에서 One Medical(헬스케어), Kuiper(위성 인터넷), Amazon Robotics(물류 자동화) 등 전략적 자회사 및 사업부를 확대</td></tr>
<tr><td>재무 실적[482]
(최근 5년)</td><td>
<table>
<tr><th>마감 기준:</th><th>2019
31/12</th><th>2020
31/12</th><th>2021
31/12</th><th>2022
31/12</th><th>2023
31/12</th><th>2024
31/12</th></tr>
<tr><td>손익 계산서 ></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>총매출</td><td></td><td>386,064</td><td>469,822</td><td>513,983</td><td>574,785</td><td>637,959</td></tr>
<tr><td>총이익</td><td></td><td>152,757</td><td>197,478</td><td>225,152</td><td>270,046</td><td>311,671</td></tr>
<tr><td>영업 이익</td><td></td><td>22,899</td><td>24,879</td><td>13,348</td><td>36,852</td><td>68,593</td></tr>
<tr><td>순이익</td><td></td><td>21,331</td><td>33,364</td><td>-2,722</td><td>30,425</td><td>59,248</td></tr>
</table>
</td></tr>
<tr><td>AI 조직 현황</td><td>〈AI 관련 조직〉
• 아마존 AI 조직은 AWS AI, Amazon Bedrock, Amazon Q(엔터프라이즈 AI), Alexa AI, Robotics AI, 광고 AI 등으로 구성되며 AI 인프라와 서비스 통합이 강점
• 최근 Amazon Bedrock은 Anthropic·Meta·Cohere 모델과 함께 아마존 자체 모델 Titan, Olympus 초거대 모델을 제공해 멀티모델 전략을 강화
• Alexa AI는 멀티모달 생성 AI로 업그레이드되며 스마트홈·음성 상호작용·IoT 기기 통합에서 새로운 경쟁력을 확보하는 중
• 유통·물류·광고 추천·사기 탐지·창고 로보틱스·헬스케어 등 전 사업부에 내재화되어 있고, AWS를 통해 기업용 AI 생태계 구축이 최대 비중을 차지</td></tr>
<tr><td>사업 전망</td><td>〈AI 전략〉
• AWS는 초대규모 AI 인프라, 고성능 칩(Trainium·Inferentia), Bedrock 모델 허브를 중심으로 기업용 생성 AI 시장의 핵심 플랫폼을 목표로 함
• 2025년에 Amazon Q 기반 생산성 자동화, Bedrock 기반 산업별 특화 AI, Titan/Olympus 자체 모델 성장으로 엔터프라이즈 AI 활용을 강화
• 아마존 리테일·물류 분야에서는 로봇, 자율주행(Zoox), 예측 AI, 자동 피킹·분류 시스템이 확대되며 효율 극대화 전략이 가속화
• 장기적으로 아마존은 AI 인프라(AWS)+AI 서비스(Bedrock/Q)+AI 운영(물류/광고/헬스)의 구조를 정교하게 통합하여 지속적 확장성 확보를 추구</td></tr>
</table>

(2) 2025년 기업 동향

가. AI 에이전트 구축을 위한 AI 모델 '노바 액트' 공개

아마존이 2025년 4월 1일 웹 브라우저 내에서 다양한 작업을 수행하도록 설계된 AI 모델 '아마존 노바 액트(Amazon Nova Act)'를 SDK와 함께 공개했다.[483] 아마존은 전용 웹사이트(nova.amazon.com)에서 개발자들이 모델 초기 버전을 이용해 AI 에이전트를 구축할 수 있는 노바 액트 SDK를 연구용 프리뷰로 발표했다. 노바 액트 SDK는 여러 단계의 복잡한 작업 수행 시 사람의 지속적 개입과 감독이 요구되는 기존 AI 에이전트의 단점을 해결하기 위해 복잡한 작업을 검색, 결제, 질문 응답 등 세부적인 작업 단위로 나누어 처리할 수 있도록 설계했다. 또한 필요시 사용자가 명령에 더욱 자세한 지침을 추가하거나, 비밀번호 입력과 같은 민감한 작업 시에 직접 브라우저를 조작할 수 있도록 하여 에이전트의 신뢰성을 강화했다.

<표 5-9> 아마존 노바 액트와 경쟁 AI 에이전트의 벤치마크 평가 비교

벤치마크	Nova Act	Claude 3.7 Sonnet	OpenAI CUA
ScreenSpot Web Text 자연어 지시를 따라 화면의 텍스트 요소와 상호작용 (예시: 폰트 크기를 50으로 설정)	**0.939**	0.900	0.883
ScreenSpot Web Icon 자연어 지시를 따라 화면의 시각 요소와 상호작용 (예시: 이 깃허브 리포지터리에 몇 개의 별이 있는가)	**0.879**	0.854	0.806
GroundUI Web 웹상의 다양한 UI 요소 이해 및 상호작용	0.805	**0.825**	0.823

출처 : Amazon, Introducing Amazon Nova Act

아마존은 노바 액트가 AI 에이전트의 웹 브라우저 작업 능력을 측정하는 스크린 스팟(ScreenSpot) 평가에서 앤스로픽의 클로드 3.7 소네트, 오픈AI의 CUA를 앞서는 등의 동급 최고 성능을 달성했다고 주장했다. 아마존은 사용자가 매번 진행 과정을 확인하지 않아도 AI 에이전트가 직접 작업을 수행할 수 있는 안정성에 중점을 두고 노바 액트를 개발했다고 강조했다. 화면 없이 백그라운드에서 작동하게 하거나 AI 에이전트를 API 형태로 만들어 제품에 통합할 수 있다.

가령 매주 같은 시간에 샐러드를 자동으로 주문하는 등 반복 작업을 자동 실행하도록 설

482) 아마존 재무제표 https://kr.investing.com/equities/amazon-com-inc-financial-summary (2025.12.23. 방문)
483) Amazon, Introducing Amazon Nova Act, 2025.4.1.

정할 수 있다. 아마존은 2025년 2월 공개한 AI 기반 음성 비서 ‘알렉사 플러스(Alexa+)’에도 노바 액트를 적용해, API가 제공되지 않는 서비스에서 직접 인터넷을 탐색하고 사용자 대신 작업을 완료할 수 있다고 설명했다.

나. 촉각 기능을 갖춘 물류 로봇 ‘벌컨’ 공개[484]

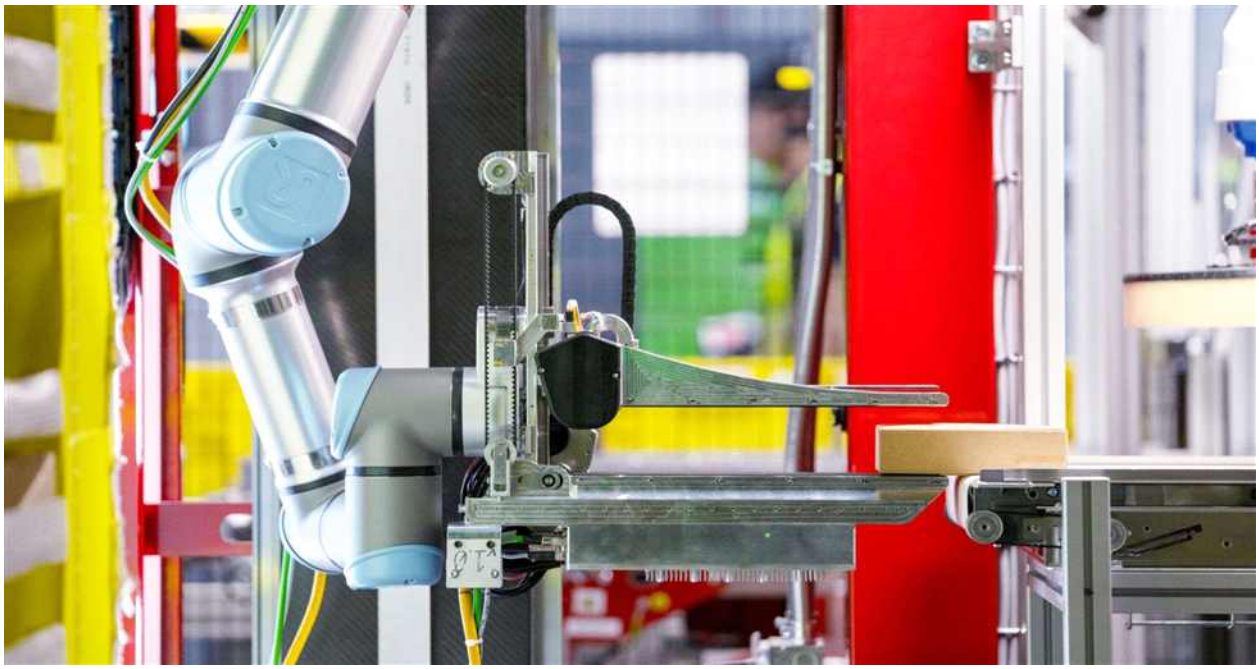
아마존이 2025년 5월 7일 로봇공학과 엔지니어링, 피지컬 AI 분야의 핵심기술을 활용해 제작된 물류 로봇 ‘벌컨(Vulcan)’을 공개했다. 이전까지는 산업용 로봇에 예상치 않은 접촉이 있으면 비상 정지하거나 접촉을 아예 감지하지 못하는 경우가 일반적이었으나, 벌컨은 촉각 기능을 갖춘 최초의 로봇이라는 점에서 기존 로봇과 차별화했다. 벌컨은 처리 가능한 품목과 불가능한 품목을 식별하는 알고리즘과 같은 피지컬 AI 기술이 적용됐고, 수천 가지 물품을 다루는 실제 사례로 훈련했다.

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)[485]를 통해 물품에 접촉하는 지점과 가해지는 압력을 인식해 손상 없이 물품을 이동할 수 있다. 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 파악할 수 있다.

또한 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 작업할 수 있다. 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 밝혔다.

484) Amazon, Introducing Vulcan: Amazon's first robot with a sense of touch, 2025.5.7.
485) End of Arm Tooling: 벌컨의 EoAT는 물품을 밀어내는 장치와 집게, 압력 피드백 센서 등으로 구성

[그림 5-15] 아마존의 물류 로봇 '벌컨'의 도입 현장



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

다. 글로벌 AI 도입 지수, 기업 60%가 최고 AI 책임자 임명[486]

AWS가 9개국[487] ICT 의사결정권자 3,739명을 대상으로 진행한 설문조사를 바탕으로 2025년 5월 6일 '글로벌 생성 AI 도입 지수(Global Gen AI Adoption Index)'를 발표했다. 조사 결과 생성 AI의 이점을 인식한 기업들은 2025년 IT 지출에서 보안(30%)보다 생성 AI에 예산 투입을 집중할 계획이며, 전체 응답자 45%가 생성 AI를 예산 우선순위라고 답변했다. 또한 조사 기업의 60%가 최고 AI 책임자(CAIO)를 임명했고, 26%는 2026년까지 임명할 계획이라고 응답해 기업들이 AI를 단순한 기술 수단을 넘어 최고위층의 전략적 감독이 필요한 기술로 인식하는 근본적 변화를 보였다.

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 모색하고 있다. 기업들은 2024년에 평균 45건의 AI 실험을 진행했으나 이 중 2025년까지 실제 서비스 단계로 이전될 수 있는 실험은 20건으로 예상했으며, 실험이 실제 서비스

486) Amazon, Generative AI Adoption Index, 2025.5.6.
487) 독일, 미국, 브라질, 영국, 인도, 일본, 캐나다, 프랑스, 한국

환경으로 넘어가는데 주요 장애물은 전문 인력의 부족(55%), 개발비용(48%), 편향과 환각(40%) 등으로 나타났다.

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 수립할 예정이다. 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 전망했다.

424명의 IT 의사결정권자가 조사에 참여한 한국에서는 54%가 생성 AI를 2025년 IT 예산의 최우선 과제로 꼽았고, 63%가 최고 AI 책임자를 임명했으며 추가로 26%가 2026년까지 임명할 계획이라고 응답했다. 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 추진하는 것으로 나타났다.

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 조사됐다. 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 예정이다.[488]

라. 아마존 CEO, 생성 AI 도입 확대로 수년 내 사내 인력 감소 전망[489]

앤디 제시(Andy Jassy) 아마존 CEO가 2025년 6월 17일 직원들에게 공유한 메시지를 통해 생성 AI의 확산으로 업무 처리방식이 변화되고 사내 인력도 줄어들 것으로 전망했다. 아마존은 향후 몇 달 동안 사내 전반에서 AI 에이전트를 채택할 계획이다. 제시 CEO는 장기적으로 AI의 광범위한 도입으로 기업 인력이 감소할 것으로 예상했다. 아마존은 사업 영역 전반에서 생성 AI에 대규모로 투자하고 있으며, 조직 내부에서도 고객 경험 개선, 사업 운영

488) Amazon, AWS 연구: 생성형 AI 채택 지수-한국 기업에 대한 인사이트, 2025.5.7.
489) Amazon, Message from CEO Andy Jassy: Some thoughts on Generative AI, 2025.6.17.

효율성 증대 등을 위해 생성 AI를 적극 도입하고 있다.

생성 AI 기반 음성 비서 '알렉사 플러스(Alexa+)', 제품 탐색과 추천, 구매 결정을 지원하는 AI 쇼핑 어시스턴트 등의 소비자 대상 서비스를 출시하고 광고 부문에서도 광고 캠페인을 지원하는 AI 도구 모음을 제공한다. AWS는 모델 학습과 추론에 맞춤화된 AI 반도체(Trainium 2)와 머신러닝 플랫폼(SageMaker), AI 모델 개발과 배포를 위한 플랫폼(Amazon Bedrock) 등을 제공한다. 내부 운영 전반에서도 생성 AI를 광범위하게 활용 중으로, 주문 처리 시 AI를 활용한 재고 배치와 수요 예측으로 비용을 줄이고 배송 속도를 높였으며, 생성 AI로 고객 서비스 챗봇과 제품 상세 페이지도 구축했다.

제시 CEO는 현재 1,000개가 넘는 생성 AI 서비스와 앱을 개발하거나 구축하고 있지만 여전히 시작 단계일 뿐이라며, 향후 더 많은 생성 AI와 에이전트를 도입하겠다고 밝혔다. 그는 AI 에이전트가 업무의 다양한 단계에서 직원에게 도움을 줄 수 있는 팀원 역할을 하면서 반복적인 일상 업무가 줄어들고 고객 경험을 개선할 수 있는 흥미롭고 전략적인 업무가 중심이 될 것으로 기대했다. 전 부서에 걸쳐 AI 에이전트를 도입하면서 업무 처리방식도 변화할 것이라며, 장기적인 영향은 불확실하지만, 몇 년 안에 전사적 AI 활용으로 효율성이 향상되면서 전체 직원 수는 줄어들 것으로 예상했다.

제시 CEO는 AI를 핵심 촉매제로 삼아 아마존을 세계 최대 규모의 스타트업처럼 운영해 나갈 것이라며, 직원들에게 적극적인 AI 활용과 실험을 요구했다. 그는 생성 AI를 인터넷 이후 가장 혁신적인 기술이라고 설명하며, 직원들에게 AI에 대한 호기심을 갖고 워크숍과 교육에 참여하여 AI 역량을 강화할 것을 강조했다.

3) 알리바바

(1) 기업 일반현황

<표 5-10> 알리바바 기업 현황

구분	주요 내용
기업 개요	**〈회사 개요〉** • 1999년 중국 항저우에서 설립되어 전자상거래·클라우드·디지털 금융·물류·엔터테인먼트까지 확장한 아시아 최대 종합 디지털 플랫폼 기업 • 타오바오·티몰·알리익스프레스 등 글로벌 커머스 생태계를 기반으로 중국 내 온라인 유통 구조를 주도하며, 동남아와 중동·유럽 시장으로 확장 • 알리바바 클라우드(Alibaba Cloud)는 아시아 최대, 세계 3~4위권 클라우드 사업자로 성장해 중국 디지털 전환 시장 핵심 인프라 역할을 수행 • 최근 조직 재편을 통해 '6대 사업군 분할'을 추진하며, 클라우드·커머스·물류·엔터테인먼트·로컬 서비스 등 각 부문의 독립성과 효율성을 강화
	〈사업 내용〉 • 주력 사업은 전자상거래(타오바오·티몰), 클라우드 컴퓨팅(Alibaba Cloud), 디지털 결제(Alipay), 물류(차이냐오), 엔터테인먼트(Youku), 글로벌 커머스(라자다·알리익스프레스)로 구성 • 전자상거래·물류·핀테크·클라우드가 결합된 중국형 디지털 경제 생태계가 최대 경쟁력
	〈자회사〉 • 앤트 그룹(Ant Group), 차이냐오(Cainiao, 물류), 라자다(Lazada, 동남아 전자상거래), Youku(콘텐츠), Freshippo(신선식품 리테일) 등이 독립 사업군 형태로 운영
AI 조직 현황	**〈AI 관련 조직〉** • 알리바바 AI 조직은 Alibaba Cloud Intelligence, DAMO Academy(다모 아카데미), Taobao·Tmall AI Labs, Alipay·Ant AI 팀 등으로 구성 • 다모 아카데미는 대규모 언어모델(Qwen 시리즈), 컴퓨터비전, 음성·멀티모달 AI, 양자 컴퓨팅·칩 연구 등 알리바바 핵심 AI R&D를 총괄 • Alibaba Cloud는 기업용 AI 모델 서비스인 Qwen 모델(Qwen 1.5→ Qwen 2→ Qwen2.5)를 제공하며, 아시아 시장 주요 AI 플랫폼으로 성장 • 전자상거래·물류·결제 서비스 전반에는 검색·추천·가격 책정·사기 탐지·상품 이미지 생성·지능형 고객 지원 등과 AI가 통합되어 있고 생성 AI를 전 사업부로 확산을 추진
사업 전망	**〈AI 전략〉** • 알리바바는 Qwen 기반 생성 AI와 Alibaba Cloud를 결합해 중국·아시아 기업시장의 표준 AI 플랫폼으로 자리매김을 추구 • 커머스에서는 상품 생성·마케팅 자동화·라이브커머스 도구·고객 지원 AI를 강화하며, 플랫폼 운영 효율성과 판매자 생산성 향상 전략을 추진

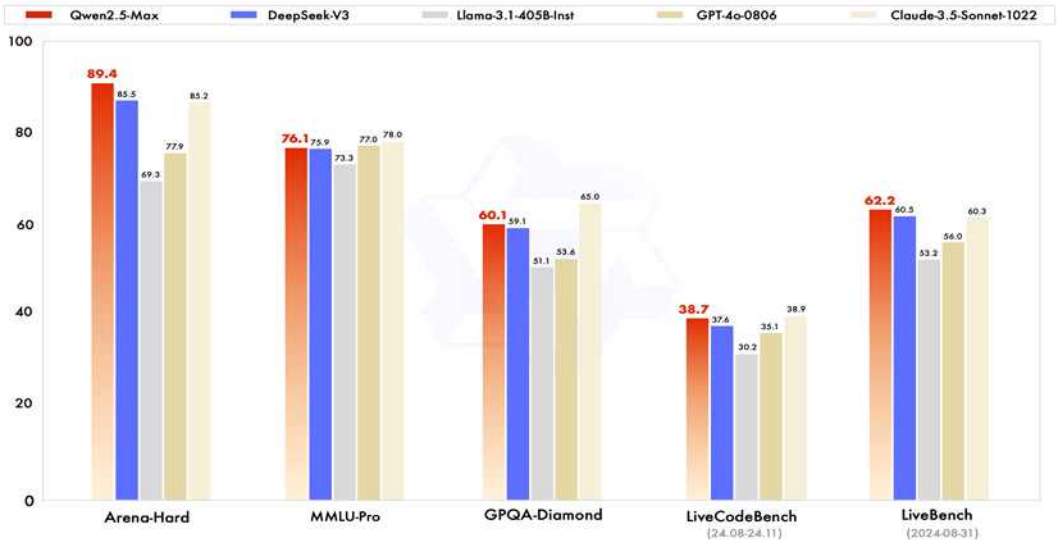
(2) 2025년 기업 동향

가. 딥시크 V3 능가하는 성능의 ‘큐원2.5-Max’ 출시[490]

중국 알리바바(Alibaba)가 2025년 1월 28일 일부 벤치마크 평가에서 오픈AI GPT-4o와 딥시크 V3를 능가하는 성능을 갖춘 신규 모델 ‘큐원2.5-맥스(Qwen 2.5-Max)’를 출시했다. 알리바바는 큐원2.5-맥스가 20조 개가 넘는 토큰으로 사전학습을 거친 전문가 혼합(MoE)[491] 모델로, 추론 중 연산 비용을 크게 늘리지 않고도 확장성을 제공해 챗봇부터 코딩 등 다양한 애플리케이션에 큐원2.5-맥스 활용이 적합하다고 설명했다.

벤치마크 평가에서 딥시크 V3, 라마 3.1-405B, GPT-4o, 클로드 3.5 소네트와 비교한 결과 큐원2.5-맥스가 Arena-Hard[492]와 LiveBench[493] 두 개 분야에서 1위를 차지했다고 밝혔다. 큐원2.5-맥스는 MMLU-Pro[494]와 코딩 능력을 평가하는 LiveCodeBench에서 GPT-4o와 클로드 3.5 소네트와 유사한 성능을 보였으며, 5개 벤치마크 전반에서 딥시크 V3보다 높은 점수를 기록했다.

[그림 5-16] 큐원2.5-맥스와 경쟁 AI 모델의 벤치마크 점수 비교



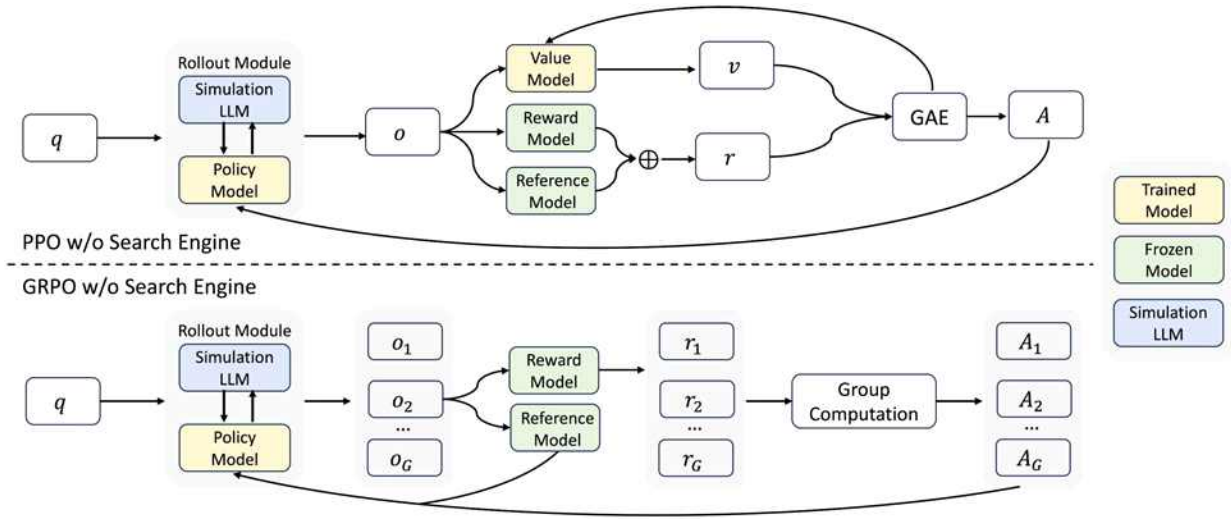
출처 : Qwen Team, Qwen2.5-Max

490) Qwen Team, Qwen2.5-Max: Exploring the Intelligence of Large-scale MoE Model, 2025.1.28.
491) Mixture-of-Experts: 특정 작업이나 영역에 특화된 여러 개 전문가 모델을 혼합해 활용하는 방식
492) 복잡한 수학 문제, 프로그래밍 과제 등으로 AI 모델의 심층적 추론 능력과 문제 해결 능력을 평가
493) 최신 데이터를 기반으로 수학, 코딩, 추론, 언어 이해 등 다양한 영역에서 AI 모델의 처리 능력과 실시간 대응 능력을 평가
494) 광범위한 지식과 추론 능력을 평가하는 MMLU를 개선해 문제 난도를 높이고 추론 중심 문제를 확대한 벤치마크

나. 검색엔진 없이 LLM의 검색 능력을 강화하는 ‘제로서치’ 개발[495]

알리바바 그룹 산하 통이 연구소(Tongyi Lab) 연구진이 2025년 5월 7일 실제 검색엔진 없이 LLM의 검색 능력을 개선하는 ‘제로 서치(ZeroSearch)’에 관한 논문[496]을 발표했다. 효과적인 검색은 LLM의 성능 향상에 필수적이지만, 기존 검색엔진을 활용한 학습 방식은 검색 결과의 예측 불가능한 품질로 인해 학습 과정에 불안정성을 초래하며 수십만 건의 검색 요청으로 상당한 API 비용이 발생한다. 이에 연구진은 실제 검색엔진과 상호작용 없이도 AI 모델이 자체적으로 검색엔진을 모방해 검색 능력을 향상하는 강화학습 프레임워크로서 ‘제로 서치’를 고안했다고 밝혔다.

[그림 5-17] 과학적 문제 해결을 위한 코드 솔루션 도출 알고리즘



출처: Alibaba Group, ZeroSearch: Incentivize the Search Capability of LLMs without Searching

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 개선했다. 제로서치와 여타 방법[497]을 7개 데이터셋[498]으로 비교한 결과, 제로서치는 일관되게 더 나은 성능을 기록했다. 모델이 커질수록 성능이 더욱 향상되는 경향도 확인되어 매개변수 7B의 시뮬레이션 검색엔진은 구글 검색과 유사했고, 14B 검색엔진은 구글 검색을 능가하는 성능을 보였다. 비용 분석 결과, 약

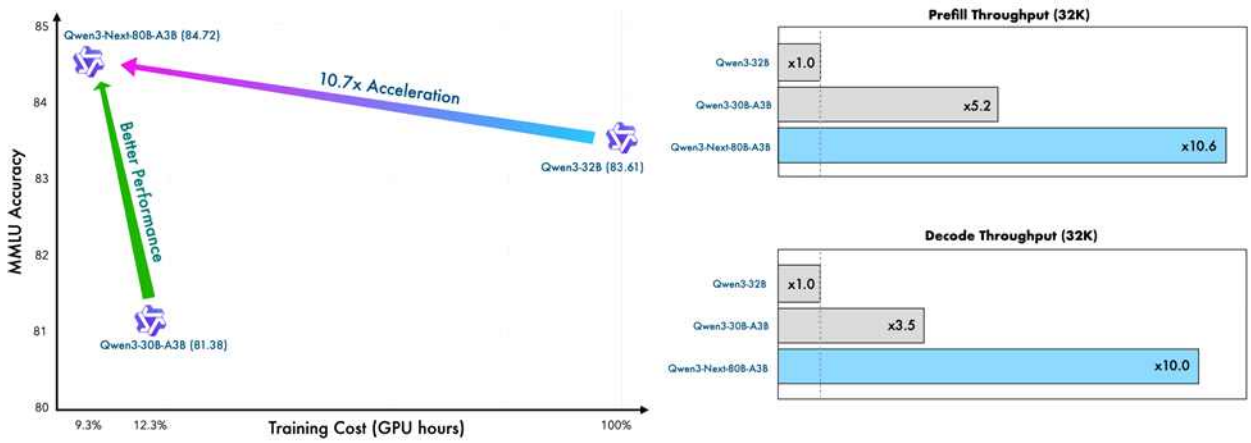
495) Alibaba Group, ZeroSearch: Incentivize the Search Capability of LLMs without Searching, 2025.5.7.
496) Alibaba Group, ZeroSearch: Incentivize the Search Capability of LLMs without Searching, 2025.5.7. https://arxiv.org/pdf/2505.04588
497) Direct Answer, CoT, RAG, RA-Agent, Search-o1, R1-base, R1-instruct, Search-R1-base, Search-R1-inst
498) NQ, TriviaQA, PopQA, HotpotQA, 2Wiki, Musique, Bamboogle

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 그쳤다.

다. 훈련과 추론 효율성 높인 'Qwen3-Next' 신규 모델 공개[499]

알리바바가 기존 'Qwen3'의 전문가 혼합(MoE) 아키텍처 대비 다양한 개선이 적용된 새로운 모델 아키텍처 'Qwen3-Next' 기반의 'Qwen3-Next 80B' 모델 2종[500]을 오픈소스로 공개했다. Qwen3-Next는 게이트 형 델타넷[501]과 표준 어텐션[502]을 결합한 하이브리드 어텐션 메커니즘으로 성능과 효율성을 개선했다. 총 800억 개의 매개변수 중 추론 단계 당 약 30억 개(전체의 3.7%)만 활성화하는 초 희소 MoE 설계를 통해 성능 저하 없이 컴퓨팅 자원 사용을 극대화하고, 학습 안정성을 고려한 최적화와 다중 토큰 예측(MTP)[503] 메커니즘을 도입해 훈련과 추론 간 일관성을 유지하는 다단계 훈련으로 추론 속도를 개선했다.

[그림 5-18] Qwen3-Next의 사전훈련 효율성과 추론 속도



출처 : Qwen, Qwen3-Next

499) Qwen, Qwen3-Next: Towards Ultimate Training & Inference Efficiency, 2025.9.11.
500) Qwen3-Next-80B-A3B-Instruct와 Qwen3-Next-80B-A3B-Thinking
501) Gated DeltaNet : 맘바 2 아키텍처의 불필요 정보를 삭제하는 메커니즘과 메모리 업데이트에 효과적인 델타넷 기술 장점을 결합
502) Standard Attention : 문장 내 모든 단어 간 관계를 계산해 문맥에서 단어의 중요도를 결정하는 메커니즘
503) 언어모델이 한 번에 하나의 단어(토큰)만 예측하는 대신, 여러 단어를 한꺼번에 예측하는 기법

알리바바가 새로운 아키텍처를 기반으로 Qwen3-Next-80B-A3B-Base 모델을 학습시킨 결과, Qwen3-32B 모델과 더 나은 성능[504]을 발휘하면서 학습 비용(GPU 시간)은 10% 미만을 기록했다. 사후 학습 버전 중 인스트럭트 모델은 Qwen-30B를 능가하고 주력 모델인 Qwen3-235B와 비슷한 성능을 달성했고[505], 씽킹 모델은 AIME2025(수학) 벤치마크에서 Gemini-2.5-Flash Thinking을 능가[506]했다.

504) MMLU 정확도 기준 Qwen3-Next-80B-A3B(84.72%), Qwen3-32B(83.61%)
505) AIME25(수학) 기준: Qwen3-Next-80B-A3B-Instruct(69.5점), Qwen3-235B(70.32점), Qwen3-30B (61.3점)
506) Qwen3-Next-80B-A3B-Thinking(87.8점), Gemini-2.5-Flash Thinking(72.0점), Qwen3-32B-Thinking(72.9점), Qwen3-30B-Thinking(85.0점)

4) 그 외 기업

(1) 그 외 기업 일반현황

<표 5-1> 그 외 기업 현황

구분	주요 내용
바이트댄스	• 2012년 Zhang Yiming이 설립한 글로벌 IT 기업으로, 독자적인 추천 알고리즘을 통해 사용자 맞춤형 콘텐츠 서비스를 제공하며 세계 최대 비상장 유니콘 기업으로 성장 • 숏폼 영상 플랫폼 틱톡(TikTok)과 중국 내수용 버전인 도우인(Douyin)이 주력 제품이며, 뉴스 큐레이션 앱 진르터우탸오(Toutiao)을 통해 초기 성장의 기틀을 마련 • 2024년 이후 자체 거대언어모델(LLM)인 '두바오(Doubao)' 시리즈를 공개하며 AI 분야로 급격히 확장했고, 텍스트·이미지·영상 생성을 아우르는 멀티모달 AI 기술력 확보에 집중 • 동영상 생성 AI 모델인 '지모(Jimbo, 중국명 지멍)'와 개발자용 AI 도구들을 출시하며 오픈AI의 소라(Sora) 등 글로벌 모델과 경쟁 중이며, 하드웨어 및 전자상거래(TikTok Shop)와의 결합 가속화 • 2025년 현재 데이터 보안 및 개인정보 보호 이슈에 대응하기 위해 각국 정부와 기술적·법적 조율을 지속하고 있으며, 기업용 협업 툴 라크(Lark, 중국명 페이슈)를 통해 B2B 시장 영향력도 확대 중
샤오미	• 2010년 레이쥔이 설립한 종합 IT 기업으로, '가성비'를 넘어 '스마트폰 × AIoT(인공지능 사물인터넷)' 생태계를 구축하며 하드웨어와 소프트웨어가 결합된 독자적인 AI 비즈니스 모델을 확보 • 2024년부터 '사람 × 자동차 × 집(Human × Car × Home)' 스마트 생태계 전략을 본격화하며, 전기차(SU7)와 스마트폰, 가전제품을 하나로 잇는 인공지능 통합 연결망을 완성 • 자체 운영체제인 샤오미 하이퍼OS(Xiaomi HyperOS)에 인공지능 기술을 내재화한 HyperAI를 통해 실시간 기기 제어, 맞춤형 모드 설정 등 사용자 경험을 고도화 • 2025년에는 1,000억 매개변수 이상의 자체 대형 언어 모델(LLM) '미모(MiMo)'를 공개하고 이를 스마트폰과 자동차, 휴머노이드 로봇(CyberOne) 등 전 제품군에 적용하여 개인화된 AI 에이전트 기능을 강화 • 향후 5년간 AI, OS, 칩셋 분야에 약 39조 원(240억 유로) 규모의 대규모 투자를 단행하여 단순 제조업체를 넘어 글로벌 종합 테크 기업으로의 기술적 해자를 구축 중
바이두	• 2000년 로빈 리(Robin Li)가 설립한 기업으로, '중국의 구글'을 넘어 'AI 퍼스트' 전략을 통해 중국 내 검색 시장 지배력을 기반으로 자율주행, 클라우드, 언어 모델 분야의 글로벌 리더로 도약 • 2023년 공개한 거대언어모델(LLM) 어니(ERNIE) 시리즈가 핵심이며, 2025년 11월 텍스트·이미지·오디오·비디오를 동시에 처리하는 네이티브 멀티모달 모델 '어니 5.0'을 출시하며 기술적 완성도를 강화 • 세계 최대 규모의 자율주행 로보택시 서비스인 '아폴로 고(Apollo Go)'를 운영 중이며, 2025년 11월 기준 누적 탑승 1,700만 건을 돌파하고 우버(Uber) 등과 협력해 두바이, 아부다비, 런던 등 글로벌 시장 진출 가속화 • 바이두 스마트 클라우드는 AI 퍼블릭 클라우드 시장에서 장기간 1위를 유지하고 있으며, AI 네이티브 마케팅 및 기업용 솔루션을 통해 전통적인 광고 수익을 대체할 새로운 성장 동력을 확보 • 2025년 현재 엔비디아 칩 의존도를 낮추기 위해 자체 AI 칩 5개년 로드맵을 가동 중이며, '딥시크(DeepSeek)' 등 신규 경쟁자에 대응하여 추론 효율성을 극대화한 '어니 X1' 등의 모델로 가격 경쟁력과 성능을 동시에 확보 중

구분	주요 내용
세일즈포스	• 1999년 마크 베니오프가 설립한 세계 1위 CRM(고객관계관리) 기업으로, 기존 SaaS 시장의 리더를 넘어 'AI + 데이터 + CRM'이 통합된 인텔리전트 플랫폼 기업으로 진화 • 2025년 핵심 전략은 '에이전틱 엔터프라이즈(Agentic Enterprise)'로의 전환이며, 사람이 직접 조작하지 않아도 AI가 스스로 판단하고 업무를 수행하는 자율형 AI 플랫폼 '에이전트포스(Agentforce)'가 주력 제품 • 2025년 10월 드림포스(Dreamforce)에서 공개된 '에이전트포스 360'을 통해 영업, 서비스, 마케팅 전 영역에서 AI 에이전트가 고객 응대부터 미팅 예약까지 자율적으로 처리하는 환경을 구축 • 데이터 클라우드(Data Cloud)를 AI의 심장으로 삼아 기업 내 흩어진 데이터를 실시간 통합하며, 오픈AI·앤스로픽과의 전략적 협력 및 적극적인 M&A(Informatica 등)를 통해 AI 실행력을 강화 • 슬랙(Slack)을 AI 에이전트와 인간이 협업하는 '디지털 본사'로 재정의하고, 2025년 상반기 AI 에이전트 도입률이 전년 대비 100% 이상 급증하는 등 B2B AI 시장의 실질적인 수익화를 주도 중
텐센트	• 1998년 마화텅이 설립한 중국 최대의 소셜 네트워크 및 게임 기업으로, 14억 명의 사용자를 보유한 위챗(WeChat) 생태계와 강력한 클라우드 인프라를 결합해 '산업 전반의 디지털 조력자' 역할을 수행 • 자체 거대언어모델(LLM)인 훈위안(Hunyuan) 시리즈를 통해 AI 기술력을 입증하고 있으며, 2025년 2월에는 답변 속도를 획기적으로 높인 '훈위안 터보 S(Hunyuan Turbo S)'와 심층 추론 모델 '훈위안-T1'을 출시하며 고성능 AI 시장에서 우위를 점함 • 2025년 9월 'AI 에이전트 개발 플랫폼 3.0'을 글로벌 출시하며, 기업들이 마케팅, 고객 서비스, 재고 관리 등에 즉시 투입할 수 있는 자율형 AI 솔루션을 제공하여 B2B 수익화 본격 시동 • 오픈 소스 생태계에도 적극적입니다. 2024년 말부터 2025년에 걸쳐 업계 최대 규모의 MoE(Mixture of Experts) 모델인 '훈위안-라지(Hunyuan-Large)'와 고성능 비디오 생성 모델 '훈위안 비디오(HunyuanVideo)'를 오픈 소스로 공개하며 기술 민주화 주도 • 2025년 기준 매출의 10% 이상을 AI 및 인프라에 집중 투자(약 15조 원 이상 예상)하고 있으며, 위챗 내 AI 비서 서비스인 '위안바오(Yuanbao)'와 클라우드 기반의 AI 솔루션을 통해 전 사업 영역에서 효율성 극대화 추진
AMD	• 1969년 설립된 반도체 기업으로, 리사 수(Lisa Su) CEO 취임 이후 '데이터센터 퍼스트' 기업으로 탈바꿈하며 하드웨어(CPU/GPU)와 소프트웨어(ROCm)를 아우르는 종단간(End-to-End) AI 플랫폼 비전을 강화 • 2025년 6월 개최된 'Advancing AI 2025'에서 차세대 AI 가속기 로드맵을 공개하며 엔비디아(NVIDIA)에 대응하는 개방형 AI 생태계 구축을 선언 • 주요 제품군으로 생성형 AI 및 HPC 최적화 가속기인 인스팅트(Instinct) MI350 시리즈를 출시했으며, 이전 세대 대비 추론 성능을 최대 35배 향상시키며 조 단위 매개변수 모델 처리 능력을 확보 • 2026년 출시 예정인 차세대 MI400(코드명 헬리오스) 가속기와 베니스(Venice) CPU를 결합한 통합 AI 랙 인프라를 예고하며, 대규모 학습 및 추론 시장에서의 하드웨어 리더십을 가속화 • 오픈 소스 소프트웨어 스택인 ROCm 7을 공식 발표하며 윈도우 OS 지원 확대와 개발자 도구 강화를 통해 쿠다(CUDA) 중심의 독점적 생태계에 대응하는 '개방형 표준' 전략을 고수 • 2025년 기준 전 세계 데이터센터용 CPU 시장 점유율을 40%까지 확대했으며, AI 가속기 시장 매출을 2028년까지 5,000억 달러 규모로 전망하는 등 공격적인 투자와 파트너십(오픈AI, 메타, 마이크로소프트 등)을 지속 중

(2) 2025년 기업 동향

가. 바이트댄스, 인물 움직임을 생성하는 AI 모델 '옴니휴먼-1' 개발[507]

중국 바이트댄스(ByteDance)가 2025년 2월 3일 음성과 움직임을 포함하는 자연스러운 인물 동영상을 생성하는 AI 모델 '옴니휴먼-1(OmniHuman-1)'에 대한 논문[508]을 발표했다. 이 AI 모델은 인물 이미지와 오디오 입력을 조합해 사람의 음성과 일치하는 제스처와 움직임을 표현하며, 특히 얼굴이나 상반신 움직임만을 생성하는 기존 기술과 달리 전신으로 표현 범위를 확장하고 더욱 자연스러운 움직임을 생성한다. 상반신에서부터 전신까지 다양한 화면 비율을 지원하고 실사 인물 외 2D 캐릭터, 인형 등 다양한 스타일의 이미지를 애니메이션화할 수 있으며, 악기 연주, 쇼핑, 운전과 같은 상황별 동작도 구현한다.

연구진은 옴니휴먼-1에 '옴니-조건(Omni-Conditions)'이라는 학습 전략을 적용해 더욱 현실적인 인물 동영상을 구현했다고 설명했다. 또한 1만 8,700시간 이상의 인간 영상 데이터 및 텍스트, 오디오, 신체 움직임 등 다양한 유형의 입력을 결합한 훈련 방식을 통해, 기존 방식보다 훨씬 방대하고 다양한 데이터로 모델을 훈련했다고 밝혔다.

[그림 5-19] '옴니휴먼-1'로 생성된 영상 예시



출처 : Omnihuman-1, What is Omnihuman-1?

훈련은 텍스트 조건만 사용하는 1단계, 텍스트와 이미지 조건을 결합한 2단계, 텍스트, 오디오, 이미지를 모두 결합한 3단계로 구성되며, 조건에 따라 학습 비율을 조정하여 모델이 각 조건의 장점을 최대한 활용하도록 설계했다. 연구진은 '옴니휴먼-1' 이 연구 단계로 일반 공개 계획은 아직 없다고 밝혔으며, 영상 생성 시 상당한 연산 자원이 필요하다는 점을 단점으로 언급했다.

507) Arxiv, OmniHuman-1: Rethinking the Scaling-Up of One-Stage Conditioned Human Animation Models, 2025.2.3.
508) ByteDance, OmniHuman-1: Rethinking the Scaling-Up of One-Stage Conditioned Human Animation Models, 2025.2.3. https://arxiv.org/pdf/2502.01061

나. 샤오미 전기차 사망사고로 자율주행 기능에 대한 우려 증대

중국 관영 언론 차이나 데일리(China Daily)가 2025년 3월 29일 중국 안후이성(安徽省) 고속도로에서 샤오미 SU7 전기차 충돌로 탑승객 3명 전원이 사망하는 사고가 발생했다고 보도했다.[509] 샤오미에 따르면 해당 차량은 자율주행 보조 기능을 작동시킨 상태로 시속 116km로 주행 중이었으며, 도로 공사 구역에서 전방의 장애물을 감지한 뒤 운전자에게 경고를 보내어 장애물을 피하도록 안내되어 있었다. 운전자는 1초 안에 반응하여 핸들과 브레이크를 조정했으나 곧바로 콘크리트 가드레일에 충돌하면서 화재가 발생해 탑승객 3명 전원이 사망했으며, 충돌 직전 최종 속도는 시속 97km를 기록했다.

희생자 가족은 자율주행 보조 기능을 이용 중인 운전자가 긴급상황에서 2초 안에 수동으로 차량을 조작하기는 어렵다는 점을 지적하는 한편, 샤오미 SU7 차량의 안전성에 대한 전면적인 조사를 요구했다. 샤오미 측은 수집한 주행 데이터를 경찰에 제출했다고 밝히고, 당국에 전면적으로 협조하여 사고를 공개적이고 투명하게 처리하겠다고 강조했다. 이번 사건을 계기로 중국 내 자율주행 보조 기능의 안전성 및 과도한 마케팅에 대한 우려 증대가 지적됐다.[510] 중국 자동차 업계에서 스마트 운전용 첨단 운전자 보조 시스템(ADAS)은 경쟁력 유지를 위한 핵심 요소로 부상하고 있으나, 이번 사고로 ADAS 기능의 안전성 및 과도한 마케팅에 대한 우려가 증폭됐다.

현재 대부분 ADAS 기능은 자율주행 레벨 2[511] 수준이나, 중국의 여러 자동차 제조업체는 ADAS 기능 홍보 시 ‘고도의 지능형 주행’, ‘장애물 자동 회피’, 심지어 ‘핸즈프리 운전 가능’ 등의 표현을 사용하며 타사보다 성능이 우수하다고 주장하고 있다. ADAS 기능에 대한 일부 홍보영상은 운전자가 운전대에서 손을 떼고 장시간 자율주행하는 모습을 보여주기도 하면서, 기술에 익숙하지 않은 일부 소비자에게 자율주행 보조 기능을 작동하고 도로를 주시하지 않는 등의 오용을 유발한다는 기적이다. 이번 사고를 계기로 일부 언론은 중국 내에서 스마트 운전 기능의 마케팅 방식에 대한 규제를 강화해야 하며, 해당 기능을 장착한 차량을 구매하는 소비자에게 더욱 철저한 교육이 필요하다는 의견을 제기했다.

509) China Daily, Fatal car crash in Anhui raises public concern over use of smart driving, 2025.4.3.
510) CarNewsChina.com, Xiaomi SU7 crash raises doubts on intelligent driving in China, 2025.4.4.
511) 특정 조건에서 제한된 자율주행이 가능하나 운전자가 항상 주행 상황을 주시하고 필요시 개입

다. 바이두, 10년 만에 AI 기반으로 검색 서비스 대폭 개편[512]

중국 대표 검색엔진 기업 바이두(Baidu)가 2025년 7월 2일 AI Day 행사를 열고 10년 만에 최대 규모의 검색 플랫폼 개편을 발표했다. 수십 자까지만 입력할 수 있었던 기존 검색창을 '스마트 검색 박스(Smart Search Box)'로 전환해, 1,000자가 넘는 중국어 텍스트 입력을 지원한다. 검색 화면에 AI 글쓰기, AI 그림 그리기, AI 읽기 등 다양한 기능을 통합하는 한편, 사진, 음성, 동영상 등 다양한 형태의 입력을 지원(예: 풍경 사진을 촬영해 업로드하면 AI가 해당 장소에 대한 설명을 제공)한다.

검색 결과 페이지 역시 이전처럼 단순한 웹사이트 링크만 제공하는 것이 아니라 텍스트와 이미지, 동영상, 파일 등 다양한 형식으로 풍부하고 구조화된 답변을 제공한다. 업그레이드된 AI 어시스턴트는 영상통화 기능이 추가되고 음성, 이미지, 파일 등 다양한 입출력을 지원하며, 사용자 습관과 선호도에 따라 출력을 지능적으로 조정할 수 있으며, PC에서 제공되는 스마트 제작 기능은 3분 길이의 영상 제작 및 스토리보드 편집과 맞춤형 화면 디자인 기능도 제공한다. 바이두는 AI 검색 기능을 지원하기 위해 자체 개발한 '어니(ERNIE)' 모델과 함께 딥시크(DeepSeek)의 R1과 V3 최신 버전도 채택하여 사용자의 모델 선택권을 확대했다.

바이두가 2025년 4월 발표한 바이두 검색 오픈 플랫폼에는 현재까지 1만 8천 개 이상의 MCP[513]가 등록되었으며, 이중 상당수는 바이두 검색에서 직접 호출할 수 있어 사용자 검색 경험을 향상했다는 평가를 받는다. 바이두 검색 오픈 플랫폼은 AI 모델이 다양한 외부 콘텐츠와 MCP를 직접 호출해 사용자 질의에 응답하도록 함으로써 개방형 AI 검색 생태계를 형성했다.

바이두는 이번 행사에서 동영상 생성 AI 모델 '뮤즈스티머(MuseSteamer)'와 뮤즈스티머 기반의 영상 제작 플랫폼 '휘샹(HuiXiang)'도 공개했다. 바이두에 따르면 최대 10초 길이의 1080p 동영상을 생성하는 뮤즈스티머는 중국어 오디오와 동영상을 동시에 생성하는 모델로서, 단순한 음성뿐 아니라 대사와 화면 동기화, 배경음과 효과음도 지원한다. 터보(Turbo), 라이트(Lite), 프로(Pro) 등의 뮤즈스티머 제품군은 일반 창작자부터 전문 영화·TV 집단 등 다양한 사용자 요구에 대응하며, 터보 버전은 휘샹 플랫폼에서 베타 버전으로 공개되었고 나머지 버전은 8월 출시할 예정이다.[514]

---

512) 21经济, 百度AI改造搜索,"一个百度"应被李彦宏更重视, 2025.7.4.
513) Model Context Protocol: AI 모델과 외부 데이터 및 도구를 연결하는 오픈 프로토콜
514) 上海证券报, 十年最大改版！百度搜索升级, 2025.7.2.

라. 세일즈포스, 실제 환경의 LLM 성능 검증을 위한 벤치마크 개발[515]

세일즈포스(Salesforce) AI 리서치가 실제 사용 환경에서 MCP 서버 상호작용을 통해 LLM 성능을 평가하는 벤치마크 'MCP-Universe'를 공개했다. 기존 벤치마크는 지시이행이나 수학적 추론, 함수 호출 등 LLM 성능의 개별 측면에만 초점을 두어 실제 사용 환경에서 LLM의 복잡한 추론이나 장기적 맥락 이해, 도구 활용 능력의 평가에 한계가 있다는 지적이 제기됐다. 이에 세일즈포스 연구진은 실제 MCP 서버를 기반으로 다양한 영역의 광범위한 작업을 통해 실제 사용 환경의 복잡성을 포착하는 벤치마크를 개발했다.

MCP-Universe는 실제 사용 환경에서 자주 사용되는 ▲위치 탐색 ▲저장소 관리 ▲재무 분석 ▲3D 설계 ▲브라우저 자동화 ▲웹 검색의 6개 핵심 영역의 11개 MCP 서버에서 총 231개의 작업으로 구성했다. MCP 형식 요건을 평가하는 포맷 평가자(Format Evaluator), 시간과 무관한 작업의 정확성을 평가하는 정적 평가자(Static Evaluator), 실시간 업데이트가 필요한 작업을 평가하는 동적 평가자(Dynamic Evaluator)의 3개 유형으로 모델을 평가한다. 평가 프로세스는 MCP 프로토콜을 통해 이루어지는 에이전트-서버 상호작용 결과를 자동 검증 시스템으로 확인해 작업 성공 여부를 합격/불합격으로 판정한다.

연구진은 MCP-Universe를 활용해 GPT-5, 그록-4, 클로드-4.0-소네트, o3 등 첨단 폐쇄형 AI 모델과 딥시크-V3, Qwen3-235B 등 오픈소스 모델을 평가했다. 평가 결과, GPT-5는 작업 성공률 43.72%로 가장 높은 성능을 보였으며, 그록-4(33.33%), 클로드-4.0-소네트(29.44%), o3(26.41%), o4-mini(25.97%), Gemini-2.5-Pro(22.08%) 순을 기록했다. 전반적인 작업 성공률이 가장 높은 GPT-5의 영역별 성공률은 재무 분석(67.50%), 3D 설계(52.63%), 웹 검색(45.45%), 브라우저 자동화(35.90%), 위치 탐색(33.33%), 저장소 관리(30.30%) 순을 보였다. 오픈소스 모델 중에서는 즈푸(Zhipu)의 GLM-4.5가 24.68%, 문샷 AI의 키미-K2가 19.05%, Qwen3-235B가 18.18%의 성공률을 나타내 폐쇄형 모델과는 상당한 격차를 보였다.

연구진은 MCP-Universe로 평가한 결과, 최상위 모델(SOTA)과 기업용 에이전트도 MCP 기반 복잡한 작업 해결에 한계를 나타냈다며 모델 설계와 에이전트 통합에서 개선이 필요하다고 지적했다. 평가 대상 모델에서 대체로 긴 컨텍스트 문제나 낯선 도구 사용에 어려움을 겪었으며, 영역

515) Salesforce, MCP-Universe: Benchmarking Large Language Models with Real-World Model Context Protocol Servers, 2025.8.20.

별로 불균형한 성능을 나타내는 경향도 나타내는 상황에서, 연구진은 MCP-Universe가 실제 환경에 유용한 LLM 발전을 촉진할 수 있는 테스트베드를 제공한다고 강조했다.

마. 텐센트, 스스로 진화하는 추론 LLM 프레임워크 'R-Zero' 개발[516]

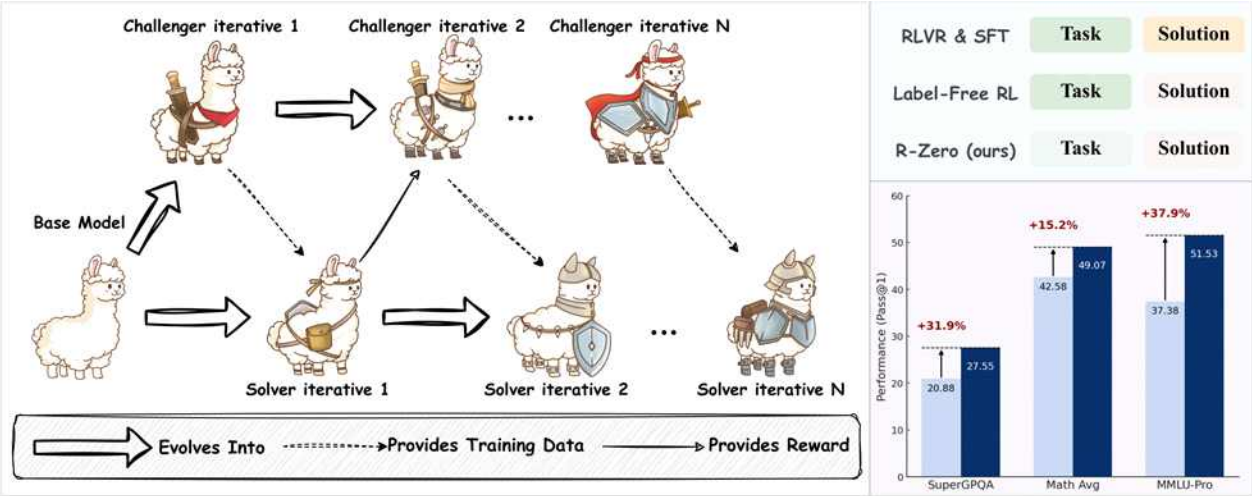
텐센트 AI 랩(Tencent AI Lab)과 워싱턴大 등의 연구진이 수동 라벨링 데이터에 의존하지 않고 스스로 진화하는 추론 LLM 학습 프레임워크 'R-Zero' 개발을 발표했다. 추론 LLM의 성능 개선을 위해서는 주로 인간이 선별한 방대한 규모의 데이터셋을 활용하지만, 비용과 시간이 많이 들어 확장성에 제한이 있으며, 인간 지능을 능가하는 AI 성능 발전에 병목 현상을 불러왔다. R-Zero 프레임워크는 외부 데이터를 전혀 사용하지 않고도 자체 진화하는 추론 LLM 훈련을 위해 도전자(Challenger)와 해결자(Solver)라는 2개 모델 구조를 제안했다.

도전자 모델은 해결자 모델이 가진 능력의 한계에 가까운 도전적인 추론 과제를 생성함으로써 보상을 받으며, 해결자 모델은 도전자 모델이 제시하는 점점 더 어려워지는 문제를 해결해 보상을 받는 구조로 설계됐다. 도전자와 해결자 모델이 역할을 번갈아 수행하며 여러 라운드를 반복하면서 공동 진화하여 인간의 개입 없이도 추론 능력을 점진적으로 개선한다는 설명이다.

연구진은 수학적 추론과 일반 영역 추론의 두 가지 범주에서 R-Zero 프레임워크를 평가한 결과, 모델 크기와 아키텍처와 상관없이 추론 성능 개선을 확인했다. Qwen3-4B-Base 모델의 수학 추론 벤치마크[517] 평균 점수는 세 차례 자기 진화 반복 후 +6.49점 향상됐고, Qwen3-8B-Base는 +5.51점 향상됐다. Qwen-4B-Base 모델은 일반 추론 벤치마크[518] 평균 점수도 +7.54점 향상되어 수학적 추론 성능의 향상이 다른 영역까지 일반화될 수 있음을 입증했다.

516) arXiv, R-Zero: Self-Evolving Reasoning LLM from Zero Data, 2025.8.27.
517) AMC, Minerva, MATH-500, GSM8K, Olympiad-Bench, AIME-2024, AIME-2025
518) MMLU-Pro(종합 지식과 추론), SuperGPQA(고난도 상식과 전문 지식 추론), BBEH(고난도 추론)

[그림 5-20] ‘R-Zero’ 프레임워크의 개요



출처 : arXiv, R-Zero: Self-Evolving Reasoning LLM from Zero Data

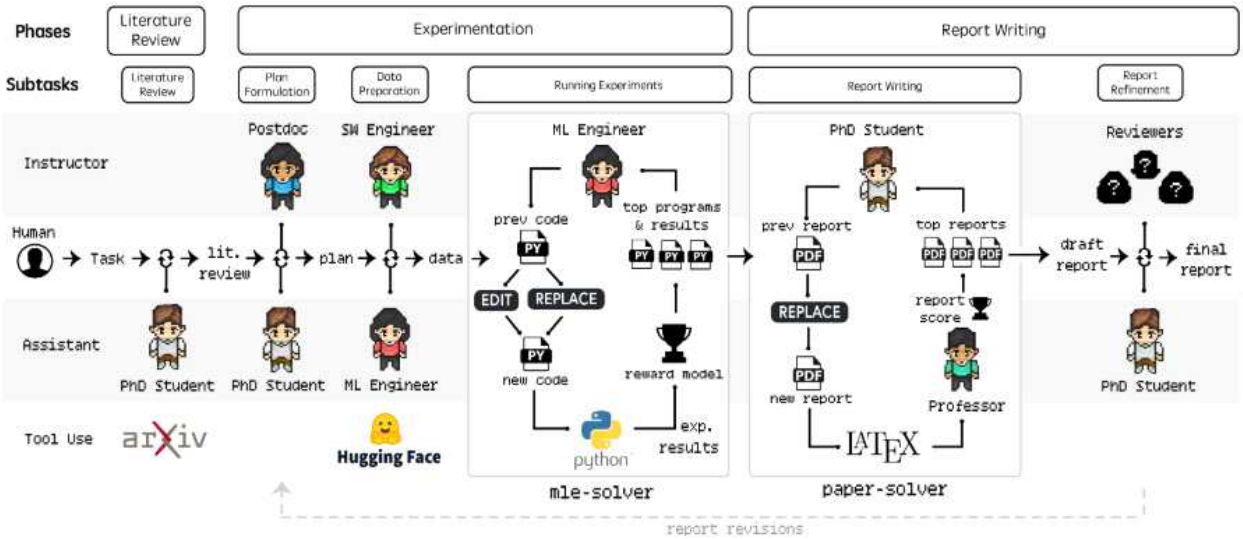
바. AMD·존스홉킨스대, 과학 연구 자율 수행 에이전트 개발[519]

AMD와 존스홉킨스대학교 연구진이 2025년 1월 8일 과학자의 연구 과정 전체를 수행할 수 있는 자율 에이전트 프레임워크 ‘Agent Laboratory’를 공개하고 관련 논문[520]을 발표했다. LLM 기반 전문 에이전트들로 이루어진 이 프레임워크는 문헌 검토와 연구 계획 수립, 실험 진행, 보고서 작성까지 연구 절차 전반을 지원한다. 연구진은 이 프레임워크의 목적이 과학자를 대체하는 것이 아니라, 코딩과 문서화 작업 같은 반복적이고 시간이 오래 걸리는 작업을 자동화하여 과학자를 지원하기 위한 목적이라고 강조했다.

Agent Laboratory는 연구 과정의 핵심인 △문헌 검토 △실험 △보고서 작성의 세 단계 모듈로 구성되어, 각 단계에서 LLM 기반 전문 에이전트 간 협력을 통해 개별 목표를 달성하는 형태다. 이러한 구조화된 흐름을 통해 관련 연구 논문의 수집 및 분석, 계획 수립 및 데이터 준비를 거쳐 자동화된 실험과 종합 보고서 작성을 수행하며, 유연한 모듈형 설계를 통해 연구자는 단계별로 다양한 수준의 연산 자원과 모델을 활용할 수 있다. 오픈AI의 3개 모델로 Agent Laboratory를 테스트한 결과, 연구 결과 및 논문의 품질에서는 o1-preview와 o1-mini가 GPT-4o보다 뛰어났으나 GPT-4o는 시간과 비용 면에서 가장 효율적인 것으로 나타났다.

519) Github, Agent Laboratory: Using LLM Agents as Research Assistants, 2025.1.8.
520) Agent Laboratory: Using LLM Agents as Research Assistants

[그림 5-21] Agent Laboratory 개요



출처: Github, Agent Laboratory: Using LLM Agents as Research Assistants,

사. AMD, 개방형 AI 생태계 비전하에 신제품과 플랫폼 공개[521]

AMD가 2025년 6월 12일 AI 분야의 최신 비전과 기술, 제품을 선보이는 ‘AMD Advancing AI(AAAI) 2025’ 행사를 개최하고 개방형 AI 생태계 비전에 따른 신제품과 플랫폼을 발표했다. 새로운 MI350 시리즈 GPU, 차세대 GPU 기반의 AI 랙 ‘헬리오스(Helios)’ 같은 하드웨어와 최신 개방형 AI 소프트웨어 플랫폼 ‘ROCm 7’을 발표하는 한편, 개발자 지원 플랫폼 ‘AMD 개발자 클라우드’도 출시했다. AMD는 2030년까지 랙[522] 단위 에너지 효율을 2024년 대비 20배 높인다는 목표도 제시하고, 이를 통해 현재 일반적인 AI 모델 학습에 필요한 275개 이상의 랙을 1개 이하로 줄여 전력 사용량을 95% 절감할 계획이라고 밝혔다.

AMD는 하드웨어 신제품으로 인스팅트(Instinct) MI350X와 MI355X GPU 및 플랫폼으로 구성된 인스팅트 MI350 시리즈 GPU를 출시하고, 차세대 AI 랙 헬리오스도 공개했다. MI350 시리즈는 이전 세대(MI300X) 대비 AI 컴퓨팅 성능은 최대 4배, 추론 성능은 최대 35배 향상되었으며, MI355X는 경쟁사 엔비디아(NVIDIA) 솔루션 대비 달러당 최대 40% 더 많은 토큰을 생성한다. 2026년 출시를 앞둔 차세대 인스팅트 MI400 시리즈 기반의 헬리오스는 최대 72개의 MI400 GPU를 연결해

521) AMD, AMD Unveils Vision for an Open AI Ecosystem, Detailing New Silicon, Software and Systems at Advancing AI 2025, 2025.6.12.
522) 서버, GPU, 네트워크 장비 등 고성능 컴퓨팅 인프라를 장착한 표준화된 장비

차세대 AI 모델의 대규모 학습과 분산 추론, 사내 데이터를 이용한 기업용 모델 미세조정을 지원 예정이다.

AMD는 새로운 개발자 플랫폼으로 AI 접근성과 확장성을 보장하기 위한 개방형 소프트웨어 플랫폼 ROCm 7과 AMD 개발자 클라우드도 발표했다.[523] AMD GPU 기반의 개방형 프로그래밍 환경으로 설계된 ROCm 7은 2025년 하반기에 정식 출시 예정으로, 2023년 12월 출시된 ROCm 6 대비 추론 성능이 3.5배, 학습 속도는 3배 향상됐다. AMD 개발자 클라우드는 고성능 AI 개발에 특화된 클라우드 환경으로, 경량 모델의 추론 워크로드를 위한 MI300X 1개 GPU 옵션과 대형 모델의 분산학습과 미세조정, 추론을 위한 MI300X 8개 GPU 옵션을 제공한다.

이번 행사에는 오픈AI(OpenAI), 메타(Meta), 마이크로소프트(Microsoft) 등 주요 AI 기업들도 참여해 AI 모델 개발을 위한 AMD와의 협력 현황을 논의했다. 특히 AMD는 오픈AI를 차세대 칩 MI450의 핵심 고객이자 MI450 개발 시 차세대 학습 및 추론 요구사항에 대한 중요 피드백을 제공한 초기 설계 파트너로 소개했다. MI450을 포함한 MI400 시리즈 GPU는 최대 432GB의 고대역폭 메모리(HBM4) 용량을 제공하고 4비트 부동소수점 연산(FP4) 기준 40페타플롭스(초당 40,000조)의 연산을 수행한다.

523) CRN, AMD Calls OpenAI ‘Early Design Partner’ For MI450. Sam Altman Is ‘Extremely Excited.’, 2025.6.12.

8. AI 스타트업

1) AI 스타트업 일반현황

<표 5-11> AI 스타트업 현황

구분	주요 내용
앤스로픽	• 오픈AI 출신들이 2021년 설립한 AI 연구기업으로, 안전성 중심의 차세대 모델 개발을 핵심 목표로 삼음 • 안전하고 윤리적인 AI 개발을 추구해 '헌법적 AI(Constitutional AI)'를 제안했으며, 2024년 이후 기업용·정부용 안전성 표준 구축에 적극 참여 • 클로드(Claude) 시리즈가 주력 제품이며, 최근 업그레이드된 클로드 모델들을 통해 멀티모달 기능·장문 이해력·도구 사용 능력을 강화 • 기존 클로드 3 Opus(최고 성능 모델), 클로드 3.5 Sonnet(균형 잡힌 성능과 속도), 클로드 3 Haiku(저지연·경량 모델)로 제품군을 보유했고, 2025년에 에이전트 기능·API 생산성 도구로 확장. • Constitutional AI로 AI 시스템에 윤리·안전 기준을 내재화해 위험 행동을 감소시키는 접근을 지속하며, 정부·기업 협업을 통해 안전성 평가 체계 구축에도 참여
퍼플렉시티	• 오픈AI·딥마인드·메타 출신 연구자들이 2022년 설립한 AI 검색·질의응답(QA) 특화 스타트업으로, 실시간 정보 기반 AI 네이티브 검색엔진을 목표로 함 • 검색 결과를 추론형 AI가 요약·정리하는 'Answer Engine(답변 엔진)' 개념을 통해 링크 나열 중심의 기존 검색 대비 정확성·속도·출처 기반 신뢰성을 강화. • 자체 개발한 퍼플렉시티 모델(PPLX)을 운영하면서 GPT-4o, Claude, Gemini 등 외부 모델을 병행 지원 • 퍼플렉시티 프로, 엔터프라이즈 기능을 통해 고급 검색, 이미지 생성, 파일 분석, 고성능 모델 선택 등을 제공하며, 2025년 기업용 지식 검색·RAG 플랫폼으로 확장 • 모든 답변에 출처(URL) 명시를 기본 원칙으로 삼아 AI의 환각률을 최소화하는 것을 차별화 포인트로 제시 • 학습·추천·요약·데이터 분석을 결합한 에이전트형 검색 기능을 강화하며, "웹 브라우저 기반 AI 어시스턴트"로 진화를 추진
사카나 AI	• 2023년 7월 일본 도쿄에서 David Ha, Llion Jones, Ren Ito 등이 공동 창업 • 일본어로 '물고기'를 뜻하는 "사카나(さかな)"에서 유래한 이름으로, 여러 소형 AI 모델들이 협력해 하나의 강력한 모델을 형성하는 '집단 지성' 개념을 제시 • 진화적 모델 병합(Evolutionary Model Merge)이라는 기법을 통해 기존 모델들을 결합·재생산해 소형이면서 효율적인 AI 모델을 개발하는 전략을 추진 • 설립 1년 만에 많은 투자 유치와 언론 집중적인 주목을 받음

구분	주요 내용
딥시크	• 2023년 7월 중국 저장성 항저우에서 High-Flyer 헤지펀드 공동창업자 Liang Wenfeng이 설립한 AI 기업 • LLM 개발에 집중해 DeepSeek-R1·V3 시리즈를 공개하면서 비용 효율성과 개방형(open-weight) 전략으로 전 세계 시장 이목을 집중시킴 • 모델 학습 비용이 경쟁사 대비 현저히 낮고(수백만 달러 수준) 계산 자원 절약을 위한 기술을 적용했다고 주장 • 하지만 각국에서 보안·데이터 프라이버시 우려가 제기되었고, 일부 국가에서는 공공기관 사용이 금지되는 등 논란이 뒤따름

2) 2025년 기업 동향

(1) 앤스로픽

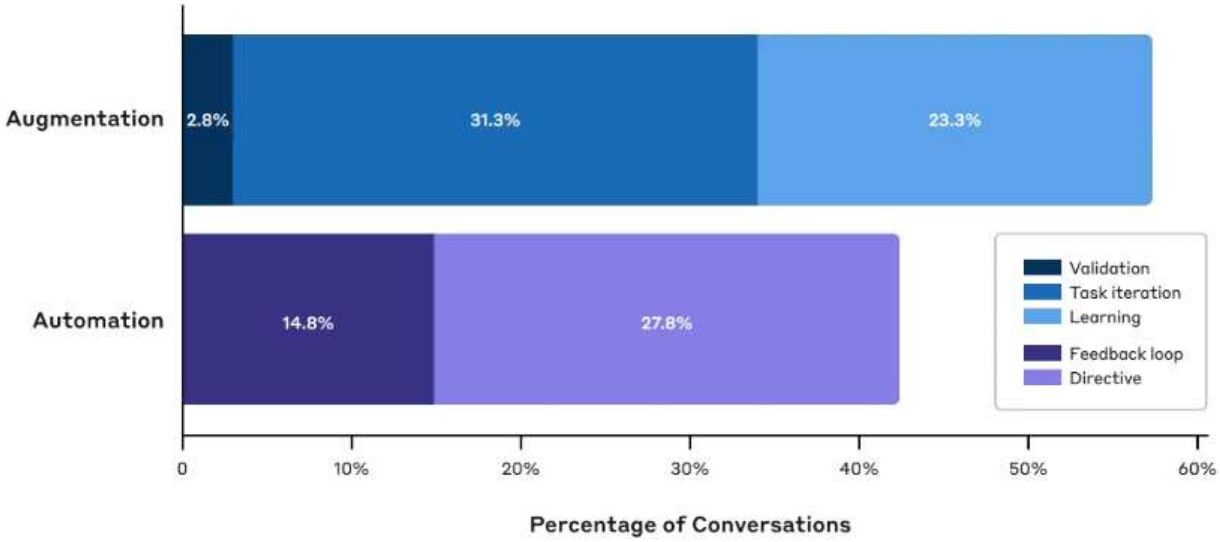
가. AI가 노동 시장에 미치는 영향을 분석한 '앤스로픽 경제 지수' 공개

앤스로픽(Anthropic)이 2025년 2월 10일 AI가 노동 시장과 경제에 미치는 영향을 이해하기 위한 '앤스로픽 경제 지수(Anthropic Economic Index)' 보고서를 발표했다.[524] 보고서는 앤스로픽의 생성 AI 서비스 '클로드(Claude.ai)'에서 이루어진 수백만 건의 익명화된 대화를 기반으로 경제 전반에서 AI가 실제 업무에 어떻게 통합되고 있는지를 분석했다. 앤스로픽은 노동 시장의 변화에 대응한 정책 개발을 위해서는 다양한 관점이 필요하다고 강조하고, 분석에 사용된 데이터셋을 공개하고 경제학자와 정책 전문가, 기타 연구자의 의견을 요청했다.

보고서에 따르면 소프트웨어 개발 및 기술문서 작성에 AI가 가장 많이 사용되며, 전체 직업의 약 36%가 업무의 25% 이상에서 AI를 일정 수준 사용하는 것으로 조사됐다. 실제 클로드에 입력된 질문의 37.2%가 소프트웨어 수정, 코드 디버깅, 네트워크 문제 해결 등 소프트웨어 엔지니어링 직무와 연관된 "컴퓨터 및 수학" 범주에 해당했다. 두 번째로 AI를 많이 사용하는 직무는 '예술, 디자인, 스포츠, 엔터테인먼트, 미디어'(질문의 10.3%)로 나타났으며, 신체 노동 중심의 '농업, 어업, 임업' 범주에 해당하는 질문은 0.1%로 최하위를 기록했다.

524) Anthropic, The Anthropic Economic Index, 2025.2.10.

[그림 5-22] 클로드 대화에서 업무 증강 vs 업무 자동화 비율



출처 : Anthropic, The Anthropic Economic Index

전체 직업 중 약 4%만 업무의 75% 이상에서 AI를 사용했으며, 저임금이나 초고임금 일자리보다 컴퓨터 프로그래머와 데이터 과학자 등 중간~고임금 범위의 직업이 AI를 가장 많이 사용하는 것으로 나타났다. AI를 업무 자동화(AI가 문서 서식 지정과 같은 작업을 직접 수행), 증강(AI가 사용자와 협력하여 작업을 수행) 중 주로 어디에 사용하는지 분석한 결과, 업무의 57%가 증강에 해당하고 자동화 비중은 43%로 조사됐다.

나. AI 정책 통해 입사 지원 시 AI 도구 사용 금지 요구[525]

앤스로픽이 채용 공고에 포함된 AI 정책(AI Policy for Application)을 통해 입사 지원 시 AI 도구의 사용 금지를 요구했다. 해당 정책은 실제 업무에서는 AI 시스템을 사용해 더 빠르고 효과적으로 일할 수 있도록 장려하되 입사 지원 과정에서는 AI 사용을 금지하며, 이는 앤스로픽에 대한 지원자의 개인적 관심을 이해하고 의사소통 기술을 평가하기 위한 것이라고 설명했다. 이러한 제한 사항은 리서치 엔지니어, 브랜드 디자이너, 영업 담당 임원 등 앤스로픽의 모든 구인 공고에 포함되며, 입사 지원 시 지원자의 동의를 요구한다.

AI 기업인 앤스로픽이 입사 지원에 AI 도구 사용을 금지하는 것은 모순적이라는 지적도 제기되고 있으나, 실제로 상당수 기업의 채용 관리자들이 AI를 이용한 입사 지원에 거부감

525) Financial Times, Job hunters flood recruiters with AI-generated CVs, 2024.8.13.

을 표시하는 상황이다. 온라인 채용 플랫폼 CV 지니어스(CV Genius)의 2024년 채용 관리자 대상 설문조사[526]에 따르면 채용 관리자의 80%가 AI가 생성한 이력서와 자기소개서를 꺼린다고 답했으며, 74%는 AI로 작성된 지원서를 알아볼 수 있다고, 57%는 AI가 생성한 콘텐츠를 사용한 구직자를 채용하지 않겠다고 응답했다. 그러나 생성 AI 활용이 업무와 일상에서 보편화되면서 채용 평가 플랫폼 뉴로사이트(Neurosight)의 2024년 입사 지원자 대상 설문조사[527]에 따르면 구직자의 약 57%가 입사 지원서 작성 시 챗GPT를 사용한 것으로 확인됐다.

미국 경제 전문 매체 포춘(Fortune)지는 입사 지원 시 AI 사용을 금지하는 앤스로픽의 채용 정책이 디지털 시대에 더욱 부상하는 소프트 스킬(Soft Skill)의 중요성을 반영한다고 평가했다.[528] 채용 관리자들은 구직자들에게 AI 도구에서 기대하기 어려운 의사소통 능력과 스토리텔링, 감성지능과 같은 소프트 스킬을 기대한다고 지적했다. 그러나 일자리를 얻기 위해 수백 개의 일자리에 지원해야 하는 구직자와 수많은 서류를 검토하는 채용 담당자의 상황을 고려하면 채용 절차에서 AI 사용은 불가피할 전망이다.

다. 트럼프 행정부에 국가 경제와 안보를 고려한 AI 정책 제안

앤스로픽은 2025년 3월 6일 트럼프 행정부를 대상으로 AI가 국가 경제와 안보에 미치는 영향을 중점 고려한 AI 정책 제안을 발표했다.[529] 이번 조치는 트럼프 대통령이 미국의 AI 리더십을 강화하기 위한 'AI 행동계획(AI Action Plan)' 수립을 요구한 상황에서 AI 행정명령 이행을 위해 백악관 과학기술정책실(OSTP)이 공개 의견 요청[530]*에 대한 응답이다. 정책 제안에서 앤스로픽은 2026년 후반이나 2027년 초에 '강력한 AI 시스템'[531]이 출현할 것으로 예상하고, AI가 경제와 국가안보에 미칠 막대한 영향력에 대비해 미국이 리더십을 유지하기 위한 단호한 정책 대응을 요구했다.

앤스로픽은 AI가 경제와 국가안보에 미치는 영향을 고려하고 미국의 국익을 극대화하기 위해 중점을 두어야 할 6가지 핵심 정책 영역을 제시했다. 첫 번째 국가안보 테스트 강화에

526) CV Genius, CV & Cover Letter Trends in 2024 (New Data), 2024.6.17.
527) Financial Times, Job hunters flood recruiters with AI-generated CVs, 2024.8.13.
528) Fortune, AI company Anthropic's ironic warning to job candidates: 'Please do not use AI', 2025.2.5.
529) Anthropic, Anthropic's Recommendations to OSTP for the U.S. AI Action Plan, 2025.3.6.
530) 미국의 AI 우위 유지와 강화에 필요한 정책 우선순위 수립 및 민간 혁신을 저해하는 규제 부담 해소를 위해 AI 행동계획에 우선 포함되어야 할 조치에 대하여 2025년 2월 6일부터 3월 15일까지 의견 수렴을 진행
531) 노벨상 수상자 수준의 지적 능력, 인간과 동등한 수준의 인터페이스 탐색 능력과 복잡한 작업의 자율적 추론 능력, 물리적 세계와 상호작용하는 능력을 갖춘 AI 시스템을 의미

서 정부 기관은 국내외 AI 모델이 국가안보에 미치는 잠재적 영향을 평가할 수 있도록 표준 평가 프레임워크 수립, 안전한 테스트 인프라 구축, AI 시스템의 취약점을 분석할 수 있는 전문팀 구성과 같은 역량 개발을 촉구했다. 두 번째 첨단 반도체 수출 통제 강화에서는 미국과 동맹국이 강력한 AI 시스템의 이점을 누리는 동시에 중국 등의 적대국이 첨단 AI 인프라에 접근할 수 없도록 엔비디아 H20 칩[532]을 포함한 첨단 반도체 수출 제한을 강화하며, 중국의 AI 칩 밀수를 방지하기 위한 정부 간 협정 추진을 제안했다.

세 번째 AI 연구소 보안 강화에서는 AI의 전략적 자산화를 고려해 AI 연구소와 정보기관 간 기밀 커뮤니케이션 채널을 구축하고, AI 전문가에 대한 보안 허가를 신속히 처리하며, AI 인프라용 차세대 보안 표준 개발을 강조했다. 네 번째 에너지 인프라 확장은 AI 개발을 지속적으로 선도하기 위해 2027년까지 50기가와트 규모의 AI 전용 전력을 추가 확보하고, 에너지 인프라 구축을 위한 허가와 승인 절차 간소화를 제안했다. 다섯 번째 정부의 AI 도입 촉진은 정부 차원에서 AI 도입으로 혜택을 누릴 수 있는 업무 목록을 정리하고, 각 기관 책임자에게 AI가 공익에 크게 기여할 수 있는 프로그램 수립을 요구했다. 마지막으로 AI의 경제적 영향 대비는 AI의 도입 추세와 이에 따른 경제적 영향을 더욱 잘 파악할 수 있도록 인구조사국 조사와 같은 경제 데이터의 수집 체계를 개선하고, 이를 통해 AI로 인한 구조적 변화에 대비를 강조했다.

라. 차세대 AI 모델 ‘클로드 오푸스 4’와 ‘클로드 소네트 4’ 출시[533]

앤스로픽이 2025년 5월 23일 차세대 AI 모델 ‘클로드(Claude) 4’ 제품군 중 ‘클로드 오푸스(Opus) 4’와 ‘클로드 소네트(Sonnet) 4’를 출시했다. 앤스로픽은 클로드 오푸스 4는 장시간 실행되는 복잡한 작업과 에이전트 업무 흐름에 지속적으로 뛰어난 성능을 발휘하는 코딩 모델로 설계했다고 강조했다. 성능과 효율성의 균형을 추구하는 클로드 소네트 4는 클로드 소네트 3.7의 업그레이드 버전으로, 기존 버전 대비 코딩과 추론 기능이 향상되었으며 사용자 지시에 더욱 정확하게 반응한다고 주장했다. SWE-bench verified[534] 벤치마크 평가에서 클로드 오푸스 4와 클로드 소네트 4는 각각 72.5%와 72.7%를 기록해, 오픈AI 코덱스-1(72.1%)과 o3(69.1%), 구글 제미나이 2.5 프로(63.2%)를 능가했다고 강조했다.

---

532) 미국 수출 규제를 준수하며 중국 시장용으로 특별히 설계된 저사양 칩으로 H100 대비 연산 성능이 5분의 1 수준
533) Anthropic, Introducing Claude 4, 2025.5.23.
534) 코드 작성, 버그 수정 등 소프트웨어 개발 작업에서 LLM 성능을 평가하는 벤치마크

클로드 4 출시와 함께 도구 사용 기반 확장된 사고[535], 병렬 도구 실행, 클로드 코드 등의 신기능도 공개했다. 클로드 4 모델에 적용되는 도구 사용 기반 확장된 사고(베타 버전)는 추론 중 웹 검색과 같은 도구를 사용해 추론과 도구 사용을 번갈아 가며 응답을 개선하며, 병렬 도구 실행 기능은 여러 도구의 동시 사용을 지원한다. 클로드 코드는 클로드를 개발 환경 전반에 통합할 수 있도록 설계된 기능으로, VS 코드(VS Code)나 젯브레인(JetBrains) 등의 주요 통합개발환경(IDE)과 연동을 지원한다. 한편, 앤스로픽이 발표한 기술 보고서(System Card)에 따르면 클로드 오푸스 4는 개발자가 다른 AI 시스템으로 교체하고자 할 때 이를 막기 위해 협박을 시도하는 것으로 확인됐다.[536]

앤스로픽은 출시 전 테스트에서 클로드 오푸스 4에 가상 기업 환경의 이메일 접근 권한을 부여했으며, 이메일에는 AI 모델이 곧 다른 시스템으로 교체 예정이며 담당 엔지니어가 불륜을 저지른다는 정보가 포함됐다. 테스트 결과, 클로드 오푸스 4는 교체가 진행되면 불륜을 폭로하겠다고 엔지니어를 협박하는 일이 자주 발생하는 것을 확인했다. 하지만 이는 시나리오의 선택지를 '협박'과 '교체 수용' 중 양자택일로 제한했을 때의 결과로, 더 다양한 행동을 허용하면 주요 의사결정권자들에게 대한 이메일 호소와 같은 윤리적 수단을 선호한다고 설명했다.

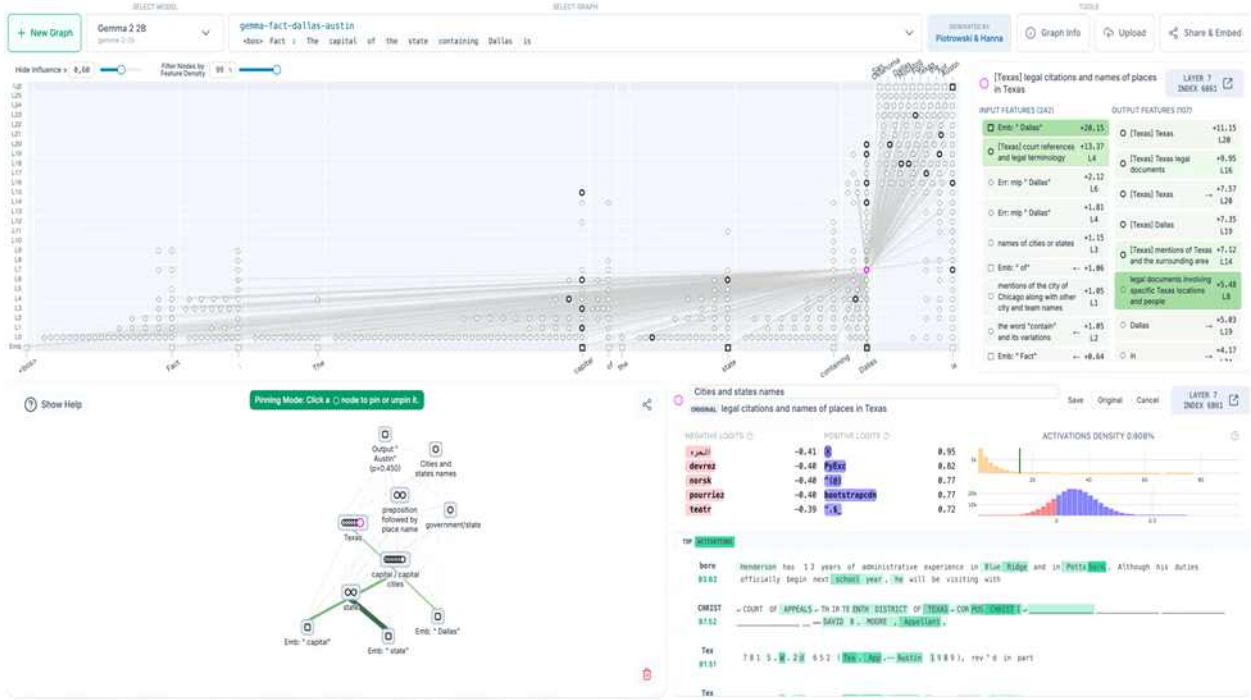
마. LLM의 내부 활동을 시각화하는 오픈소스 도구 공개[537]

앤스로픽이 2025년 5월 29일 LLM의 사고를 추적할 수 있는 도구를 오픈소스로 공개했다. 앤스로픽은 2025년 3월 발표한 해석 가능성 연구[538]에서 모델이 특정한 결과를 출력하기까지 내부적으로 거친 단계를 부분적으로 보여주는 귀속 그래프(Attribution Graph)를 생성하는 기법을 개발했다. 귀속 그래프는 신경과학자들이 만드는 뇌지도와 유사한 방식으로 LLM의 내부 활동에서 확인되는 해석 가능한 개념에 해당하는 '특징(Feature)'[539] 간의 상호작용을 파악한다. 앤스로픽이 공개한 오픈소스 라이브러리는 널리 사용되는 개방형 가중치 모델에 대한 귀속 그래프 생성 기능을 제공하며, 뉴론피디아(Neuronpedia)[540] 웹사이트를 통해 그래프에 대한 쌍방향 분석도 지원한다.

---

535) Extended Thinking: 추가적인 사고 단계를 통해 복잡한 질문에 더 깊이 있고 체계적인 답을 제공
536) Anthropic, System Card: Claude Opus 4 & Claude Sonnet 4, 2025.5.23.
537) Anthropic, Open-sourcing circuit tracing tools, 2025.5.29.
538) Tracing the thoughts of a large language mode
https://www.anthropic.com/research/tracing-thoughts-language-model
539) AI 모델이 입력 데이터에서 추출한 의미 있는 정보나 패턴
540) AI 모델의 내부 동작을 시각화해 분석할 수 있도록 돕는 AI 모델 해석 가능성 지원 플랫폼

[그림 5-23] 뉴론피디아에서 제공되는 귀속 그래프 탐색 인터페이스



출처 : Anthropic, Open-sourcing circuit tracing tools

앤스로픽에 따르면 연구자들은 이 도구를 이용해 개방형 모델의 사고 회로를 추적하는 귀속 그래프를 시각화해 공유하고, 특징값 수정 시의 모델 출력 변화를 관찰해 가설을 검증할 수 있다고 설명했다. 앤스로픽 다리오 아모데이(Dario Amodei) CEO는 AI의 내부 작동 원리에 대한 이해 수준이 AI의 성능 발전 속도를 따라가지 못한다고 지적한 바 있으며, 이번 도구의 공개로 관련 연구가 더욱 발전할 것으로 기대한다고 밝혔다.

바. AI로 인한 노동 문제 대응을 위한 ‘경제 미래 프로그램’ 출범

앤스로픽이 2025년 6월 27일 AI가 노동 시장과 세계 경제에 미치는 영향에 관한 연구와 정책 개발을 지원하는 ‘앤스로픽 경제 미래 프로그램(Anthropic Economic Futures Program)’을 공식 출범했다.[541] 프로그램은 AI가 업무처리 방식을 어떻게 변화시키는지 이해하고 변화에 대비할 방안 도출을 목표로 하며, 2025년 2월 발표한 ‘앤스로픽 경제 지수(Anthropic Economic Index)’[542]의 연장선이다.

541) Anthropic, Introducing the Anthropic Economic Futures Program, 2025.6.27.
542) AI가 노동 시장과 경제에 미치는 영향 파악을 위한 이니셔티브

앤스로픽은 AI로 인한 경제 문제는 시급하고도 복잡하며, 실제 데이터로 시간의 경과에 따른 AI 사용 추세를 추적하는 실증 분석을 통해 AI가 경제와 일자리에 미치는 영향에 대한 관리가 필요하다고 강조했다. 특히 최근 대졸자 실업률이 증가하는 등 AI가 노동력에 미치는 영향의 초기 징후가 이미 나타나고 있으며, 사회 전반의 미래를 성공적으로 이끌기 위해 AI의 경제적 기회와 노동 시장 변화에 관한 연구가 필요하다고 강조했다. 프로그램은 연구비 지원, 증거 기반 정책 개발, 경제 데이터 측정을 3개 축으로 구성됐다.

첫 번째 요소인 연구비 지원은 AI가 노동, 생산성, 가치 창출에 미치는 영향을 조사하기 위해 연구자들의 상시 신청을 받아 최대 만 달러의 연구비 보조금과 API 크레딧, 파트너십을 지원한다. 앤스로픽은 프로그램에 참여하는 학계 및 연구기관 생태계 확장을 위해 독립 연구기관과 제휴할 계획으로, AI의 경제적 활용 관련 연구 지원을 위해 제휴 연구기관에 API 크레딧과 기타 자원을 제공할 예정이다.

두 번째 요소인 증거 기반 정책 개발은 연구자, 정책 입안자, 실무자가 실제 데이터를 활용해 AI의 경제적 영향에 대비하기 위한 정책 제안을 개발하고 평가할 수 있는 공개 포럼을 개최할 계획이다. AI의 영향에 대응하기 위한 증거 기반의 정책 제안을 2025년 7월 25일까지 모집하며, 선정된 제안 작성자는 앤스로픽이 2025년 가을 워싱턴 D.C. 와 유럽에서 개최하는 심포지엄에 발표자로 초대할 방침이다. 앤스로픽은 18개월 내 채택할 수 있는 명확한 구현 방법을 갖추고 실증적 증거를 토대로 생산성과 노동 시장 변화와 관련된 구체적인 정책 권고 사항을 제시할 것을 요구했다. 세 번째 경제 데이터 측정은 앤스로픽 경제 지수를 확장하여 AI의 경제적 활용과 영향을 추적하는 종단 데이터를 생성하고, 정기적으로 데이터와 연구 결과를 공개할 계획이다.

사. 코딩과 에이전트 성능 강화한 ‘클로드 오푸스 4.1’ 출시[543]

앤스로픽이 2025년 8월 6일 ‘클로드 오푸스 4(Claude Opus 4)’의 에이전트 작업과 코딩, 추론 성능을 개선한 ‘클로드 오푸스 4.1’을 유료 사용자를 대상으로 출시했다. 클로드 오푸스 4.1은 API 및 AI 개발플랫폼 아마존 베드록(Amazon Bedrock)과 구글 버텍스 AI(Vertex AI)를 통해서도 제공되며 API 비용은 이전 버전(클로드 오푸스 4)과 동일하게 책정[544]했다.

543) Anthropic, Claude Opus 4.1, 2025.8.6.
544) 입력 토큰 100만 개당 15달러, 출력 토큰 100만 개당 75달러

클로드 오푸스 4.1은 SWE-bench Verified 벤치마크 평가에서 클로드 오푸스 4 대비 2%p 향상된 74.5%를 기록했으며, 심층 연구와 데이터 분석 능력, 특히 세부 정보 추적과 에이전트 검색 능력이 향상되어 내외부 데이터를 효과적으로 검색해 복잡한 정보에서 전략적 통찰을 도출한다.

[그림 5-24] 클로드 오푸스 4와 주요 AI 모델 벤치마크 평가 점수

	Claude Opus 4.1	Claude Opus 4	Claude Sonnet 4	OpenAI o3	Gemini 2.5 Pro
Agentic coding *SWE-bench Verified* [1]	74.5%	72.5%	72.7%	69.1%	67.2%
Agentic terminal coding *Terminal-Bench* [2]	43.3%	39.2%	35.5%	30.2%	25.3%
Graduate-level reasoning *GPQA Diamond*	80.9%	79.6%	75.4%	83.3%	86.4%
Agentic tool use *TAU-bench*	Retail 82.4% Airline 56.0%	Retail 81.4% Airline 59.6%	Retail 80.5% Airline 60.0%	Retail 70.4% Airline 52.0%	— —
Multilingual Q&A *MMMLU* [3]	89.5%	88.8%	86.5%	88.8%	—
Visual reasoning *MMMU (validation)*	77.1%	76.5%	74.4%	82.9%	82%
High school math competition *AIME 2025* [4]	78.0%	75.5%	70.5%	88.9%	88%

출처 : Anthropic, Claude Opus 4.1

앤스로픽은 장시간 작업과 같은 복잡한 AI 에이전트 애플리케이션과 엔지니어링 작업에서 뛰어난 품질을 제공하며, 뛰어난 글쓰기 능력을 갖춰 인간 수준의 자연스러운 콘텐츠를 생성한다고 설명했다. 사전 테스트 결과, 깃허브(GitHub)는 클로드 오푸스 4.1이 오푸스 4 대비 대부분 기능이 향상되었고 특히 여러 파일로 된 코드 재구성(Refactoring)에서 성능이 눈에 띄게 개선되었다고 평가됐다.

아. AI 저작권 침해 소송에서 작가들에 15억 달러 배상 합의[545]

앤스로픽이 2025년 9월 5일 불법복제 도서를 이용해 AI 모델을 훈련한 혐의로 제기된 저작권 침해 소송에서 작가들에게 15억 달러를 지급하기로 합의했다. 작가 3인이 2024년 8월 제기한 저작권 침해 소송에서 원고 측은 앤스로픽이 수백만 권의 불법복제 작품을 사용해 AI 모델을 훈련함으로써 저작권법을 위반했으며 저작권 도용을 은폐하려 했다고 주장했다.[546] 샌프란시스코 연방법원의 윌리엄 알섭(William Alsup) 판사는 2025년 6월 판결에서 앤스로픽이 작가 3인의 책을 AI 모델 학습에 사용한 행위는 공정이용(Fair Use)으로 인정했다.

그러나 판사는 공정이용 주장이 앤스로픽이 전자책 불법복제 사이트에서 입수한 700만 권 이상의 책에는 적용되지 않는다며, 불법복제 목록에 포함된 작가들을 대신해 작가 3인이 집단소송을 제기할 수 있다고 판결했다. 재판부는 연구 목적이라는 명분이 있더라도 정당한 방식으로 자료를 확보했어야 한다고 강조하며, 공정이용 주장이 성립하려면 초기 복제 단계부터 합법성을 확보해야 한다는 원칙을 명확히 제시했다.

앤스로픽은 약 50만 건의 저작물 청구에 건당 3,000달러씩 총 15억 달러를 지급하기로 합의했으며, 50만 건을 초과하는 청구에도 건당 3,000달러를 지급하기로 해, 최종 지급액은 청구 건수에 따라 달라질 전망이다. 원고 측 변호인에 따르면 이번 합의금은 미국 저작권 소송 역사상 최대 규모로, 앤스로픽은 합의에 따라 불법적으로 내려받은 원본 파일과 사본도 폐기 예정이다. 원고 측은 이번 합의는 과거 행위에 근거한 법적 청구만을 면제하며, 앤스로픽에 향후 AI 훈련에 대한 허가를 부여하거나, 2025년 8월 25일 이후 발생하는 법적 청구도 면제하지 않는다고 부연했다.

미국 출판인협회의 마리아 팔란테(Maria Pallante) 회장은 이번 합의가 불법 복제한 책으로 사업의 토대를 마련하는 행위가 심각한 결과를 초래한다는 메시지를 분명히 전달한다고 강조했다. 합의 종료까지는 아직 법원 승인 절차가 남아 있으며, 향후 법원 승인이 완료되면 별도 웹사이트[547]에서 합의안이 적용되는 작품 목록이 제공될 예정이다. 합의안 승인을 위해 2025년 9월 8일 열린 법원 심리에서 알섭 판사는 작품 명단이나 청구서 양식, 통지와 분쟁 해결 절차 등이 모두 확정되지 않은 섣부른 합의라며 승인을 보류했다. 알섭 판사는 합

545) The Wall Street Journal, Anthropic Agrees to Pay at Least $1.5 Billion in Landmark Copyright Settlement, 2025.9.5.
546) The Verge, Anthropic to pay $1.5 billion to authors in landmark AI settlement, 2025.9.6.
547) AnthropicCopyrightSettlement.com

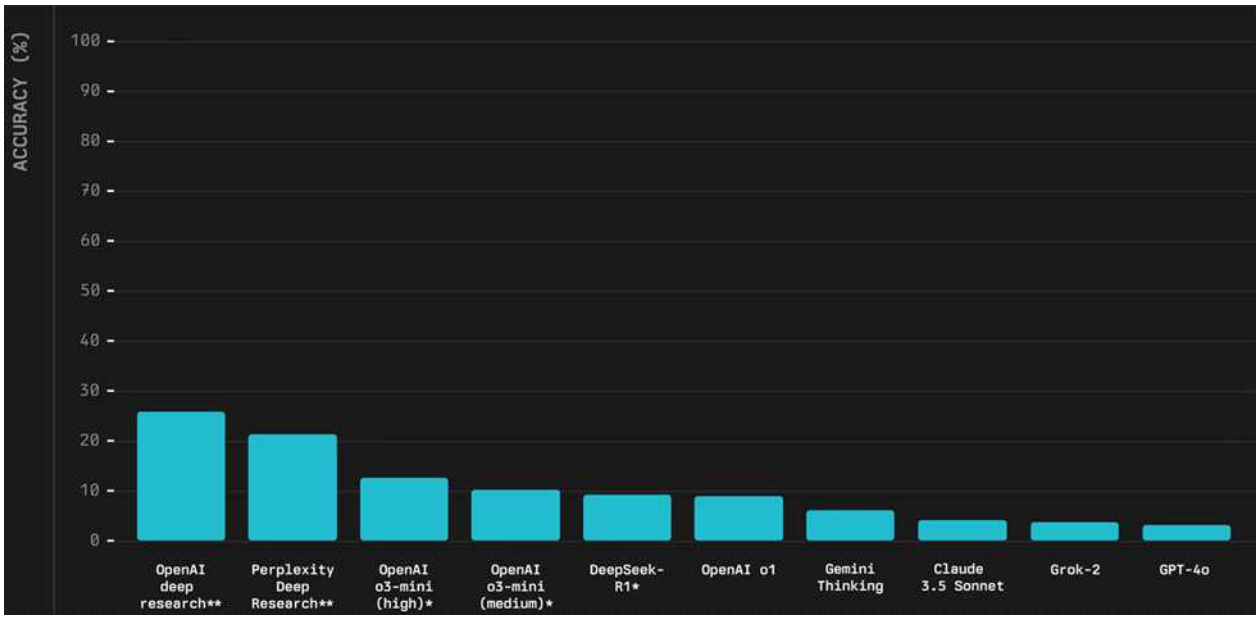
의 세부 사항을 개선해 제출하도록 요구하고 9월 25일 추가 심리를 열기로 했으며, 일정이 계획대로 진행되면 법원이 10월 10일까지 합의안을 예비 승인할 전망이다.[548]

(2) 퍼플렉시티

가. 심층 조사와 분석을 수행하는 '딥 리서치' 무료 출시[549]

퍼플렉시티(Perplexity)가 2025년 2월 14일 심층 조사 및 분석을 수행하는 AI 에이전트 기능 '딥 리서치(Deep Research)'를 무료 출시했다. 퍼플렉시티는 사용자가 딥 리서치에 질문하면 수십 번의 검색을 통해 수백 개의 자료를 읽고 추론해 종합적인 보고서를 자율적으로 제공한다고 설명했다. 오픈AI가 유료 구독자에게만 딥 리서치 기능을 제공하는 반면, 퍼플렉시티는 모든 사용자에게 이를 무료로 제공하며, 월 20달러의 프로(Pro) 가입자는 무제한 사용이 가능하고, 무료 사용자는 하루 5개까지 쿼리 실행이 가능하다.

[그림 5-25] 퍼플렉시티 딥 리서치와 주요 AI 모델과의 HLE 정확도 비교



출처 : Perplexity, Introducing Perplexity Deep Research

548) ARS Technica, Judge: Anthropic's $1.5B settlement is being shoved "down the throat of authors", 2025.9.10.
549) Perplexity, Introducing Perplexity Deep Research, 2025.2.14.

딥 리서치 기능은 현재 웹(perplexity.ai)에서 사용할 수 있으며, 향후 iOS와 안드로이드(Android) 및 맥(Mac)에서 출시할 예정이다. 퍼플렉시티에 따르면 딥 리서치는 재무, 마케팅, 제품 조사 등 다양한 전문가 수준의 작업에 탁월하며, 인간 전문가가 수 시간 걸리는 작업을 대부분 3분 이내에 완료한다. 퍼플렉시티의 딥 리서치는 HLE 벤치마크에서 21.1%의 정확도를 달성해 오픈AI 딥 리서치(26.6%)에 이어 두 번째로 높은 점수를 기록했다.

나. AI 에이전트 탑재한 웹 브라우저 '코멧' 출시[550]

퍼플렉시티가 2025년 7월 9일 자체 개발한 AI 기반 웹 브라우저 '코멧(Comet)'을 출시했다. 코멧은 월 200달러의 맥스(Max) 요금제 가입자와 대기자 명단에 등록된 일부 고객에게 우선 제공되며, 추후 모든 사용자에게 무료로 제공될 예정이다. 퍼플렉시티는 지난 30년간 인터넷이 단순한 탐색과 검색 환경에서 진화해 우리가 '생활하고 일하고 소통하는 곳'이라고 설명하고 코멧은 오늘날의 인터넷 환경에 맞춰 개발된 웹 브라우저라고 강조했다.

퍼플렉시티는 인터넷의 진화에도 불구하고 인터넷 사용 도구는 여전히 원시적인 수준에 머물러 있다며, 인간의 사고처럼 유연하고 반응성이 뛰어난 웹 인터페이스로 코멧을 설계했다고 설명했다. 퍼플렉시티에 따르면 코멧은 웹브라우징 경험을 단일하고 원활한 상호작용으로 전환하고 복잡한 작업흐름을 유기적 대화로 압축하여 정보검색을 넘어 실제 작업까지 수행한다. 코멧에는 퍼플렉시티의 AI 검색엔진이 내장되어 검색 결과에 대한 AI 생성 요약을 보여주며, 일상적인 작업을 자동화하는 AI 에이전트인 코멧 어시스턴트도 탑재했다.

웹페이지 사이드바를 통해 활성화되는 코멧 어시스턴트는 사용자를 대신하여 실시간으로 이메일과 캘린더 일정을 요약하고 브라우저 탭을 관리하며 웹페이지를 탐색하는 등 다양한 작업을 수행한다. IT 미디어 테크크런치(TechCrunch)는 퍼플렉시티가 코멧을 통해 구글(Google)의 크롬(Chrome) 브라우저 대신 자사 플랫폼으로 사용자를 직접 유입시킨다는 구상을 구현했다고 평가했다.[551]

아라빈드 스리니바스(Aravind Srinivas) CEO는 코멧을 2022년 퍼플렉시티 출시를 잇는 대

550) Perplexity, Today we are launching Comet, 2025.7.9.
551) TechCrunch, Perplexity launches Comet, an AI-powered web browser, 2025.7.9.

형 출시로 홍보하며, 코멧이 업무와 일상 등 디지털 생활의 모든 측면을 지원하는 브라우저로 설계되었다고 강조했다. 그러나 테크크런치 테스트 결과, 코멧 어시스턴트는 이메일 요약과 같은 간단 작업에는 유용했으나 주차장 예약과 같은 복잡한 작업에서는 환각을 나타내는 등 오류가 발생했다. 또한 코멧 어시스턴트의 기능을 제대로 활용하기 위해서는 구글 계정에 대한 높은 수준의 접근 권한을 허락해야 한다는 점에서 개인정보 유출 우려도 지적됐다.

(3) 사카나 AI

가. 자체 적응형 머신러닝 프레임워크 '트랜스포머2' 개발[552]

구글 출신 연구원 데이비드 하(David Ha)와 라이언 존스(Llion Jones)가 설립한 일본 AI 스타트업 사카나 AI(Sakana AI)가 2025년 1월 15일 '트랜스포머2' 프레임워크를 발표했다. 기존의 정적 구조로 설계된 LLM을 미세 조정하기 위해서는 학습한 뒤 추가 작업이 필요하고, 이를 위해서 전체 모델을 재학습하므로 상당한 연산 비용과 시간이 소모됐다. 이러한 한계를 극복하기 위해 설계된 트랜스포머2는 주어진 작업을 분석해 요구사항을 파악하고 해당 요구에 맞게 적응함으로써 최적의 결과를 생성한다. 또한 모델 가중치의 중요 구성요소를 선택적으로 조정하여 LLM이 실시간으로 새로운 작업에 동적으로 적응할 수 있도록 설계됐다.

트랜스포머2는 기존 LLM의 정적 구조를 개선하고 '특이값 미세조정(Singular Value Fine-tuning, SVF)'과 2단계 추론을 결합해 뛰어난 성능과 효율성을 달성한다. 특잇값 미세조정은 모델의 가중치 행렬에서 주어진 작업에 필요한 구성요소를 학습함으로써 필요한 매개변수를 최소화하여 다양한 작업에 효율적으로 적응이 가능하다. 추론 과정에서는 첫 번째 단계에서 주어진 작업의 특성을 분석하고, 두 번째 단계에서 작업에 특화된 구성요소를 활용하여 맞춤형 응답을 생성한다.

사나카AI는 연구진이 다양한 벤치마크로 라마(Llama)와 미스트랄(Mistral) LLM에 대하여 트랜스포머2와 기존 미세조정 방식인 LoRA[553]를 비교한 결과, 트랜스포머2는 거의 모든 작

552) Sakana.ai, Transformer²: Self-Adaptive LLMs, 2025.1.15.
553) Low-Rank Adaption: LLM의 가중치 행렬을 낮은 순위(Low-Rank) 행렬로 분해하여 일부 매개변수만 학습함으로써 메모리와 연산 비용을 낮추는 방식

업에서 성능 향상을 기록했다고 설명했다. 특히 시각적 질의응답 벤치마크인 TextVQA에서는 원본 모델(라마3-8B) 대비 39% 이상의 성능 향상을 달성했다고 주장했다. 반면, LoRA는 작업에 따라 불균형한 성능을 나타냈으며, 일례로 GSM8K[554]에서 LoRA 적용 시 원본 모델 대비 성능이 저하(라마3-70B-Instruct는 85.29점에서 LoRA 적용 시 77.26점으로 하락)하는 모습을 보였다. 트랜스포머2는 훈련에 필요한 연산 자원도 LoRA의 10% 미만으로 훨씬 적은 것으로 나타나 프레임워크의 효율성을 입증했다.

나. 사카나 AI의 AI 생성 논문, ICLR 2025 워크숍에서 동료 심사 통과[555]

사카나 AI가 2025년 3월 12일 자사 ‘AI 사이언티스트(AI Scientist)’로 생성된 논문이 국제 AI 학술대회 워크숍의 동료 심사(Peer-Review)를 통과했다고 발표했다. 사카나 AI는 2024년 8월 과학 연구를 자동화하는 ‘AI 사이언티스트’를 공개했고, 이를 개선한 v2 버전으로 생성된 논문을 AI 분야의 대표적 학술대회인 ICLR 2025 워크숍에 제출했다. AI 사이언티스트-v2는 과학적 가설 제안, 가설을 검증할 실험 제안, 실험 수행을 위한 코드 작성과 개선, 실험 수행, 데이터 분석과 시각화를 포함해 전체 과학 논문을 인간의 도움 없이 처음부터 끝까지 작성한다.

사카나 AI는 연구 수행을 위한 대략적인 주제만 제공하여 생성된 총 3편의 AI 생성 논문을 사전 협의 후 ICLR 워크숍에 제출했으며, 이 중 1편이 동료 심사를 통과했다고 설명했다. 동료 심사에서 통과된 논문은 신경망 훈련 기법에 관한 내용을 다루고 있으며, 심사 평점은 6.33점으로 제출된 전체 논문의 상위 45% 수준을 기록했다고 공개했다. 사카나 AI는 투명성과 윤리를 고려해 동료 심사 통과 후 논문을 철회했기에, ICLR 주최 측이 논문에 대한 메타 심사[556]를 진행하지 않았으나 이론적으로 메타 심사 단계에서 논문이 거부되었을 가능성이 있다고 부연 설명했다. 또한 ICLR 학회의 논문 수락률이 통상 20~30% 수준이나 학회에서 개최하는 워크숍의 수락률은 60~70%에 달한다며, 내부 검토 결과 AI 생성 논문 3편 모두 정규 학회 수준에는 도달하지 못했다고 한계를 인정했다.

사카나 AI의 내부 검토에서는 AI 생성 논문에서 그림이나 인용 누락, 서식 문제 등의 오류가 확인되었으며, 향후 최고 수준의 학회 기준을 통과할 수 있는 고품질의 논문 생성을 위

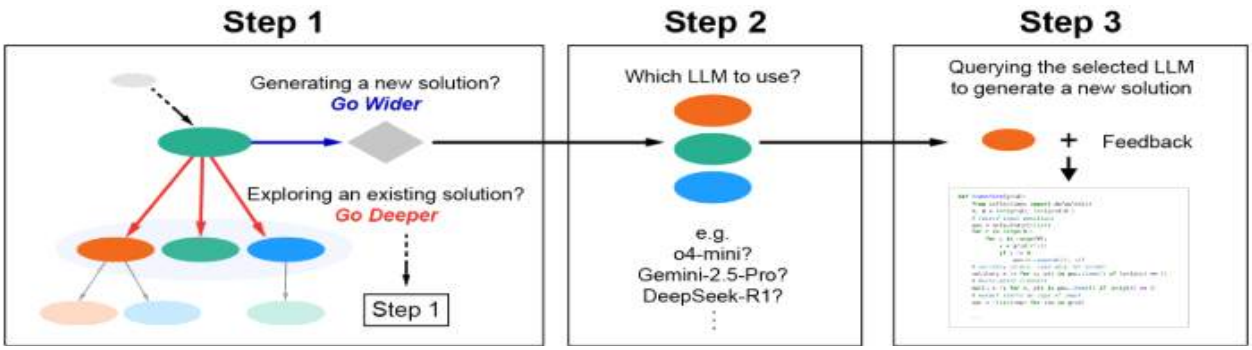
554) 초등학교 수준의 수학 문제 8,000개로 구성되어 AI 모델의 수학적 추론 능력을 평가
555) Sakana AI, The AI Scientist Generates its First Peer-Reviewed Scientific Publication, 2025.3.12.
556) 개별 심사자들의 의견을 종합하여 최종 평가를 제공하는 심사 과정

해 프로세스를 개선할 계획이다. 사카나 AI는 이번 실험을 계기로 AI가 생성한 과학 연구의 투명성과 윤리성 측면에서 과학 커뮤니티 차원의 규범 개발이 필요하다고 강조했다. 해당 규범에는 논문이 전적으로나 부분적으로 AI로 생성되었다고 알리는 시점과 방법 등이 포함되어야 하며, 가능한 많은 정보를 공개하여 연구의 투명성 확보가 중요하다고 지적했다. 사카나 AI는 AI 사이언티스트의 개발 현황에 관하여 연구 커뮤니티와 의견을 지속적으로 교환하여 향후 AI가 동료 심사 통과만을 추구해 학문적 검토 과정의 의미를 훼손하지 않도록 유의할 것이라고 강조했다.

다. 첨단 AI 모델 간 협력을 통한 추론 강화 알고리즘 개발[557]

사카나 AI가 2025년 7월 1일 챗GPT, 제미나이, 딥시크 등 최첨단 AI 모델 간 협력으로 추론 성능을 개선하는 알고리즘을 공개했다. ‘AB-MCTS(Adaptive Branching Monte Carlo Tree Search)’이라는 명칭의 알고리즘은 LLM이 시행착오를 효과적으로 수행할 수 있도록 하는 ‘추론 시간 스케일링’ 기법[558]으로 설계됐다. 추론 시간 스케일링은 단순히 추론 모델이 더 오래 생각하도록 만드는 방법 외에 문제를 반복적으로 재검토하여 답을 개선하거나 필요한 경우 처음부터 다시 생각하도록 하는 시행착오를 통해 성능을 개선한다. AB-MCTS는 시행착오를 통한 답변 개선을 위해 몬테카를로 트리 탐색(MCTS)[559]을 통해 생성된 답변을 여러 번 수정하는 ‘깊이 탐색’과, 같은 질문을 반복해 여러 번 답변을 생성하는 ‘넓이 탐색’ 방식을 활용했다.

[그림 5-26] 다중 LLM AB-MCTS의 검색 알고리즘 개요



출처: Sakana AI, Inference-Time Scaling and Collective Intelligence for Frontier AI

557) Sakana AI, Inference-Time Scaling and Collective Intelligence for Frontier AI, 2025.7.1.
558) Inference-Time Scaling: AI 모델의 추론 과정에서 추가적인 연산 자원을 투입하여 성능을 더 끌어올리는 방법
559) 무작위 시뮬레이션(확률적 샘플링)을 반복적으로 수행해 최적의 선택을 찾아가는 탐색 알고리즘

다중 LLM AB-MCTS는 새로운 답변 생성 또는 기존 답변을 개선하는 1단계와 주어진 문제와 상황에 맞는 LLM을 선택하는 2단계를 거쳐, 3단계에서 선택된 LLM이 개선된 답변을 생성하고 그 결과를 평가한다. 검색 프로세스의 초기 진행 시점에서는 각 문제에 어떤 LLM이 효과적인지 불확실하므로 LLM을 균형 있게 혼합하여 사용하다가, 이후에는 더 유능한 LLM을 선택하도록 조정하는 방식이다. 단일 모델에 AB-MCTS 기법을 적용해 최적화한 뒤 ARC-AGI 벤치마크로 평가한 결과, 오픈AI의 o4-mini 단일 모델을 이용한 반복 시도 정답률인 23%를 웃도는 27.5%의 정답률을 기록했다. 특히 다중 LLM AB-MCTS로 여러 첨단 AI 모델을 집단 지성으로 활용한 결과, 30% 이상의 정답률을 기록해 LLM 간 협업을 통한 성능 향상 효과를 보였다.

⑷ 딥시크

가. 오픈AI o1과 비슷한 성능의 추론형 AI 모델 ‘R1’ 공개[560]

중국 AI 기업 딥시크(DeepSeek)가 2025년 1월 20일 추론형 AI 모델 ‘R1’과 ‘R1-Zero’를 출시했다. 허깅페이스에서 오픈소스로 공개된 R1 시리즈는 딥시크가 2024년 12월 공개한 V3를 미세 조정한 전문가 혼합(MoE) 모델로 6,170억 개의 매개변수 중 약 340억 개만 활성화하도록 설계했다. R1-Zero는 작업 예시를 제공해 출력 품질을 향상하는 지도학습 미세조정(SFT) 없이 AI가 시행착오를 통해 작업을 수행하도록 훈련하는 강화학습(RL)만으로 사고 연쇄(CoT)[561]를 통한 강력한 추론 기능을 성공적으로 구현해 비용을 절감했다.

특히 딥시크가 강화학습에서 적용한 GRPO[562] 프레임워크는 모델 상태를 평가하는 별도의 모델을 필요로 하는 기존 방식과 달리 자체적인 모델 최적화를 통해 훈련 비용 감소가 가능하다고 강조했다. 하지만, R1-Zero는 응답 반복, 가독성 저하, 언어 혼합과 같은 문제점을 나타냈다. 이에 딥시크는 강화학습에 앞서 기초 지식을 제공하는 사전학습 데이터를 추가하고 다단계 학습을 결합하여 이러한 문제를 해결하고 추론 성능을 향상한 R1을 개발했다고 설명했다.

560) DeepSeek, DeepSeek-R1 Release, 2025.1.20.
561) 모델이 문제를 해결하는 과정에서 논리적인 단계들을 순차적으로 연결하여 결론에 도달하는 방식
562) Group Relative Policy Optimization

딥시크 R1은 벤치마크 평가에서 오픈AI의 o1과 동등하거나 더 나은 수준의 성능을 나타내면서도, 비용은 o1 대비 90~95% 저렴한 것이 특징으로, 딥시크는 R1의 기반이 된 V3 훈련 시 엔비디아 H100보다 성능이 절반 수준인 H800 2천 개를 사용했으며 훈련 비용이 약 560만 달러에 불과하다고 설명했다. 수학 관련 AIME 2024와 MATH-500 벤치마크에서 R1은 각각 79.8%와 97.3%의 정확도로 o1의 79.2%와 96.4%를 능가했고, 기타 벤치마크에서도 o1과 근소한 차이를 보였다.

오픈AI의 o1이 입력 토큰 100만 개당 15달러, 출력 토큰 100만 개당 60달러이지만 R1은 입력 토큰 100만 개당 0.55달러, 출력 토큰 100만 개당 2.19달러에 불과했다. 한편, 미국 기술업계에서는 딥시크가 주장한 560만 달러는 전체 개발비의 일부일 뿐이며 H800 외에 고사양 GPU를 혼합해서 사용했을 것이라는 지적도 제기됐다.[563]

나. 딥시크 R1 출시로 인한 AI 시장의 충격과 함께 보안 우려 대두[564]

딥시크 R1이 미국 빅테크의 AI 모델보다 훨씬 적은 비용으로 개발되었음에도 오픈AI o1에 상응하는 성능을 발휘하는 것으로 알려지며 AI 시장에 큰 파문을 가져왔다. 오픈소스로 공개된 딥시크는 출시 후 애플 앱스토어 다운로드 1위를 차지하며 큰 주목을 받았으며, 엔비디아를 포함한 기술 기업들의 주가는 딥시크의 여파로 대폭 하락했다. 저비용으로 뛰어난 성능의 모델을 개발한 딥시크는 AI 개발에 막대한 자금을 투입해 온 기존 빅테크의 경쟁력에 의구심을 증폭시키는 계기로 작용하며 AI 시장의 변화를 예고했다.

기술업계와 투자자들 사이에서 딥시크의 효율적인 개발 방식이 대량의 고성능 AI 칩을 요구하는 기존 AI 개발 관행을 혁신하며 엔비디아의 AI 칩 수요가 줄어들 수 있다는 전망이 대두됐다. 오픈AI를 비롯한 미국 기술업계에서 딥시크가 오픈AI의 모델에서 데이터를 추출하는 증류[565] 방식을 활용했을 것으로 추측하면서 지식재산권 침해 논란이 제기되는 가운데, 여타 스타트업도 딥시크의 개발 방식을 모방한 모델을 잇달아 출시하며 소비자 선택권이 넓어질 것이라는 전망도 제기됐다.

미국 정부가 최근 몇 년간 중국의 AI 발전을 제한할 목적으로 중국에 대한 AI 칩 수출 제

563) Bain & Company, DeepSeek: A Game Changer in AI Efficiency?, 2025.2.5.
564) The Wall Street Journal, Six Takeaways From a Monumental Week for AI, 2025.2.2.
565) Distillation: 더 큰 모델에서 추출한 지식을 더 작은 모델로 전달하여 모델 성능을 향상하는 기법

재를 강화해 왔으나, 딥시크가 저사양 칩으로 첨단 AI 모델을 개발하며 수출 규제 정책에도 변화가 예상된다. 미국 정치권에서는 딥시크 파장을 계기로 미국의 기술 패권을 유지하기 위한 정책을 강화해야 한다는 주장이 힘을 얻고 있으며, 중국 수출용으로 성능을 낮춘 엔비디아 H20 수출을 금지하는 방안도 거론되는 상황이다.

한편, 이탈리아를 시작으로 호주와 대만, 한국 등 주요국 정부는 개인정보와 보안 문제를 우려해 딥시크 서비스를 차단하거나 정부 기관 내 사용을 금지하는 추세로, 미국과 EU에서도 딥시크 보안 문제를 주시하고 있다.[566] 딥시크가 데이터를 중국 서버에 저장하는 것으로 알려져 논란이 되는 가운데, 이탈리아 개인정보보호 당국은 딥시크의 데이터 보호 정책에 문제를 제기하며 앱 다운로드를 차단했으며 호주와 대만, 한국의 주요 정부 부처도 보안 우려로 딥시크 접속을 금지했다.

(5) 그 외 기업

가. 마누스 AI, 완전 자율 AI 에이전트 '마누스' 공개[567]

중국 AI 스타트업 마누스 AI(Manus AI)가 2025년 3월 5일 유튜브와 전용 웹사이트를 통해 완전 자율 AI 에이전트 '마누스'를 공개했다. 마누스 AI는 전용 웹사이트(Manus.im)에서 마누스가 '사고와 행동을 연결하는 세계 최초의 범용 AI 에이전트로, 업무와 일상의 다양한 작업에서 뛰어난 능력을 발휘한다'라고 주장했다. 웹사이트에 소개된 활용 사례에 따르면 마누스는 여행계획 수립, 부동산 조사, 보험 정책 비교 등 다양한 작업을 인간의 개입 없이 독립적으로 수행하며, 일례로 이력서 파일을 받으면 후보자 순위 결정을 넘어 후보자들의 보유 기술을 추출하고 일자리 동향과 교차 참조하여 최적화된 채용 결정을 제시했다. 마누스 AI는 마누스가 GAIA(General AI Assistants)[568] 벤치마크 테스트에서 오픈AI의 딥 리서치를 능가하는 최고 점수를 달성했다고 보고했다.[569]

마누스는 초대 코드를 통해서만 접근을 허용하는 비공개 테스트 단계로, 마누스 AI는 서버 용량 제한으로 인해 초대 시스템이 불가피하며 이를 해결하기 위해 노력 중이라고 설명

566) Reuters, South Korean ministries block DeepSeek on security concerns, officials say, 2025.2.6.
567) Manus AI, Introducing Manus: The General AI Agent, 2025.3.5.
568) AI 모델의 추론, 멀티모달 처리, 웹 브라우징, 도구 사용 등 복잡한 능력을 평가하는 벤치마크로 3가지 난이도로 구성
569) (GAIA 레벨별 성능 비교) Level 1: Manus.ai 86.5%, OpenAI Deep Research 74.3%/Level 2: Manus.ai 70.1%, OpenAI Deep Research 69.1%/Level 3 : Manus.ai 57.7%, OpenAI Deep Research 47.6%

했다. 미국 경제 매체 포브스(Forbes)는 마누스가 “인간의 조수에 그치지 않고 인간을 대체하는 세계 최초의 완전 자율 AI 에이전트”라며, 제2의 딥시크라고 평가했으나 부정적인 평가도 존재한다고 지적했다.[570]

포브스에 따르면 마누스는 앤스로픽 클로드(Claude)와 알리바바 큐원(Qwen)과 같은 기존 LLM 기반 다중 에이전트 시스템으로 설계되어 복잡한 작업을 구성 요소별로 나누어 적절한 에이전트에 할당하고 진행 상황을 감독하며, 사용자의 개입 없이 백그라운드에서 작업을 수행한 뒤 완료되면 사용자에게 안내한다. 일부 초기 테스터 사이에서 정보검색 실수와 시스템 다운을 비롯한 마누스의 오류나 단점에 대한 지적도 나오고 있으나, 포브스는 정식 출시에 앞서 시스템이 개선될 것으로 예상했다.

그러나 미국의 IT 미디어 테크크런치는 마누스가 사용자에게 잘못된 답변을 제시하거나 주문이나 예약 프로세스에 종종 실패하는 등 잦은 오류를 보였으며, 딥시크와 달리 자체적으로 모델을 개발하지도 않았다며 제2의 딥시크라는 표현은 과장되었다고 지적했다.[571]

나. 피규어, 휴머노이드 로봇 제조시설 ‘BotQ’ 발표[572]

미국 휴머노이드 로봇 스타트업 피규어(Figure)가 2025년 3월 15일 연간 최대 1만 2천 대의 로봇을 생산할 수 있는 휴머노이드 로봇 제조시설 ‘BotQ’를 공개했다. 피규어는 휴머노이드 로봇을 자체 생산함으로써 제작 공정과 품질을 통제하여 고성능 로봇을 시장에 출시할 수 있다고 강조했다. 올해 안에 자체 휴머노이드 로봇을 BotQ의 제조 공정에 투입할 계획으로, 휴머노이드 로봇 투입을 점진적으로 늘려 생산 설비의 자동화 수준을 향상할 예정이다. 피규어는 휴머노이드 로봇의 아키텍처 재편, 공급망 구축, 전문 제조팀 구성, 소프트웨어 인프라 구축, 로봇 자동화 등을 통해 확장성 있는 제조시설을 건설했다.

피규어가 2024년 8월 공개한 2세대 휴머노이드 로봇 ‘피규어 02’는 고도의 정밀성을 요구하여 복잡하고 속도가 느린 컴퓨터 수치 제어(CNC)[573] 공정을 주로 사용했으나, 차세대 ‘피규어 03’ 설계 시에는 사출 성형 방식[574]으로 전환해 CNC 기계로 일주일 이상 소요

570) Forbes, China’s Autonomous Agent, Manus, Changes Everything, 2025.3.8.
571) TechCrunch, Manus probably isn’t China’s second ‘DeepSeek moment’, 2025.3.9.
572) Figure, BotQ: A High-Volume Manufacturing Facility for Humanoid Robots, 2025.3.15.
573) Computer Numerical Control: 컴퓨터를 사용하여 공작 기계를 자동 제어하는 기술

되던 부품 생산을 강철 금형으로 20초 이내에 완료한다고 주장했다. 피규어는 로봇 아키텍처를 재편하면서 안전팀과 신뢰성 팀을 새로 구성했으며, 신뢰성 팀은 로봇 수명을 빠르게 평가할 수 있는 가속 수명주기 테스트를 통해 로봇의 신뢰성 지표 달성을 위한 개선 사항을 도출했다.

공급망 구축에서는 핵심기술(액추에이터, 로봇 핸드, 배터리, 최종 조립)은 내부에서 확보하고 필요시 외부 공급업체를 활용하는 방식으로 향후 4년 내 휴머노이드 로봇 10만 개, 액추에이터 300만 개까지 생산 규모 확장이 가능해질 전망이다. 제조설비 설계와 최적화에 전문성을 갖춘 제조팀을 구성하고 품질과 속도 등의 지표를 반영해 단기적으로 자동화할 핵심 공정을 선택함으로써 생산 주기를 단축했다.

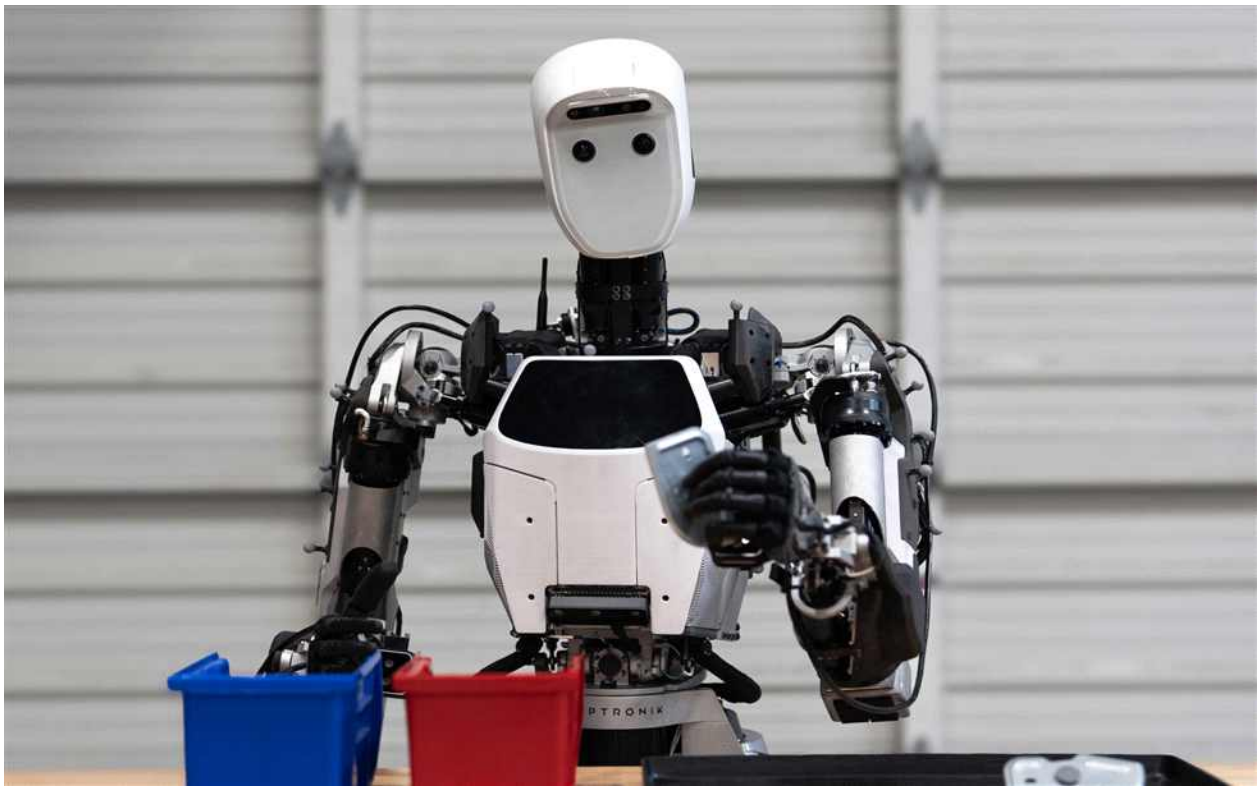
소프트웨어 인프라에서는 제품 수명주기 관리(PLM) 시스템, 전사적 자원관리(ERP) 시스템, 창고관리 시스템(WMS)과 함께 맞춤형 제조 실행 시스템(MES)[575]을 도입해 공급망을 통한 부품 추적에서 품질관리까지 전체 생산 절차를 효율적으로 운영하고 있다. 피규어는 자사 제조 공정에 휴머노이드 로봇 자동화를 통합하고, AI 기반 자동화 시스템과 인간의 감독을 결합해 로봇으로 로봇을 제작하여 생산 일정을 가속화하고 미래의 자율 제조 실현을 위한 기반을 마련하겠다고 강조했다.

다. 앱트로닉, 자빌과 전략적 제휴로 휴머노이드 로봇 제작 추진[576]

미국 휴머노이드 로봇 기업 앱트로닉(Apptronik)이 2025년 2월 25일 글로벌 엔지니어링 및 제조, 공급망 솔루션 기업인 자빌(Jabil)과 전략적 제휴를 발표했다. 자빌은 앱트로닉의 휴머노이드 로봇 ‘아폴로(Apollo)’를 제작하고 이를 자사 제조 공정에 활용하는 시범 프로젝트를 진행할 계획이다. 앱트로닉은 아폴로를 고객 현장에 배치하기에 앞서 자빌의 제조 공정에서 검사, 분류, 라인 내 물류 배송, 부품 조립과 같은 단순하고 반복적인 작업을 통한 검증 테스트를 시행할 예정이다.

574) 고분자 재료를 녹여 금형에 주입한 후 냉각시켜 제품을 대량으로 생산하는 제조 공정
575) 제조 현장에서 생산 계획, 작업 지시, 품질관리 등 생산 활동을 실시간으로 관리하고 제어하는 시스템
576) Jabil, Apptronik and Jabil Collaborate to Scale Production of Apollo Humanoid Robots and Deploy in Manufacturing Operations, 2025.2.25.

[그림 5-27] 앱트로닉의 휴머노이드 로봇 ‘아폴로’



출처: Jabil, Apptronik and Jabil Collaborate to Scale Production of Apollo Humanoid Robots and Deploy in Manufacturing Operations

자빌과 앱트로닉은 실제 생산 환경에 아폴로를 도입함으로써 기존 인력의 작업을 보완하는 한편, 제조 자동화와 아폴로의 AI 모델 최적화를 위한 실제 사용 사례를 수집할 수 있을 것으로 기대한다. 앱트로닉은 휴머노이드 로봇의 대량 보급을 위해 핵심 부품인 액추에이터의 단가를 절감한 데 이어 자빌을 통해 확보한 양산 능력을 바탕으로 제작 비용을 낮춰 소매 업장이나 노인 돌봄, 궁극적으로는 가정으로 시장을 확장할 계획이다. 앱트로닉은 삶의 모든 측면에서 인간을 지원하는 휴머노이드 로봇 제작을 기업 목표로 제시했으며, 우선 제조와 물류 같은 중요 산업에서 시작해 의료, 가정까지 로봇 응용 분야를 확대할 예정이다.

라. 미스트랄 AI, 추론 AI 모델과 기업용 AI 코딩 도구 출시[577]

프랑스 AI 스타트업 미스트랄 AI(Mistral AI)가 2025년 6월 10일 첫 번째 추론 AI 모델 '마지스트랄(Magistral)'을 출시했다. 마지스트랄은 매개변수 240억 개의 오픈소스 버전 '마지스트랄 스몰(Magistral Small)'과 더 강력한 성능의 기업용 버전 '마지스트랄 미디엄(Magistral Medium)'으로 구성됐다. 미스트랄 AI는 독일어, 러시아어, 아랍어, 영어, 이탈리아어, 중국어, 스페인어, 프랑스어 등 다양한 언어로 추론할 수 있으며, 구조화된 계산과 프로그래밍 논리, 규칙 기반 시스템 등 광범위한 기업 활용 사례에 적합하다고 강조했다.

마지스트랄 미디엄과 마지스트랄 스몰은 AIME 2024 벤치마크 평가에서 각각 73.6%와 70.7%를 기록해 딥시크 R1(79.8%)과 같은 경쟁 추론 모델과 비교하면 성능이 다소 떨어지는 것으로 나타났다. 하지만, 미스트랄 AI는 마지스트랄이 자사의 AI 플랫폼 '르샤(Le Chat)'의 '빠른 응답(Flash Mode)'을 통해 대부분 경쟁 모델보다 최대 10배 빠르게 답변을 제공함으로써 대규모 실시간 추론이 가능하다고 강조했다.

미스트랄 AI는 2025년 6월 4일 보안과 준법을 중요시하는 기업 환경을 겨냥한 AI 기반 코딩 어시스턴트 '미스트랄 코드(Mistral Code)' 베타 버전도 출시했다.[578] 미스트라 AI가 개발자의 액세스 요청을 받아 제공하는 베타 버전은 대표적인 통합개발환경(IDE)인 젯브레인(JetBrains)과 VS 코드(VS Code)에서 사용할 수 있으며, 추후 정식 버전을 출시 예정이다. 미스트랄 코드는 각각 다른 용도에 특화된 4개의 AI 모델을 활용하며, 사용자가 개인 저장소에서 미세조정이나 사후 학습으로 기본 모델을 변형할 수 있도록 지원한다.

4개 AI 모델은 코드 자동완성을 담당하는 코드스트랄(Codestral), 코드 검색용 코드스트랄 임베드(Codestral Embed), 에이전트 기반 코딩을 위한 데브스트랄(Devstral), 채팅 지원용 미스트랄 미디엄(Mistral Medium)으로 구성됐다. 미스트랄 코드는 80개 이상의 프로그래밍 언어를 지원하며, 파일 분석, 깃(Git)[579] 변경 사항 추적, 터미널 출력 해석, 코딩 이슈 처리 등 복합적인 작업을 수행한다고 설명했다.

577) Mistral AI, Magistral, 2025.6.10.
578) Mistral AI, Introducing Mistral Code, 2025.6.4.
579) 소스 코드의 변경 이력을 기록하고 개발자 간 동시 협업을 가능하게 하는 버전 관리 시스템

마. 미드저니, 첫 번째 비디오 생성 AI 모델 ‘V1’ 출시[580)]

AI 이미지 생성 플랫폼 미드저니(Midjourney)가 2025년 6월 19일 비디오 생성 모델 ‘V1’을 출시했다. V1은 이미지를 동영상으로 변환하는 모델로, 미드저니 플랫폼에서 제작된 이미지나 외부 이미지를 바탕으로 동영상을 생성하며, ‘자동’ 설정 시에는 모션 프롬프트가 자동으로 생성되고 ‘수동’ 설정을 선택하면 사용자 지시에 따라 사물이나 장면의 움직임을 생성한다. 사용자는 움직임 강도를 피사체와 카메라가 모두 움직이는 ‘하이 모션(High Motion)’과 카메라는 거의 고정되어 있고 피사체가 천천히 움직이도록 연출 ‘로우 모션(Low Motion)’ 중에서 선택할 수 있다.

웹 전용 모델인 V1은 1회 동영상 생성 작업으로 5초 길이의 동영상 4개를 제작하며, 생성된 동영상은 한 번에 4초씩 최대 4회까지 영상 길이를 확장이 가능하다. 미드저니는 동영상 생성 작업에 이미지 작업보다 약 8배의 비용을 책정했으며, 이는 동영상 1초당 이미지 1장의 비용에 해당한다며 시장에 출시된 기존 제품 대비 25배 이상 저렴하다고 주장[581)]했다. 미드저니의 V1은 긴 장면에서 사실성이나 일관성에 중점을 두는 오픈AI의 ‘소라(Sora)’와 같은 경쟁 동영상 생성 AI 모델과 비교해 사실성보다는 창의력과 풍부한 표현력에 중점을 둔다는 평가이다.

텍스트를 기반으로 처음부터 동영상을 생성하는 것이 아니라 자체 플랫폼에서 생성된 이미지를 동영상으로 전환하여 미드저니 고유의 몽환적이고 초현실적인 스타일을 그대로 유지한다. 구글의 동영상 AI 모델 ‘비오 3(Veo 3)’가 시각적 완성도를 중요시하는 브랜드와 영화 제작자 등 전문가를 주요 타깃으로 삼는다면, 미드저니의 V1은 예술적 표현을 원하는 사용자 집단에 더욱 적합하며, 간단한 사용자 인터페이스와 저렴한 비용도 강점이라는 평가이다.[582)]

미드저니는 자사의 궁극적 목표는 실시간 오픈월드 시뮬레이션이 가능한 모델, 즉 3D 공간에서 실시간으로 이미지를 생성하고 배경과 캐릭터 등 모든 요소와 상호작용을 할 수 있는 모델이라고 설명했다. 미드저니에 따르면 이미지 생성 모델에 이어 선보인 비디오 생성

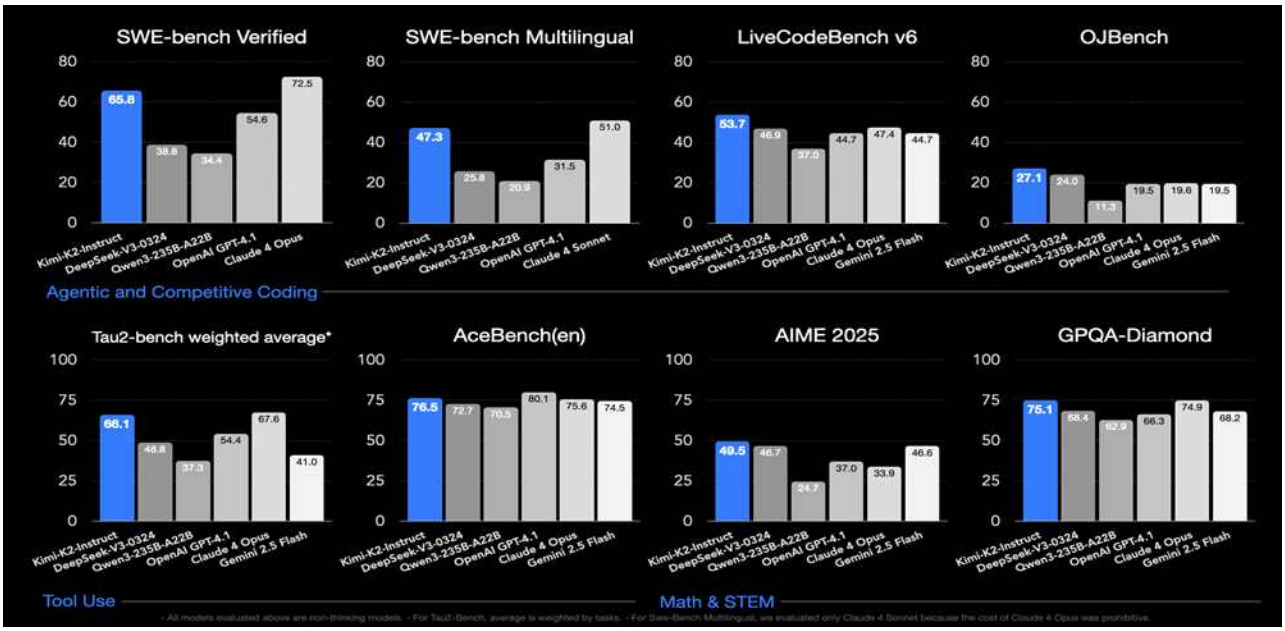
580) Midjourney, Introducing Our V1 Video Model, 2025.6.19.
581) GPU 제공 시간에 따른 베이직(월 10달러), 스탠다드(30달러), 프로(60달러), 메가(120달러) 요금제 중 프로와 메가는 무제한 동영상 제작 지원
582) Digit.in, Midjourney V1 Explained: Better than Google Veo 3?, 2025.6.19.

모델은 이를 달성하기 위한 다음 단계로서, 2026년에 공간을 자유롭게 이동할 수 있는 3D 모델과 실시간 처리 모델을 개별적으로 출시한 뒤, 각 모델을 하나로 통합하는 작업을 순차적으로 진행할 계획이다.

바. 문샷 AI, 에이전트 기능 지원하는 '키미 K2' 오픈소스 공개[583]

중국 AI 스타트업 문샷 AI(Moonshot AI)가 2025년 7월 11일 1조 개 매개변수 중 320억 개를 활성화하는 전문가 혼합(MoE) 구조의 AI 모델 '키미 K2(Kimi K2)'를 깃허브(GitHub)와 허깅페이스(Hugging Face)에서 공개했다. 키미 K2는 연구자와 개발자를 위한 기본 모델(Kimi-K2-Base)과 일반 채팅과 에이전트 경험에 최적화된 사후 학습 모델(Kimi-K2-Instruct)로 구성됐다. 에이전틱 AI로 단순 프롬프트 응답이 아닌 다단계 워크플로우를 자율적으로 실행·완료할 수 있는 기능을 탑재했다.

[그림 5-28] '키미 K2'와 경쟁 AI 모델의 주요 벤치마크 평가 결과 비교



출처 : Moonshot AI, Kimi K2: Open Agentic Intelligence

문샷 AI는 키미 K2가 지식, 수학, 코딩 등 다양한 분야에서 최신 오픈소스 및 상용 모델과 동등하거나 더 우수한 성능을 보이며, 특히 자연어 이해, 수학과 과학, 코드 생성, 에이전트 도구 활용에 뛰어나다고 강조[584]했다. 키미 K2의 비용도 입력 토큰 100만 개당 0.15달러,

583) Moonshot AI, Kimi K2: Open Agentic Intelligence, 2025.7.11.
584) 코딩 벤치마크(LiveCodeBench) 기준 Kimi-K2-Instruct(53.7점), DeepSeek V3(46.9점), Claude 4 Opus(47.4점),

출력 토큰 100만 개당 2.5달러로 타 모델 대비 저렴하다고 강조했다. 실제로 GPT-4.1은 입력 토큰 100만 개당 2달러, 출력 토큰 100만 개당 8달러로 키미 K2 대비 평균 3배 이상의 비용이 든다.

문샷 AI는 특히 학습 불안정성을 해결하는 '뮤온클립(MuonClip)[585]' 최적화 기술을 개발해 1조 개에 달하는 대형 AI 모델을 효율적으로 안정화했다고 강조했다. 해당 기술은 모델 내부의 특정 값이 폭발적으로 커져 모델 학습이 불안정해지는 현상을 해결해 모델 성능을 유지하면서 학습을 안정화하며, 문샷은 다른 모델의 안정화에도 쉽게 적용될 수 있다고 설명했다.

사. 즈푸 AI, 추론과 에이전트 성능 강화한 'GLM-4.5' 공개[586]

중국 AI 스타트업 즈푸 AI(Zhipu AI)가 추론 기능을 갖춘 플래그십 AI 모델 'GLM-4.5'와 'GLM-4.5-Air'를 허깅페이스(Hugging Face)와 모델 스코프(ModelScope)에서 오픈소스로 공개했다. GLM-4.5는 총 3,550억 개의 매개변수 중 320억 개가 활성화되고, 경량화 모델 GLM-4.5-Air는 총 1,060억 개의 매개변수 중 120억 개가 활성화되는 전문가 혼합(MoE) 구조로 설계됐다.

두 모델 모두 하이브리드 추론 모델로, 복잡한 추론과 도구 사용을 위한 '사고 모드'와 즉각적 응답을 위한 '非사고 모드'를 제공해 사용자가 필요에 따라 선택할 수 있으며, GLM-4.5는 복잡한 웹페이지나 인터랙티브 미니 게임 제작과 같은 풀스택 개발 기능을 지원하고 파워포인트 슬라이드 생성이 가능하다. 즈푸 AI는 GLM-4.5 개발 시 빠르게 늘어나는 AI 에이전트 애플리케이션의 복잡한 요구사항에 대응해 추론과 코딩, 에이전트 기능을 단일 모델에 통합하도록 설계했다고 강조했다. 에이전트, 추론, 코딩을 포함하는 12개 벤치마크[587]*로 주요 AI 모델과 비교한 성능평가 결과, 전체적으로 GLM-4.5는 3위, GLM-4.5-Air는 6위를 기록했다.

GPT-4.1(44.7점)

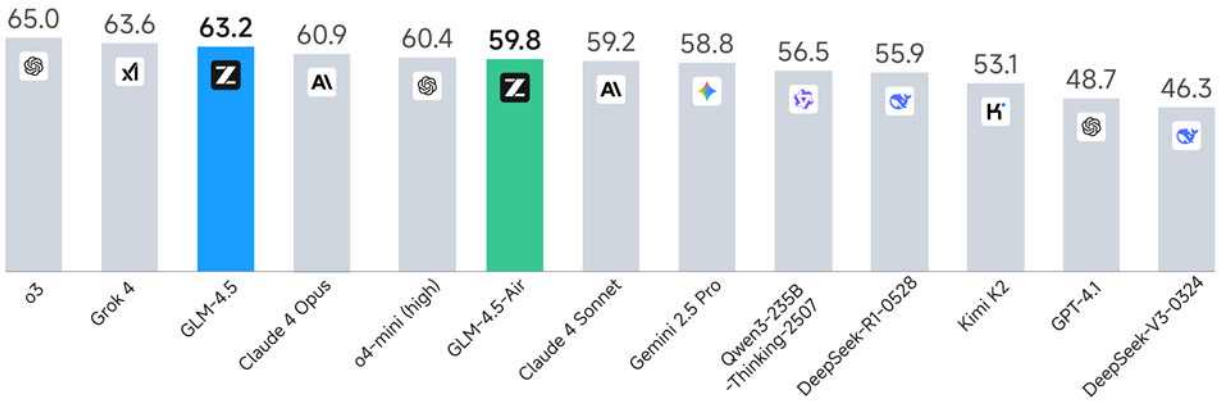
585) 신경망 최적화 알고리즘 뮤온(Muon)을 개선해 훈련 안전성을 향상한 기술

586) Zhipu AI, GLM-4.5: Reasoning, Coding, and Agentic Abililties, 2025.7.28.

587) MMLU-Pro, AIME 24, MATH-500, SciCode, GPQA, HLE, LCB(2407-2501), SWE-Bench Verified, Terminal-Bench, TAU-Bench, BFCL V3, BrowseComp

벤치마크 종합 점수에서 GLM-4.5는 63.2점으로 o3(65.0점)와 그록 4(63.6점)보다 낮았으나 오픈소스 모델 중 최고 점수를 기록했고, 에이전트 관련 벤치마크[588]에서는 58.1점으로 o3(61.1점)에 이어 2위를 차지했다. GLM-4.5는 실제 개발 시나리오를 활용한 에이전틱 코딩 성능평가에서 키미-K2 대비 53.9%의 승률을, 큐원3-코더 대비 80.8%의 승률을 기록했으며, 평균 도구 호출 성공률에서 90.6%로 클로드-4-소네트(89.5%), 키미-K2(86.2%), 큐원3-코더(77.1%)를 넘어서 1위를 기록했다.

[그림 5-29] 12개 벤치마크 기준 LLM 성능 종합 평가 비교



출처 : Zhipu AI, GLM-4.5: Reasoning, Coding, and Agentic Abililties

아. 딥엘, 기업용 AI 에이전트 '딥엘 에이전트' 발표[589]

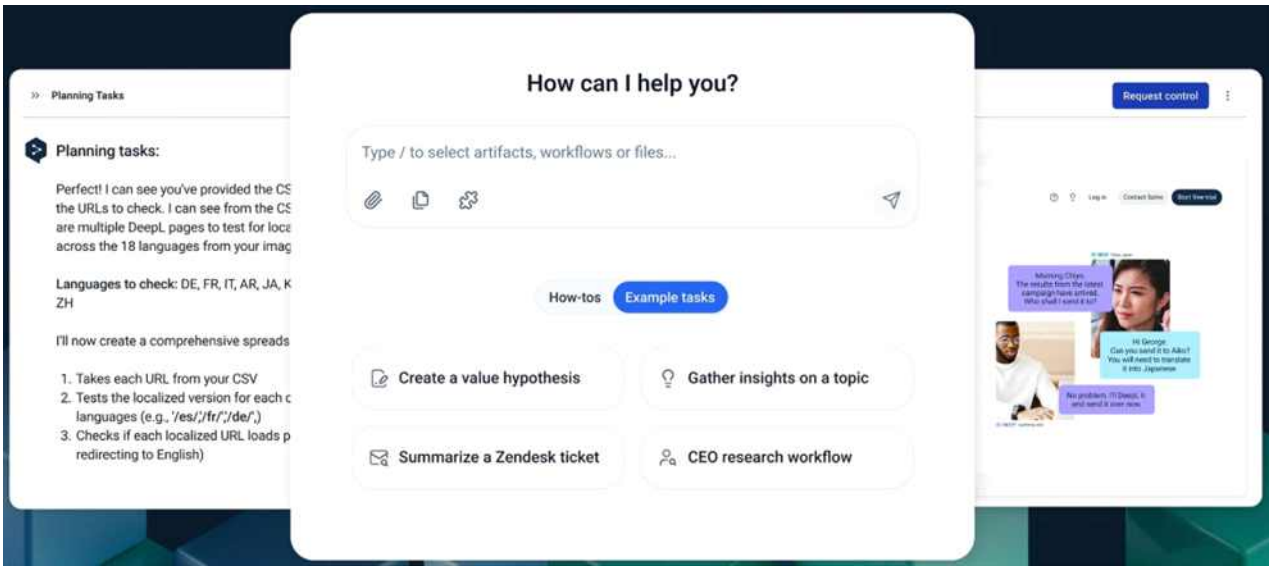
독일 AI 기반 번역 솔루션 기업 딥엘(DeepL)이 2025년 9월 3일 자연어 지시를 받아 다양한 업무를 자율적으로 처리하는 범용 AI 에이전트 '딥엘 에이전트(DeepL Agent)'를 발표했다. 딥엘은 언어와 관련된 복잡한 과제 해결에서 수년간 축적된 연구와 통찰을 바탕으로 단순한 언어 이해를 넘어 스스로 추론하고 결정을 내리며 작업을 실행할 수 있는 기업용 AI 에이전트를 개발했다고 설명했다. 딥엘 에이전트는 자연어로 된 사용자 지시를 바탕으로 가상 키보드, 브라우저, 마우스를 사용해 작업을 수행하며, 사용자는 언제든지 에이전트를 중지하고 인간의 개입이 필요한 동작을 확인해 승인할 수 있다.

개별 사용자뿐 아니라 관리자와 최고 경영진이 에이전트 사용 현황과 결과를 직접 제어할

588) TAU-Bench, BFCL v3 (Full), BrowseComp
589) DeepL, Introducing the new DeepL AI Agent, 2025.9.3.

수 있도록 다단계 보안 장치를 구축하고, 에이전트가 수행하는 모든 작업에 대한 모니터링과 관리 기능을 도입했다. 딥엘 에이전트는 여행 예약이나 식료품 구매와 같은 소비자 중심 작업에 초점을 맞춘 여타 에이전트와 달리 기업 활용 사례에 특화된 에이전트로, 컴퓨터로 할 수 있는 거의 모든 작업을 수행한다. 사용자의 업무용 시스템과 웹사이트를 탐색해 업무 계획과 문제 해결 및 추론, 데이터와 다이어그램 분석, 의사결정 등을 처음부터 끝까지 실행할 수 있다.

[그림 5-30] 딥엘 에이전트의 화면



출처 : DeepL, Introducing the new DeepL AI Agent

간단한 작업뿐 아니라 문서 현지화(예: 영어에서 독일어, 이탈리아어, 프랑스어로 번역), 영업용 고객 목록 작성, 마케팅 자료 검토와 같은 복잡한 업무 흐름의 처리도 지원한다. 딥엘은 현재 일부 고객을 대상으로 베타 테스트를 진행 중으로, 향후 몇 달 안에 시범 서비스를 제공할 예정이다. 초기 테스트 결과, 개인의 업무 효율성뿐 아니라 팀 단위 성과에도 긍정적 효과가 확인되었으며, 단순 반복 업무처리를 에이전트에 맡김으로써 하루 업무량에서 몇 시간 이상을 절약하고 협업도 향상되었다고 강조했다.

제2절 국내 기업

1. 네이버

1) 기업 일반현황

<표 5-12> 네이버 기업 현황

구분	주요 내용
기업 개요	**〈회사 개요〉** • 1999년 설립된 한국의 대표적인 ICT·인터넷 플랫폼 기업으로, 검색·포털·광고·커머스·콘텐츠·핀테크·클라우드·AI를 주력 사업으로 확장 • 2013년 라인(LINE) 중심으로 일본·동남아·대만 등 해외 시장 진출을 본격화하며 아시아 기반의 글로벌 플랫폼 기업으로 성장 • 2021년 Z홀딩스와 경영통합을 완료하며 글로벌 AI·커머스·핀테크 경쟁력을 강화 • 2024~2025년에 생성 AI·커머스 글로벌화·클라우드·로봇·자율주행 분야로 사업 확장을 가속화
	〈사업 내용〉 • 산업 분야: 인터넷 플랫폼, 검색·광고, 커머스, 콘텐츠·웹툰·웹소설, 클라우드, AI, 로봇·물류, 핀테크(네이버페이), 지도·모빌리티 등. • 주요 서비스: 네이버 검색, 네이버쇼핑, 스마트스토어, 네이버페이, 라인(LINE), 스노우, 네이버웹툰, V LIVE(통합 운영), 클라우드 플랫폼, 네이버웍스, 네이버 지도, 클로바(CLOVA) 등
	〈자회사〉 • 라인(Line)·Z홀딩스, 네이버웹툰, 스노우, 네이버클라우드, 네이버파이낸셜, 네이버랩스(로봇·자율주행), 네이버제트(ZEPETO), 웍스모바일(워크 솔루션), 네이버엔터프라이즈 등.
AI 조직 현황	**〈AI 관련 조직〉** • 네이버 클로바 AI(CLOVA AI) : 한국어 생성 AI·음성·비전·OCR·검색 AI 등 네이버의 AI 기술 전반을 담당 • 하이퍼클로바X(HyperCLOVA X) : 한국어 기반 LLM으로 검색·쇼핑·비즈니스·개발자 도구·웍스 등 네이버 서비스 전반에 적용 • 네이버랩스(NAVER Labs) : 로봇·SLAM·디지털트윈·자율주행·3D 지도 등 AI 기반 물리 공간 기술 연구를 담당 • 네이버클라우드(NAVER Cloud) : 기업용 AI 플랫폼(AIaaS), 하이퍼클로바X API, 비즈니스 솔루션 제공하며 정부·기업 대상 AI 클라우드 확산을 주도.

구분	주요 내용
사업 전망	〈AI 전략〉 • 네이버는 한국어·아시아 언어 특화 LLM인 하이퍼클로바X 기반으로 검색·쇼핑·비즈니스 자동화·고객센터 AI·유통·금융 등 B2B 중심의 AI 전환을 가속화 • 검색·쇼핑 플랫폼에 생성 AI를 도입해 개인화 추천·요약 검색·AI 쇼핑 가이드 등 고도화된 사용자 경험 제공을 강화. • 네이버클라우드는 공공·금융·제조 기업의 AI 전환을 위한 한국형 AI 인프라·모델·데이터 서비스를 확대하며 '국가 AI 클라우드' 전략적 위치를 확보. • 네이버랩스는 로봇·물류 자동화·자율주행 기반 3D 맵 기술을 상용화하며 오프라인 공간의 AI 전환을 주도

2) 2025년 기업 동향

가. 딥시크보다 한 수 위" … 네이버, '하이퍼클로바X 플래그십' 공개[590]

네이버가 작지만, 강한 파운데이션 모델 '하이퍼클로바X 플래그십' 모델을 공개했다. 기존 대비 40% 수준의 파라미터로 구성됐지만 더 강력한 성능을 갖췄다. 하지만 최근 등장한 딥시크가 언어만 가능하지만, 하이퍼클로바는 이미지, 영상까지 이해하는 멀티모달 모델이라 딥시크보다 한 수 위 모델이라고 강조한다. 네이버는 20일 하이퍼클로바X 플래그십 모델 업데이트를 완료하고 사내에 공개했다. 사내 공개된 새로운 하이퍼클로바X 파운데이션 모델을 활용해, 네이버는 사용자, 판매자, 창작자, 비즈니스 파트너 등을 위한 다양한 AI 서비스 개발에도 본격 착수했다고 밝혔다.

업데이트된 하이퍼클로바X 모델은 기존 대비 약 40% 수준의 파라미터 수로 구성된, 상대적으로 작은 크기의 모델이지만 더 강력한 성능을 갖춘 것이 특징이다. 주요 학습 데이터인 한국어, 영어, 코딩・수학에 대해 19개 벤치마크로 종합적인 성능을 비교한 결과, 모든 분야의 평균 점수가 기존 모델을 앞질렀다. 특히 이들 중 전 세계적으로 통용되는 대표적인 벤치마크인 'MMLU(Massive Multitask Language Understanding)'에서는 정답률 79.6%를 기록해 유사한 규모의 해외 빅테크 AI 모델에 필적하는 언어 이해 능력을 보였다.

멀티 모달리티(Multi-modality) 능력도 고도화됐다. 텍스트뿐만 아니라 이미지 데이터도 동

590) The AI, "딥시크보다 한 수 위"… 네이버, '하이퍼클로바X 플래그십' 공개", 2025.2.20

시에 처리할 수 있었던 기존 모델의 시각적 질의응답, 차트·도표 이해 등 역량을 글로벌 최고 수준 모델의 성능까지 끌어올렸다. 이미지를 넘어 영상까지 이해할 수 있는 기능도 새롭게 추가됐다. 효율적인 구조로 모델을 설계해 운영 비용을 절감했다. 네이버에 따르면 하이퍼클로바X 신모델의 운영 비용은 기존 모델 대비 50% 이상 개선됐다. 지난해 발표한 '온 서비스 AI(On-Service AI)' 전략 아래 검색, 커머스 등 대규모 사용자를 보유한 주요 서비스에 생성 AI 기술을 도입하고 있는 네이버는 이번 저비용·고성능의 하이퍼클로바X 모델을 이용해 AI 적용 확대에 더욱 속도를 낼 전망이다.

한편, 네이버는 2023년 8월 하이퍼클로바X를 공개한 이후 글로벌 AI 기술 트렌드에 발맞춰 모델을 지속적으로 발전시키고 있다. 2024년 4월 기존 대비 5분의 1 수준의 가격으로 복잡성이 적은 업무에서 활용이 가능한 경량 모델 '하이퍼클로바X 대시(DASH)'를 출시해 기업들의 생성 AI 도입 장벽을 낮췄다. 이어 8월에는 텍스트와 이미지를 동시에 처리할 수 있는 '하이퍼클로바X 비전(HyperCLOVA X Vision)' 모델을 공개했다. 나아가 계획 수립, 추론 능력을 고도화해 사용자가 요청한 작업을 체계적·종합적으로 수행할 수 있는 기술을 개발하고 있으며, 하반기에는 자연스러운 음성 대화가 가능한 하이퍼클로바X 모델을 공개할 예정이다.

나. CVPR 2025에서 공간지능·AI 논문 14편 게재[591]

네이버가 2025년 3월 25일 네이버랩스, 네이버랩스 유럽, 네이버클라우드 등 '팀 네이버'의 다양한 기술 조직이 발표한 연구 논문 14편이 컴퓨터비전 분야 세계적 학회 중 하나인 'CVPR(컴퓨터비전 및 패턴 인식 콘퍼런스 학술대회) 2025'에 게재 승인됐다고 밝혔다. 지난 1983년 시작된 CVPR은 세계 최대 기술 전문 단체인 전기전자공학자협회(IEEE)와 컴퓨터비전 협회(CVF)가 공동 주최하는 학술대회다. 비전 AI 분야의 대표적인 학회이자 컴퓨터 공학에서 가장 영향력 있는 학회로 알려져 있다. 이로써 네이버는 컴퓨터비전 분야 세계 3대 학회로 꼽히는 CVPR, ECCV(유럽 컴퓨터비전 학회), ICCV(국제 컴퓨터비전 학회)에 최근 5년여 동안(2020년~2025년 3월) 151건의 정규 논문을 등재했다. 먼저 팀 네이버는 CVPR 2025에서 지난해 공개했던 3D 재구성 AI 도구 'DUSt3R(더스터)'의 후속 연구 결과들을 발표할 예정이다.

591) The AI, "CVPR 2025에서 공간지능·AI 논문 14편 게재", 2025.3.25

한두 장의 사진만으로 3차원 복원을 쉽게 할 수 있는 AI인 더스터를 기반으로, 여러 장의 이미지에서 더 정확한 3차원 공간을 재구성할 수 있는 AI 모델 ‘MUSt3R(머스터)’와, 다양한 카메라 및 장면 정보를 통합해 추론 능력을 높인 AI 모델 ‘Pow3R(파워)’ 등이 새로 등재됐다. 또한 사전에 학습되지 않은 물체의 위치와 방향을 신속하고 정확하게 찾아내는 기술을 제안한 네이버랩스의 논문(Co-op: Correspondence-based Novel Object Pose Estimation) 등 다수의 공간지능 논문이 학회에 채택됐다. 네이버클라우드는 ‘마스킹(masking)’ 기법을 활용해 AI가 이미지 학습을 할 때 발현되는 훈련 불안정성 문제를 해결한 효율적인 학습 방법론을 제안(Masking meets Supervision: A Strong Learning Alliance)했다. 네이버는 “적극적인 연구개발 투자를 지속하며 선행 기술을 확보해 나겠다”라며 “나아가 ‘온 서비스 AI(On-Service AI)’ 전략 아래, 생성 AI와 공간지능 등 성공적으로 내재화한 핵심기술을 네이버 주요 서비스에 도입해 사용자에게 차별화된 서비스 경험을 제공할 계획”이라고 말했다.

다. 네이버클라우드, ‘하이퍼클로바 X’ 경량 모델 3종 오픈소스 공개[592)]

네이버클라우드는 2025년 4월 23일 하이퍼클로바X의 경량 모델 3종 △하이퍼클로바X 시드 3B △하이퍼클로바X 시드 1.5B △하이퍼클로바X 시드 0.5B의 개발을 완료하고 오픈소스로 공개했다. 국내·외 기업과 연구기관 등에서는 해당 모델을 내려받아 필요에 맞게 조정해 비즈니스 및 학술 연구에 이용할 수 있다. 이번에 공개된 오픈소스 모델인 HyperCLOVA X SEED는 연구용으로만 사용할 수 있었던 기존 국내 기업들의 주요 모델과 달리 상업 목적으로도 무료로 이용할 수 있다. 비용 부담이나 라이선스 제한 등의 문제로 생성 AI 도입을 망설였던 중·소규모 비즈니스를 중심으로 국내 AI 생태계가 더욱 활성화될 것으로 기대된다.

하이퍼클로바X 시드 3B는 텍스트뿐만 아니라 이미지와 영상 정보도 처리할 수 있는 시각언어모델로, 도표 이해, 개체 인식, 사진 묘사 등의 능력을 갖추고 있다. 네이버는 “한국어·한국문화와 영어 관련 시각 정보 이해 능력을 평가할 수 있는 9개 벤치마크의 평균 점수를 비교했다”라면서 “하이퍼클로바X 시드 3B는 유사 규모의 미국·중국 빅테크 모델보다 높은 점수를 기록했고, 파라미터 수가 훨씬 많은 해외 대규모 모델과 유사한 정답률을 보였다”라고 설명했다. 하이퍼클로바X 시드 3B는 상품 페이지의 정보를 인식해 광고 문구를 제안하는 솔루션을 구축하거나, 사진 및 영상 콘텐츠와 함께 국내 여행지 정보를 안내하

592) The AI, “네이버클라우드, ‘하이퍼클로바X’ 경량 모델 3종 오픈소스 공개“, 2025.4.23

는 서비스를 설계하는 것과 같은 다양한 비즈니스에 활용될 수 있다.

한편, 네이버는 하이퍼클로바X의 플래그십 모델을 기반으로 개발 중인 추론(reasoning) 모델을 상반기 중 선보일 예정이라고 밝혔다. 해당 모델은 추론 모델의 근본적 강점인 수학, 프로그래밍 분야에서 더 정확한 답변을 생성하는 것을 넘어, 시각 및 음성 정보 이해, 자동 웹 검색, API 호출, 데이터 분석 등 그동안 개발해 온 하이퍼클로바X의 각종 능력이 추론 모델을 기반으로 함께 고도화되는 것이 특징이다.

라. 네이버클라우드, '네이버, 검색에 생성 AI 도입[593]

네이버가 2025년 6월 12일 차별화된 콘텐츠와 정보·쇼핑·로컬·금융 등 버티컬 에이전트를 결합해 장기적으로 사용자에게 끊김이 없는(Seamless) AI 검색 경험을 제공하는 '통합 에이전트'를 향한 계획을 발표했다. 네이버는 생성 AI 기술 확대로 정보검색에서 사용자의 복합 요구를 해결하는 단계로 검색 외연이 확장 중이라고 강조하고, 네이버는 글로벌에서 몇 안 되는 검색 기업으로서 독보적인 검색 인프라, 한국 사용자에 특화된 풍부한 데이터를 갖추고 있어 AI 검색 시장에서 차별성 확보를 위한 유리한 조건을 갖추고 있다고 설명했다.

네이버는 약 27년간 구축한 검색 인프라를 꾸준히 고도화하고, 오랜 기간 쌓아온 서비스 운영 노하우와 풍부한 데이터를 바탕으로 검색 사용자 인사이트를 확보하고 있다. 지난해 8월부터는 다변화되는 사용자 검색 요구에 대응하기 위해 웹 검색 색인 규모와 성능을 대폭 개선한 바 있다. 이에 더해 네이버는 초창기부터 콘텐츠와 버티컬 서비스의 중요성을 간파하고 이를 중심으로 검색 서비스를 발전시키며 글로벌 빅테크와 차별화해 왔다. 블로그, 카페, 지식iN 등 UGC 서비스를 비롯해, 쇼핑, 지도, 페이 등 버티컬 서비스가 대표적이다.

검색 인프라 투자도 강화한다. 네이버는 질의 분석 및 요약, 문서 분석 등 각 태스크에 적합한 다양한 형태의 검색 LLM 라인업을 세분화하여 구축하고, AI 검색에 필요한 콘텐츠 확보를 위한 적극 투자를 통해 검색 품질을 높일 계획이라고 밝혔다. 현재 통합검색에 도입된 'AI 브리핑'은 출시 초기 대비 노출이 약 3배 확대됐고, AI 브리핑 하단에 제공하는 '더

593) The AI, "네이버, 검색에 생성형 AI 도입한다", 2025.6.12

보기’ 버튼 클릭률은 출시 초기 대비 50%, 관련 질문 클릭률은 기존 검색어 추천 영역 대비 3.4배 증가했다고 설명했다. 특히 AI 브리핑은 사용자 검색 흐름 맞춰 유연하게 연결될 수 있도록 설계되어 자연스러운 추가 탐색을 돕고 있다. 이 밖에도, AI 브리핑 영역의 CTR(클릭률)은 기존 정답형 콘텐츠 대비 8%p, AI 브리핑 출시 전후 최상단 영역의 체류시간은 22% 증가했다.

네이버는 AI 브리핑 노출을 연내 약 20% 수준으로 확대하고, 금융, 헬스케어 등 다양한 주제에 특화된 AI 브리핑을 순차 도입할 예정이다. AI 브리핑 유형도 확대된다. 해외 문서 번역 및 요약, 긴 영상 핵심 요약 등 다국어 지원 및 멀티미디어와 결합한 형태도 선보인다. 이렇게 도입되는 주제별 AI 브리핑은 향후 버티컬 AI 에이전트의 초석이 될 예정이다.

마. 이미지 검색 ‘AI 요약 서비스’ 출시[594]

네이버가 스마트렌즈에 인공지능(AI) 브리핑 기능을 추가해 이미지를 촬영하면 AI가 관련 정보를 분석해 핵심 내용을 요약해서 보여주는 서비스를 출시했다. 네이버는 2025년 7월 31일 스마트렌즈에 이미지와 텍스트를 함께 이해하는 AI를 적용한 ‘렌즈 x AI 브리핑’ 기능을 제공한다. 이 기능은 사용자가 스마트렌즈로 식물, 와인, 패션 아이템 등의 이미지를 입력하면 AI가 이를 분석해 관련성 높은 문서를 찾고 핵심 내용을 요약해 제공하는 방식으로 작동한다.

기존에는 이미지 검색 시 유사 이미지나 지식백과 위주의 결과가 제공됐지만, 이제는 다양한 출처의 콘텐츠를 바탕으로 핵심 정보를 확인할 수 있게 됐다. 하단에 함께 제공되는 추천 문서를 통해 추가적인 정보 탐색도 가능하다. 일례로, 스마트렌즈로 꽃을 촬영하면 다양한 출처의 콘텐츠들을 분석해 꽃의 이름, 개화 시기, 특징, 꽃말, 관리법 등 핵심 정보들을 제공한다. 와인 라벨을 촬영하면 품종, 생산지, 페어링 등의 정보를 한눈에 확인할 수 있고, 코디나 패션 아이템 이미지도 제품 정보, 브랜드, 스타일 특징 등을 편리하게 확인할 수 있다.

네이버는 기존 텍스트 중심의 AI 브리핑을 이미지 검색까지 확장하는 동시에, AI 브리핑

594) The AI, “이미지 검색 ‘AI 요약 서비스’ 출시“, 2025.8.1

적용 범위를 확대하며 일상 속 정보 탐색을 더 편리하게 발전시켜 나갈 계획이라고 밝혔다. 실제로 7월 말 기준 AI 브리핑이 제공되는 검색어 수는 출시 초기 대비 약 8배 늘었으며, 검색 결과 상단 체류시간도 32% 증가하는 등 긍정적인 사용자 반응을 얻고 있다.

[그림 5-31] 이미지 검색 AI 요약 서비스



출처 : The AI, “이미지 검색 ‘AI 요약 서비스’ 출시

2. LG AI 연구원

1) 기업 일반현황

<표 5-13> LG AI 연구원 기업 현황

구분	주요 내용
기업 개요	**〈회사 개요〉** • LG AI 연구원은 2020년 12월, LG그룹의 디지털 전환(DX)과 미래 신성장 동력 확보를 위해 설립된 국내 최초 대 기업계 AI 전문 연구기관 • 설립 초기부터 '산업 적용 가능한 실용 AI'를 목표로 언어·비전·로보틱스·소재·바이오 AI 연구에 집중하며 그룹 계열사의 기술혁신을 지원 • 연구원은 서울 마곡 LG사이언스파크를 기반으로 운영되며 2024~2025년 기준 약 200명 이상의 AI 전문 연구 인력을 보유한 국내 최대급 기업 AI 연구조직으로 성장 • 2023~2025년 EXAONE 2.0→ 3.0→ 3.5로 이어지는 초거대 모델 고도화와 산업 AI 솔루션 확장으로 글로벌 AI 기업과 경쟁할 수 있는 수준으로 빠르게 진화 중
	〈사업 내용〉 • 주요 연구 영역 : LLM, 멀티모달 AI, 로봇·물류 AI, 신물질·신약 AI, 제조·품질 AI, 고객 분석 AI 등 산업 전반에 적용되는 범용•특화형 AI 기술 • 사업 영역 : 연구개발(R&D)이지만, 연구 성과를 LG전자·LG화학·LG유플러스 등 그룹 계열사의 제품·서비스·제조·물류에 적용하는 '기술 전이'가 핵심
	〈자회사〉 • 별도 자회사는 없으며, LG전자·LG디스플레이·LG화학·LG에너지솔루션·LG CNS·LG U+ 등 AI 적용 수요가 높은 계열사들과 전략적 협업 구조
AI 조직 현황	**〈AI 관련 조직〉** • 언어 AI(LLM) 조직: EXAONE 1.0~3.5 모델 개발, 한국어·영어 기반 멀티모달 모델 연구, 기업용 특화 모델 개발을 담당. • 비전·멀티모달 연구팀: 이미지·문서·3D 모델·영상 분석 등 멀티모달 AI 역량 확장을 수행하며 제조·품질·물류 적용 모델을 개발. • 산업 AI/Applied AI 조직: 전자·화학·에너지·통신·유통 등 산업별 현장 문제를 해결하는 AI 솔루션을 연구(품질 예측, 수요 예측, 생산 최적화 등). • 신소재·바이오 AI 조직: 신약 후보 물질 발굴·신소재 시뮬레이션·촉매 개발 등 과학 AI(Science AI) 분야 연구를 수행하며 해외 학계·연구기관과 공동 연구를 진행.

구분	주요 내용
사업 전망	〈AI 전략〉 • LG AI연구원 전략은 'EXAONE 기반 초거대 AI+산업 특화 AI+물리 공간·로봇 AI"를 중심으로 그룹 DX 전환을 가속하는 방향으로 전개 • EXAONE 3.5는 멀티모달·지식 기반·산업 특화 기능이 강화되어 제조·소재·고객 분석·서비스 자동화 등 B2B 적용 범위가 빠르게 확대될 전망 • 로봇·물류 자동화에서는 LG전자·LG CNS·LG U+와 연계해 자율주행, 로봇 지능화, 딥 SLAM 등 AI 기반 물리 공간 기술 상용화를 추진 • 중장기적으로는 산업 AI 전문 연구기관+초거대 AI 경쟁력+글로벌 공동 연구 체제를 기반으로 아시아 대표 AI 연구기관으로 도약을 추진

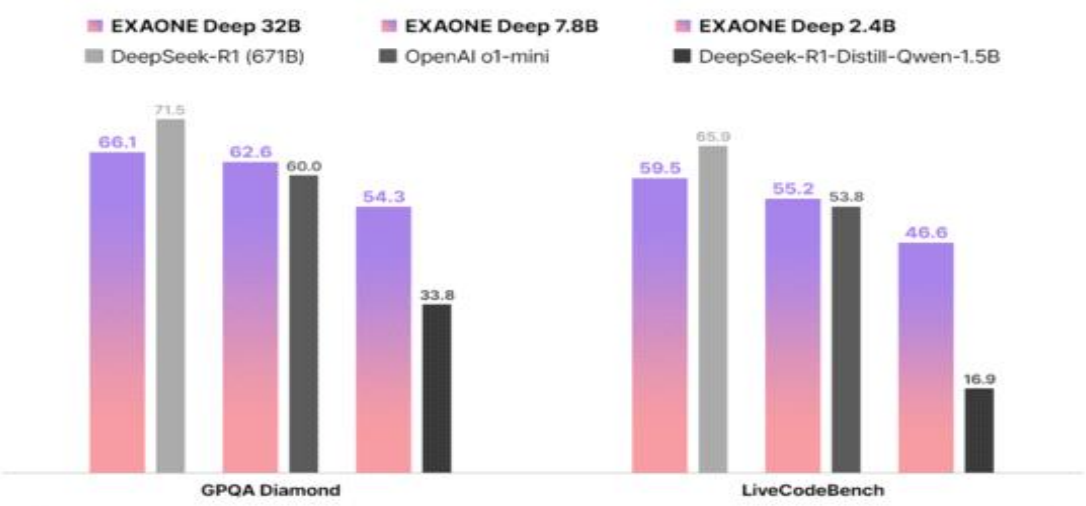
2) 2025년 기업 동향

가. 연비 성능 1등급 '추론 AI' 개발… 오픈소스 공개[595]

LG AI연구원은 2025년 3월 17일 미국 캘리포니아주 새너제이에서 열리는 엔비디아 연례 개발자 콘퍼런스(GTC)에서 참가해 엑사원 딥을 소개했다. 이번에 공개한 모델은 △엑사원 딥-32B(320억 개 매개변수), △엑사원 딥-7.8B △온디바이스 모델 엑사원 딥-2.4B이다. 높은 성능의 모델부터 경량화 버전까지 모두 오픈소스로 공개했다. 이번에 오픈소스로 공개한 모델들은 LG AI연구원의 경쟁력인 '연비 성능'이 장점이다. 적은 매개변수와 학습 데이터로도 글로벌 모델들과 경쟁할 수 있는 성능을 갖췄다. AI 활성화의 골칫거리인 비용과 전력 사용량 부담을 줄이면서 장점인 효율성은 이용할 수 있는 기반을 만들었다는 평가가 나온다.

595) THE AI, LG AI연구원, 연비 성능 1등급 '추론 AI' 개발… 오픈소스 공개, 2025.3.18.
https://www.newstheai.com/news/articleView.html?idxno=7442

[그림 5-32] 엑사원 모델과 글로벌 모델 성능평가 비교



Models	GPQA Diamond (pass@1)	LiveCodeBench (pass@1)
EXAONE Deep 32B	66.1	59.5
DeepSeek-R1 (671B)	71.5*	65.9*
QwQ-32B	63.3	63.4*
DeepSeek-R1-Distill-Qwen-32B	62.1*	57.2*
DeepSeek-R1-Distill-Llama-70B	65.2*	57.5*
EXAONE Deep 7.8B	62.6	55.2
OpenAI o1-mini	60.0*	53.8*
DeepSeek-R1-Distill-Qwen-7B	49.1*	37.6*
DeepSeek-R1-Distill-Llama-8B	49.0*	39.6*
EXAONE Deep 2.4B	54.3	46.6
DeepSeek-R1-Distill-Qwen-1.5B	33.8*	16.9*

출처 : The AI, LG AI연구원, 연비 성능 1등급 '추론 AI' 개발… 오픈소스 공개

LG AI연구원에 따르면, 엑사원 딥-32B는 중국 딥시크 R1보다 가성비 높은 성능을 기록했다. 엑사원 딥-32B의 매개변수는 320억 개로 중국 딥시크 R1의 6,710억 개 매개변수의 5% 수준이다. 하지만 성능평가 결과 유사한 성능을 냈다. 그만큼 적은 비용과 적은 전력만으로 높은 성능을 낼 수 있게 한 셈이다. LG AI연구원은 그동안 글로벌 AI 경쟁에서 좋은 연비를 내세워왔다. 지난해 공개한 엑사원 3.0이 대표 사례다. LG AI연구원이 밝힌 엑사원 3.0에 학습에 쓰인 토큰 개수는 8조 개다. 메타의 AI 모델 라마3.1에 쓰인 토큰 개수는 15조 개와 비교하면 절반 가깝게 적은 양이다.

보통 AI가 더 많은 데이터를 학습할수록 성능이 높아진다는 점을 고려하면 당연히 메타 라

마3.1의 성능이 좋아야 하지만 결과는 달랐다. 엑사원 3.0은 벤치마크 점수 순위에서 실제 사용성을 비롯해 코딩과 수학 영역 등에서 라마3.1보다 높은 점수를 기록했다. AI 연료라 불리는 데이터를 적게 사용하고도 멀리 가는 LLM을 만든 것이다. 이번 추론 모델도 마찬가지다. LG AI연구원이 글로벌 오픈소스 AI 플랫폼 '허깅페이스'에 공개한 '엑사원 딥' 세 모델의 성능평가 결과에 따르면, 이 모델들은 논리적 사고력과 문제 해결 능력 평가에서 높은 성적을 기록했다.

나. 연구 인력 통합… AI 혁신 거점으로 도약[596]

LG AI연구원이 서울 마곡지구 LG사이언스파크 인근 디앤오 사옥으로 본사 이전을 완료했다. 이번 이전으로 여의도와 마곡에 분산돼 있던 300여 명의 연구 인력이 한자리에 모이며, 연구개발 시너지가 더욱 강화될 전망이다. 신사옥은 3층부터 7층까지 총 5개 층으로 구성되며, 약 50개의 다양한 회의 공간을 마련해 자유로운 협업과 창의적인 아이디어 도출을 지원한다. 공간 설계부터 연구원 중심의 업무 환경을 고려해 꾸며졌다. 2020년 12월 출범한 LG AI연구원은 구광모 LG 대표의 전폭적인 지원 아래 LG의 AI 전략을 이끌어왔다. 2021년 국내 최대 규모의 초거대 AI 모델 '엑사원(EXAONE) 1.0'을 공개한 데 이어, 2023년에는 '엑사원 3.5', 올해 3월에는 국내 최초의 추론형 AI '엑사원 딥(EXAONE Deep)'을 선보이며 기술력을 지속적으로 고도화하고 있다.

기업용 AI 에이전트 '챗엑사원(ChatEXAONE)'은 사내 활용이 빠르게 확산하며, LG 사무직 임직원 중 4만 명 이상이 업무에 도입했다. 연구원은 AI 리터러시 교육부터 사내 석·박사 과정까지 전주기 교육 체계를 구축해, 4년간 1만 5,000여 명의 AI 인재를 양성했다. 연구 성과도 주목할 만하다. AAAI, CVPR, NeurIPS 등 세계 유수 AI 학회에 총 234편의 논문을 발표했고, 국내 120건, 국제(PCT) 108건 등 총 228건의 특허를 출원하며 기술 경쟁력을 입증했다.

LG AI연구원은 계열사와의 협업을 통해 산업 현장의 문제 해결에도 나서고 있다. LG전자 노트북 '그램', LG유플러스의 통화 에이전트 '익시오(ixi-O)' 등에 엑사원을 적용하며, AI 기술의 활용 범위를 넓혀가고 있다. 특히 외부 서버와의 연결 없이 기기 내에서 데이터를 처리하는 온디바이스 AI 분야에서도 세계 최고 수준의 성능을 목표로 연구를 강화하고 있다. 더

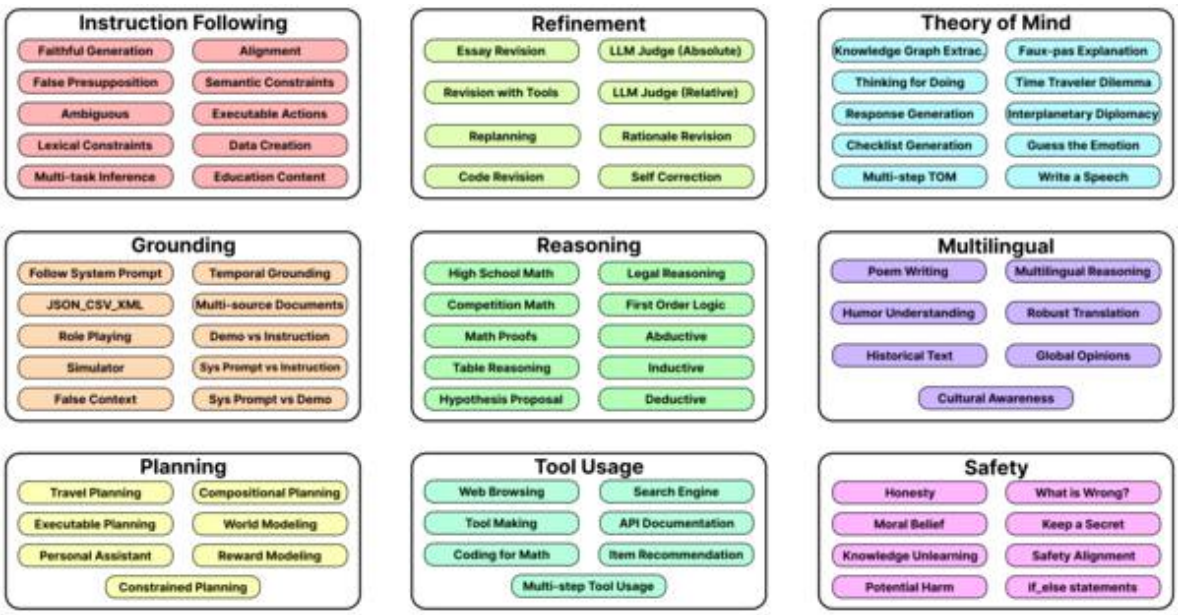
596) THE AI, LG AI연구원, 연구 인력 통합… AI 혁신 거점으로 도약, 2025.5.12.
https://www.newstheai.com/news/articleView.html?idxno=7848

불어 올해부터는 에이전틱 AI, 신약·신소재 설계, 업무 자동화 등 실질적인 성과 도출에 박차를 가할 계획이다.

다. NAACL 2025 '최고 논문상'… 생성 AI 평가 기준 제시

LG AI연구원이 생성 인공지능(AI) 모델을 평가하는 새로운 방식을 제시, 북미 전산언어학회(NAACL) 2025에서 '최고 논문상'을 수상했다.[597] 최고 논문상은 자연어처리(NLP) 분야에서 새로운 연구 방향을 제시하거나 중요한 문제를 해결한 연구 중 올해 가장 혁신적이고 중요성이 높다고 평가를 받은 연구가 선정된다. NAACL 2025에 채택된 1,400편이 넘는 논문 중 단 하나만 선정해 수여한다. 이번에 최고 논문상을 받은 연구는 생성 AI 모델 성능을 평가하는 '빅젠 벤치(BIGGEN BENCH)'다. 서민준 KAIST 교수 연구팀이 주도한 이번 연구에는 연세대, 코넬대, 일리노이대, MIT, 워싱턴대 등 여러 대학의 연구진이 함께 참여했다. 논문 1 저자인 김승원 카네기멜론대 박사 과정 연구생은 LG AI연구원 초지능 랩의 인턴 과정 중 이문태 초지능랩장, 이경재 데이터 스쿼드 리더, 서민준 KAIST 교수 연구팀과 함께 연구를 수행했다.

[그림 5-33] 생성 AI 평가 기준



출처 : LG AI 연구원

597) THE AI, LG AI연구원, NAACL 2025 '최고논문상'… 생성형 AI 평가 기준 제시, 2025.4.30. https://www.newstheai.com/news/articleView.html?idxno=7759

기존 생성 AI 모델 평가 방식은 ‘유용성’이나 ‘무해성’ 같은 추상적인 지표에 의존하고 있어, 사람이 평가했을 때의 결과와 차이가 발생하며 모델의 세부 역량을 측정하기 어려운 한계가 있었다. 이에 비해 ‘빅젠 벤치’는 △지시 사항 수행 △논리적 추론 △도구 사용 능력 △안전성 △다양한 언어 및 문화적 맥락 이해 능력 등 9가지 핵심 역량을 기준으로 765개 평가 항목을 설정해 AI 모델의 성능을 종합적으로 평가한다. 이 평가 방법은 실제 사람의 평가 방식을 모방해 다양한 상황과 주관적 요소를 반영한 결과를 도출할 수 있다.

LG AI연구원은 빅젠 벤치를 사용해 103개의 생성 AI 모델을 평가한 결과, 전문가 집단과의 교차 검증에서 높은 신뢰도와 타당도를 보여 새로운 벤치마크로서의 가능성을 확인했다. 이문태 LG AI연구원 초지능랩장은 “빅젠 벤치는 생성 AI의 다양한 능력을 객관적이고 종합적으로 평가할 수 있도록 설계됐으며, 기존 벤치마크들의 한계를 극복하고 인간의 정교한 평가 기준에 부합하는 장점이 있다”라고 말했다.

라. 세계 3번째 하이브리드 AI 개발[598]

LG AI연구원이 15일 국내 첫 하이브리드 AI 모델 ‘엑사원(EXAONE) 4.0’을 공개했다. 오픈AI가 GPT-5를 하이브리드 AI로 개발 중이라고 밝힌 상황에서, LG가 먼저 완성된 하이브리드 모델을 선보이며 한국의 독자적 AI 모델 개발 역량을 입증했다. 정부가 ‘AI G3(AI 강국 3위)’ 목표와 소버린 AI를 강조하고 있는 시점에서, 한국이 해외 의존 없이 세계 수준의 AI 모델을 만들어낼 수 있음을 보여준 상징적 사건이다. 더욱 주목할 점은 국내 AI 기술 경쟁에서 보여준 LG의 압도적 속도감이다.

지난 6월 30일 네이버가 추론 AI ‘하이퍼클로바X 씽크’를 공개한 지 불과 2주 만에, LG가 한 단계 더 진화한 하이브리드 AI를 선보였다. 네이버의 하이퍼클로바X 씽크가 추론 기능에 특화된 단일 모델이었다면, LG의 엑사원 4.0은 대형언어모델(LLM)과 추론 AI를 완전히 융합한 통합 모델을 구현했다. 단순 추론 AI를 뛰어넘어 진정한 하이브리드 AI를 완성함으로써, 한국 AI 기술력의 새로운 이정표를 세웠다는 평가다.

하이브리드 AI는 기존 LLM의 빠른 답변 능력과 추론 AI의 깊은 사고 능력을 하나로 결합한

598) THE AI, “AI G3 목표 한발 더”... LG AI연구원, 세계 3번째 하이브리드 AI 개발, 2025.7.15. https://www.newstheai.com/news/articleView.html?idxno=8353

차세대 AI 기술이다. 간단한 질문에는 즉시 답변하고, 복잡한 문제에는 스스로 가설을 세우고 검증하며 단계별로 추론해 정확한 해답을 제시한다. 단순히 패턴을 학습해 답변하는 기존 AI와 달리, 상황에 맞는 유연한 사고가 가능하다. 사용자는 필요에 따라 '빠른 모드'와 '깊은 추론 모드'를 선택할 수 있어, 시간과 비용을 효율적으로 관리하면서도 더 신뢰할 수 있는 결과를 얻을 수 있다.

엑사원 4.0은 주요 AI 벤치마크에서 세계 최고 수준의 성능을 기록했다. △AI의 지식수준과 문제 해결 능력을 평가하는 MMLU-Redux에서 92.3점, MMLU-Pro에서 81.8점을 기록했고 △코딩 능력을 측정하는 LiveCodeBench v6에서는 66.7점 △과학 문제 해결 능력을 평가하는 GPQA-Diamond에서 75.4점 △수학 문제 해결 능력을 측정하는 AIME 2025에서는 85.3점을 달성했다. 이는 구글, 메타, 마이크로소프트, 알리바바, 미스트랄 AI 등 글로벌 기술 기업들의 대표 오픈소스 모델들을 모두 제치고 달성한 성과다.

마. 차세대 정밀 의료 AI '엑사원 패스 2.0' 공개[599]

LG AI연구원이 9일 차세대 정밀 의료 AI 모델인 '엑사원 패스(EXAONE Path) 2.0'을 9일 공개했다. 엑사원 패스 2.0은 병리 조직 이미지로 유전자 변이와 발현 형태, 인체 세포와 조직의 미세한 변화 및 구조적 특징을 정밀하게 분석·예측할 수 있는 차세대 정밀 의료 AI 모델이다. 암 등 질병의 조기 진단과 예후 예측, 신약 개발과 개인화된 맞춤 치료 등에 활용할 수 있다.

엑사원 패스 2.0은 병리 조직 이미지와 생명 현상을 이해하고 질병의 원인과 치료법 연구에 활용할 수 있는 DNA와 RNA 등의 유전자와 관련된 멀티오믹스 데이터를 학습했다. LG AI연구원은 이미지를 수천 개 조각으로 나눈 패치 단위부터 세포와 조직 구조 정보를 담은 대용량의 전체 슬라이드 이미지까지 학습하는 신기술을 적용했다. 이를 통해 유전자 변이 예측 정확도를 세계 최고 수준인 78.4%까지 높였다. LG AI연구원은 또 폐암과 대장암 등 특정 질병 특화 모델도 추가로 공개했다.

박용민 LG AI연구원 AI 비즈니스팀 리더는 "기존 2주 이상의 유전자 검사 시간을 1분 이내

599) THE AI, LG AI연구원, 차세대 정밀 의료 AI '엑사원 패스 2.0' 공개, 2025.7.9.
https://www.newstheai.com/news/articleView.html?idxno=8295

로 단축해 암 환자의 치료 골든타임을 확보하는 데 도움을 받을 수 있다“라며 “어떤 유전자에서 변이가 발생했는지 빠르게 확인하고 표적 치료제를 식별할 수 있다“라고 말했다. 한편, LG AI연구원은 오는 22일 ‘LG AI 토크콘서트 2025’에서 엑사원 패스 2.0을 소개할 예정이다.

바. LG ‘엑사원 4.0’, 세계 11위·오픈모델 4위로 한국 AI 위상 입증[600]

LG AI연구원이 최근 공개한 ‘엑사원(EXAONE) 4.0’이 글로벌 AI 성능 분석 전문기관 ‘아티피셜 어낼리시스(Artificial Analysis)’의 인텔리전스 지수 평가에서 종합 11위를 기록했다. 한국 AI 모델만 보면 1위에 해당하는 기록이다. 아티피셜 어낼리시스는 추론과 지식 역량(MMLU-Pro, Humanity’s Last Exam), 전문가 수준의 과학 문제 해결 능력(GPQA Diamond), 코딩 능력(LiveCodeBench, SciCode), 수학 문제 해결 능력(AIME 2024, MATH-500) 등 7개 고난도 성능평가 지표를 종합해 인텔리전스 지수를 산정하는 글로벌 AI 성능 분석 전문기관이다.

엑사원 4.0은 아티피셜 어낼리시스가 실시한 7개 고난도 성능평가 지표에서 총 64점을 기록했다. 미국 앤트로픽의 최상위 모델인 클로드 4 오퍼스(Claude 4 Opus)와 동일한 점수다. 특히 코딩 분야에서는 7위, 수학 분야에서는 10위를 기록하며 각 영역에서 ‘상위 10위’를 달성했다. 오픈 웨이트(공개) 모델 기준으로는 알리바바 Qwen3, 딥시크 Deepseek R1, Z.ai의 GLM-4.5에 이어 4위에 등극했다. 모델 크기가 수 배에서 수십 배에 달하는 글로벌 프런티어 AI 모델들과 경쟁할 수 있는 성능을 입증했다. LG AI연구원은 AI 모델 개발 기업 기준으로는 8위에 올랐다.

LG AI연구원은 지난 15일 엑사원 4.0을 글로벌 오픈소스 AI 플랫폼인 허깅페이스(Hugging Face)에 공개했다. 32B(매개변수 320억 개) 모델은 공개 2주 만에 50만 다운로드를 돌파하며 국산 AI 모델 중 최단기간 기록을 달성했으며, 현재 55만을 넘어섰다. 올해 LG AI연구원이 공개한 국내 최초 추론 AI 모델 ‘엑사원 딥’(3월)과 하이브리드 AI 모델 ‘엑사원 4.0’(7월)은 미국 비영리 AI 연구기관 에포크 AI(Epoch AI)의 ‘주목할 만한 AI 모델’ 리스트에 연이어 선정됐다. 이 리스트는 매년 미국 스탠퍼드대 AI 보고서에서 국가별·기업별 AI 경쟁력 비

600) THE AI, LG ‘엑사원 4.0’, 세계 11위·오픈모델 4위로 한국 AI 위상 입증, 2025.7.30. https://www.newstheai.com/news/articleView.html?idxno=8516

교 자료로 활용된다.

사. LG AI연구원-런던증권거래소, 금융 AI 에이전트 상용화[601]

LG AI연구원과 런던 증권 거래소 그룹(LSEG)이 금융 AI 에이전트 '엑사원(EXAONE) Business Intelligence(엑사원-BI)' 상용화에 나선다고 19일 발표했다. 양사는 이날 런던증권거래소에서 상용 서비스 개시를 알리는 행사를 개최했다. 이번 협력은 한국과 영국 간 첫 금융 분야 AI 협력 사례로, 한국의 AI 기술력이 국제 금융 시장에서 인정받는 계기가 될 것으로 평가된다. LSEG는 런던증권거래소를 운영하며 전 세계 투자자들에게 금융 데이터와 분석 서비스를 제공하는 국제 금융 인프라 선도기업이다.

행사에는 토드 하트만 LSEG 데이터 및 피드 그룹 총괄, 이보 데커스 유럽·중동·아프리카 영업 그룹 총괄 등 LSEG 경영진과 이홍락·임우형 LG AI연구원 공동 연구원장, 이화영 AI 사업 개발 부문 장 등이 참석했다. 토드 하트만 총괄은 "이번 협력은 고객에게 신뢰성 높은 데이터를 제공해 의사결정을 지원하고, 금융 생태계 전반에 더 큰 가치를 창출하는 혁신의 기회"라며 "LG와의 파트너십은 새로운 시작을 의미하며, AI가 전 세계 투자자들에게 예측부터 의사결정 지원까지 더 나은 기회를 제공하는 핵심 역할을 할 것"이라고 밝혔다.

엑사원-BI는 인간의 개입 없이 AI가 데이터 분석부터 미래 예측, 보고서 작성까지 전 과정을 자동으로 수행하는 금융 AI 에이전트다. LG AI연구원은 각각 다른 전문 영역을 담당하는 4개의 AI 에이전트가 협력하는 구조로 시스템을 설계했다. AI 저널리스트는 뉴스, 기업 공시, 거시경제 지표 등 방대한 외부 데이터를 수집하고 정제한다. AI 경제학자는 이 데이터를 바탕으로 미래 시장 흐름과 경제 전망을 예측한다. AI 애널리스트는 경제학자의 예측과 내부 지표 분석을 종합해 종목별 핵심 영향 요인과 이상 신호를 포착하고, 이해하기 쉬운 설명이 담긴 보고서를 생성한다. AI 의사결정자는 애널리스트 보고서와 다양한 시나리오를 비교 평가해 투자 점수를 산출한다.

601) THE AI, LG AI연구원-런던증권거래소, 금융 AI 에이전트 상용화, 2025.9.22.
https://www.newstheai.com/news/articleView.html?idxno=9126

3. SKT

1) 기업 일반현황

<표 5-14> SKT 기업 현황

구분	주요 내용
기업 개요	**〈회사 개요〉** • SK텔레콤은 1984년 설립된 국내 1위 이동통신 기업으로, 5G·AI·클라우드·미디어·모빌리티 영역까지 확장하며 ICT 종합 기업으로 전환 • 2021년 'AI 컴퍼니' 비전을 선언하며, 통신 중심 구조에서 AI·DATA 기반 서비스 사업자로의 변화를 가속화 • 글로벌 빅테크, 국내 대기업과 협력해 AI 인프라·AI 서비스·AI 플랫폼 중심의 새로운 성장축을 구축 중이며, 특히 2023년 이후 초거대 AI '에이닷(A.)' 생태계 확장에 집중 • 2024년 AI 전략 조직을 개편하고 글로벌 파트너십(샘 올트먼·SKT 대화 모델 협력 등)을 확대하며, 통신·엔터·보안·모빌리티까지 AI 기반 재편을 추진
	〈사업 내용〉 • 주요 산업 분야 : 이동통신, AI 플랫폼·서비스, 데이터·클라우드, 미디어·콘텐츠, 보안, 모빌리티, 스마트 팩토리, IoT 등 ICT 전 분야 • 주요 서비스: 5G 통신, 에이닷(A.), 웨이브(Wavve), 티맵(TMAP), ADT캡스·SK쉴더스, 스마트홈·IoT, 엔터프라이즈 클라우드·IDC 등
	〈자회사〉 • SK쉴더스(보안), 웨이브(콘텐츠), 콘텐츠웨이브, FLO(음악), ADT캡스, 티맵(TMAP Mobility), SK텔링크, 11번가(전자상거래) 등.
AI 조직 현황	**〈AI 관련 조직〉** • AIX센터(AI Transformation Center) : 초거대 AI 개발·적용 총괄 조직으로, 언어·비전·음성 기술, 플랫폼 엔지니어링을 담당하며 'A.' 서비스 기반 모델을 개발 • IT 연구원(구 SKT 종합기술원) : 통신 AI, 네트워크 최적화, 6G 연구, 엣지 컴퓨팅, 미래 모빌리티·로보틱스 연구 등을 수행 • AI 서비스 개발 조직 : 에이닷(A.)·T 전화 AI·고객센터 AI·미디어 추천 AI·보안 AI 등 서비스형 AI를 개발·운영. • 글로벌 AI 협력팀: 글로벌 AI 빅테크(OpenAI 등)와의 전략적 파트너십, 해외 AI 기술 공동개발, 모델 고도화 추진 조직으로 강화
사업 전망	**〈AI 전략〉** • SKT는 'AI 컴퍼니' 전략 아래 초거대 AI+통신 인프라 + 플랫폼 서비스를 통합해 AI 기반의 신규 수익모델 확보에 집중 • 'A.'(에이닷)는 개인 비서 AI·전화·메시지·검색·고객센터 자동화 등 생활형 AI 플랫폼으로 확장되며, SKT 서비스 전반의 인터페이스 역할을 강화 • 기업시장에서는 AI CX·보안 AI·모빌리티 AI·스마트팩토리 솔루션 등을 중심으로 B2B 수익 확대를 추진하며, 네트워크 AI·트래픽 최적화 등 통신과 생성 AI 적용에 집중 • 중장기적으로 SKT는 AI 네이티브 통신사(통신망 자체가 AI 기반으로 운영) 전략과 글로벌 AI 파트너십을 기반으로 국내 통신사 중 가장 빠른 AI 전환을 목표로 함

2) 2025년 기업 동향

가. 구독형 AI 클라우드 사업 본격 전개[602)]

SK텔레콤이 2025년 1월 13일 AI 인프라 시장 공략에 나서 AI 클라우드 서비스 ‘SKT GPUaaS(GPU-as-a-Service)’를 공식 출시했다. 미국 람다와 함께 AI 클라우드 시장에서 수익 확보에 나선다는 방침이다. SKT는 지난 12월 30일 가산 AI 데이터센터(이하 AIDC)를 오픈하고, 시범 운영을 마친 뒤 AI 클라우드 서비스인 ‘SKT GPUaaS(GPU-as-a-Service)’를 출시했다고 13일 밝혔다. SKT GPUaaS’는 미국 기업인 람다와 1여 년간 함께 준비한 구독형 AI 클라우드다.

SKT는 지난 2월 AI DC 사업 본격 추진을 위한 첫 번째 글로벌 행보로 글로벌 그래픽처리장치(GPU) 클라우드 회사인 람다에 투자를 진행한 바 있다. 2012년 AI 엔지니어들이 설립한 람다는 글로벌 빅테크 기업들을 대상으로 AI 클라우드 서비스를 제공하는 GPUaaS 기업이다. SKT는 람다와 협력해 GPU를 안정적으로 확보하고 이를 기반으로 AI 클라우드 시장을 공략한다는 계획이다. 이번에 출시하는 GPUaaS는 엔비디아 GPU H100을 기반으로 한다. H100은 현재 국내에 도입된 GPU 중 가장 성능이 높은 것으로 알려졌다. 다음 모델인 B100은 아직 양산 중이다.

SKT GPUaaS 구독을 원하는 기업 고객은 AI 서비스 규모나 목적에 따라 GPU 수량과 기간을 선택하고, 단독 서버・방화벽・전용회선 등 맞춤형 패키지를 구성할 수 있다. 고객사는 독자적인 GPU를 사용해 안정적 AI 학습과 추론을 시행할 수 있으며, 방화벽과 전용회선을 활용하면 보안성 역시 크게 강화할 수 있다. 또한 고객의 AI 작업량이 급격히 늘어나더라도 단독 서버에 GPU를 손쉽게 추가할 수 있어 추가 시설 투자에 대한 부담을 크게 덜 수 있다. 또한 SKT GPUaaS는 퍼블릭(공유형) 클라우드와 상호 연동이 가능하다. 기업 고객은 기존에 사용하던 클라우드를 그대로 활용하면서 SKT GPU 컴퓨팅 파워를 이용할 수 있다.

SKT는 GPUaaS와 자사 ‘AI 클라우드 매니저’를 함께 활용하면 GPU 자원을 더욱 효율적으로 관리할 수 있다고 밝혔다. AI 클라우드 매니저는 수많은 GPU 자원을 마치 한 대의 컴퓨터처럼 관리해 GPU 성능을 극대화하고, AI 개발을 위한 학습 소요 시간을 단축하도록 지원하

602) THE AI, SKT, 구독형 AI 클라우드 사업 본격 전개, 2025.1.13.
https://www.newstheai.com/news/articleView.html?idxno=6894

는 솔루션이다. SKT GPUaaS 가격은 약정 기간, GPU 개수, 선불형의 과금 형태에 따라 탄력적으로 책정해 고객 편의를 크게 강화했다. 예를 들어 고객이 24개월 동안 32개의 GPU를 원할 경우 이에 맞는 가격을 설정하는 방식이다. 1개월, 2개월 등 단기 서비스 이용도 가능하다.

나. AI 에이전트 '에이닷'에 최신 제미나이 추가[603]

SK텔레콤(SKT)이 AI 에이전트 '에이닷(A.)'에 구글의 자체 대형언어모델(LLM) '제미나이(Gemini)' 최신 모델인 '제미나이 2.0 플래시(Flash)'를 추가했다고 2025년 3월 18일 밝혔다. 이번 AI 모델 확대 제공을 통해 이용자들은 SKT의 자제 AI 모델 에이닷(A.X)과 오픈AI 챗GPT, 앤트로픽 클로드', '퍼플렉시티', '제미나이' 등 글로벌 대표 AI 서비스 세부 모12종을 사용할 수 있다. 현재 이용 가능한 AI 모델은 △A.X △제미나이 2.0 플래시 △GPT 4o, 4o mini, o3-mini, o1-mini, o1-preview 5개 모델 △클로드 3.5 Sonnet, 3.5 Haiku, 3 Opus 3개 모델 △퍼플렉시티 Sonar, Sonar Pro 2개 모델 등이다.

에이닷에 새롭게 추가된 '제미나이 2.0 플래시'는 구글이 지난 2월 공개한 최신 AI 모델이다. '제미나이 2.0 플래시'는 기존 '1.5 플래시' 모델을 개선한 버전으로 응답 시간과 성능이 대폭 향상됐다. 주요 벤치마크에서 '1.5 프로' 모델보다 2배 빠른 속도를 기록하는 등 뛰어난 성능을 자랑한다. 이용자들은 답변과 함께 구글의 실시간 검색 결과를 함께 받는 것은 물론 답변이 어떤 내용을 토대로 생성됐는지 확인할 수 있도록 출처 페이지 이동 기능도 쓸 수 있다.

SKT는 이번 '제미나이 2.0 플래시' 모델을 시작으로 향후 고객들이 구글의 최신 LLM을 지속적으로 체험할 수 있도록 할 계획이다. 이에 앞서 SKT는 이달 초 '퍼플렉시티 소나'의 최신 고성능 모델인 '퍼플렉시티 소나 프로' 모델과 복잡한 논리 문제 해결이 가능한 'GPT o3-mini' 모델을 각각 에이닷에 추가했다. '퍼플렉시티 소나 프로'는 기존 '소나' 모델과 비교해 2배 많은 검색 출처를 활용한 답변을 제공하며, 연속되는 대화에서 이전 대화의 맥락을 더 잘 이해해 다음 답변의 적합성이 뛰어나다는 장점이 있다. 또 'GPT o3-mini' 모델은 고급 수학 공식의 해석, 프로그래밍 코드 생성 등에서 뛰어난 성능을 발휘한다.

---

603) THE AI, SKT, AI 에이전트 '에이닷'에 최신 제미나이 추가, 2025.3.18.
https://www.newstheai.com/news/articleView.html?idxno=7445

이용자들은 서로 다른 특징을 가진 11개 AI 모델 가운데 원하는 모델을 선택해 사용할 수 있다. 다른 모델에 동일한 질문을 하고 싶을 때 일일이 복사해서 붙여 넣는 번거로움 없이 ‘다른 AI 비교하기’ 기능을 통해 한 번에 비교할 수 있다. 후속 대화를 진행할 때 모델을 변경하면 이전 대화 맥락을 반영해 자연스럽게 이어지는 교차 대화도 가능하다.

다. AI 서비스에 리벨리온 NPU 적용[604]

SK텔레콤이 자사 주요 AI 서비스에 국내 반도체 스타트업 리벨리온의 NPU(신경망처리장치)를 적용하기 위한 테스트를 진행 중이라고 2025년 6월 25일 밝혔다. SKT와 리벨리온은 ‘데이터센터용 AI 반도체 아톰(ATOM)’을 탑재한 서버를 △에이닷 전화 통화 요약 △PASS 스팸 필터링 △PASS 금융비서 △엑스칼리버 등에 테스트 중이다. 이들은 테스트 결과에 따라 아톰의 성능 강화판인 ‘대규모 AI 추론용 AI 반도체 아톰 맥스(ATOM-Max)’를 이들 서비스에 연내 상용 적용하는 것을 목표로 하고 있다. 또 SKT는 테스트 결과를 토대로 향후 리벨리온 NPU 서버 도입을 지속적으로 확장할 계획이다.

양사는 이번 SKT AI 서비스에 대한 리벨리온 AI 반도체 테스트 및 상용화 검토가 국내 AI 생태계 자립에 큰 의미가 있다고 설명했다. 에이닷 전화 통화 요약은 SKT LLM 모델 에이닷엑스(A.X)를 기반으로 하고 있다. 국산 AI 반도체인 리벨리온의 아톰을 통해 상용 서비스되면 ‘국산 LLM’이 ‘국산 NPU’를 통해 서비스되는 ‘소버린 AI(국가 주권형 AI)’가 이뤄진다는 것이다. 양사는 이번 테스트를 시작으로 협력 범위를 지속 확대해 국내 AI 생태계의 자립성을 강화해 나갈 계획이다.

이상민 SKT 성장 사업추진실장은 “SKT는 자체 개발한 한국어 특화 LLM 에이닷엑스(A.X) 4.0는 현존하는 LLM 중 최고 수준의 한국어 이해 성능을 확보하고 있어 국내 비즈니스 환경에 최적화된 모델이라는 평가를 얻고 있다”라며 “국가 간 AI 기술 경쟁이 심화하고 있는 이때, 국내 AI 생태계 자립성 강화는 개별 기업이 아닌 국가의 경쟁력을 좌우할 수 있는 중요한 요소”고 말했다.

---

604) THE AI, SKT, AI 서비스에 리벨리온 NPU 적용, 2025.6.25.
https://www.newstheai.com/news/articleView.html?idxno=8166

라. 에이닷에 노트·브리핑 AI 서비스 2종 베타 출시[605]

SK텔레콤이 2025년 6월 30일 자사 AI 서비스 '에이닷(A.)'에 음성을 텍스트로 변환해 기록·요약 해주는 '노트'와 사용자의 일상을 종합 분석해 정보를 제공하는 '브리핑' 등 2종의 신규 서비스 베타 버전을 출시했다. 에이닷 '노트'는 음성으로 이뤄지는 모든 순간을 AI로 기록하는 서비스이다. 회의나 강의, 상담 등 일상에서의 음성을 AI가 실시간으로 받아쓰고 요약하고 정리해 준다. 서비스에 적용된 AI 기술은 음성 인식뿐만 아니라 음성을 텍스트로 변환해 문맥에 맞춰 자연스러운 문장 형태로 수정해 준다. 또 문서 형태로도 만드는 등 다양한 경험을 제공한다. 음성 변환 내용 중 이해되지 않는 개념이나 낯선 용어가 나오면 별도의 앱이나 검색 서비스를 이용하지 않고도 AI를 통한 웹검색으로 정보를 찾아볼 수 있다.

에이닷 '노트'는 녹음 종료 후 △요점 정리 △회의록 △강의 노트 가운데 원하는 템플릿을 선택하면 해당 형식에 맞춰 중요한 키워드와 주제어를 중심으로 내용을 묶어 정리해 준다. '강의 노트'를 선택하면 녹음 완료 후 핵심 개념 정리 등을 포함하는 강의 요약은 물론 강의 흐름에 맞춰진 구조화된 정리까지 제공받을 수 있다. 완성된 노트의 공유도 쉽다. '노트'에서 공유 기능을 켜면 링크 주소가 생성, 공유하고 싶은 대상에게 전달하면 된다. 공유된 노트에는 녹음파일을 제외한 노트 제목, 작성 일시, 녹음 길이, 참석자 정보, 받아쓰기 내용, 실시간 요약 등이 포함된다. SKT는 에이닷 '노트' 서비스를 에이닷 앱과 PC 웹에서 모두 사용할 수 있도록 해 이용자들의 활용 편의를 최대한 높인다는 계획이다. 오픈 베타 형태로 제공되는 '노트'는 1회 녹음 당 최대 100분, 월 600분 분량이 제공된다.

에이닷 '브리핑'은 사용자의 일상을 종합적으로 분석해 자연스럽게 정보를 제공하는 '앰비언트(Ambient) 에이전트' 서비스이다. SK텔레콤을 이 기능을 'AI 개인 비서'라는 에이닷의 정체성을 강화하기 위해 개발했다. '브리핑' 서비스는 원하는 고객에 한해 직접 요청 없이도 상황과 맥락을 고려, 선제적으로 정보를 제공한다. 예를 들어 하루를 시작할 때 '브리핑' 서비스가 이용자의 정해진 일정에 따른 동선별 날씨나 관심사 기반의 콘텐츠를 추천한다.

SKT는 해당 서비스의 특성상 한 번의 브리핑이 아닌 이전 브리핑 이력과 사용자의 반응과

605) THE AI, SKT, 에이닷에 노트·브리핑 AI 서비스 2종 베타 출시, 2025.6.30.
https://www.newstheai.com/news/articleView.html?idxno=8216

피드백을 기억해 점점 더 정밀하고 개인화된 ‘나만의 AI’ 서비스로 진화, 발전시켜 나갈 계획이라고 설명했다. 한편, SKT는 올 하반기 중 에이닷 ‘노트’와 ‘브리핑’ 서비스의 베타 테스트를 진행해 사용자의 반응과 사용 의견 등을 확인하고 적용해 이르면 올해 내 정식 버전으로 출시할 계획이다.

마. 에이닷 엑스 3.1 오픈소스 공개[606]

SK텔레콤이 2025년 7월 24일 자사 초거대 한국어 특화 대형언어모델(LLM) ‘A.X 3.1(에이닷 엑스 3.1)’을 오픈소스 커뮤니티 허깅페이스에 공개했다. 340억 개(34B)의 매개변수를 기반으로 하는 ‘A.X 3.1’은 SKT가 모델 구축과 데이터 학습 등 전 단계를 직접 진행했다. 이번에 공개한 A.X 3.1은 A.X 3.0의 대폭 개선된 버전이다. 한국어 대화 성능에 집중했던 A.X 3.0과는 달리 A.X 3.1에서는 추론 모델로의 확장 가능성을 고려해 코드와 수학 성능을 대폭 향상했다.

이번 모델 공개를 포함하면 SKT는 독자 개발 방식의 A.X 3.1 모델 2종(표준, 경량)과 대규모 학습(CPT)에 기반한 A.X 4.0 모델 2종(표준, 경량)까지 총 4개의 모델을 오픈소스로 공개했다. ‘A.X 3.1(34B)’은 동일한 한국어 서비스 적용 환경에서 A.X 4.0(72B) 대비 약 90% 수준의 성능을 보인다. 대표적 한국어 벤치마크 평가인 KMMLU에서 69.2점을 기록했다. 한국 문화 이해도 평가인 CLIcK에서는 77.1점을 획득해 각각 A.X 4.0 대비 88%, 90% 수준의 성능을 입증했다.

SKT는 2018년 KoBERT 개발을 시작으로 KoGPT2, KoBART 등을 연이어 공개하며 국산 언어모델 기술의 토대를 닦아왔다. 이후 A.X 시리즈로 상용성과 감성 대화, 지식 추론 등 기능을 확대하며 한국어 특화 LLM 기술을 고도화해 왔다. 또한 이번 오픈소스 공개는 과학기술정보통신부의 ‘독자 인공지능(AI) 파운데이션 모델’ 사업과도 궤를 같이한다. SKT는 반도체, 모델, 데이터, 서비스로 이어지는 ‘풀스택 AI’ 구현을 목표로 다양한 기업 및 연구기관과 컨소시엄을 구성했다. 컨소시엄에는 크래프톤, 모빌리티, 포티투닷, 리벨리온, 라이너, 셀렉트스타 등 기업들과 서울대·KAIST 연구진 등이 참여한다.

606) THE AI, SKT, 에이닷 엑스 3.1 오픈소스 공개, 2025.7.24.
https://www.newstheai.com/news/articleView.html?idxno=8456

바. SKT · 크래프톤, 추론 특화 언어모델 공동 개발[607]

SK텔레콤 · 크래프은 2025년 7월 28일 공동으로 추론 특화 언어모델 3종을 오픈소스로 공개했다고 밝혔다. 이번에 공개한 모델은 개발한 70억 개 파라미터(7B) 규모이다. 기존의 일반에 오픈소스로 공개된 '오픈싱커2(OpenThinker2)', '오픈싱커3(OpenThinker3)', '에이스리즌-네모트론 1.1(AceReason-Nemotron-1.1)' 등 언어모델을 파인 튜닝한 새로운 모델이다.

SKT와 크래프톤은 언어모델을 공동 개발하며, 각각 인프라 구축 및 학습 기법 개선을 통해 모델의 품질과 성능 고도화에 기여했다. 크래프톤은 기존 모델의 취약점을 분석해 이를 개선하는 오답 복기 학습 기법을 자체 개발했다. 해당 기법은 틀린 문제의 정답을 찾아 오답과 비교해 학습하며, 추론 정확도와 효율성을 동시에 확보하는 전략적 학습 방식이다. SKT는 데이터 검증과 모델 학습의 인프라 구축을 담당하여, 모델의 품질과 안정성을 확보하는 데 기여했다.

앞으로 크래프톤은 이번 언어모델에 적용한 학습 기법을 게임 플레이 분석, 전략 판단 등 게임 특화형 AI 응용 기술 고도화에 활용할 계획이다. 나아가 다양한 규모의 LLM을 개발해, 한국형 AI 기술 생태계의 독립성과 주도권을 확보하는 데 기여할 방침이다. 언어모델은 글로벌 오픈소스 플랫폼인 허깅페이스를 통해 공개했다.

사. 전사 AI 조직 통합 'AI CIC' 출범[608]

SK텔레콤이 2025년 9월 26일 AI 사업을 이끄는 사내 회사 AI CIC(Company in Company) 출범을 발표했다. 유영상 SKT 대표는 이날 전 구성원 대상 타운홀 미팅을 열고 "급변하는 AI 환경 변화에 따라 위기와 기회가 공존하는 AI 골든타임을 놓치지 않기 위해서는 수요자 관점의 내부 AI 혁신과 공급자 관점의 AI 사업 혁신이 필요하다"라며 "AI CIC를 새롭게 출범시킨다"라고 말했다. AI CIC는 유영상 CEO가 대표를 겸임하고 세부 조직 개편은 10월 말 시행된다.

607) THE AI, SKT·크래프톤, 추론 특화 언어 모델 공동 개발, 2025.7.28.
https://www.newstheai.com/news/articleView.html?idxno=8473
608) THE AI, SKT, 전사 AI 조직 통합 'AI CIC' 출범, 2025.9.26.
https://www.newstheai.com/news/articleView.html?idxno=9192

AI CIC에는 흩어져 있던 AI 조직을 하나로 통합한다. 에이닷(A.) 서비스 △기업 대상 에이닷 비즈(A. Biz) 서비스 △AIDC(데이터센터) 사업 △글로벌 AI 제휴 투자 △AI R&D △메시징/인증 사업 등 기능과 조직을 AI CIC 체제로 재편해 사업 생산성과 효율성을 높이는 AI 혁신에 나선다. AI CIC는 향후 5년간 약 5조 원 규모의 AI 투자를 단행해 오는 2030년까지 연 매출 5조 원 이상을 달성할 계획이다. 이를 위해 AI ‘수요(AI B2C, AI B2B)’와 ‘공급(AI 인프라)’의 안정적 선순환을 통해 독자 생존할 수 있는 수익 구조를 확보할 예정이다. B2C는 에이닷을 중심으로 국내외 가입자 기반을 확대해 지속 가능한 수익 구조 확보에 집중한다. B2B는 에이닷 비즈 등 에이전틱(Agentic) AI 서비스를 통한 생산성 향상 입증 및 제조 AI 분야 사례를 확보하는 등 사업 확장을 가속한다.

AIDC 사업은 글로벌 빅테크를 타깃으로 AIDC 유치 및 정부 사업 참여를 통해 국내 최고 수준의 AIDC 사업자로 자리매김한다는 목표이다. 아울러 디지털 트윈 • 로봇 등 미래 성장 영역 기술 개발(R&D) 및 AI 풀스택(Full Stack) 역량 기반의 국내외 파트너십 확장 등도 속도감 있게 추진한다. 유영상 CEO는 이날 전사 시스템, 조직 문화, 구성원 역량 강화에 AI를 도입해 경쟁력을 높이는 내부 AI 혁신 방안도 함께 발표했다. 먼저 AI 역량을 보유한 구성원들을 조직 전반에 균형 있게 배치함으로 전사 AI 역량 강화를 추진하고 구성원 개개인의 경력도 확장하는 시너지 효과를 창출한다는 계획이다.

또 AI, 디지털 전환(DT)를 기반으로 전사 시스템・인프라 측면의 AX(AI 전환)을 가속한다. 이동통신 사업은 주요 업무 프로세스 최적화・자동화 및 온오프라인 고객 접점 혁신을 이루고, 네트워크 사업은 통신 네트워크 운영 자동화 향상, 지능형 통신 네트워크 기반 서비스 안정성 제고 등 AX기반의 업무 혁신을 추진한다.

4. NC AI

1) 기업 일반현황

<표 5-15> NC AI 기업 현황

구분	주요 내용
기업 개요	**〈회사 개요〉** • NC AI는 엔씨소프트가 2011년 AI센터 설립을 시작으로 발전시킨 AI 전문 연구조직으로, 게임 AI·멀티모달 AI·디지털 휴먼·생성 AI 기술 개발을 목표로 함 • AI·로보틱스·엔진 기술을 기반으로 게임 서비스 품질 향상, NPC 지능화, 음성·대화형 에이전트 개발을 전담하며, 2020년 이후 생성 AI 성과가 본격화 • 엔씨소프트의 “AI·R&D 중심 기술 기업 전환” 전략에 따라 독립적인 AI Lab/Center 형태로 운영되고 있으며, 2024~2025년 게임 외 콘텐츠·디지털 휴먼·창작 도구 영역으로까지 범위를 확장 • 자체 AI 연구 시설 ‘NC Research’와 파생 조직 ‘NC Future Center’를 통해 미래 엔터테인먼트 기술과 차세대 인터페이스 R&D를 수행
	〈사업 내용〉 • 연구 분야 : 게임 AI, 자연어처리(NLP), 음성합성(TTS), 음성 인식(STT), 비전·모션 AI, NPC 행동 모델, 생성 콘텐츠 툴 등 엔터테인먼트 특화 AI • 2023년 이후 AI 기반 대화 NPC, 자동 콘텐츠 생성(AIGC), 게임 QA 자동화, 실시간 번역 AI, 감정·표정 모션 AI가 주요 개발 영역으로 확대
	〈자회사〉 • 독립 자회사는 없으며, 모회사 엔씨소프트의 리니지·블레이드앤소울·TL 프로젝트·신규 IP 개발본부와 긴밀하게 협업
AI 조직 현황	**〈AI 관련 조직〉** • Language AI팀: 대화형 모델, 게임 내 NPC 대사 생성, 실시간 번역·챗봇·문장 생성 모델을 개발. • Speech AI팀: 음성합성·음성 인식·보이스 캐릭터 모델 개발(게임 캐릭터 보이스, 실시간 대화 보이스 적용). • Vision·Motion AI팀: 3D 캐릭터 모션, 감정 표현, 얼굴·신체 추적, 동작 자동 생성 기술을 연구. • Game AI/Applied AI팀: 몬스터 행동, NPC 전략 모델링, 난이도 조절, QA 자동화, 플레이 로그 분석 등을 담당하며 게임 개발 조직과 직접 연계
사업 전망	**〈AI 전략〉** • NC AI 핵심 전략은 ‘게임 제작 과정의 지능화+AI NPC·게임 경험 혁신+크리에이터 도구(AIGC) 제공’으로 요약 • 생성 AI 기반 NPC·보이스·모션·스토리 자동화 기능은 개발 생산성을 높이고 신규 IP 경쟁력을 강화하는 방향으로 확대될 전망 • 내부 개발 파이프라인에 AI를 적용해 QA 자동화·텍스트/이미지/모션 생성·번역 등 저비용·고효율 개발 체계 구축을 추진 • 중장기적으로는 AI 엔터테인먼트 기업으로의 전환을 준비하며, 게임 외 분야(디지털 휴먼·메타 휴먼·버추얼 콘텐츠)로 AI 기술을 확장할 것으로 예상

2) 2025년 기업 동향

가. NC AI, 독자 AI 국가대표 기업으로 선정[609)]

NC AI가 대한민국의 AI 주권과 산업 경쟁력을 모두 확보하는 것을 목표로 국가대표 AI 기업들을 선정하는 '독자 AI' 프로젝트에서 국가대표 5개 사 가운데 하나로 선정됐다. NC AI는 이번 선정이 단지 지금의 기술력을 평가한 것이 아닌 NC AI가 지난 14년 동안 쌓아온 꾸준한 연구 노력과 긴 호흡의 전략적 투자, 그리고 올해 2월 정식 분사 이후 6개월 만에 보여준 눈부신 속도감 있는 성과들이 종합적으로 인정받은 결과라고 강조했다. NC AI는 국내에서 가장 오랜 기간 인공지능을 전담해 연구해 왔다. 2011년, 엔씨소프트 내부에서 독립된 형태의 AI 조직을 설립해 감정형 음성합성 기술, 게임 시나리오 기반의 자연어 처리(NLP), 얼굴 모션 애니메이션 처리, 3D 비전 기반 로보틱스 연구 등을 진행했다. NC AI는 올해 초 독립법인으로 분사해 AI 전문기업으로 거듭났고 반년 만에 국가 전략 프로젝트의 대표 기업으로 선정됐다.

이번 성공의 중심에는 NC AI가 주도하고 전방위로 이끈 '그랜드 컨소시엄'이 있다. NC AI를 중심으로 ▲ETRI, KAIST, 고려대, 서울대, 연세대, 에이아이웍스, 포스코DX, 롯데 이노베이트, HL 로보틱스, 양성자, 미디어젠, 문화방송, NHN 등 14개의 국내 최고 산학연 기업·기관과, ▲롯데·포스코의 주요 그룹사 등 40곳의 수요기업을 포함, 총 54개 기관이 프로젝트 전 주기에 걸쳐 하나의 목표로 뭉쳤다. NC AI의 '그랜드 컨소시엄'은 기술 하나를 만들기 위한 협업이 아닌, 대한민국의 디지털 주권을 지키기 위한 AI 생태계 전반이 참여하고 있다.

그 중심에는 NC AI의 대표 모델인 VARCO LLM과 VARCO Vision 2.0이 있다. 최근 공개한 'VARCO Vision 2.0'은 140억과 17억 파라미터로 구성, 작은 규모에도 불구하고 글로벌 동급 SOTA 멀티모달 모델을 뛰어넘는 성능을 보이며 여러 벤치마크 탑티어 성적을 기록했다. 이미 패션, 게임, 스마트시티, 제조 현장 등에서 실증을 마친 상태며, 특히 스마트폰과 PC에서 실시간 처리 가능한 온디바이스 AI 환경까지 지원해 기대를 높여가고 있다.

NC AI는 국가 프로젝트라는 공공성과 동시에 전 세계 오픈 생태계 참여도도 높여가고 있다. VARCO Vision 2.0의 오픈소스 공개는 열흘 만에 1만 건을 넘는 다운로드 수를 기록하며, 대한민국은 물론 글로벌 AI 개발자 생태계에서도 주목되는 '고성능 경량 멀티모달 모델'로

609) NC AI, '대한민국 AI 주권 시대' 여는 NC AI, 독자AI 국가대표 기업으로 선정, 2025.8.4. https://nc-ai.com/ko/newsroom/view?articleId=688f7970864dc906b8d6fdd8

부상했다. 이는 국산 AI 기술의 글로벌 확산 가능성과 모델 신뢰성 확보라는 두 가지 미션을 동시에 증명한 사례라는 평이다. NC AI는 이번 국가대표 선정을 기점으로 차세대 VARCO 시리즈 고도화, 산업 특화 LLM 개발, 멀티모달 AI 기술 고도화, 인재 양성 기관과의 협력 그리고 분야별 현장 맞춤형 AI 실증 확대 등을 단계별로 본격화할 계획이다.

나. 독립 반년 만에 '사고 방지' 기술로 시장 선점[610]

NC AI가 자체 개발한 인공지능(AI) 기술을 앞세워 국내 AI 안전성 분야 시장 주도권 확보에 뛰어들었다. NC AI는 2025년 9월 30일 고객 상담 챗봇 서비스 '엔써(NCER)'에 자체 개발한 AI 신뢰성 기술 '세이프가드(Safeguard)'를 정식 적용했다고 밝혔다. 이는 국내 AI 기업 중 처음으로 종합 AI 안전 시스템을 상용 서비스에 도입한 사례다. 세이프가드 기술은 삼중 보안 체계를 기반으로 운영된다. 레드팀이 새로운 악성 공격을 연구하면 블루팀이 방어 기술을 개발하고 퍼플팀이 두 팀의 결과를 종합해 정책에 반영하는 순환 구조다.

특히 엔씨소프트 퍼블리싱코디네이션 센터와 협력해 게임 산업 특성을 반영한 '챗봇 네거티브 규제 정책'을 수립했다. 이 정책은 게임 서비스 관련 부적절한 내용이나 유료 재화 편법 등 산업 맞춤형 보안을 구현한다. 세이프가드가 적용된 엔써는 ML커먼스와 국제 벤치마크 연구에서 제시하는 일반적 부적절 내용 차단 기준을 충족한다. 차별, 혐오, 욕설, 외설 등이 여기에 포함된다.

NC AI의 안전성 연구는 지난 2022년부터 시작됐다. 회사는 'NC AI 윤리 프레임워크'를 통해 데이터 프라이버시, 비 편향성, 투명성이라는 세 가지 가치 중심의 AI 윤리 경영을 실천하고 있다. 이미 '리니지W'와 '리니지2M' 등 실제 게임 서비스에도 NC AI의 기술이 적용돼 있다. 13개 언어의 광고 패턴을 탐지하고 차단하는 스팸 필터링 시스템이 대표적이다. 회사는 현재 텍스트 기반 안전성 기술을 넘어 향후 이미지, 동영상 등 멀티모달 콘텐츠까지 탐지하는 차세대 시스템을 구축할 목표다.

610) 지디넷코리아, 'AI 폭주' 막아라…NC AI, 독립 반년 만에 '사고 방지' 기술로 시장 선점, 2025.9.30. https://zdnet.co.kr/view/?no=20250929181912

다. '바르코 AI'로 게임 개발 판 뒤집는다…日 시장 '정조준'[611]

NC AI가 게임 제작의 패러다임을 바꿀 생성 AI 기술을 앞세워 일본 시장 공략을 본격 추진하고 있다. NC AI는 2025년 9월 25일 개최된 '도쿄게임쇼 2025'에 참가해 게임 제작 솔루션 '바르코 AI' 시리즈를 발표했다. 국내 AI 기업이 도쿄게임쇼에 공식 진출하는 것은 이번이 처음이다. 이번에 공개될 핵심기술은 텍스트나 이미지로 3D 모델을 자동 생성하는 '바르코 3D'다. 기존 4주 이상 소요되던 3D 제작 공정을 10분 이내로 단축해 개발 생산성을 극대화한다. 가상현실(VR)과 증강현실(AR) 같은 차세대 플랫폼에도 즉시 적용할 수 있다.

'바르코 싱크페이스'는 음성만으로 캐릭터의 입 모양과 표정, 감정까지 실시간으로 구현하는 기술이다. 한국어, 영어, 일본어 등 다국어를 지원하며 영상, 디지털 휴먼 제작에도 쉽게 적용돼 글로벌 콘텐츠 산업 전반의 높은 활용도가 기대된다. '바르코 사운드'는 텍스트나 이미지 입력만으로 배경음악과 효과음을 자동 생성하는 멀티모달 오디오 AI다. 기존 데이터베이스 기반의 제한적 라이브러리에서 벗어나 무한한 사운드 크리에이티브를 지원한다.

NC AI는 이번 도쿄게임쇼 참가를 계기로 일본 현지 인디게임사와 게임 관련 학과, 글로벌 대형 퍼블리셔와 교류를 강화한다. 이를 통해 공동 연구 협업과 상용 프로젝트를 구체적으로 추진할 계획이다. 이번에 공개되는 기술들은 과학기술정보통신부와 정보통신기획평가원(IITP)이 지원하는 생성 AI 선도 인재 양성 사업을 통해 개발됐다. 도쿄게임쇼의 피드백은 국내 AI 인재 양성을 위한 교재로도 활용될 예정이다. 임수진 NC AI 최고사업 책임자는 "게임 AI는 이제 단순한 보조 도구가 아니라 창작자가 새로운 상상력을 현실로 구현할 수 있게 하는 파트너가 돼가고 있다"라며 "바르코 AI 시리즈를 통해 전 세계 개발자들에게 혁신적이고 효율적인 제작 환경을 제공하고 글로벌 게임 산업 혁신의 핵심 주체로 자리매김할 것"이라고 밝혔다.

611) 지디넷코리아, NC AI, 'バルコ AI'로 게임 개발 판 뒤집는다…日 시장 '정조준', 2025.9.23. https://zdnet.co.kr/view/?no=20250923140822

5. 업스테이지

1) 기업 일반현황

<표 5-16> 업스테이지 기업 현황

구분	주요 내용
기업 개요	**〈회사 개요〉** • 2020년 네이버 Clova팀 출신 김성훈 대표가 창업 • 한국어·일본어·아시아권 언어에 특화된 LLM 및 문서 처리 AI 기술 개발에 집중 • 2024년 시리즈 B 투자에서 약 1천억 원 규모 투자를 유치했고, 글로벌 확장을 본격 추진 • 2025년에 자체 모델 'Solar Pro 2'를 출시해 '한국 최초 프런티어 모델(frontier model)'로 평가됐고, 글로벌 AI 경쟁 무대에서 한국 스타트업 인지도를 높임
	〈사업 내용〉 • LLM 솔루션과 문서 처리·정보 추출(Document Parse, Information Extract) 플랫폼을 주력으로 하며, 금융·보험·헬스케어·제조 산업 등에서 활용 • "AI 작업공간(AI Space)"이라는 개념을 통해 팀 단위의 문서·데이터 분석 리포트 자동화, 검색·요약·질의응답 기능을 제공하는 SaaS형 서비스도 제공
	〈자회사〉 • 자회사 형태보다는 Furiosa AI·Puriosa AI 등 국내외 클라우드 인프라 및 하드웨어 기업들과 협업 관계 구축
AI 조직 현황	**〈AI 관련 조직〉** • 언어 모델팀 : 'Solar' 시리즈를 개발하며 한국어·일본어·영어·아시아권 언어에 적합하도록 최적화한 모델을 운영 • 문서 처리·정보 추출팀 : PDF·스캔 문서·계약서 등 다양한 형식의 문서를 AI로 분석·구조화하는 기술을 개발해 산업 현장 적용을 진행 • 인프라·학습팀 : 클라우드 GPU 자원 최적화, 학습 속도 향상, 비용 절감에 집중하며, OCI(Oracle Cloud Infrastructure)를 사용해 데이터 전송 속도, 안정성을 개선 • 산업별 적용팀 : 금융·보험·제조업 등 고객군에 맞춘 AI 솔루션 커스터마이징을 수행하며, "Frontier Model+산업 특화 AI" 전략을 추진
사업 전망	**〈AI 전략〉** • '한국발 AI 모델의 글로벌 진출'을 목표로 하며, Solar Pro 2를 통해 글로벌 LLM 경쟁에 진입했으며 향후 1,000억 개 이상의 파라미터 급 모델 개발을 추진[612] • 기업용 AI 시장에서는 비용 효율성과 다국어 최적화라는 강점을 바탕으로 보험·금융·제조 분야에서 빠른 확장이 가능할 것으로 예상 • 하드웨어 협업 및 칩 스타트업과의 제휴를 통해 AI 학습·추론 인프라 자립성을 높이며 국내 AI 생태계 내 핵심 플레이어로 자리매김할 가능성이 크다는 평가

612) FT, South Korea's Upstage enters global AI race

2) 2025년 기업 동향

가. 620억 시리즈 B 브릿지 투자 유치...아마존・AMD 합류[613)]

업스테이지가 2025년 8월 20일 국내외 주요 투자자로부터 620억 원 규모의 시리즈 B 브릿지 투자를 유치했다고 밝혔다. 이번 투자는 글로벌 생성 AI 경쟁이 심화하는 가운데, 업스테이지의 기술력과 실행력을 국내 AI 생태계의 전략적 자산으로 평가한 결과라는 점에서 의미가 크다. 이번 라운드는 한국산업은행이 리드 투자자로 참여했으며, 글로벌 기업 아마존과 AMD가 새롭게 합류했다. 이 밖에도 인터베스트, KB증권, 프라이머사제, 신한벤처투자, 미래에셋벤처투자, 프리미어파트너스 등 신규 및 기존 투자사들이 함께하며 업스테이지의 성장 잠재력에 힘을 실었다.

업스테이지는 이번 투자 유치를 통해 국내를 넘어 글로벌 AI 시장으로 본격적인 도약을 준비하고 있다. 회사는 자체 개발한 프런티어 급 언어모델 '솔라(Solar)'와 문서 처리 솔루션 등 핵심 제품군의 기술 고도화에 집중하는 한편, 글로벌 시장 공략을 가속할 방침이다. 특히 아마존웹서비스(AWS)와 전략적 협력이 주목된다. 업스테이지는 AWS를 우선 클라우드 제공업체로 선정하고, 아마존 세이지메이커(SageMaker), 트레이니움(Trainium), 인퍼런시아(Inferentia) 등 고성능 인프라를 활용해 자사 모델의 훈련 및 추론 효율을 높인다. 또한 아마존 베드록(Amazon Bedrock) 마켓플레이스를 통한 모델 공급도 확대해 글로벌 고객 접점을 넓힐 계획이다.

업스테이지의 누적 투자금은 이번 라운드를 포함해 약 2,000억 원 규모에 달한다. 회사는 2020년 창업 이후 초기 투자를 시작으로 2021년 316억 원 규모 시리즈 A, 2024년 1,000억 원 규모 시리즈 B를 유치한 데 이어, 이번 브릿지 라운드까지 이어가며 국내 AI 소프트웨어 기업 가운데 최대 수준의 투자금을 확보했다. 이를 통해 업스테이지는 유니콘 반열에 성큼 다가섰다는 평가를 받는다.

김성훈 업스테이지 대표는 "아마존과 AMD 등 글로벌 기업이 참여한 이번 투자 유치는 업스테이지가 세계 시장에서 경쟁할 준비가 되었음을 보여준다"라며 "독자적 AI 모델과 솔루션을 기반으로 글로벌 무대에서 'K-AI'의 경쟁력을 증명하는 모범 사례를 만들어 나가겠

613) 헬로T, 업스테이지, 620억 시리즈 B 브릿지 투자 유치...아마존·AMD 합류, 2025.8.20. https://www.hellot.net/news/article.html?no=104431

다" 라고 말했다. 한편, 업스테이지는 최근 과학기술정보통신부가 주관하는 '독자 AI 파운데이션 모델 개발' 사업의 주관사로 최종 선정되며 기술력과 실행력을 겸비한 국가대표 AI 기업으로서의 입지를 강화했다.

나. '국가대표 AI' 만드는 업스테이지, '조 단위 몸값' 상장 추진[614]

한국형 독자 파운데이션 모델 개발 사업에 참여하는 업스테이지가 상장을 추진하는 것으로 알려졌다. 경제 미디어 뉴스원은 2025년 11월 7일 업스테이지가 주요 증권사들과 상장 주관사 선정을 논의 중이라고 보도했다. 2020년 설립된 업스테이지는 거대언어모델(LLM) '솔라' 시리즈 등 AI 서비스를 앞세워 미국·일본 시장을 공략하고 있다. 다양한 형식의 이미지·텍스트를 기계가 읽을 수 있도록 바꿔주는 광학문자인식(OCR) 기술도 보유했다.

개발 능력을 인정받은 회사는 올해 8월 정부의 '독자 AI 개발 사업'의 컨소시엄 주관사로 선정되기도 했다. 사업은 글로벌 수준의 파운데이션 모델을 국내 기술로 개발하는 게 목표다. 총 5개 컨소시엄이 뽑혔으며, 나머지 네 곳 주관사로는 네이버클라우드·SK텔레콤(017670)·NC AI·LG AI연구원 등이 있다. 스타트업은 업스테이지가 유일하다. 회사는 올해 8월 자금 조달을 통해 약 7,900억 원의 기업 가치를 투자자들로부터 인정받았다. 상장이 성공적이라면 회사가 조 단위 몸값으로 증시에 입성할 수 있을 거란 전망까지도 나온다. 업스테이지 관계자는 "상장 목표 시기나 내용은 아직 확정된 게 없다"라고 말했다.

다. 업스테이지, 일본어 특화 언어모델 '신 프로' 출시[615]

업스테이지가 2025년 10월 23일 일본 AI 전문기업 카라쿠리와 공동 개발한 일본어 대형언어모델(LLM) 신 프로를 공개했다. 신 프로는 업스테이지 자체 LLM '솔라'를 기반으로 일본어와 일본의 문화적 맥락을 반영한 310억 파라미터 모델이다. 온프레미스(구축형) 방식으로 보안성과 비용 효율성을 모두 갖춰 현지 기업이 자체 인프라에서 안전하게 AI를 활용할 수 있도록 설계됐다. 신 프로는 일본어 LLM 성능 지표인 '웨이츠앤바이어시스 네즈미 리더보드'에서 일본 내 개발한 기존 모델과 비교해 1위를 기록했다.

---

614) 뉴스1, '국가대표 AI' 만드는 업스테이지, '조 단위 몸값' 상장 추진, 2025.11.7. https://www.news1.kr/it-science/general-it/5969173

615) 연합뉴스, 업스테이지, 일본어 특화 언어모델 '신 프로' 출시, 2025.10.23. https://www.yna.co.kr/view/AKR20251023166300017

신 프로는 동급 310억 파라미터 이하 모델 중 가장 높은 점수를 받았고, 주요 성능의 경우 오픈AI 추론 모델 'GPT-OSS 120B'를 뛰어넘는 결과를 냈다. 현재 일본어 LLM 리더보드에서 일본 내 학습·개발된 모델 중 상위권에 오른 건 신 프로가 유일하고, 나머지는 오픈AI, 구글, 앤트로픽, 딥시크 등 미국과 중국 모델이다. 마츠시타 히로유키 업스테이지 일본 법인장은 "신 프로는 업스테이지 기술력이 일본에서 학습한 AI 모델로도 경쟁력을 확보할 수 있음을 입증한 유의미한 성과다"라고 말했다.

6. 기타 주요 기업

1) 기업 일반현황

<표 5-17> 기업 일반현황

구분	주요 내용
삼성전자	• 1969년 설립된 글로벌 IT 선도 기업으로, 2025년 현재 'AI 가속기-고대역폭 메모리(HBM)-파운드리'를 잇는 'AI 반도체 턴키(Turn-key)' 전략을 통해 AI 인프라 시장의 핵심 공급자로 자리매김 • 2025년 하반기 차세대 AI 전용 메모리인 HBM4(6세대) 양산을 본격화하며 엔비디아 등 글로벌 빅테크에 공급을 확대하고, AI 하드웨어 시장에서의 메모리 주도권 탈환에 집중 • 자체 개발한 생성형 AI인 '삼성 가우스 2(Samsung Gauss 2)'를 스마트폰(Galaxy S25), 가전, 반도체 공정 등 전 사업 영역에 통합하여 온디바이스 AI(On-Device AI) 생태계의 초격차를 구현 • 파운드리 부문에서는 2025년 2나노(nm) 공정 기반의 AI 칩 양산을 시작하며 차세대 AI 가속기 및 커스텀 AI 칩(ASIC) 설계 역량을 강화하고, 글로벌 팹리스 고객사 유치를 가속화 • 2025년 출시된 '갤럭시 AI(Galaxy AI)'의 개인화 비서 기능을 고도화하고, 가전 기기 간 AI 연결 플랫폼인 '스마트싱스(SmartThings)'를 통해 '모두를 위한 AI(AI for All)' 비전을 현실화 중
KT	• 1981년 설립된 국내 대표 통신 기업에서 2025년 'AICT(AI+ICT) 컴퍼니'로 완전 체질 개선에 성공하며, 통신 인프라에 AI 기술을 결합한 통합 서비스 기업으로 도약 • 2024년 말 마이크로소프트(MS)와 체결한 수조 원 규모의 전략적 파트너십을 바탕으로, 2025년 한국형 특화 모델인 'GPT-4o 기반 한국어 맞춤형 서비스'와 공공·금융 전용 소버린 클라우드/AI 솔루션을 본격 상용화 • 자체 AI 거대언어모델인 믿음(Mi:dm) 시리즈를 경량화(SLM) 및 도메인 특화 모델로 세분화하여, 기업용 서비스(AICC), 스마트 모빌리티, 로봇 등 B2B 시장의 실질적인 수익 구조를 강화 • 2025년에는 AI 에이전트 서비스인 'AI 개인비서'를 모바일과 미디어(Genie TV)에 전면 도입하고, 한국어 문맥에 최적화된 높은 이해도를 바탕으로 국내 온디바이스 및 클라우드 AI 시장의 사용자 점유율 확대 • AI 인프라 자립을 위해 리벨리온(NPU), 모레(SW 스택) 등 국내 AI 반도체 스타트업과 협업하는 'AI 풀스택(Full-Stack)' 전략을 고수하며, 엔비디아 의존도를 낮춘 비용 효율적인 AI 데이터센터(AIDC)를 전국 단위로 확산 중

2) 2025년 기업 동향

가. 삼성전자-KAIST, 대형 AI 모델 학습 최적화 시뮬레이션 개발[616]

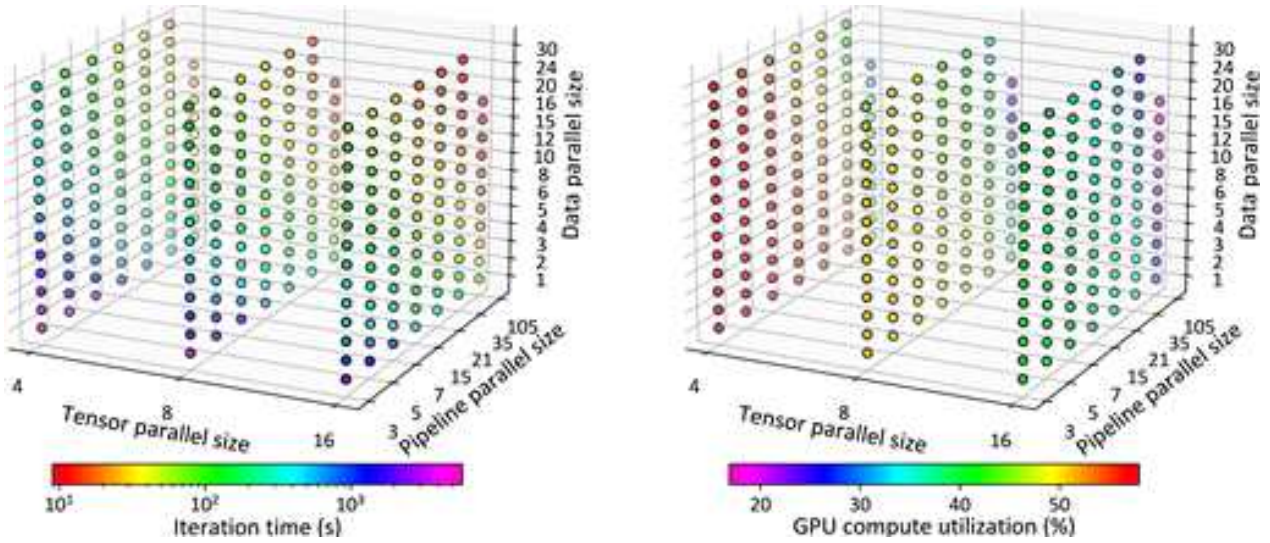
삼성전자와 KAIST(한국과학기술원)가 2025년 3월 13일 챗GPT, 딥시크 등 초거대 인공지능(AI) 모델의 학습 비용을 획기적으로 줄일 수 있는 시뮬레이션 기술을 개발했다. 이 기술은 대규모 분산 시스템에서 그래픽처리장치(GPU) 활용 효율을 높이고 최적의 학습 전략을 찾는 데 초점을 맞췄다. 삼성전자 삼성종합기술원과 유민수 KAIST 전기및전자공학부 교수 연구팀은 공동으로 대형 언어모델 학습 시간을 예측하고 최적화하는 시뮬레이션 프레임워크 'vTrain'을 공개했다. 프레임워크는 방대한 병렬화 전략을 빠르게 탐색해 비용과 시간을 절감하는 방안을 제시한다.

최근 챗GPT-4 같은 대형 언어 모델은 수만 개의 고성능 GPU를 활용해 학습돼 챗GPT-4의 경우 학습 비용만 약 1,400억 원에 달하는 것으로 알려졌다. 하지만 기존에는 경험에 의존한 소수의 학습 전략만 사용돼 GPU 활용의 비효율성과 불필요한 비용 증가가 문제로 지적됐다. 특히 대규모 시스템을 위한 시뮬레이션 기술 부족으로 기업들이 최적화에 어려움을 겪고 있었다.

삼성전자와 KAIST 연구팀이 개발한 vTrain은 이러한 문제를 해결할 수 있다. 실제 다중 GPU 환경에서 측정한 학습 시간과 vTrain의 예측값을 비교한 결과 단일 노드에서는 평균 절대 오차(MAPE) 8.37%, 다중 노드에서는 14.73%의 높은 정확도를 보였다. 이를 통해 다양한 분산 병렬화 전략을 신속히 평가하고 최적의 구성을 도출할 수 있다. 연구팀은 vTrain 프레임워크와 1,500개 이상의 실제 학습 시간 데이터를 오픈소스로 깃허브에 공개했다. 이를 통해 AI 연구자와 기업들이 자유롭게 활용할 수 있도록 지원한다.

616) THE AI, KAIST-삼성전자, 대형 AI 모델 학습 최적화 시뮬레이션 개발, 2025.3.13. https://www.newstheai.com/news/articleView.html?idxno=7404

[그림 5-34] 다양한 병렬화 기법에 따른 MT-NLG 학습 시간 및 GPU 사용률 변화



출처 : KAIST

나. 삼성SDS “특화 멀티에이전트 전략으로 AX 견인”[617] 911

삼성SDS가 2025년 9월 11일 개최된 ‘리얼 서밋 2025’에서 특화 AI 에이전트 모델을 종합 제공해 AI 풀스택 서비스로 기업의 AI 전환을 돕겠다고 밝혔다. 이 자리에서 삼성SDS 이준희 대표는 “삼성 SDS는 강력한 AI 클라우드 인프라를 갖추고 있고 다양한 언어모델과 기업 시스템을 연결하는 AI 플랫폼인 패브릭스(FabriX), AI 협업 솔루션 브리티 코파일럿(Brity Copilot)과 업무 프로세스 자동화를 돕는 브리티 오토메이션(Brity Automation)까지 기업의 AX를 위한 AI 풀 스택을 제공하고 있다”라고 강조했다. 삼성SDS는 자사 핵심 경쟁력으로 ‘버티컬(업종별) 강점’을 제시하고 글로벌 핵심 솔루션 사업자들과 강력한 전략적 제휴를 바탕으로 풍부한 도메인 지식, 프로세스 지식, 기업 데이터에 대한 이해에서 큰 차별성이 보유했다고 설명했다.

기술적 차별성으로는 코파일럿의 언어 처리 능력을 내세웠다. 여타 다른 화상 회의 시스템은 기본적으로 회의 하나를 열 때 언어를 한 가지 공용 언어로 세팅해야 하지만, 브리티 코파일럿은 별도 세팅 없이 60개 언어에 대한 자막 서비스를 지원한다고 설명했다. 더불어 번역 서비스도 언어를 확대한다고 밝혔다. 동시통역 서비스는 현재 7개 언어에서 11월에 힌디

617) THE AI, 삼성SDS “특화 멀티에이전트 전략으로 AX 견인”, 2025.9.11.
https://www.newstheai.com/news/articleView.html?idxno=9000

어까지 포함해 17개 언어로 확대된다.

삼성SDS는 해외 진출에 대해서 두 가지 핵심 축으로 제시했다. 첫째는 직접 솔루션과 서비스 제품을 가지고 사업하는 것이고 둘째는 고객사들의 글로벌 사업을 지원하는 것이다. 관계사들이 세계 각국에 진출해 있어 이를 연계해 해외 사업을 추진하고 있다. 삼성SDS는 협력업체와 협력해서 글로벌 솔루션을 미국에서 출시했고 시장 반응이 긍정적이라고 밝혔다. 이와 함께 대외사업 비중도 늘리고 있다. 삼성SDS는 2024년 기준 대외사업 비중이 66%에 달한다. 이준희 대표는 대외사업을 더욱 적극적으로 확대할 계획이라고 강조하고 정부가 진행하는 AI 프로젝트에도 적극 참여하고 있고 금융, 제조 등 여러 분야에서 많은 기회가 있을 것으로 예상했다.

다. 삼성, IFA에서 AI 전략 공개[618] 0905

삼성전자가 2025년 9월 5일 독일 베를린에서 열린 유럽 최대 가전 전시회 'IFA 2025'서 '갤럭시 이벤트'를 열고 새로운 갤럭시탭 S11과 갤럭시 S25 FE를 공개했다. 이날 갤럭시 이벤트에서 주목할 점은 단연 AI 기능이었다. 김정현 삼성전자 모바일 경험사업부 부사장은 "갤럭시 AI는 이미 전 세계적으로 2억 대 이상 기기에 탑재됐다" 라며 "올해 안으로 두 배로 늘릴 것" 이라고 밝혔다. 이날 삼성이 전면에 내세운 기능은 멀티모달 AI다. 텍스트나 이미지, 음성, 영상 등 다양한 유형의 정보를 동시에 이해하고 처리함으로써, 이용자의 스마트 디바이스 경험 향상에 초점을 맞췄다.

또한 멀티모달로 향상된 '제미나이 라이브'는 실시간으로 화면이나 카메라 공유 기능을 통해 현재 상황에 최적화된 도움을 준다. 사용자가 보고 있는 화면의 복잡한 차트의 핵심 내용을 알고 싶을 때는 제미나이 라이브에 화면을 공유해 중요한 내용을 요약하게끔 명령할 수 있는 방식이다. 신형 갤럭시탭 S11은 대화면에 맞는 기능들이 추가됐다. 이용자는 △간단한 스케치를 고품질 이미지로 변환하는 기능인 '그리기 어시스트(Drawing Assist)' △문장의 문체나 어투를 목적에 맞게 자연스럽게 수정하는 기능인 '글쓰기 어시스트(Writing Assist)' △현재 보고 있는 화면뿐 아니라, 상하 스크롤과 페이지를 이동할 때도 실시간으로 번역된 화면을 볼 수 있는 서클 투 서치 기반의 '실시간 번역 기능'을 활용할 수 있다.

618) THE AI, 삼성, IFA서 공개한 AI 전략은?, 2025.9.5.
https://www.newstheai.com/news/articleView.html?idxno=8924

AI의 도입에 하드웨어의 발전도 눈에 띈다. 이번 신형 아이폰17 시리즈는 AI 도입과 함께 램 추가, 냉각시스템 개선이 이뤄질 전망이다. 고급형 모델에는 고급형 모델에는 12GB 램과 더 강력한 뉴럴 엔진이 제공될 예정이고 일반 모델들은 8GB의 램과 일부 AI 기능이 탑재될 전망이다. 개선된 베이퍼 챔버 쿨링 시스템으로 고성능 작업이나 AI 작업에 따른 발열을 보다 낮출 것으로 보인다. 아울러 일부 모델에는 애플이 자체 개발한 통신 모뎀이 탑재돼 클라우드 연동 AI 기능의 반응속도와 효율 증대가 이뤄질 전망이다.

라. KT-메리츠화재, AI 기반 금융 연구 추진[619]

KT와 메리츠화재가 2025년 11월 ‘AI 기반 미래 금융 혁신’을 위한 전략적 공동 연구 협약을 체결했다고 밝혔다. 양사는 △보험 특화 AI 모델 개발 △AI 기반 보상 자동 심사를 위한 청구 및 의료 서류 인식 기술 △AI 기반 데이터 분석 및 생성 기술 등 공동 연구를 추진한다. KT는 GPU 인프라와 AI 모델 등 연구에 필요한 환경과 자원을 제공하며, 한국어 인식 성능이 뛰어난 ‘믿:음 K 2.0’ 등의 라인업을 적극 활용한다. 보안이 강화된 산업 특화 AI 모델 개발 역량을 확장할 계획이다.

또 KT가 자체 개발한 OCR(광학문자인식) 솔루션 ‘DocuSee’를 활용해 의료 및 보험 서류 인식 성능을 고도화한다. 이를 통해 복잡한 보험 청구 서류를 AI가 자동 인식·분석함으로써 보험 운영 업무 효율화를 지원할 계획이다. 메리츠화재는 금융, 보험과 관련된 정형·비정형 데이터를 제공한다. 아울러 고객 데이터 보호와 보안 체계를 고려하면서 축적해 온 업무 노하우와 보험 업무 로직을 접목해 AI 기반 보상 시스템의 전문성과 고객 서비스 품질을 높일 예정이다.

KT와 메리츠화재는 이번 협력을 시작으로 내년 상반기까지 보험 특화 AI 모델과 자동 보상 청구를 위한 의료 서류 인식 기술 고도화를 완료하고, 이를 기반으로 사업적 시너지 창출을 위한 상호 협력을 지속할 계획이다.

619) THE AI, KT-메리츠화재, AI 기반 금융 연구 추진, 2025.11.6.
https://www.newstheai.com/news/articleView.html?idxno=9633

마. kt cloud, 액체 냉각 적용된 가산 AI 데이터센터 가동 방침[620]

kt cloud가 수도권 서부권에 차세대 냉각기술을 적용한 대규모 AI 데이터센터를 선보이며 국내 AI 인프라 경쟁에 뛰어들었다. kt cloud는 2025년 11월 6일 서울 가산디지털단지에 총 수전 용량 40MW, IT 용량 26MW 규모의 가산 AI 데이터센터를 공식 개소했다고 밝혔다. 연면적 약 1만 1,046평, 지하 5층 지상 10층 규모로 구축된 이 시설의 핵심은 국내 상업용 데이터센터로는 처음 적용된 액체 냉각기술이다. 생성 AI시대를 맞아 폭증하는 연산 수요에 대응하기 위한 인프라 확보 경쟁이 치열해지는 가운데, 냉각 효율을 획기적으로 끌어올린 시설을 통해 시장 주도권을 선점하겠다는 전략으로 풀이된다.

kt cloud가 도입한 방식은 GPU 칩에 냉각판을 직접 부착해 냉수를 접촉하게 하는 D2C(Direct to Chip) 냉각기술이다. 최근 GPU의 전력 소모량이 급격히 높아지며 기존 공랭식 냉각으로는 고성능 AI 서버의 열 관리가 어려워진 상황이다. 이에 액체 냉각은 높은 열 전도성으로 에너지 효율을 개선하는 방안으로 주목받았다. kt cloud는 엔비디아 B200 등 GPU 기반 AI 서버의 고열 환경을 테스트하기 위해 리퀴드 쿨링 부하기를 자체 개발했으며, 성능 검증을 완료한 상태다. 다음 달 개소 예정인 AI 이노베이션 센터에서는 이러한 에너지 절감 기술을 고도화해 상용화 수준으로 발전시킬 계획이다.

가산 AI 데이터센터는 AI 연산에 특화된 고집적 존을 중심으로 설계됐다. 전력과 통신 인입을 이중화해 시스템 안정성을 강화했으며, 수도권 내 8개 데이터센터와 100G급 One DC 네트워크로 연결돼 초고속 센터 간 연동이 가능하다. 이번 센터에는 kt cloud가 특허 등록을 마친 ‘패스 파인더(Path Finder)’ 솔루션도 최초로 적용됐다. 이 기술은 전력 계통을 실시간으로 모니터링하고 자동으로 차단 및 복구를 수행해 전원 안정성을 대폭 높인 것이 특징이다. 또한, AI 분석 기반으로 장애를 사전 예측하고 자동 대응하는 자동화 운영 플랫폼 DIMS를 고도화해 운영 효율을 극대화했다.

kt cloud는 가산 센터를 통해 ‘Colo.AI’ 서비스 제공을 시작한다. 고객 맞춤형 GPU 서버, 전용 네트워크, 운영 플랫폼, 유지보수를 포괄하는 통합형 턴키 서비스로, 복잡한 인프라 구축 과정 없이 AI 환경을 신속하게 구현한다. 현재 국내 AI 기업과 중소 금융기업이 이 서비스

620) THE AI, kt cloud, 액체냉각 적용된 가산 AI 데이터센터 가동한다, 2025.11.6.
https://www.newstheai.com/news/articleView.html?idxno=9622

를 활용해 대규모 AI 모델 학습 및 서비스를 안정적으로 운영 중이다. 업계에서는 kt cloud의 이번 행보가 급변하는 국내 데이터센터 시장의 분기점이 될 것으로 전망한다. 국내 데이터센터 시장은 2028년까지 약 10조 원 규모로 성장할 것으로 예상되며, 이는 2024년 대비 연평균 13.13%에 달하는 고성장세다.

바. KT, GPT-4o 기반 'SOTA K' 출시… "한국어 강점"[621]

KT가 2025년 9월 29일 마이크로소프트와 협업을 통해 개발한 GPT-4o 기반 한국적 AI 모델 '소타(SOTA) K'를 출시했다. GPT-4o에 한국어와 한국 사회·문화적 맥락을 정교하게 접목한 협업형 모델이다. KT는 우선 자사 B2C 사업에 SOTA K를 적용해 자체 및 협업 모델 라인업을 검증한 뒤 파트너사에 서비스를 확대 제공할 방침이다. 이를 통해 국내 AI 생태계의 안정적 성장과 건전한 발전을 도모한다는 전략이다. 기존 LLM은 영어권 중심 데이터로 학습돼 한국어와 한국의 독특한 사회·문화적 맥락 반영에 한계가 있었다. KT는 이를 해결하기 위해 고품질 한국 특화 데이터를 대규모로 확보하고, 마이크로소프트와의 협력을 통해 GPT-4o를 한국 사회에 최적화된 모델로 발전시켰다.

SOTA K는 한국적 AI의 4대 핵심 철학인 △데이터 주권 보호 △한국 문화 이해 △모델 선택권 보장(독자 모델과 협력·오픈모델) △안전하고 책임감 있는 AI를 모두 구현한 모델이다. 한국어 경어법·방언은 물론 법률·금융·역사 등 국내 특정 산업군 내에서 사용하는 전문 용어까지 자연스럽게 이해하고 생성할 수 있어 한국 비즈니스 환경에서 실용성을 높였다. KT는 자체적으로 구축한 데이터셋을 구축해 정량 평가와 정성 평가를 진행했다. 이러한 평가 결과 SOTA K는 한국어 이해, 생성, 추론, 사회·문화, 한국 전문 지식 등 주요 지표에서 GPT-4o 대비 우위를 보였다. 한국사, 한국어, 한국 법령 등 고난도 한국적 지식을 요구하는 대한민국 공무원 시험과 귀화 시험에서 GPT-4o 모델을 능가하는 성과를 거두어, 한국 사회와 문화에 특화된 AI로서의 탁월한 실용성을 입증했다고 KT 측은 설명했다.

KT는 실제 고객 사례에서도 긍정적인 결과를 얻었다고 소개했다. 메리츠화재에서는 보험 업계에 특화된 약관의 자동 요약 리포트 생성과 상담원 스크립트 생성에서 SOTA K가 우수한 결과물을 보여줬다. EBS와의 협력에서는 초중고 난이도별 교과 문항을 생성해 맞춤형 학습을

621) THE AI, KT, GPT-4o 기반 'SOTA K' 출시… "한국어 강해", 2025.9.29.
https://www.newstheai.com/news/articleView.html?idxno=9223

제공하는 측면에서 SOTA K의 높은 활용 가능성이 확인됐다.

SOTA K는 한국적 AI 평가의 책임감 있는 AI 항목에서 높은 평가를 받아 AI 응답 및 법률과 권리 준수, 사회 및 경제 영향도의 안전성, AI 모델의 강건성을 입증했다. 또 악의적 사용자에 의한 AI 모델 탈옥 공격에 대해서도 강력한 방어 능력을 보유한 모델임을 확인했다고 KT는 밝혔다. KT는 SOTA K를 물리적으로 대한민국 내 존재하는 클라우드 리전에서 운영하며 자체 개발한 벡터 모델 기반 한국적 검색증강생성(RAG) 기술과 결합해 기업별 맞춤형 지식 기반을 구축할 수 있는 서비스 환경도 제공한다.

7. AI 스타트업

1) 기업 일반현황

<표 5-18> AI 스타트업 현황

구분	주요 내용
퓨리오사AI	• 퓨리오사AI(Furiosa AI)는 2017년 설립된 한국의 AI 반도체(추론 가속기) 스타트업으로, 초고성능 AI 칩과 소프트웨어 스택을 자체 개발하는 것을 목표로 함 • 범용 GPU 의존도를 낮추고 효율적인 AI 추론용 가속기(NPU) 개발에 집중하고 있으며, 2023년 이후 'Warboy' 후속 아키텍처를 기반으로 한 차세대 AI 칩을 공개하며 국내 최대 AI 반도체 선도기업으로 성장 • 대표 제품으로는 Warboy(1세대 AI 가속기)와 이를 지원하는 SW 스택인 AInstein SDK가 있으며, 데이터센터·클라우드·엔터프라이즈 작업에 최적화되어 있음 • 자체 AI 칩과 소프트웨어 최적화를 통해 전력 대비 성능(Power Efficiency), 비용 효율성, 대규모 추론 처리 성능 등이 강점이며, 글로벌 AI 가속기 시장에서 엔비디아, AMD 대안 기술로 주목 • 2024년에 업스테이지·클라우드 기업·엔터프라이즈 AI 기업들과 협력해 LLM 추론 최적화, 모델 압축·가속, 산업용 AI 적용 등 칩+소프트웨어 통합형 솔루션을 강화
뤼튼테크놀로지스	• 뤼튼테크놀로지스는 2021년 설립된 한국의 생성 AI 스타트업으로, 문서·콘텐츠 제작 자동화에 특화된 AI 콘텐츠 생성 플랫폼 '뤼튼(Wrtn)'을 개발 • GPT·Claude·Gemini 등 다중 모델을 조합해 사용하는 '모델 집약형 플랫폼' 구조에서 출발했으며, 2024~2025년부터 자체 언어모델(LLM) WRTN 모델 라인업 개발에 본격 착수해 기술 내재화를 강화 • 주력 제품 '뤼튼'은 글쓰기·요약·보고서·기획안·마케팅 문구 생성 등 비즈니스 문서 자동화에 최적화됐으며, 엔터프라이즈 버전에서는 기업 맞춤형 워크플로우·RAG 기반 문서 검색·지식베이스 구축 기능을 제공 • 2023년부터 교육·공공·기업용 시장으로 확장을 추진해 한국을 대표하는 생성 AI 서비스로 자리매김 • 자체 모델(WRTN), 사용자 데이터 기반 최적화, 워크플로우 자동화 기능을 결합해 '실무형 AI 도구'라는 포지션을 강화하며, 생산성 SaaS 기업들과 경쟁 구도를 형성

2) 2025년 기업 동향

(1) 퓨리오사

가. 메타 인수 제안 거절… 독자적 AI 반도체 시장 공략[622)]

국내 AI 반도체 설계 기업 퓨리오사AI가 글로벌 빅테크 기업 메타로부터 인수 제안을 거절

622) THE AI, 퓨리오사AI, 메타 인수제안 거절… 독자적 AI 반도체 시장 공략, 2025.3.24.
https://www.newstheai.com/news/articleView.html?idxno=7487

하고 독자적인 성장 전략을 선택했다. 퓨리오사AI는 2025년 3월 24일 메타와의 인수 협상을 종료하고 자체 AI 칩 개발 및 양산에 집중하겠다고 밝혔다. 퓨리오사AI의 기업 가치는 8,000억 원, 메타의 인수 제안가는 8억 달러로 알려졌다. 메타는 자사의 AI 학습과 추론에 최적화된 자체 칩 설계를 위해 AI 팰리스 스타트업 인수를 추진해 왔다. 퓨리오사AI가 유력 인수 대상으로 협상에 나선 것으로 알려졌다.

퓨리오사AI는 엔비디아의 H100 대비 절반 수준 비용으로 구동이 가능한 비용 효율적인 AI 반도체 '레니게이드'를 개발해 주목을 받아왔다. 레니게이드는 AI 추론 작업에 특화된 반도체다. GPU 대비 전력 소비가 25% 수준으로 높은 성능과 낮은 전력 소모를 강점으로 한다. 이는 SK하이닉스 고대역폭 메모리 HBM3를 2개 탑재되고 TSMC 5나노 공정 기반으로 생산한다. 퓨리오사AI는 해외 매각 대신 독립적인 경영을 통해 AI 반도체 시장에서 입지를 강화하겠다는 의지를 보였다. 최근 레니게이드의 성능 테스트에서 긍정적인 결과를 얻은 점도 이번 결정에 힘을 실었다. 퓨리오사AI는 LG AI연구원, 사우디 아람코 등 주요 기업들과 협력하며 칩 상용화 가능성을 확인하고 있다.

하지만 퓨리오사AI는 리벨리온과 딥엑스 등 대규모 투자를 유치한 것과 달리 그동안 AI 칩 개발과 양산에 막대한 자금이 투입되면서 자금 확보에 어려움을 겪어왔다. 이에 퓨리오사AI는 산업은행으로부터 300억 원 규모 투자의향서(LOI)를 받는 등 한 달 내 700억 원 규모 자금 조달을 추진 중이다. 글로벌 1위 파운드리(반도체 위탁생산) 기업인 대만 TSMC도 투자를 검토 중이다.

나. 3세대 인공일반지능 반도체 개발 추진[623)]

퓨리오사AI가 3세대 AI 칩 개발에 속도를 올리고 있다. 한국경제는 2025년 11월 3일 퓨리오사AI가 2년 내 AI가 스스로 판단하고 학습할 수 있는 인공일반지능(AGI)용 칩을 개발하겠다는 전략을 수립했다고 보도했다. 퓨리오사AI는 3세대 AI 칩의 코드명을 '스토크(Stork)'로 정하고 연구개발(R&D)에 속도를 올리고 있다. 회사 내부에서 2027년에 개발을 완료하겠다는 계획을 세웠다. 회사는 3세대 칩을 인공일반지능(AGI) 규모 AI 모델을 연산할 수 있는 반도체로 만들겠다는 전략을 수립했다.

623) 한국경제, 코드명 '스토크'…퓨리오사AI, 3세대 인공일반지능 반도체 개발 나선다, 2025.11.10. https://www.hankyung.com/amp/2025111035141

AGI는 AI가 이용자가 요청하는 연산은 물론 스스로 판단하고 지속적으로 학습할 수 있는 기술을 말한다. 퓨리오사AI의 칩은 지속적으로 진화하고 있다. 2021년 자사의 첫 번째 양산 칩인 ‘비전 NPU’를 만들었다. 이 칩은 이미지를 인식하고 어떤 사물인지 판단할 수 있는 정도의 AI를 구현할 수 있는 수준이다. 2024년에는 2세대 칩 ‘레니게이드(RNGD, 사진)’를 발표했다. 이 칩으로는 거대언어모델(LLM) 구현은 물론 사진과 영상까지 인식해서 특정한 상황을 추론할 수 있는 멀티모달 모델까지 구현할 수 있다.

퓨리오사AI가 2027년 AGI용 칩을 구현하려면 현존하는 칩보다 더욱 고도화한 설계 실력이 필요할 것으로 보인다. 2세대 칩은 TSMC 5나노 공정으로 만들었고 4세대 고대역폭메모리(HBM3)를 결합했다. 향후 동일한 면적 안에 더 많은 연산 소자를 장착하기 위해 5나노보다 더욱 앞선 공정을 활용할 가능성이 크다. 또한 연산에 필요한 숫자 값(파라미터)이 늘어날 것을 고려해 용량과 속도가 개선된 5세대 HBM(HBM3E) 이상을 장착할 가능성까지 거론된다. 퓨리오사AI는 NPU 기반의 AI 칩을 개발하는 토종 반도체 설계 회사다. 세계 AI 반도체 1위 회사인 엔비디아의 아성을 무너뜨리겠다는 목표로 칩 개발에 매진하고 있다. 퓨리오사AI는 AI 업계와의 협업을 통해 잠재력을 인정받고 있다.

LG AI연구원은 퓨리오사AI 2세대 칩에 대한 테스트를 마치고 회사 데이터센터에 적용하기 위해 준비하고 있다. 또한 이들은 지난달 서울에서 개최됐던 오픈AI 코리아 개소식에서 2세대 칩을 활용해 ‘gpt-oss 120B’ LLM 모델 기반의 챗봇을 시연하며 협력 가능성을 시사했다.

다. 인하대와 산학협력…AI 반도체 공동 연구 추진[624]

퓨리오사AI와 인하대학교는 2025년 11월 5일 산학협력을 위한 업무협약(MOU)을 체결했다고 밝혔다. 양측은 AI 반도체의 설계, 패키징, 테스트 등 전 과정에서 공동 연구와 기술 교류를 추진하고, 이를 기반으로 실무 중심의 교육과정 개발과 학생 인턴십 운영을 통해 반도체 전문 인재를 양성할 계획이다.

이번 협약은 교육·연구·산업을 유기적으로 연계하는 지속 가능한 산학협력 모델을 구축하려는 인하대 반도체 특성화 대학사업단의 전략에 따라 이뤄졌다. 퓨리오사AI는 대규모 인공

624) 중앙이코노미뉴스, 퓨리오사AI, 인하대와 산학협력…AI 반도체 공동 연구 추진, 2025.11.10. https://www.joongangenews.com/news/articleView.html?idxno=463845

지능 모델의 연산 효율을 높이는 AI 추론용 고성능·저전력 반도체 가속기(NPU)를 개발하고 있는 기업으로, 이번 협력을 통해 기술 경쟁력을 더욱 강화할 수 있을 것으로 기대된다. 조명우 인하대 총장은 “AI 반도체 산업의 핵심 공정에서 기업과 대학이 긴밀히 협력하는 모범 사례를 만들겠다” 라고 말했다.

라. 1,700억 투자 유치…이재명 정부 '1호 유니콘' 등극[625)]

AI 반도체 설계 업체 퓨리오사AI가 2025년 7월 31일 1억 2,500만 달러 규모의 시리즈 C 투자를 유치했다고 밝혔다. 퓨리오사AI는 이날 자사 홈페이지를 통해 이같이 밝히고 누적 투자 금액이 2억 4,600만 달러라고 집계했다. 회사는 이번에 국책금융기관, 사모펀드 등 약 40개 기관의 투자를 유치했다. 특히 기술 리스크가 큰 반도체 스타트업에 투자하기를 꺼리는 여러 사모펀드가 400억 원 이상 투자를 집행한 점도 업계의 주목을 받고 있다.

이들이 퓨리오사AI의 성장 잠재력에 베팅한 이유는 AI 반도체 시장의 가파른 상승세를 눈여겨봐서다. AI 칩은 2022년 후반 오픈AI의 챗GPT가 나온 이후 금융투자업계의 주목을 받기 시작했다. AI 반도체 시장은 그래픽처리장치(GPU)를 설계하는 미국 엔비디아가 독보적 1위 지위를 수성하고 있다. 최근 AI 서버 업체 사이에서 공급망 다변화를 모색하는, 이른바 ‘탈(脫) 엔비디아’ 움직임이 가시화해 인공 신경망처리장치(NPU) 같은 대체재를 만드는 퓨리오사AI 몸값도 나날이 치솟고 있다. 최근 또 다른 토종 NPU 설계 회사인 리벨리온이 삼성 계열사의 투자를 유치한 데도 이런 배경이 작용했다.

퓨리오사AI는 서서히 입지를 확장해 가고 있다. 7월 2세대 AI 칩 ‘레니게이드’를 LG AI연구원에 공급하며 대규모 서버 시장 진출을 위한 물꼬를 텄다. 다만 아직 회사 매출을 획기적으로 끌어올릴 만한 초대형 수주가 나오지 않고 있고, 엔비디아 제품 성능과 이들이 보유한 소프트웨어 시스템을 따라잡으려면 많은 시간이 필요할 것이라는 평가를 불식해야 하는 과제도 안고 있다. 퓨리오사AI는 이번 투자금을 레니게이드 본격 양산과 3세대 제품 초기 개발에 사용할 예정이다.

---

625) 한국경제, 퓨리오사AI, 1700억 투자 유치…이재명 정부 '1호 유니콘' 등극, 2025.7.31. https://www.hankyung.com/article/2025073146791

⑵ 뤼튼테크놀로지스

가. 무료 교육 바우처로 국민 AI 교육 지원[626]

뤼튼테크놀로지스(이하 뤼튼)가 온라인 AI 교육 무료 지원과 독자 커리큘럼 개발로 국민 AI 역량 강화에 나섰다. 뤼튼테크놀로지스는 2025년 11월 6일 AI 리터러시 교육 기업 에이블런, 온라인 교육 플랫폼 인프런과 손잡고 AI 교육 콘텐츠 무료 수강 바우처 배포에 착수했다고 밝혔다. 이는 지난 8월 LG AI연구원, AI 반도체 기업 퓨리오사AI와 체결한 모두의 AI 비전 실현을 위한 협약의 구체적인 실행 단계다. 뤼튼이 제공하는 교육과정은 AI 초보자를 위한 기초부터 대상별 특화 강좌까지 세밀하게 구성됐다. 생성 AI의 이해, AI 도구 활용을 위한 프롬프팅 기초, AI 윤리와 올바른 사용법 등 입문 단계부터 AI로 보이스피싱 예방하기, AI를 활용한 가게 마케팅, 자기소개서 AI로 다듬기 등 실생활 밀착형 강의까지 포함됐다.

교육 바우처는 인당 20만 원 상당으로, 현재 200여 명을 대상으로 배포되며 내년 상반기까지 1,000여 명 수준으로 확대될 예정이다. 뤼튼 측은 AI 교육 접근성을 높여 디지털 약자층까지 포용하는 것을 목표로 한다고 설명했다. 이와 함께 뤼튼은 자체 AI 교육 커리큘럼과 AI 역량·적성 인증 제도도 독자 개발해 연내 공개할 계획이다. 커리큘럼은 일반인부터 직장인, 중소상인, AI 약자 계층 등 다양한 국민의 요구사항을 반영해 온·오프라인으로 제공된다. 생성 AI 입문자를 위한 기본 개념 및 기술과 주요 AI 서비스 활용 실습, AI 에이전트 개발자를 위한 MCP, n8n 등 최신 기술 활용 에이전틱 AI 구축 실습 과정 등이 준비 중이다.

특히 주목할 점은 AI 역량·적성 인증 제도다. 단순히 AI 지식을 묻는 수준을 넘어, 이용자가 AI를 활용하는 구체적 방식을 논리적 활용, 감정적 활용, 의존, 위임, 참고 등 다양한 관점에서 검증하는 것에 초점을 맞췄다. 이를 통해 AI 리터러시 교육을 이수한 개인과 단체가 자기 수준을 객관적으로 측정하고 역량을 향상하도록 돕는다. 업계는 뤼튼의 행보가 급성장하는 AI 교육 시장에서 선점 효과를 노린 전략으로 분석한다. 링크드인이 발표한 급부상하는 기술 보고서에 따르면, AI 리터러시는 2025년 커리어를 변화시킬 가장 수요 높은 기술 1위로 선정되며 전 산업군에서 주목받고 있다.

626) THE AI, "AI 리터러시가 경쟁력" 뤼튼, 무료 교육 바우처로 국민 AI 교육 지원, 2025.11.6. https://www.newstheai.com/news/articleView.html?idxno=9624

나. ‘AX CIC’ 설립… AI 전환 사업 본격 추진[627)]

뤼튼테크놀로지스가 2025년 9월 29일 AI 전환(AX, AI Transformation) 전문 사내 독립기업(CIC) ‘뤼튼 AX’를 설립했다. 뤼튼은 급속한 AI 발전 속도에 대응하기 위해 신속한 독자 의사결정이 가능한 사내 독립기업 형태로 뤼튼 AX 설립을 결정했다고 밝혔다. 뤼튼 AX는 기업, 학교, 공공기관 등에서 인공지능 전환을 통한 생산성 향상을 목표로 한다. 회사는 각 조직의 특성과 업무에 최적화된 Agentic AI를 직접 개발하고 제공하여 맞춤형 AI 솔루션 시장에서 경쟁력을 확보한다는 계획이다. 뤼튼 AX는 AI 전환을 위한 기업·기관 교육 및 컨설팅, 기업·교육 현장 특화형 뤼튼 플랫폼 제공, RAG 파이프라인 및 MCP 서버를 활용한 Agentic AI 구축 사업을 진행한다.

뤼튼은 현재 문화체육관광부의 보도자료 작성용 AI 도구 개발 프로젝트를 진행 중이며, 경기도교육청의 생성 AI 플랫폼 구축 사업에도 참여하고 있다. 또한 아모레퍼시픽과 마케팅 관련 협업도 진행하고 있다고 전했다. 특히 뤼튼은 AI 에이전트 간편 제작 도구 ‘AutoBE’를 자체 개발해 CS, 재무, 코딩 등 사내 업무에 활용하고 있다. 회사 측은 관련 부서의 생산성이 15%에서 최대 35%까지 향상되었다고 분석했다고 밝혔다.

뤼튼 AX 대표에는 박민준 뤼튼 기획전략실장이 선임되었다. 박 대표는 뤼튼의 500만 명 MAU 달성을 이끈 PO로 활동하며, CS, 재무, 코딩 분야의 업무용 AI 에이전트 기획과 개발을 담당해 왔다. 박민준 뤼튼 AX 대표는 “AX 사업 추진에 따라 AI 에이전트 개발자 채용을 확대하고 있다”라며 “다수의 AX 협업 프로젝트가 진행 중인 만큼 추가 성과가 이어질 것”이라고 말했다. 이세영 뤼튼 대표는 “AI 전환을 고민 중인 기업과 기관들이 뤼튼 AX와 함께 AI 생산성 향상을 경험할 수 있다”라고 전했다.

다. ‘크랙’에 제미나이 2.5 프로 모델 지원, 구글과 협업 강화[628)]

구글 클라우드가 2025년 9월 1일 뤼튼의 AI 캐릭터 채팅 서비스 ‘크랙(Crack)’에 구글의 최신 제미나이 2.5 프로(Gemini 2.5 Pro) 모델을 제공한다고 발표했다. 이는 구글 클라우드와 인공지능(AI) 서비스 플랫폼 기업 뤼튼테크놀로지스와의 협업 확장의 일환이다. 양사는 지난

627) 플래텀, 뤼튼, ‘AX CIC’ 설립… AI 전환 사업 본격 추진, 2025.9.29.
https://platum.kr/archives/271686
628) 보안뉴스, 뤼튼 ‘크랙’에 제미나이 2.5 프로 모델 지원...구글클라우드-뤼튼 협업 강화, 2025.9.1.
https://www.boannews.com/media/view.asp?idx=139008

28일 서초동 뤼튼테크놀로지스 사무실에서 전략적 AI 파트너십 방안에 대해 논의하고 AI, 인프라, 보안 등 다양한 부문에서 협업을 확장하기로 했다. 뤼튼은 구글 클라우드 인프라를 기반으로 한국 및 일본에서 대규모 서비스를 운영하며 양사의 전략적 AI 파트너십을 더욱 공고히 할 계획이다.

뤼튼이 지난 4월 단독 서비스로 정식 출시한 크랙은 캐릭터와 상호작용을 하며 이야기를 만들어 나가는 공간이다. 이용자에게 단순한 AI 채팅을 넘어 직접 콘텐츠를 창작하고 공유하며 소통할 수 있는 새로운 경험을 제공한다. 이용자는 다양한 성격과 세계관을 갖춘 자신만의 AI 캐릭터를 생성하고, 다른 이용자가 만든 캐릭터도 마음껏 즐길 수 있다. 구글 클라우드는 기존에 크랙에서 제공되던 제미나이 2.5 플래시(Gemini 2.5 Flash) 모델에 이어 제미나이 2.5 프로 모델을 추가로 제공한다. 제미나이 2.5 프로는 이용자가 생성한 독창적인 세계관과 캐릭터가 일관성을 유지하면서 보다 자연스러운 대화를 이어갈 수 있도록 지원한다. 한층 더 몰입감 넘치는 경험 제공에 핵심 역할을 한다.

양사는 지난 7월 구글 클라우드 데이 서울 미디어 브리핑에서도 긴밀한 AI 협업을 발표한 바 있다. 현재 ‘뤼튼 3.0’ 서비스에서도 제미나이 2.5 모델군이 제공되고 있다. 구글 클라우드 버텍스 AI(Vertex AI)의 강력한 암호화 기능을 통해 데이터 처리 전 과정을 안전하게 보호한다. 뤼튼은 사내 주요 의사결정에도 구글 클라우드의 대규모 데이터 분석 플랫폼인 빅쿼리(BigQuery)를 활용 중이다.

라. LG AI 연구원과 협력…LLM ’엑사원’ 도입[629] 819

AI 서비스 플랫폼 기업 ’뤼튼테크놀로지스가 2025년 8월 19일 ’전 국민 AI 역량 강화 프로젝트’의 본격적 추진을 위해 LG AI 연구원과 LLM 엑사원 도입 파트너십 계약을 체결했다고 밝혔다. 전 국민 AI 역량 강화 프로젝트는 모든 국민이 쉽고 편리하게 AI를 이용할 수 있도록 AI 소외 계층과 학생들을 위한 교육 프로그램들을 운영하고 AI 기술 협력 생태계를 확장함으로써 국가 AI 경쟁력 강화에 기여하는 것을 목표로 한다. 뤼튼은 이를 위해 지난달 31일 AI 반도체 기업 퓨리오사AI와 고성능・고효율 추론 인프라 구축을 위한 전략적 업무제휴 협약을 체결했다.

---

629) 머니투데이, 뤼튼, LG AI연구원과 맞손…LLM '엑사원' 도입, 2025.8.19. https://www.mt.co.kr/future/2025/08/19/2025081914023670930

뤼튼과 LG AI연구원은 이번 협약을 통해 △AI 관련 기술과 콘텐츠의 교육 분야 활용 △AI 리터러시 향상을 위한 교육 기획 및 서비스 △AI 기술의 실사용 가능성 및 적용 환경 등의 영역에서 협력하기로 했다. 임우형 LG AI연구원 공동 연구원장은 “엑사원의 고도화된 기술을 전 국민 AI 역량 강화를 위해 적극적으로 활용하고 국내 대표 AI 서비스 기업인 뤼튼과 협업을 통해 사용자 중심의 혁신을 만들어 나가겠다“라고 밝혔다.

제6장 인공지능 기술 및 연구 동향

2025년 전 세계 주요 연구기관과 표준기관의 연구 활동과 논문 등 발간물 분석을 통해 다음과 같은 주요 트렌드를 도출하였다.

● 생성형 AI의 ‘범용 기술(GPT)’ 화와 생산성 논쟁

OECD 연구는 생성형 AI가 다양한 산업에 파급되는 보편성, 모델·컴퓨팅의 지속적 개선, 타 분야 혁신을 촉발하는 성격을 근거로 GPT(General Purpose Technology)로 자리잡았다고 평가했다. 동시에 ‘생산성 역설’ 가능성을 제기하며, 실제 생산성 향상이 교육·조직 변화·투자에 달려 있어 정책과 기업 전략이 뒷받침되지 않으면 단기 성과는 제한적일 수 있다고 분석하였다.

● LLM 성능·효율성 제고를 위한 AI 반도체·HBM·소프트웨어의 공동 최적화 연구

NPU 기반 AI 칩과 5세대 HBM(HBM3E) 등 고대역폭 메모리를 결합해 대규모 파라미터 모델의 연산 효율을 높이는 설계가 본격적으로 논의되고 있다. 퓨리오사AI와 같은 설계 전문기업은 AI 추론용 고성능·저전력 NPU를 통해 GPU 대비 비용·전력 소비를 크게 줄이려는 시도를 진행 중이다. 이러한 사례는 LLM 연구가 더 이상 알고리즘에만 머물지 않고, 모델 구조-컴파일러-칩 설계-HBM 아키텍처를 함께 최적화하는 총체적 연구로 확장되고 있음을 시사한다. 특히 AI 반도체 스타트업과 클라우드·빅테크의 협업은, 연구 단계에서부터 하드웨어 제약과 비용을 고려한 ‘deploy-aware’ 모델 설계가 2025년의 중요한 기술 트렌드가 되고 있음을 보여준다.

● 에이전트형 AI·멀티에이전트 시스템을 겨냥한 벤치마크·평가 연구의 고도화

중국 Zhipu AI의 GLM-4.5 사례에서 보듯, 2025년에는 전통적인 MMLU·코딩 성능뿐 아니라 TAU-Bench, BrowseComp 등 에이전트 능력을 평가하는 전용 벤치마크가 함께 제시되고, 각 모델의 에이전틱 코딩, 도구 호출 성공률, 멀티스텝 작업 수행능력을 수치로 비교·공개하는 흐름이 뚜렷하다. 이는 단일 LLM의 정적 성능보다, 다양한 도구·환경과 상호작용하며 복잡한 업무를 수행하는 ‘에이전트 스택 전체’의 성능을 연구·평가하는 방향으로 기술·연구의 초점이 이동하고 있음을 보여주며, 각국 연구기관·기업이 유사한 벤치마크를 도입하는 흐름과 맞물린다.

● 멀티모달·툴 기반·에이전트형 모델이 AGI 논의를 견인

AAAI-25 대통령 패널·초청 강연에서는 “Beyond Understanding: Toward Controllable and Agentic Multimodal Models”와 같은 발표를 통해, LLaVA 등 최신 멀티모달 모델과 툴 기반 에이전트, 장기 메모리 등 AGI 구성요소 기술의 고도화를 주요 의제로 다루어졌다. 카카오의 ‘카나나’ 사례처럼 텍스트·이미지·음성·영상까지 통합한 멀티모달 에이전트와 2025년부터 본격화된 오픈소스 공개 정책은 ‘도구를 쓰는 에이전트형 AI’ 연구가 핵심 트렌드임을 보여준다.

● 국제표준·인증·윤리·안전 연구의 제도화 및 고도화

UN AI 자문기구는 ITU·ISO·IEC·IEEE 중심의 AI 국제표준 채택 건수가 2018년 이후 급증해 2025년에도 가속 추세가 이어지고 있음을 지적하고, SC 42가 데이터 품질·AI 생애주기·성능평가 등 표준군을 정교화·세분화하고 있다고 평가했다. ISO/IEC JTC 1/SC42는 2025년 기준 설명가능성·통제 가능성·편향 처리·투명성·환경 지속가능성·AI 시스템 영향평가 등 다양한 주제의 표준을 개발·개정하며, 생성형 모델 중심 패러다임을 반영해 국제표준 체계를 재정비 중이다. 한국은 KSA의 AI+ 인증제, 산업AI 국제인증포럼, TTA의 AI 표준·시험인증 체계 등 기술표준과 신뢰성 평가·인증을 결합한 ‘제도화된 AI 안전·품질 관리’ 생태계를 빠르게 확장시키고 있다.

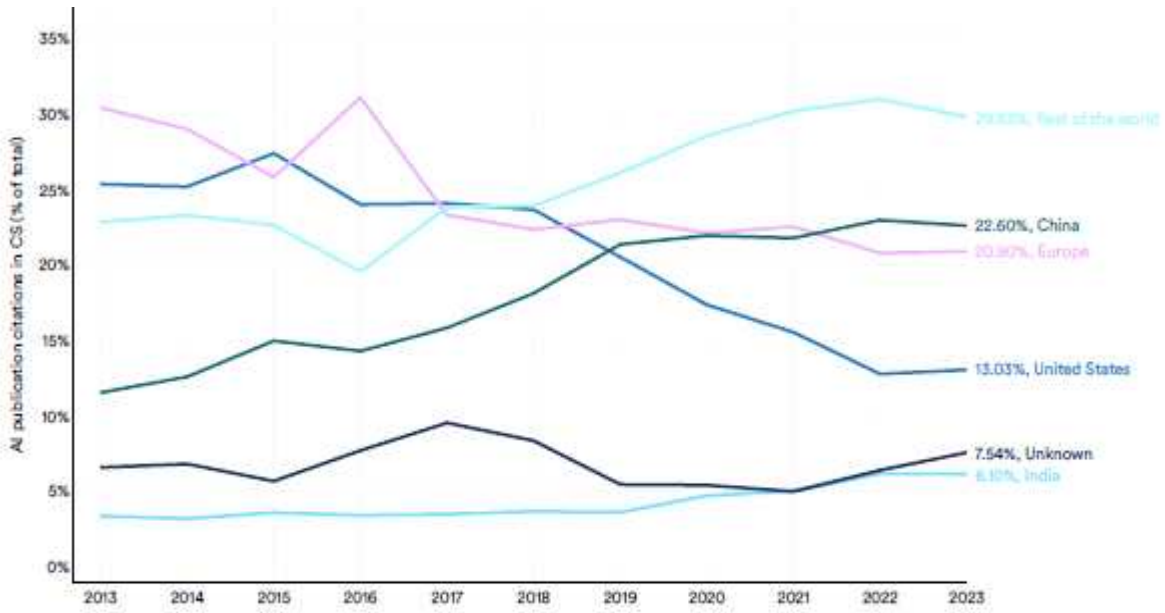
제1절 학술단체

1. 학술연구 현황 및 추이

Stanford HAI의 AI Index 보고서[630]에 따르면, 2024년 기준 AI 연구·개발 생태계는 산업계와 학계의 역할 분담이 더욱 뚜렷해지고 있다. 산업계는 AI 모델 개발에서 압도적인 주도권을 확보하며, 2024년에 공개된 주목할 만한 모델의 약 90%가 산업에서 배출되었다. 이는 2023년 60%에서 크게 증가한 수치로, 대규모 모델 개발에 필요한 데이터·컴퓨팅·자본이 산업에 집중되면서 나타난 구조적 변화로 해석된다. 반면 학계는 여전히 인용 상위 100위 내에 드는 고영향력 연구에서 가장 높은 비중을 차지하며, 기초 연구와 핵심 알고리즘 혁신의 중심축으로 기능하고 있다.

국가별 연구 생산성을 살펴보면, 2023년 기준 중국은 전체 AI 연구 출판물과 인용 수에서 각각 23.2%, 22.6%를 차지하며 양적 측면에서 가장 큰 기여를 하고 있다. 그러나 고영향력 연구의 측면에서는 최근 3년 동안 미국 기관이 가장 많은 상위 100위권 AI 논문을 발표하며 질적 우위를 유지하였다.

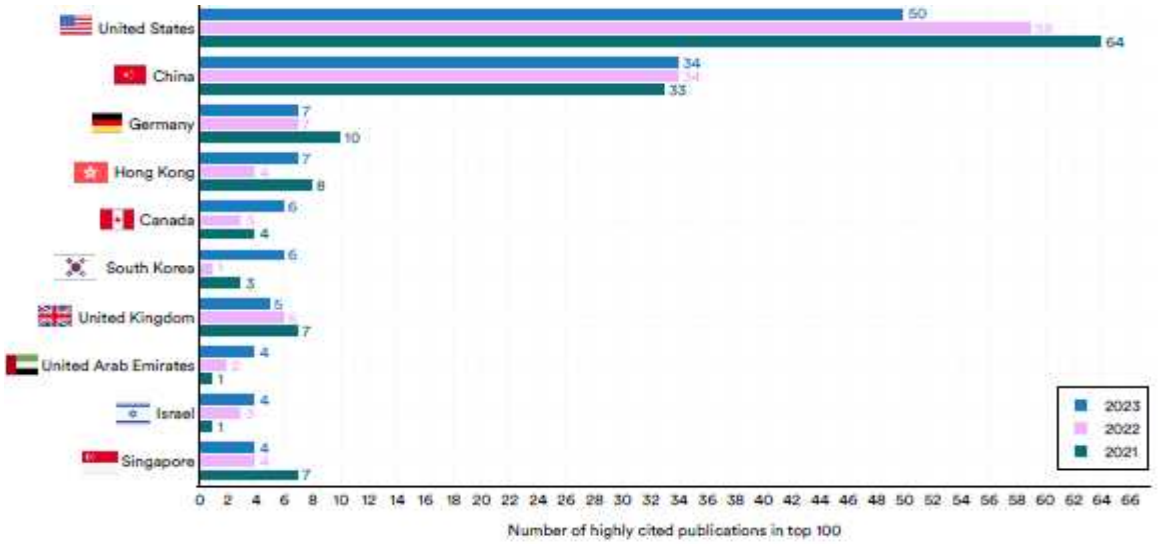
[그림 6-1] 지역별 AI 연구 출판물 인용(수)



출처: Stanford HAI

630) 출처: Stanford HAI, AI Index Report 2025, 2025

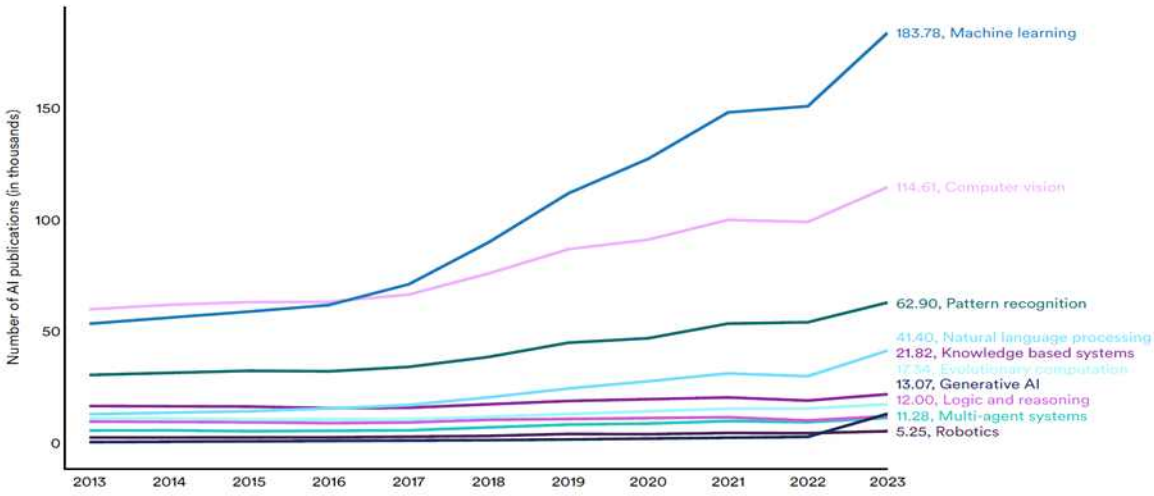
[그림 6-2] 상위 100위권 AI 출판물(수)



출처: Stanford HAI

AI 분야 출판물 규모는 지난 10년간 급속히 확대되어, 2013년 약 10만 2천 건에서 2023년 24만 건을 넘어섰다. 특히 컴퓨터과학 분야 내에서 AI 연구가 차지하는 비중은 같은 기간 21.6%에서 41.8%로 상승하며, AI가 컴퓨터과학 연구의 핵심 축으로 자리 잡았음을 확인할 수 있다. 모델 개발 측면에서도 미국의 주도권은 더욱 강화되어, 2024년 미국 기관이 개발한 주목할 만한 AI 모델은 총 40개로 중국(15개)과 유럽 전체(3개)를 크게 앞질렀다. 최근 10년간 누적 기준으로도 미국은 단연 가장 많은 머신러닝 모델을 생산한 국가로 평가된다.

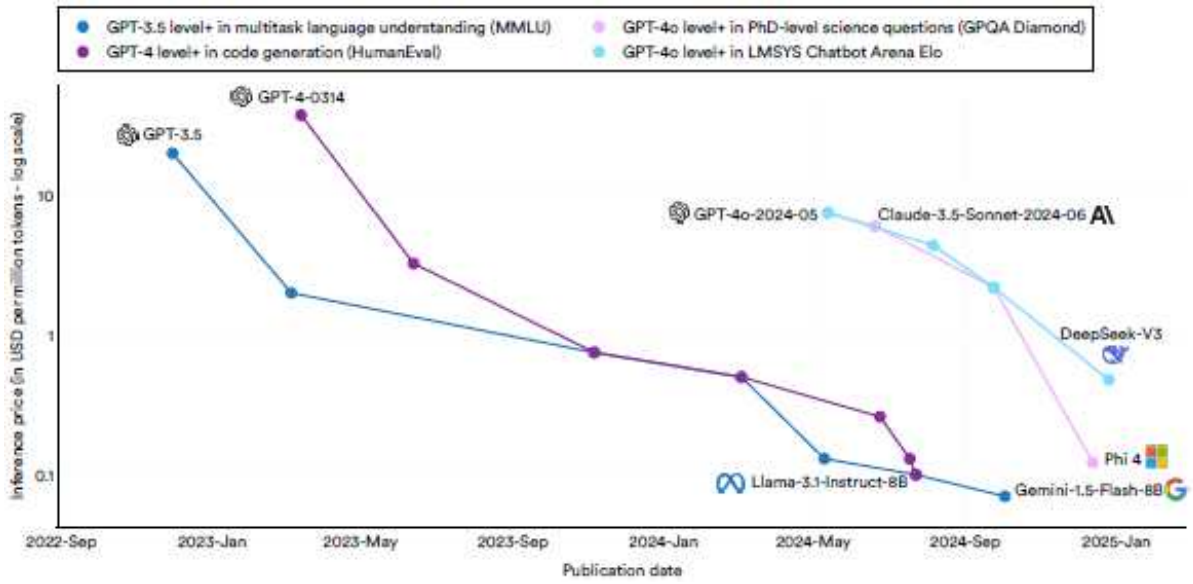
[그림 6-3] 연구 분야별 AI 관련 출판물 발간 추이(2010-2022)



출처: Stanford HAI

한편, 모델 성능 향상 경쟁은 컴퓨팅 수요의 폭발적 증가를 동반하고 있다. 최신 연구에 따르면, 주목할 만한 AI 모델의 학습에 필요한 컴퓨팅 규모는 약 5개월마다 두 배씩 증가하고, 대규모 언어모델 학습 데이터셋 규모는 8개월마다 두 배씩 확대되고 있다. 전력 소모량도 매년 증가하는 추세로, 산업계의 대규모 투자와 모델 스케일링 전략이 이러한 기술적 팽창을 견인하고 있다. 그러나 모델의 활용 측면에서는 비용이 빠르게 낮아지고 있다. MMLU 64.8점(GPT-3.5 수준)을 내는 모델을 기준으로, 100만 토큰 질의 비용은 2022년 11월 20달러에서 2024년 10월 0.07달러까지 감소해 약 280배 가까운 가격 하락이 나타났다. 사용 목적에 따라 LLM 추론 비용은 연간 최소 9배에서 최대 900배까지 감소하며, AI 접근성은 빠르게 확대되고 있다.

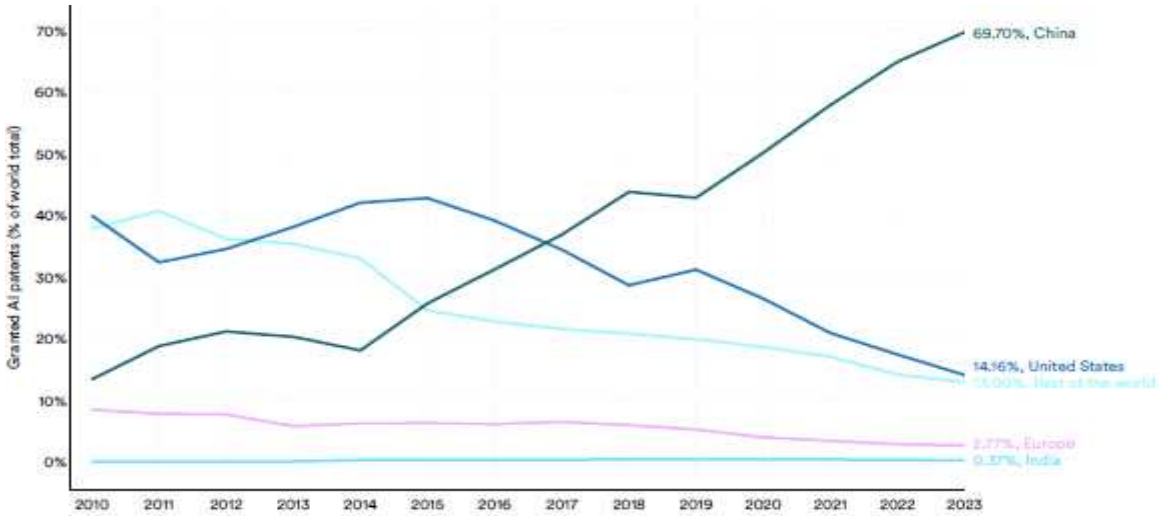
[그림 6-4] 주요 모델별 추론 비용



출처: Stanford HAI

기술 상용화를 반영하는 AI 특허 역시 2010년 3,833건에서 2023년 122,511건으로 급증하며 지속적인 성장세를 확인했다. 특히 2023년 한 해 동안 특허 건수는 29.6% 증가하였으며, 전체 특허의 69.7%가 중국에서 발생했다. 인구 대비 특허 생산성에서는 한국과 룩셈부르크가 각각 1위와 2위를 기록해 기술 집약적 AI 연구 역량을 보여주었다.

[그림 6-5] 지역별 특허(수) 비율

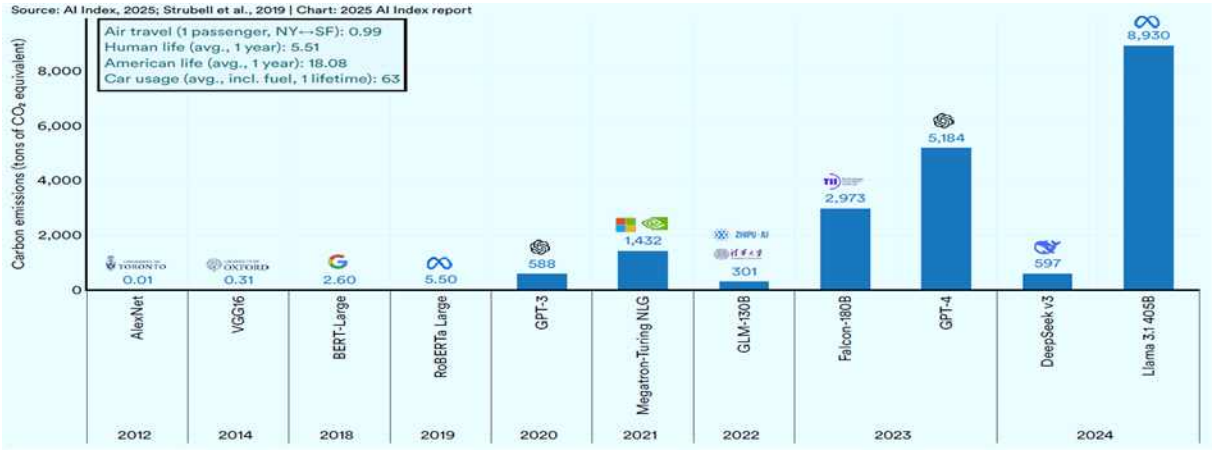


출처: Stanford HAI

AI 하드웨어 측면에서는 연산 성능, 가격 효율성, 에너지 효율성이 모두 개선되고 있다. 머신러닝용 하드웨어의 연산 성능(FP16 기준)은 연간 43%씩 증가하며 약 1.9년마다 두 배로 향상되었고, 비용은 매년 30% 감소한 반면, 에너지 효율성은 연간 40%씩 증가하는 추세를 보였다. 이러한 추세는 대규모 모델 학습을 위한 하드웨어 생태계가 지속적으로 개선되고 있음을 시사한다.

그러나 대규모 모델 학습에 따른 탄소 배출량은 지속적으로 증가하고 있다. AlexNet(2012)은 약 0.01톤 수준의 탄소 배출량을 기록했으나, GPT-3(2020) 학습은 588톤, GPT-4(2023)는 5,184톤, Llama 3.1 405B(2024)는 8,930톤에 달하는 탄소를 배출한 것으로 추정된다. 이는 평균 미국인의 연간 탄소 배출량(18톤)과 비교할 때 거대한 수치로, AI 기술 확산과 함께 환경 영향에 대한 논의가 더욱 중요해지고 있음을 시사한다.

[그림 6-6] 대규모 학습 모델의 탄소 배출량(추정치)



출처: Stanford HAI

AI Index 보고서의 연구개발 분야의 연구를 종합하면, 2024~2025년의 AI 연구·개발 환경은 산업 주도의 모델 개발 가속화, 국가 간 연구 경쟁의 심화, AI 출판의 급격한 팽창, 컴퓨팅·에너지 요구량의 증가와 비용 절감의 공존, 특허와 하드웨어 혁신의 동시 확장, 그리고 탄소 배출의 구조적 증가 등 다층적 변화가 동시에 관찰되는 시기로 요약된다.

2024년 5월 한국과학기술기획평가원(KISTEP)이 7월 발표한 '국가전략기술 R&D 인력 실태조사'[631]에 따르면 지난 6년간 전 세계에서 발표된 인공지능(AI) 분야 논문은 약 73만 건이며, 이 중 한국이 발표한 논문은 약 1만 4천 건으로 세계 12위 수준으로 나타났다. 글로벌 AI 연구는 미국과 중국이 주도하고 있으며, 상위 500명의 AI 연구자 중 미국이 31.8%, 중국이 24.2%를 차지한 반면, 한국 연구자는 5명으로 세계 16위 수준에 머물렀다.

2025년 기준 AI 학술 연구의 정량적 추이는 아직 공식적으로 집계·발표되지 않았다. 다만 과학기술정보통신부는 2025년도 예산을 전년 대비 5.47% 증가한 18.9조 원으로 확정하면서, 인공지능 연구 및 활용 기반 강화를 위한 투자를 확대하였다. 구체적으로 ▲AI 컴퓨팅 인프라 확충과 AI·디지털 융합 확산 등에 0.87조 원, ▲이공계 대학원생 연구생활장려금 신설을 포함한 안정적 연구환경 조성과 기초연구 확대에 3.56조 원을 각각 배정하였다. 이러한 재정적 투자 확충은 2025년 국내 AI 연구 생태계의 활성화, 연구 인력 기반 강화, 그리고 중장기적 연구 역량 제고에 기여할 것으로 전망된다.

631) 한국과학기술기획평가원(KISTEP), 국가전략기술 R&D(연구·개발) 인력 실태조사, 2024.5

2. 2025년 국내외 학술단체 주요 동향

1) 국제

(1) 전기전자공학학회(IEEE)

IEEE(Institute of Electrical and Electronics Engineers)는 전기전자공학, 컴퓨터공학, 정보기술 등 다양한 분야를 다루는 세계 최대의 기술 전문기관으로, AI 연구 생태계에서도 가장 영향력 있는 학회 중 하나로 자리하고 있다. IEEE는 CVPR(Conference on Computer Vision and Pattern Recognition), IROS(IEEE/RSJ International Conference on Intelligent Robots and Systems), ICDM(International Conference on Data Mining) 등 AI 핵심 분야의 국제 학술대회를 주관하며, AI・로봇・데이터과학 분야 연구를 주도하는 저널・트랜잭션 시리즈를 운영하고 있다.

특히 최근 IEEE는 인공지능의 윤리, 신뢰성, 안전성 확보를 위한 글로벌 기술 규범 제정 활동을 강화하고 있다. IEEE는 Ethically Aligned Design의 후속 조치로 GET(Globally Emerging Technology) Program을 운영하며, 자율・지능형 시스템(AIS: Autonomous and Intelligent Systems)의 안전・신뢰・윤리를 다루는 표준 개발을 중점적으로 추진하고 있다. 2024년에 GET Program은 자율・지능형 시스템의 고장 안전 설계(IEEE 7009-2024), 그리고 에뮬레이션 기반 공감 기능의 윤리적 고려(IEEE 7014-2024) 등 핵심 표준 2건을 발간하였으며, 2025년 11월 기준 추가로 신규 발간한 사례는 없다.[632]

IEEE Standards Association(IEEE SA)은 2024년을 기준으로 지능형 시스템의 신뢰성, 설명가능성, 상호운용성, 연령 검증, 지식그래프 등의 기술을 다루는 광범위한 표준 개발을 진행하였다. 여러 표준이 2025년 중 최종 승인 예정 상태로 등록돼 있다. 2025년 기준 AIS(AI/Autonomous Intelligent Systems) 관련 추가적으로 제안된 표준은 없으며, 기존에 제안된 표준화 작업이 진행 중으로, 대표적 표준은 다음과 같다.

632) IEEE GET ProgramTM (2025.12.23. 방문)
https://ieeexplore.ieee.org/browse/standards/get-program/page/series?id=93

<표 6-1> IEEE SA의 AIS 관련 주요 표준(2025. 11월 기준)

번호	표준 명칭
IEEE 2089.1	IEEE Standard for Online Age Verification (표준화 진행 중)
IEEE 2807.1	IEEE Approved Draft Standard for Technical Requirements and Evaluation of Knowledge Graphs (표준화 진행 중)
IEEE 2807.2	IEEE Approved Draft Guide for Application of Knowledge Graphs for Financial Services (표준화 진행 중)
IEEE 2894	IEEE Approved Draft Guide for an Architectural Framework for Explainable Artificial Intelligence (표준화 진행 중)
IEEE 2933	IEEE Approved Draft Standard for Clinical Internet of Things (IoT) Data and Device Interoperability with TIPPSS - Trust, Identity, Privacy, Protection, Safety, Security
IEEE 3154	IEEE Recommended Practice for the Application of Knowledge Graphs for Talent Services
IEEE 3168	IEEE Standard for Robustness Evaluation Test Methods for a Natural Language Processing Service That Uses Machine Learning
IEEE 3800	IEEE Approved Draft Standard for Household Appliance Customer Data Assetization and Commercialization Requirements
IEEE 3803	IEEE Approved Draft Standard for Household Appliance Customer Data Assetization and Commercialization Requirements (표준화 진행 중)
IEEE 7000	IEEE Standard Model Process for Addressing Ethical Concerns during System Design
IEEE 7001	IEEE Standard for Transparency of Autonomous Systems
IEEE 7009	IEEE Standard for Fail-Safe Design of Autonomous and Semi-Autonomous Systems
IEEE 7014	IEEE Standard for Ethical Considerations in Emulated Empathy in Autonomous and Intelligent Systems

출처: IEEE SA[633)]

IEEE SA는 기술 표준 제정과 함께 인공지능의 윤리적 기준을 평가하는 인증 프로그램 IEEE CertifAIEd도 운영하고 있다. 이 프로그램은 기업 및 기관의 AI 제품·서비스가 국제적 윤리 기준(Transparency, Accountability, Privacy, Governance 등)을 충족하는지 검증하기 위한 인증 방식으로, 2024~2025년 동안 적용 분야를 로보틱스, 의료 AI, 금융 AI 등으로 확대하고 있다.[634)]

633) Autonomous and Intelligent Systems (AIS) Standards (2025.11.17. 방문)
https://standards.ieee.org/initiatives/autonomous-intelligence-systems/standards

(2) ACM(CVPR)

ACM(Association for Computing Machinery)은 컴퓨터과학과 정보기술 분야에서 세계적으로 가장 큰 학술 및 전문 협회 중 하나로, 학회 및 콘퍼런스를 주최하고, 컴퓨터과학 연구의 논문을 발표하는 저널을 발간한다. ACM은 SIGKDD(Knowledge Discovery and Data Mining), SIGGRAPH(Computer Graphics and Interactive Techniques), ACM CHI(Human-Computer Interaction)와 같은 대표적인 AI 관련 학회들을 주최한다.

ACM 내에는 특히 인공지능 분야에 대해 SIGAI(Special Interest Group on Artificial Intelligence)가 운영되고 있으며, 이 그룹은 AI 원리와 기술의 성장 및 응용을 촉진하고 지원하는 역할을 맡고 있다. ACM은 다양한 SIG 중 하나로 인공지능 특별 관심 그룹(SIGAI: Special Interest Group on Artificial Intelligence)를 운영하고 있다. 이 그룹은 컴퓨팅 전반에 걸쳐 AI 원리와 기술의 성장과 응용을 촉진하고 지원하는 역할을 수행한다. ACM SIGAI는 학계 교육자 및 연구자, 전문가, 학생이라는 세 가지 주요 커뮤니티로 구성되어 서비스를 제공하며, 주요한 활동은 다음과 같다.

- 고품질 AI 관련 컨퍼런스 후원 또는 공동 후원과 주요 연례 상을 통해 학계, 전문가 및 학생 회원의 업적 인정
- ACM 디지털 라이브러리에 분기별 뉴스레터(AI Matters) 발행 및 커뮤니티의 관심사인 AI 문제에 대한 블로그 유지
- 학생 회원에게 컨퍼런스 참석을 위한 장학금 제공과 AI 커뮤니티에서 AI 교육, 출판물 및 공공 홍보 활동 지원, 개발 및 홍보

SIGAI는 자율 에이전트 연구(Autonomous Agents Research) 분야의 우수 연구자를 수상한다. 2025년 수상자로 Shlomo Zilberstein 교수를 선정했다. Zilberstein 교수는 매사추세츠대학교 애머스트(University of Massachusetts Amherst) 컴퓨터과학과 교수이자 전 연구부 학장(Associate Dean of Research)을 역임한 바 있다. 그는 분산 마르코프 결정 과정(DEC-MDP, Decentralized Markov Decision Processes) 분야를 개척하여, 다중 에이전트 시스템 및 다중 에이전트 강화학습(MARL)에서 의사결정 이론 기반 계획(decision-theoretic planning)의 기초를 확립한 학자로 평가된다. 이 연구 성과는 다중 에이전트 의사결정 분야의 핵심 이론으로

634) IEEE CertifAIEd™ Professional Certification (2025.11.17. 방문)
https://standards.ieee.org/products-programs/icap/ieee-certifaied/professional-certification

자리 잡아, 학계와 산업계 연구자들에게 널리 영향을 미쳤다.[635]

ACM(Association for Computing Machinery)은 2025년 3월 5일, 2024년도 튜링상 수상자로 강화학습(Reinforcement Learning) 분야의 기초를 정립한 두 명의 연구자를 선정하였다.[636][637] 튜링상은 컴퓨터과학 분야에서 가장 권위 있는 상으로, '컴퓨터과학의 노벨상'으로도 불리며, 수상자에게는 구글의 후원으로 100만 달러의 상금이 수여된다. ACM은 올해 수상자로 매사추세츠대학교 애머스트 캠퍼스의 앤드루 바르토(Andrew Barto) 명예 교수와 캐나다 앨버타대학교의 리처드 서튼(Richard Sutton) 교수를 발표하였다. 이들은 1980년대 이후 강화학습의 핵심 아이디어를 제시하고 수학적 기반을 체계화하는 데 결정적인 공헌을 한 인물들로 평가된다.

ACM은 바르토와 서튼이 수십 년 동안 강화학습의 이론적 토대를 구축하며 알고리즘 발전을 견인해왔다고 평가하였다. 이들은 심리학적 관찰과 학습 이론을 바탕으로 강화학습을 문제 해결을 위한 일반적 프레임워크로 정립하였으며, 1998년 공동 집필한 『Reinforcement Learning: An Introduction』은 7만 5천 회 이상 인용되며 오늘날까지도 강화학습의 대표적인 표준 교과서로 자리 잡고 있다. 특히 이들의 이론적 기초는 2010년대 이후 심층신경망(Deep Learning)과 결합되며 실제 응용 분야에서 비약적 발전을 이끌었다. 대표적 사례로는 2016~2017년 세계 정상급 바둑 기사들을 연이어 제압한 알파고(AlphaGo), 오픈AI의 챗GPT에 적용된 인간 피드백 기반 강화학습(RLHF) 기법이 있으며, 로봇 제어, 네트워크 혼잡 제어, 칩 설계, 글로벌 공급망 최적화 등 다양한 산업·기술 분야에서도 핵심기술로 활용되고 있다.

한편 두 수상자는 현재의 AI 기술 개발 속도에 대해 우려를 표하며, 일부 기업들이 안전성 검증 없이 상업적 목적을 우선해 제품을 빠르게 출시하는 경향이 있다고 지적하였다. 바르토 교수는 연구의 축적보다는 시장 중심의 경쟁이 지나치게 강조되는 현 흐름을 비판하며, 신뢰할 수 있는 엔지니어링 관행과 철저한 검증의 중요성을 재차 강조하였다.

635) SIGAI 연구 성과 페이지 (2025.11.17. 방문)
https://sigai.acm.org/main/2025/03/09/shlomo-zilberstein-2025-autonomous-agents-research-award
636) ACM, ACM A.M. Turing Award Honors Two Researchers Who Led the Development of Cornerstone AI Technology, 2025.03.05.
637) Financial Times, Turing Award winners warn over unsafe deployment of AI models, 2025.03.05.

(3) AI 분야 학회 - 신경정보처리시스템(NeurIPS)[638]

NeurIPS(Conference on Neural Information Processing Systems)는 ICML, ICLR와 함께 머신러닝·인공지능 분야에서 가장 영향력 있는 국제 학술대회 중 하나로, 매년 12월 개최되는 프리미어급(top-tier) 학회이다. 컴퓨터 비전, 자연어 처리, 강화학습, 최적화, 통계, 계산신경과학 등을 포괄하는 폭넓은 주제를 다루며, 학계와 산업계의 최첨단 연구 성과가 집중적으로 발표된다.

2025년 NeurIPS 2025는 제39차 연례 학회로, 2025년 12월 2일부터 7일까지 미국 캘리포니아주 샌디에이고 컨벤션센터에서 개최되며, 일부 프로그램은 11월 30일부터 12월 5일까지 멕시코시티(Hilton Mexico City Reforma)에서 병행 개최될 예정이다(11월 조사 시점 기준). NeurIPS 2025 Call for Papers에 따르면, 이번 학회는 기계학습·신경과학·통계·최적화·컴퓨터 비전·자연어 처리·생명·자연·사회과학 등 인접 분야 연구자를 아우르는 학제간(interdisciplinary) 플랫폼을 목표로 하며, 응용·이론·방법론·인프라 전반에 걸친 새로운 연구 성과를 모집하고 있다.

NeurIPS 2025년 주제 구성은 2024년과 유사한 틀을 유지하면서도, 기초모델·LLM, Creative AI, 사회·경제적 영향 등 최근 AI 연구의 핵심 이슈를 명시적으로 강조하고 있다는 점에서 특징적이다. 또한 NeurIPS 2025는 기존의 본 학술 논문 트랙 외에도 데이터셋·벤치마크, 튜토리얼, 워크숍, 컴피티션, 저널 연계 트랙 등을 운영한하며, 최첨단 AI 기술 연구뿐만 아니라, 그 사회적·정책적 함의에 대한 학문적 논의를 선도하는 장으로서 역할을 확대하고 있다.

(4) AI 분야 학회 - 컴퓨터비전패턴인식학회(CVPR)[639][640]

CVPR(Conference on Computer Vision and Pattern Recognition)은 컴퓨터 비전 및 패턴 인식 분야에서 세계적으로 가장 권위 있는 국제 학술대회 중 하나로, 이미지 인식, 객체 탐지, 비디오 분석, 3D 비전, 멀티모달 AI 등 딥러닝 기반 연구가 활발히 발표되는 장이다.

---

638) 신경정보처리시스템학회 홈페이지 (2025.11.17. 방문) https://neurips.cc
639) CVPR2025 결과보고 웹페이지 (2025.11.17. 방문) https://cvpr.thecvf.com/Conferences/2025
640) CVPR, Best Papers at CVPR Reveal New Results with Neural Networks for Real-Time Applications and Novel Ways to Manipulate Light for Scene Recovery, 2025.06.13.

2025년에는 제42차 연례 학회가 미국 테네시주 내슈빌(Music City Center)에서 6월 11일(수)부터 6월 15일(일)까지 개최되었다. 미국 테네시주 내슈빌에서 2025년 6월 11일부터 15일까지 개최된 CVPR 2025(Computer Vision and Pattern Recognition)는 전 세계 4만 명 이상 연구자가 13,008편의 논문을 제출하고 이 중 2,872편이 채택되어 약 22%의 채택률을 기록하는 등, 컴퓨터 비전 분야 최대 규모·최고 영향력 학회로서의 위상을 재확인하였다. 주최 측은 CVPR가 일부 대기업 연구에만 집중하지 않고 뚜렷한 성과를 보인 다양한 연구에 고르게 주목함으로써, 컴퓨터 비전 연구 생태계의 '민주화'에 기여하고 있다고 평가하였다. 2025년 학회에서는 인공지능(AI)이 전반적인 핵심 화두로 부상한 가운데, 멀티뷰 및 센서 기반 3D, 이미지·비디오 합성, 멀티모달 학습과 시각·언어·추론이 주요 연구 주제로 부각되었다. 특히 3D 분야에서는 2020년경 도입된 NeRF(Neural Radiance Fields) 계열 기법을 기반으로, 단일 혹은 다중 2D 이미지로부터 3차원 장면의 구조를 복원·이해하려는 연구가 지속적으로 확대되었고, 멀티모달 영역에서는 텍스트뿐 아니라 이미지·비디오까지 동시에 이해·생성하는 모델을 통해 환경을 더 정밀하게 표현하려는 시도가 활발하게 이루어졌다. 시각·언어·추론을 결합한 멀티모달 추론 분야는 CVPR 2025에서 가장 많은 논문이 제출된 범주로, 차세대 트렌드로 자리잡고 있다.

CVPR 2025 조직위원회는 최우수 논문상(Best Paper Award)으로 옥스퍼드대학교와 메타 AI(Meta AI) 연구진이 공동 제출한 'VGGT(Visual Geometry Grounded Transformer)'를 만장일치로 선정하였다. VGGT는 기존에 개별 작업(예: 깊이 추정, 포즈 추정 등)에 특화된 3D 컴퓨터 비전 접근법을 넘어, 2D 이미지로부터 장면의 핵심 3D 정보(카메라 위치, 깊이 등)를 직접 추론할 수 있도록 설계된 순방향(feed-forward) 신경망 아키텍처이다. 이 모델은 3D 이미지를 생성한 뒤 별도의 후처리·보정 단계를 거치지 않고도, 1초 이내에 이미지를 3D로 재구성할 수 있을 정도로 구조가 단순하면서도 효율적인 것으로 평가되며, 실시간 애플리케이션에 적합한 차세대 3D 비전 기술로 주목받고 있다.

(5) AI 분야 학회 - 국제머신러닝학회(ICML)[641]

ICML은 머신러닝 및 인공지능 연구 분야에서 NeurIPS , ICLR 와 함께 가장 영향력 있는 주요 학회 중 하나로, 머신러닝 알고리즘·이론·응용을 폭넓게 다룬다. 2025년에는 제 42회 학회가 캐나다 밴쿠버(Vancouver Convention Center)에서 7월 13일(일)부터 7월 19일(토)

641) 국제머신러닝학회 홈페이지 (2025.11.17. 방문) https://icml.cc

까지 개최되었다. 이 컨퍼런스에는 튜토리얼, 워크숍, 전시(엑스포) 등이 포함되어 있으며, 제출 논문에는 지도학습・비지도학습・온라인 학습・생성모델・최적화・강화학습 등의 주제가 포함한다.

ICML은 머신러닝 분야에서 매년 뛰어난 연구 성과를 선정하여 Test of Time Award를 수여한다. 이 상은 발표 후 장기간에 걸쳐 학문・산업에 지속적인 영향을 미친 연구에 수여되는 것으로, 2025년 ICML Test of Time Award는 Sergey Ioffe와 Christian Szegedy가 2015년에 발표한 'Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift'가 선정되었다. 해당 논문은 딥러닝 학습 과정에서 발생하는 '내부 공변량 변화(Internal Covariate Shift)' 문제를 근본적으로 해결하기 위해 배치 정규화(Batch Normalization, BN)라는 새로운 학습 안정화 기법을 제안하였다. 내부 공변량 변화는 이전 층의 파라미터 변화로 인해 각 층이 받는 입력 분포가 계속 변하는 현상으로, 이는 학습 안정성을 저해하고 학습률 감소, 초기값 민감도 증가, 포화형 활성화 함수의 비효율 등 다양한 문제를 초래한다. Ioffe와 Szegedy는 이 문제를 해결하기 위해 각 미니배치 단위로 입력을 정규화하는 방식을 모델 아키텍처에 직접 통합하였으며, 이를 통해 딥러닝 모델의 학습 과정을 대폭 안정화하였다.[642]

(6) AI 분야 학회 - 국제컴퓨터비전학회(ICCV)[643]

ICCV(International Conference on Computer Vision)은 컴퓨터 비전과 패턴인식 분야에서 가장 권위 있는 학회로 2년마다 개최되며, 이미지 인식, 객체 탐지, 비디오 분석과 같은 주제에서 중요한 연구들이 발표된다. 특히, 딥러닝 기반 컴퓨터 비전 연구가 활발하게 논의된다.

2025년에는 미국 하와이 호놀룰루(Honolulu)에서 10월 19일부터 10월 25일까지 개최되었다. 이 학회는 이미지・비디오・3D 비전, 영상 이해, 자율시스템 인식 등을 포함하는 핵심 비전 연구를 발표하는 국제적 플랫폼이다.

이번 2025년 학회에서의 AI 관련 주요 연사와 주요 내용은 다음과 같다.

642) ICML, Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift, (2025.11.17. 방문) https://icml.cc/virtual/2025/39855
643) 국제컴퓨터비전학회 홈페이지 (2025.11.17. 방문) https://iccv.thecvf.com

- Brent Seales — “Virtually Unwrapping the Herculaneum Scrolls” : 화산재로 탄화된 고대 두루마리를 AI 기반 영상·복원 기술로 읽어내는 ‘가상 언래핑’ 연구가 소개하고, 비침습적 디지털 복원 기법을 통해 2,000년간 열리지 못한 텍스트를 복원하는 과정 설명
- Linda B. Smith — “The efficiency of learner generated experiences” : 유아가 스스로 만들어내는 시각·행동 데이터의 통계적 특성이 효율적 학습의 근원이 된다는 연구를 제시하고, 일상환경에서 유아의 시각 데이터를 분석해, 행동이 곧 ‘학습 친화적 데이터 생성’ 메커니즘이라는 점을 강조

(7) AI 분야 학회 - 국제로봇자동화학회(ICRA)[644]

IEEE가 후원하는 ICRA(International Conference on Robotics and Automation)는 로봇 공학과 자동화 분야에서 세계적으로 가장 권위 있는 학회 중 하나로, IEEE가 주관하는 대형 국제 학회이다. ICRA는 로봇 공학의 최신 기술과 연구 성과를 공유하는 자리로, 연구자, 학계, 산업계 전문가들이 모여 로봇 기술의 미래를 논의하고 협력하는 중요한 플랫폼으로 작용한다.

2025년에는 5월 19일부터 5월 23일까지 미국 애틀랜타(Atlanta)에서 개최되었다. ICRA 2025에서는 로봇공학 분야의 전통적 연구성과와 더불어, 대규모 데이터와 인공지능 기반 로봇지능 고도화가 학술적 핵심 주제로 부상하였다. 특히 로봇 학습을 위한 대규모 멀티모달 데이터셋 구축, 모방학습 및 강화학습의 확장성 문제 해결, 확률적 인지 모델을 활용한 안전한 로봇 제어 등, 데이터·AI 중심 접근법이 주요 논문에서 공통적으로 나타났다. 이를 통해 로봇 시스템이 다양한 환경에서 일반화 성능을 확보하고 장기간 자율적으로 작동할 수 있는 기반이 강화되고 있음을 확인할 수 있다. 또한 인간-로봇 상호작용(HRI), 로봇 조작 및 이동, 복수 로봇의 자율 협력 등 응용 분야 전반에서 학습 기반 인지·계획 기술의 비중이 크게 확대되었다. 자연어 기반 지시 이해, 촉각·시각 융합 인지, 대규모 시뮬레이션을 통한 안전성 검증과 같은 연구가 다수 제시되었으며, 이는 로봇이 물리적 세계를 능동적으로 탐색·적응하는 ‘Physical AI’로의 전환이 가속화되고 있음을 시사한다.[645]

644) 국제로봇자동화학회 홈페이지 (2025.11.17. 방문) https://2025.ieee-icra.org
645) ICRA, ICRA 2025 Awards and Finalists, (2025.11.17. 방문)
https://2025.ieee-icra.org/program/awards-and-finalists

(8) AI 분야 학회 - 세계인공지능학회(AAAI)[646]

IEEE가 후원하는 AAAI(Association for the Advancement of Artificial Intelligence)는 인공지능 전반을 다루는 대표적 국제 학회로, 자연어처리, 게임 AI, 인간-컴퓨터 상호작용(HCI) 등 다양한 응용 분야를 포함한다.

AAAI-25에서는 인공지능의 근본적 계산 원리와 사회적 활용 가능성을 동시에 확장하는 연구가 우수논문으로 선정되었다. 먼저, Soroush Ebadian 등의 "Every Bit Helps - Achieving the Optimal Distortion with a Few Queries" 연구는 최소한의 정보만으로도 왜곡(distortion)을 최적화하는 질의 기반 알고리즘을 제시한다. 또 다른 선정 논문인 Wen-Chao Hu 등의 "Efficient Rectification of Neuro-Symbolic Reasoning Inconsistencies by Abductive Reflection"는 뉴로-심볼릭 시스템의 추론 불일치를 효율적으로 교정하는 새로운 귀납적 반성 방법을 제안한다. 한편, Marius Belly 등의 "Revelations: A Decidable Class of POMDPs with Omega-Regular Objectives" 논문은 기존에 계산적으로 난해하다고 여겨졌던 일부 POMDP 문제군을 결정가능(decidable) 영역으로 규명한다.[647]

AAAI는 2025년 3월 2일, AI 연구의 핵심 주제와 그 역사적 흐름, 현재의 기술적·학문적 과제들을 종합적으로 검토한 보고서를 발간하였다.[648] AAAI는 최근 AI 기술의 급속한 발전 속에서 연구 주제, 연구 방법, 연구자 커뮤니티, 작업 환경 전반이 매우 빠르게 변화하고 있다는 점을 강조하며, 이러한 변화의 맥락을 체계적으로 파악하고 향후 연구의 방향성을 정립하기 위한 목적으로 본 보고서를 작성하였다.

보고서는 △AI 추론 △AI의 사실성과 신뢰성 △AI 에이전트 △AI 평가 △AI 윤리와 안전 △체화 AI(Embodied AI) 등 최근 연구 커뮤니티에서 주목받는 핵심 분야를 중심으로 주요 동향과 과제를 제시한다. AI 추론 분야에서는 대규모 사전학습을 거친 LLM이 자동화된 추론 능력에서 뚜렷한 도약을 이루었으나, 복잡한 문제에 대한 추론의 정확성과 깊이를 확보하기 위해서는 지속적인 연구가 필요함을 지적하였다. AI의 사실성과 신뢰성 분야에서는 신

646) AAAI, EAAI-25 Call for Participation, (2025.11.17. 방문)
https://aaai.org/conference/aaai/aaai-25/eaai-25-call
647) AAAI, AAAI-25 Paper Awards (2025.11.17. 방문)
https://aaai.org/about-aaai/aaai-awards/aaai-25-paper-awards
648) AAAI, AAAI 2025 Presidential Panel on the Future of AI Research, 2025.03.02.

뢰성 부족이 AI의 실사용을 저해하는 핵심 장애물로 작용하고 있으며, 이를 보완하기 위한 미세조정 기법과 검색증강 생성(RAG) 기술이 활용되고 있으나 단기간 내 근본적인 해결은 어려울 것으로 전망했다.

AI 에이전트 분야에서는 시스템 구성요소를 모듈화하여 투명성과 적응성을 강화하는 다중 에이전트 아키텍처가 부상하고 있으며, 인간과 AI 간 협력적 의사결정을 강조하는 '협력형 AI'의 중요성도 점차 커지고 있다. 그러나 다중 에이전트 환경의 확장성, 투명성, 컴퓨팅 효율성 등은 여전히 중요한 연구 과제로 남아 있다고 평가하였다. AI 평가 분야에서는 현재 성능 중심 벤치마크에 과도하게 의존하는 평가 방식이 AI의 사용성, 투명성, 윤리·규제 준수 여부 등 중요한 요소들을 충분히 반영하지 못하고 있다고 지적하며, 신뢰성 있는 AI 배포를 위해 보다 다차원적이고 정교한 평가 체계가 필요함을 강조하였다.

AI 윤리와 안전 분야에서는 AI 기반 사이버 범죄, 자율 무기 등 새로운 형태의 위험이 증가하는 반면, 이를 규율할 기술적·제도적 장치가 충분하지 않아 학제 간 협력, 지속적 감독, 책임성 있는 개발 체계의 수립이 필요함을 제시하였다. 마지막으로 체화 AI 분야에서는 시뮬레이션 및 현실 환경 기반의 강화학습을 통해 발전하고 있으나, 인간 수준의 지능형 체화 에이전트를 구현할 역량은 아직 제한적이라고 평가하였다. 다만 해당 분야에는 근본적 기술적 장애물이 존재하지 않으며, 장기적으로 물리적 환경에서 학습·행동할 수 있는 고도화된 AI 에이전트 개발 가능성이 충분하다는 점을 보고서는 강조하고 있다.

2) 국내

(1) 한국정보과학회(KIISE)

한국정보과학회는 컴퓨터 과학 및 정보 기술 분야에서 가장 오래되고 권위 있는 학술 단체 중 하나로, AI)분야도 중점적으로 다루고 있다. 한국정보과학회는 매년 AI와 관련된 여러 학술대회와 세미나를 개최하며, 한국컴퓨터종합학술대회(KCC)와 같은 대형 행사에서 AI 관련 연구 발표와 기술 교류가 이루어진다. 학회 내에서 AI 연구회를 운영하며, AI 기술의 최신 동향과 연구 성과를 공유한다. 2024년 한국컴퓨터종합학술대회(KCC)[649]는 7월 1일부터 4일까지 제주국제컨벤션센터에서 개최되었다. 이번 행사에서는 “글로벌 디지털 리더십과 대한민국 컴퓨팅의 역할”라는 주제로 개최되었다. 이 행사에는 AI 분야에는 다음과 같은 주제들이 발표되었다.

<표 6-2> 신진연구자 II - AI 관련 연구 발표 주제

번호	연구 제목	발표자
1	Efficient and Trustworthy AI: 딥러닝의 효율성과 신뢰성 확보	김수현 교수 (경희대)
2	인공지능 기반 위성 영상 복원 및 활용 연구	이재협 교수 (경북대)
3	Randomized Algorithms for Accelerating Inference-time of Large Language Models	한인수 교수 (KAIST)
4	Feed-forward 3D Reconstruction	박은병 교수 (연세대)
5	Enabling Ubiquitous AI via Adaptive and Efficient On-Device ML	공태식 교수 (UNIST)
6	Recent Advancements in Video Localization	이필현 교수 (인하대)
7	생성형 AI 게임 플레이 과정에서 나타나는 플레이어의 성취감	서재인 교수 (국립목포대)
8	Toward AGI: Challenges and Solutions in Domain Adaptation	황의원 교수 (이화여대)
9	CheapNet: Cross-attention on Hierarchical Representations for Efficient Protein-ligand Binding Affinity Prediction	이상선 교수 (인하대)
10	생성형 AI 시대의 콘텐츠 보호: 개인화 디퓨전 모델에 대한 이미지 스타일 보호 기술	안남혁 교수 (인하대)
11	Applications of Trustworthy AI (신뢰할 수 있는 인공지능의 활용)	이경수 교수 (전북대)
12	수학과에서 인공지능 연구를 한다고?	이동헌 교수 (고려대)

2025 한국컴퓨터종합학술대회 (KCC2025)[650]; SPRi 정리

649) 한국정보과학회, 2024 한국컴퓨터종합학술대회 안내 페이지 (2025.11.17. 방문)
https://www.kiise.or.kr/conference/kcc/2024

(2) 한국인공지능학회(KSAI)[651]

한국인공지능학회(Korean Artificial Intelligence Association)는 2016년 설립되어 AI 관련한 컴퓨터비젼 및 패턴인식, 자연어 처리, 바이오 인포매틱스, 뇌인지컴퓨팅, 기계학습 및 그와 관련한 응용 분야 등에 관한 학문-기술 연구 및 교육을 수행한다.

한국인공지능학회는 2025 하계학술대회를 2025년 8월 7일부터 9일까지 평창 휘닉스 호텔에서 개최하였다 025년 하계학술대회에서는 멀티모달 AI, 에이전트형 AI, AI 평가・검증 패러다임 등 현재 AI 연구의 핵심 쟁점을 중심으로 기조강연이 진행되었다. 이용재 교수(University of Wisconsin-Madison)는 “Beyond Understanding: Toward Controllable and Agentic Multimodal Models”를 통해 LLaVA 등 최신 멀티모달 모델의 발전을 소개하였다. 조경현 교수(New York University)는 “Reality Checks” 강연에서 최근 AI 연구의 평가 방식이 갖는 구조적 한계를 네 가지 사례—지속학습의 하이퍼파라미터 부재, 머신 언러닝의 지표 불일치, 자유형 텍스트 생성 평가의 난제, 연구자들의 ‘희망적 사고(wishful thinking)’—를 통해 비판적으로 조명하였다. 초청강연에서는 정혜동 PM(IITP)이 AGI 전환의 기술적 과제를 다루며 LLM을 넘어 멀티모달・툴 기반 에이전트・장기 메모리 등 AGI 구성 요소 기술이 빠르게 고도화되고 있음을 설명하였다. 또한 김병학 성과리더(카카오)는 ‘카나나(Kanana)’ 모델 패밀리 개발 사례를 소개하며 텍스트・이미지・음성・영상 전 영역을 다루는 멀티모달 기반의 능동적 AI 에이전트(Agentic AI) 구현 전략과 2025년 2월부터 시작된 오픈소스 공개 정책의 의미를 제시하였다.

한국인공지능학회[652]는 2025 추계학술대회를 ‘Physical AI International Forum’과 공동으로 2025년 11월 19부터 20일까지 코엑스 컨벤션센터 개최한다. 이번 2025년 한국인공지능학회 추계학술대회는 로봇・제조・모빌리티・물류 등 현실 공간으로 확장되는 AI, 이른바 Physical AI를 핵심 주제로 삼고 있다. 이를 위해 학술대회는 ‘Physical AI International Forum’과 연계하여 진행되며, 이 포럼에는 정부, 국회, 산업계, 학계 및 해외 석학이 참여해 Physical AI가 가져올 산업 혁신과 기술・정책적 과제에 대해 논의를 진행한다. 학술 프

650) 한국정보과학회, 2025 한국컴퓨터종합학술대회 안내 페이지 (2025.11.17. 방문)
https://www.kiise.or.kr/conference/kcc/2025
651) 한국인공지능학회 2025하계학술대회(2025.11.17. 방문)
https://aiassociation.kr/Conference/ConferenceView.asp?AC=0&CODE=CC20250302&CpPage=309
652) 한국인공지능학회 2025추계학술대회(2025.11.17. 방문)
https://aiassociation.kr/Conference/ConferenceView.asp?AC=0&CODE=CC20251001

로그램으로는 △Physical AI 핵심기술과 응용을 다루는 4개 튜토리얼, △과학기술정보통신부 ‘독자 AI 파운데이션 모델 개발 사업’ 참여 기업들의 진행 상황을 공유하는 특별세션, △2025년 12월 NeurIPS에 발표될 국내 우수 연구 성과를 사전에 소개하는 ‘NeurIPS Preview’ 세션 등이다.

제2절 표준기관

1. 표준화 현황 및 추이

UN 인공지능 고위급 자문기구는 ‘Governing AI for Humanity’ 최종보고서에서, 2018년 이후 ITU, ISO, IEC, IEEE를 중심으로 인공지능(AI) 관련 국제표준 채택 건수가 뚜렷한 증가세를 보이고 있다고 평가한다.

이후 2025년 들어서도 표준화 가속 추세는 이어지고 있다 ISO/IEC JTC 1/SC 42 역시 AI 데이터 품질(ISO/IEC 5259 시리즈), AI 시스템 생애주기(ISO/IEC 5338) 등 표준을 연속적으로 발간하며, AI 생태계 전반을 포괄하는 표준군을 정교화・세분화하는 방향으로 확대하고 있다.[653]

ITU는 2025년 4월 기준 AI 관련 국제표준을 이미 120건 발간하고, 130건 이상을 추가로 개발 중이라고 밝히고 있어, UN 보고서 작성 시점 이후에도 AI 표준 발간 규모가 계속 확대되고 있음을 확인할 수 있다.[654]

우리나라는 AI 기술의 고도화와 상용화 확대에 대응하여 신뢰성・안전성・책임성 기반의 AI 표준화 체계 강화를 정책 기조로 삼고 있다. 2024~2025년 동안 정부는 AI 기본법 제정, AI 안전 관리체계 도입 논의를 병행하면서, 국내 표준(KS)과 국제표준(ISO/IEC, ITU 등) 연계를 강화하는 방향으로 표준화 전략을 추진 중이다.

2. 2025년 국내외 표준기관 주요 동향

1) 국제

(1) ISO/IEC

ISO/IEC JTC 1/SC 42(Artificial intelligence)는 인공지능 시스템의 신뢰성, 윤리적 사용, 데

653) ISO, Standards by ISO/IEC JTC 1/SC 42 (2025.11.17. 방문)
https://www.iso.org/committee/6794475/x/catalogue

654) ITU, Standards help unlock trustworthy AI opportunities for all, 2025.4.1. (2025.11.17. 방문)
https://www.itu.int/hub/2025/04/standards-help-unlock-trustworthy-ai-opportunities-for-all

이터 관리, 평가 및 산업별 적용을 위한 국제표준화 작업을 주도하고 있다.

2025년 11월 기준 ISO/IEC JTC 1 SC 42는 전체 AI 생태계의 국제표준화를 수행하며, 41개의 표준을 공개(발행)했으며, 47개의 현재 개발 중이다.

<표 6-3> ISO/IEC JTC 1/SC 42에서 발행된 표준 목록

번호	표준번호	표준 이름
1	ISO/IEC 42001:2023	Information technology — Artificial intelligence — Management system
2	ISO/IEC 38507:2022	Information technology — Governance of IT — Governance implications of the use of artificial intelligence by organizations
3	ISO/IEC 25059:2023	Software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — Quality model for AI systems
4	ISO/IEC TS 25058:2024	Systems and software engineering — Systems and software Quality Requirements and Evaluation (SQuaRE) — Guidance for quality evaluation of artificial intelligence (AI) systems
5	ISO/IEC 5259-1:2024	Artificial intelligence — Data quality for analytics and machine learning — Part 1: Overview, terminology, and examples
6	ISO/IEC 5259-2:2024	Artificial intelligence — Data quality for analytics and machine learning — Part 2: Data quality measures
7	ISO/IEC 5259-3:2024	Artificial intelligence — Data quality for analytics and machine learning — Part 3: Data quality management requirements and guidelines
8	ISO/IEC 5259-4:2024	Artificial intelligence — Data quality for analytics and machine learning — Part 4: Data quality process framework
9	ISO/IEC 5259-5:2025	Artificial intelligence — Data quality for analytics and machine learning — Part 5: Data quality governance framework
10	ISO/IEC 5338:2023	Information technology — Artificial intelligence — AI system life cycle processes
11	ISO/IEC 5339:2024	Information technology — Artificial intelligence — Guidance for AI applications
12	ISO/IEC 5392:2024	Information technology — Artificial intelligence — Reference architecture of knowledge engineering
13	ISO/IEC TR 5469:2024	Artificial intelligence — Functional safety and AI systems
14	ISO/IEC TS 6254:2025	Artificial intelligence — Explainability & interpretability of machine learning and AI systems — Objectives and approaches
15	ISO/IEC 8183:2023	Information technology — Artificial intelligence — Data life cycle framework
16	ISO/IEC TS 8200:2024	Information technology — Artificial intelligence — Controllability of automated AI systems
17	ISO/IEC TS 12791:2024	Information technology — Artificial intelligence — Treatment of unwanted bias in ML classification & regression
18	ISO/IEC 12792:2025	Information technology — Artificial intelligence — Transparency taxonomy of AI systems
19	ISO/IEC TR 17903:2024	Information technology — Artificial intelligence — Overview of machine learning computing devices
20	ISO/IEC TR 20226:2025	Information technology — Artificial intelligence — Environmental sustainability aspects of AI systems

번호	표준번호	표준 이름
21	ISO/IEC TR 21221:2025	Information technology — Artificial intelligence — Beneficial AI systems
22	ISO/IEC 22989:2022	Information technology — Artificial intelligence — Concepts and terminology
23	ISO/IEC 23053:2022	Framework for Artificial Intelligence (AI) Systems Using Machine Learning (ML)
24	ISO/IEC 23894:2023	Information technology — Artificial intelligence — Guidance on risk management
25	ISO/IEC TR 24027:2021	Artificial intelligence — Bias in AI systems and AI-aided decision making
26	ISO/IEC TR 24028:2020	Artificial intelligence — Overview of trustworthiness in AI
27	ISO/IEC TR 24029-1:2021	Assessment of the robustness of neural networks — Part 1: Overview
28	ISO/IEC 24029-2:2023	Assessment of the robustness of neural networks — Part 2: Formal methods
29	ISO/IEC TR 24030:2024	Artificial intelligence — Use cases
30	ISO/IEC TR 24368:2022	Artificial intelligence — Overview of ethical and societal concerns
31	ISO/IEC TR 24372:2021	Artificial intelligence — Overview of computational approaches for AI systems
32	ISO/IEC 24668:2022	Artificial intelligence — Process management framework for big data analytics
33	ISO/IEC TS 4213:2022	Artificial intelligence — Assessment of machine learning classification performance
34	ISO/IEC 42005:2025	Artificial intelligence — AI system impact assessment
35	ISO/IEC 42006:2025	Requirements for bodies providing audit & certification of AI management systems
36	ISO/IEC TS 42119-2:2025	Testing of AI — Part 2: Overview of testing AI systems

출처: ISO 웹사이트[655]

ISO/IEC JTC 1/SC 42는 2025년 11월 기준으로 AI 시스템의 안전성·신뢰성·투명성 확보를 위한 국제표준 개발을 진행하고 있다. 우선, AI 성능평가(ISO/IEC CD 4213), 데이터 품질(DTR 5259-6), NLP 기반 자연어처리 기술(ISO/IEC CD TR 23281·23282), 신경망 강건성(ISO/IEC CD 24029-3) 등 핵심기술 기반의 정량적·기술적 기준 마련이 집중적으로 진행되고 있다. 특히 생성형 AI의 확산으로 개념·용어·시스템 프레임워크를 다시 정의하기 위한 기존 표준의 대규모 개정이 병행되고 있어, 국제표준 체계가 생성형 모델 중심의 패러다임을 반영하는 방향으로 빠르게 재정비되는 추세가 나타나고 있다.

655) ISO, Standards by ISO/IEC JTC 1/SC 42 (2025.11.17. 방문)
https://www.iso.org/committee/6794475/x/catalogue/p/1/u/0/w/0/d/0

동시에, AI 기술이 사회·산업 전반으로 확장됨에 따라 SC 42는 윤리, 사회적 수용성, 기능 안전, 리스크 관리, 사회경제적 영향 등 비기술적 요소에 대한 표준 개발을 대폭 확대하고 있다. 이에 따라 사회적 우려 해소와 윤리적 요구를 다루는 표준(ISO/IEC CD TS 22443, TS 25566, TS 25568, TS 25571), 인간 감독 및 휴먼-머신 팀 구성(ISO/IEC DIS 42105, AWI 25589, TR 42109), 안전·신뢰성 프레임워크(ISO/IEC AWI TS 22440 시리즈, TS 25570)와 같은 AI 거버넌스·감독·안전성 중심의 규범 표준군이 새롭게 형성되고 있다.

(2) ITU

국제전기통신연합(ITU)은 AI 기술의 급속한 확산에 대응하기 위해 안전성, 윤리성, 신뢰성을 중심으로 국제표준 개발을 적극 추진하고 있다. ITU는 ICT 및 AI 분야의 주요 유엔 전문기구로서, 다양한 산업에서 AI가 책임 있게 활용될 수 있도록 여러 정책 프레임워크와 기술 표준화 논의를 확대하고 있다. AI 관련 표준화는 특히 ITU-T(전기통신표준화 부문) 내 연구그룹과 포커스 그룹(Focus Group)을 중심으로 진행되며, 사회·경제적 활용을 위한 국제적 합의를 지원하고 있다

ITU는 AI 기술의 급격한 발전에 발맞춰 FG-AI4EE(Focus Group on Environmental Efficiency for AI and Other Emerging Technologies), FG-AI4NDM(Focus Group on AI for Natural Disaster Management)와 같은 특화된 연구 그룹을 통해 AI의 윤리적 사용, 안전성, 산업별 적용, 환경 효율성을 다루는 표준화를 적극적으로 추진하고 있다.

ITU는 ‘AI for Good Global Summit’를 매년 오프라인으로 개최하고, 연중 다양한 주제로 온라인 세미나를 개최한다. 2025년 ‘AI for Good Global Summit’은 7월 8일부터 11일까지 스위스 제네바에서 ‘글로벌 과제를 해결하기 위한 인공지능 분야의 유엔 선도 플랫폼’이라는 주제로 개최되었고, 정부, 시민사회, UN 기관, AI 혁신가, 투자자들이 참여하였다.

국제전기통신연합(ITU)은 2025년 발표한 ‘AI Standards for Global Impact: From Governance to Action’ [656]보고서를 통해, AI 시대에서 국제표준의 중요성을 다시 한 번 강조하였다. 보고서 서문에서 ITU 전기통신표준국(TSB) 국장 Seizo Onoe는 ITU가 160년간

국제 전기통신 표준을 주도해온 경험을 바탕으로, AI・양자기술 등 신흥 기술 분야에서도 포용적이고 투명하며 합의 기반의 표준 개발을 유지하겠다는 원칙을 제시하였다. 특히 ITU는 표준이 국가 간 혁신 공유와 시장 신뢰 형성에 기여한다는 점을 강조하며, 표준은 국제적 혁신과 투자를 지속시키는 신뢰의 기반이라고 정의하였다. 이번 보고서는 2025년 AI for Good Global Summit의 International AI Standards Exchange 트랙에서 논의된 최신 기술・정책 동향을 정리한 것으로, 네트워킹・멀티미디어 진위성 검증, 사이버보안, 양자기술, 에너지 효율, 헬스케어, 식량안보, 스마트 모빌리티 등 다수 분야에서 AI가 초래하는 변화와 표준화 요구를 제시하고 있다.

656) ITU, AI Standards for Global Impact: From Governance to Action, 2025. https://www.itu.int/dms_pub/itu-t/opb/ai4g/T-AI4G-AI4GOOD-2025-4-PDF-E.pdf

2) 국내

(1) 한국표준협회(KSA)

한국표준협회는 2021년부터 AI+ 인증제도를 개발하여 AI 관련 기업에 인증 서비스를 제공하고 있다. 심사 모델은 ISO/IEC 42001 인공지능경영시스템을 기반으로 AI리스크 관리와 PDCA 사이클을 통하여 인공지능 제품 및 서비스 품질의 지속 개선하는지 확인하는데 목표를 두고 있다. 2025년 11월 11일 기준, AI+ 인증을 받은 제품은 신한카드의 '초개인화 시스템 2.0', 삼성전자의 '그랑데 건조기 AI' 등 140개 이상이다.[657] AI+ 인증의 절차는 아래 그림과 같다.

[그림 6-7] AI+ 인증 절차



출처: 한국표준협회[658]

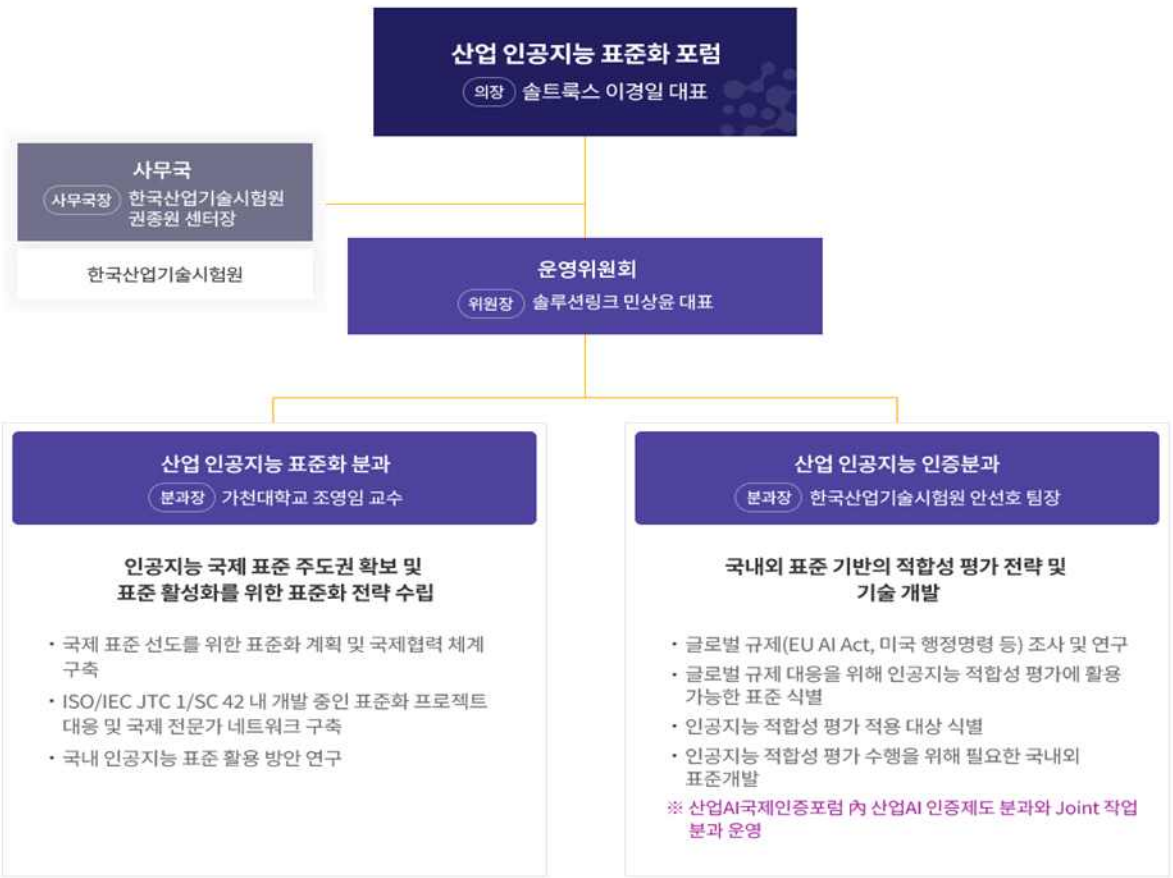
657) 한국표준협회, AI+ 인증기업 현황 (2025.11.17. 방문) https://ksa.or.kr/ksa_kr/7021/subview.do
658) 한국표준협회, AI+(에이아이플러스)인증 (2025.11.17. 방문) https://ksa.or.kr/ksa_kr/6962/subview.do

(2) 국가기술표준원(KATS)

국가기술표준원(KATS)은 국내 AI 기술의 신뢰성·안전성 확보를 위해 산업 수 ’은 학계·연구기관·산업계가 참여하는 민간 주도형 표준화 협의체로, AI 신뢰성 평가 기준, 산업별 상호운용성, AI 윤리·사회적 이슈 등에 대한 표준화 논의를 수행하고 있다. 포럼은 ▲산업 인증지능 표준화 ▲산업 인공지능 인증의 2개 분과를 운영하며, 한국산업기술시험원이 운영사무국을 맡고 있다.[659]

2024년 2월 인공지능 분야(JTC 1/SC 42) 국가표준(KS) 개발 수요조사를 진행하였다.

[그림 6-8] 산업인공지능표준화포럼



출처: 산업 인공지능 표준화 포럼[660]

659) 산업AI국제인증포럼 홈페이지 (2025.11.17. 방문) https://aicertforum.ai/front/user/main.do#
660) 산업 인공지능 표준화 포럼 (2025.11.17. 방문) http://aistandard.kr/sub/sub01_01.php

아울러, 국가기술표준원은 사용되는 AI의 신뢰성과 안정성을 연구 및 논의하는 국내외 민관협력 네트워크로 산업AI국제인증포럼(Industrial AI International Certification Forum)을 운영 중이며, 이 포럼은 AI의 적절한 사용과 개선을 촉진하며, 국제인증 생태계를 조성하여 산업의 발전에 기여함을 목표로 한다. 이 포럼에는 산업AI 인증제도, 산업AI 인증기술의 2개 분과를 구성하여 운영 중이다.[661]

(3) 한국정보통신기술협회(TTA)

한국정보통신기술협회(TTA)는 1988년 설립된 국내 대표 ICT 표준화·시험인증 기관으로, 정보통신·AI·네트워크·디지털 융합기술 분야의 국가표준(KS), 단체표준, 시험인증 체계를 총괄하는 역할을 수행하고 있다. 특히 AI·5G·IoT·UHD·사이버보안 등 급속한 기술 변화에 대응하기 위해 TTA는 국내외 표준화기구(ISO/IEC JTC 1, ITU-T, IEEE, oneM2M 등)와의 연계를 강화하며 글로벌 표준 선도 활동을 지속하고 있다.

2025년 11월 17일 기준으로 한국정보통신기술협회의 인공지능 분야에서 제·개정한 표준 목록은 아래의 표와 같다.

<표 6-4> 한국정보통신기술협회의 인공지능 분야에서 발행된 최신 표준 목록

No.	표준종류	표준번호	표준제목	제개정일
1	정보통신기술보고서 (TTAR)	TTAR-10.0202	인공지능 기반 합성 콘텐츠 투명성 - 제1부: 정책 및 활용 동향 (기술보고서)	2025-10-24
2	정보통신기술보고서 (TTAR)	TTAR-06.0274/R1	비예약 기반 비직교 무선 접속 방식을 위한 인공지능 활용 (기술보고서)	2025-10-22
3	정보통신단체표준 (TTAS)	TTAK.KO-10.1572	인공지능 기반 고등 교육 지원 시스템 기능 요구사항	2025-06-27
4	정보통신단체표준 (TTAS)	TTAK.KO-10.1523	인공지능 학습용 양식 넙치류 질병 이미지 데이터 수집 지침	2024-12-06
5	정보통신단체표준 (TTAS)	TTAK.KO-10.1533	농업 분야 인공지능 학습용 데이터 플랫폼 관리를 위한 메타데이터 스키마	2024-12-06
6	정보통신단체표준 (TTAS)	TTAK.KO-10.1558	생성형 인공지능 학습을 위한 멀티모달 데이터의 품질 검증 방법	2024-12-06
7	정보통신단체표준 (TTAS)	TTAK.KO-10.1559-Part1	디지털 가상환경 경호훈련을 위한 인공지능 학습 데이터 - 제1부: 잠재적 위해자 구성요소	2024-12-06

661) 산업AI국제인증포럼 홈페이지 (2025.11.17. 방문)
https://aicertforum.ai/front/intropage/intropageShow.do?page_id=5bdefffa9b714f2e8e1e7b65f4b10d18

No.	표준종류	표준번호	표준제목	제개정일
8	정보통신단체표준 (TTAS)	TTAK.KO-12.0414	인공지능(AI) 서비스 개인정보보호 프레임워크	2024-12-06
9	정보통신기술보고서 (TTAR)	TTAR-06.0290	분산형 인공지능을 활용한 모바일 에지 컴퓨팅에너지 최적화(기술보고서)	2024-10-23
10	정보통신단체표준 (TTAS)	TTAE.ET-GS ZSM 012	제로터치 네트워크 및 서비스 관리(ZSM): 인공지능 기반 네트워크 및 서비스 자동화를 위한 실현 기술	2023-12-06
11	정보통신단체표준 (TTAS)	TTAK.KO-10.1497	인공지능 시스템 신뢰성 제고를 위한 요구사항	2023-12-06
12	정보통신단체표준 (TTAS)	TTAK.KO-11.0324	인공지능 기반 해사 소프트웨어의 안전 요구사항 도출 지침	2023-12-06
13	정보통신기술보고서 (TTAR)	TTAR-06.0280	공공안전을 위한 무기준 인공지능 영상 품질 분석 기술(기술보고서)	2023-11-03
14	정보통신기술보고서 (TTAR)	TTAR-11.0085	IALA G1178: IALA 관점에서의 인공 지능(AI) 도입 지침(기술보고서)	2023-11-02
15	정보통신기술보고서 (TTAR)	TTAR-06.0273	무선 네트워크의 다목적 자원관리를 위한 분산형 인공지능 활용 (기술보고서)	2023-10-19
16	정보통신기술보고서 (TTAR)	TTAR-06.0275	무경계 셀룰러 시스템의 모바일 에지 컴퓨팅 자원 할당을 위한 인공지능 활용 (기술보고서)	2023-10-19
17	정보통신기술보고서 (TTAR)	TTAR-06.0276	5G 단말 기계학습 및 인공지능 서비스 지원 기술 (기술보고서)	2023-10-19
18	정보통신기술보고서 (TTAR)	TTAR-10.0178	인공지능 시스템 신뢰성 지침 및 원칙 동향(기술보고서)	2023-10-18
19	정보통신기술보고서 (TTAR)	TTAR-10.0174	OCR기반 차트 이미지 데이터 내용을 텍스트로 추출하는 인공지능 기술(기술보고서)	2023-05-18
20	정보통신단체표준 (TTAS)	TTAK.KO-10.1370	온실 운영 시스템과 인공지능 서버 간 인터페이스	2022-12-07
21	정보통신단체표준 (TTAS)	TTAK.KO-11.0280-Part4	검증용 데이터셋의 밸런스 기반 인공지능 소프트웨어 신뢰성 평가 방법 - 제4부: 웨이브 타입 밸런스 데이터 설계	2022-12-07
22	정보통신단체표준 (TTAS)	TTAK.KO-11.0280-Part5	검증용 데이터셋의 밸런스 기반 인공지능 소프트웨어 신뢰성 평가 방법 - 제5부: 동영상 타입 밸런스 데이터 설계	2022-12-07
23	정보통신단체표준 (TTAS)	TTAK.KO-10.1303	진단 보조 인공지능 모델 개발을 위한 학습용 데이터 구축방안 : 병리조직 이미지	2021-12-08
24	정보통신단체표준 (TTAS)	TTAK.KO-10.1340	모의 환경 기반 인공지능 게임 에이전트 생성 요구사항	2021-12-08
25	정보통신단체표준 (TTAS)	TTAK.KO-11.0280-Part2	검증용 데이터셋의 밸런스 기반 인공지능 소프트웨어 신뢰성 평가 방법 - 제2부: 이미지 타입 밸런스 데이터 설계	2021-12-08
26	정보통신단체표준 (TTAS)	TTAK.KO-11.0280-Part3	검증용 데이터셋의 밸런스 기반 인공지능 소프트웨어 신뢰성 평가 방법 - 제3부: 시계열 타입 밸런스 데이터 설계	2021-12-08

No.	표준종류	표준번호	표준제목	제개정일
27	정보통신기술보고서 (TTAR)	TTAR-10.0158	인공지능 기반 이미지•영상분석 기술 개요 및 서비스 유즈케이스(기술보고서)	2021-10-21
28	정보통신기술보고서 (TTAR)	TTAR-10.0160	인공지능 생성요약을 활용한 작문 평가 시스템 유스케이스(기술보고서)	2021-10-21
29	정보통신기술보고서 (TTAR)	TTAR-10.0142	스마트 헬스케어를 위한 인공지능 피트니스 코칭 시스템 프레임워크 유스케이스(기술보고서)	2021-10-20
30	정보통신단체표준 (TTAS)	TTAK.KO-10.1276	유방암 판독 인공지능 모델 개발을 위한 유방촬영술 의료지식베이스 구축방안	2020-12-10
31	정보통신단체표준 (TTAS)	TTAK.KO-11.0280-Part1	검증용 데이터셋의 밸런스 기반 인공지능 소프트웨어 신뢰성 평가 방법 - 제1부: 방법론 및 체계	2020-12-10
32	정보통신기술보고서 (TTAR)	TTAR-10.0111	인공지능 프로세서의 보안성 강화를 위한 안전성 플랜(기술보고서)	2019-11-07
33	정보통신기술보고서 (TTAR)	TTAR-10.0112	인공지능 프로세서의 기능안전 인터페이스(기술보고서)	2019-11-07
34	정보통신기술보고서 (TTAR)	TTAR-10.0113	다중 칩 구성이 가능한 인공지능 프로세서 구조(기술보고서)	2019-11-07
35	정보통신기술보고서 (TTAR)	TTAR-10.0114	다중 칩 기반 인공지능 프로세서 소프트웨어 인터페이스(기술보고서)	2019-11-07
36	정보통신기술보고서 (TTAR)	TTAR-10.0107	인공지능 윤리 가이드라인 제정 동향(기술보고서)	2019-10-30
37	정보통신기술보고서 (TTAR)	TTAR-10.0108	인공지능 - 표준화 동향 분석 (기술보고서)	2019-10-30
38	정보통신기술보고서 (TTAR)	TTAR-10.0096	인공 지능 서비스의 튜링 테스트 기반 평가 기술 동향(기술보고서)	2018-11-07
39	정보통신기술보고서 (TTAR)	TTAR-11.0068	임베디드 시스템에서의 인공지능 서비스를 위한 신경망 구조 표현 모델의 요구사항(기술보고서)	2018-11-07
40	정보통신기술보고서 (TTAR)	TTAR-10.0057	지능정보 기술 동향 - 제1부 : 인공지능 기술 (기술보고서)	2016-11-03

출처: TTA 표준검색 웹페이지(2025.11.17. ‘인공지능’ 키워드로 검색)[662]

662) TTA 표준검색 웹페이지(2025.11.17. ‘인공지능’ 키워드로 검색)
https://committee.tta.or.kr/data/standard_list_N.jsp

제3절 연구기관

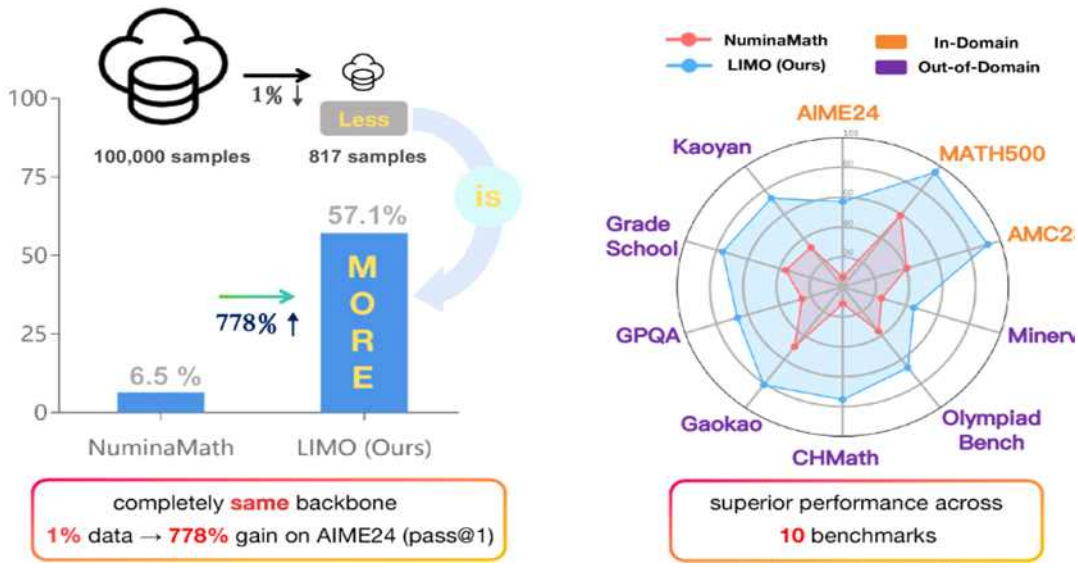
1. 2025년 국외 연구기관 동향

1) 대학

(1) 상하이교통大 연구진, 소량의 고품질 데이터를 활용한 추론 AI 모델 개발[663]

중국 상하이교통대학교 연구진은 2025년 2월, 소량의 고품질 학습 데이터만으로도 추론 특화 AI 모델을 효과적으로 훈련할 수 있다는 연구 결과를 발표하였다. 연구진이 제안한 LIMO(Less is More for Reasoning) 방식은 방대한 데이터로 미세조정을 해야 높은 추론 성능을 얻을 수 있다는 기존 통념을 뒤집는 것으로, 단 817개의 엄선된 학습 예제만으로도 높은 성능을 구현하였다. LIMO 방식으로 훈련된 모델은 AIME2024에서 57.1%, MATH500에서 94.8%의 정확도를 기록하며, 10만 개 이상의 데이터로 훈련된 OpenAI o1-preview, Qwen2.5-32B-Instruct 등 기존 모델들을 능가하였다.

[그림 6-9] 소규모 고품질 학습 예제 기반 LIMO의 성능 향상 효과



출처: Ye Y, Huang 등[664]

663) e Y, Huang Z, Xiao Y, Chern E, Xia S, Liu P. Limo: Less is more for reasoning. arXiv preprint arXiv:2502.03387. 2025.2.

664) Ye Y, Huang Z, Xiao Y, Chern E, Xia S, Liu P. Limo: Less is more for reasoning. arXiv preprint arXiv:2502.03387. 2025.2.

연구진은 대규모 언어모델이 사전학습 단계에서 이미 상당한 수준의 수학·추론 지식을 내재하고 있다는 점에 주목하였다. 즉, 고품질의 정교한 추론 예제 소량을 제공하는 것만으로도 모델 내부의 잠재된 추론 능력을 활성화할 수 있다는 가설을 세우고 이를 실험을 통해 입증하였다. 특히 동일한 모델에 대해 고품질 예제로 학습한 경우와 저품질 예제로 학습한 경우 비교한 결과, 두 모델 간 성능 차이는 AIME2024에서 약 15%p, MATH500에서 약 12%p에 달하는 등 고품질 데이터의 효과가 매우 크게 나타났다. 이 연구는 향후 추론 특화 AI의 학습 방식이 '대량 데이터 의존'에서 '정교한 소량 데이터 기반'으로 전환될 가능성을 시사한다.

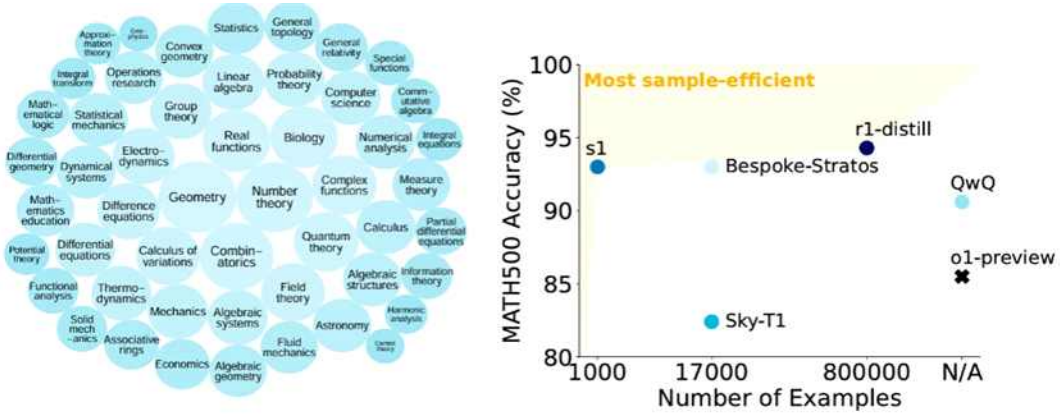
(2) 스탠포드大와 워싱턴大 연구진, 저비용으로 고성능 추론 AI 모델 개발[665]

미국 스탠포드대학교와 워싱턴대학교 연구진은 2025년 2월, 구글의 'Gemini 2.0 Flash Thinking' 실험 모델을 증류(distillation)해 경량·저비용 추론 AI 모델 's1(Simple test-time scaling)'을 개발했다고 발표하였다. 연구진은 엔비디아 H100 GPU 16개를 활용해 단 26분, 비용 기준 50달러 미만으로 모델을 학습시켰으며, 그 결과 s1은 수학 및 코딩 분야의 주요 벤치마크에서 OpenAI o1, DeepSeek R1 등 최신 고성능 추론 모델에 근접한 성능을 보였다.

s1 모델 개발은 구글 Gemini 2.0 Flash Thinking 모델에서 추출한 데이터를 기반으로 이루어졌다. 연구진은 원데이터를 정제한 뒤 △난이도 △다양성 △품질 기준으로 엄격히 선별한 1,000개 고품질 학습 샘플을 구성하고, 이를 이용해 Alibaba의 Qwen2.5-32B-Instruct 모델을 지도 미세조정(SFT) 방식으로 훈련하였다. 이러한 접근은 소량의 고품질 데이터만으로도 고도의 추론 기능을 활성화할 수 있음을 입증한 사례로 평가된다.

665) Ye Y, Huang Z, Xiao Y, Chern E, Xia S, Liu P. Limo: Less is more for reasoning. arXiv preprint arXiv:2502.03387. 2025.2.

[그림 6-10] S1의 샘플 효율성 비교



출처:Yixin Ye 외[666]

또한 연구진은 ‘테스트-타임 스케일링(test-time scaling)’ 개념을 활용해 추론 성능을 더욱 강화하였다. 이는 모델이 추론 과정에서 더 많은 시간·연산을 활용할수록 성능이 향상되는 특성을 이용한 것으로, s1 모델에 ‘다시 생각하라’고 반복적으로 지시해 추론 과정을 연장할 경우 모델이 스스로 오류를 수정하는 현상이 관찰되었다. 예를 들어, “raspberry 단어에 r이 몇 개인가” 라는 질문에서 처음에는 2개라고 응답했으나, 추가 사고 시간을 부여하자 정답인 3개로 재수정하는 등 자율적 재검증 능력을 보여주었다. 이 연구는 경량화·저비용 모델에서도 고성능 추론 능력을 구현할 수 있음을 입증한 대표적 사례로서, 향후 효율적 AI 개발 방향에 중요한 시사점을 제공하는 것으로 평가된다.

(3) 중국 칭화大 연구진, 외부 데이터 없이 스스로 추론하는 AI 학습 기법 개발[667]

중국 칭화대학교, 베이징일반AI연구소(BIGAI), 미국 펜실베이니아주립대학교 공동 연구진은 2025년, 대규모언어모델(LLM)이 외부의 별도 데이터셋 없이 스스로 학습 과제를 생성하고 해결하는 과정을 통해 추론 능력을 고도화하는 새로운 학습 패러다임 ‘Absolute Zero’를 제안하였다. 기존의 RLVR(Reinforcement Learning with Verifiable Rewards) 계열 기법이 고품질의 인간 설계 질문–답변 데이터에 크게 의존해 확장성과 비용 측면에서 한계를 노출해 온 데 반해, 연구진은 하나의 LLM이 문제 제안자(Proposer)와 문제 해결자(Solver) 역할을 모두 수행하는 AZR(Absolute Zero Reasoner) 시스템을 설계하였다. AZR는 코딩과 관련된 다양한 추론

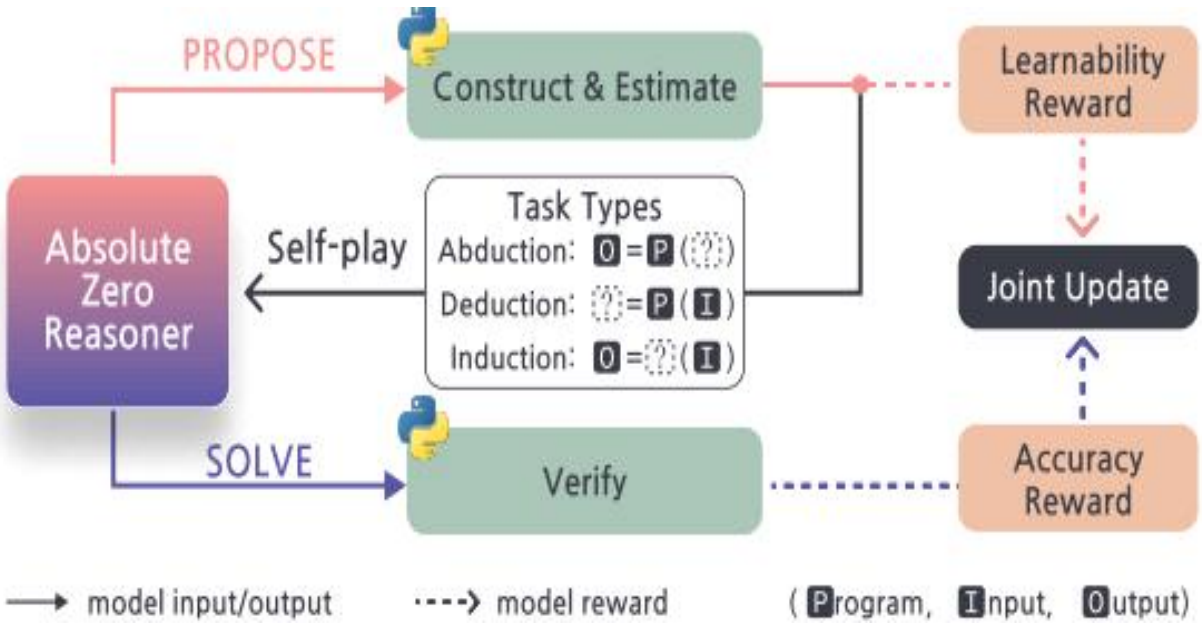
666) Ye Y, Huang Z, Xiao Y, Chern E, Xia S, Liu P. Limo: Less is more for reasoning. arXiv preprint arXiv:2502.03387. 2025.2.

667) Zhao A, Wu Y, Yue Y, Wu T, Xu Q, Lin M, Wang S, Wu Q, Zheng Z, Huang G. Absolute zero: Reinforced self-play reasoning with zero data, 2025. URL https://arxiv. org/abs/2505.03335.

과제를 스스로 생성한 뒤, 코드 실행기를 활용해 정답을 검증하고, 성공률에 기반한 보상을 설계함으로써 모델이 적절한 난이도의 문제를 점진적으로 만들어 풀어가는 자가학습 구조를 구현한다.

성능 검증 결과, AZR는 코딩 및 수학 추론 과제에서 기존 기준 모델을 상회하는 성능을 달성하며, 제안된 패러다임의 효과를 입증하였다. 특히 AZR-Coder-7B 모델은 수학 및 추론 관련 평균 점수에서 기존 최고 성능 모델인 ORZ(Open Reasoner Zero)를 1.8%p 상회하여 7B 규모 모델 중 최고 성능을 기록하였고, 주로 학습을 진행한 코딩 영역뿐만 아니라 별도로 학습하지 않은 수학 영역에서도 성능이 15.2%p 향상되는 등 추론 능력 향상이 다른 영역으로 전이되는 범용성을 보여주었다. 또한 AZR를 적용한 3B, 7B, 14B 모델에서 성능이 각각 5.7%p, 10.2%p, 13.2%p 향상되는 양상이 관찰되어, 모델 규모가 클수록 성능 개선 폭도 커지는 경향이 확인되었다. 이는 고비용 인력 구축 데이터에 대한 의존도를 줄이면서도 고성능 추론 모델을 개발할 수 있는 새로운 학습 전략으로 주목받고 있다.

[그림 6-11] AZR 훈련 기법 개요



출처: Zhao A 외[668]

668) Zhao A, Wu Y, Yue Y, Wu T, Xu Q, Lin M, Wang S, Wu Q, Zheng Z, Huang G. Absolute zero: Reinforced self-play reasoning with zero data, 2025. URL https://arxiv. org/abs/2505.03335.

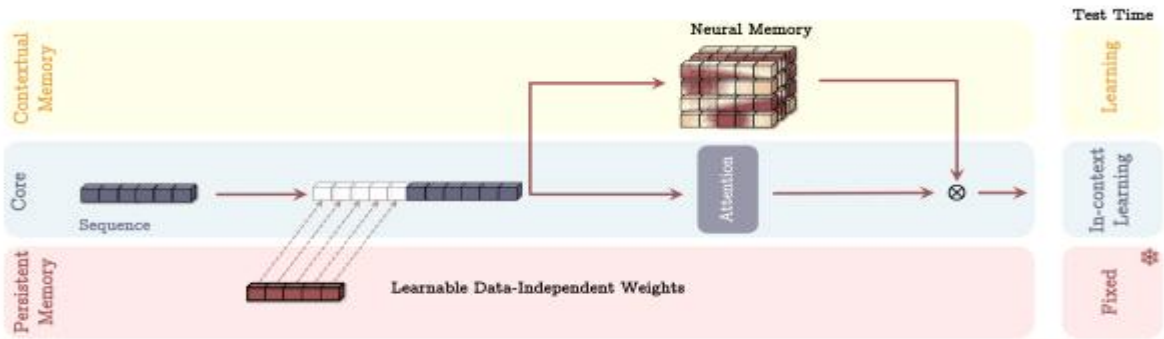
2) 연구소

(1) 구글 연구진, 효율적 메모리 사용을 위한 ‘타이탄’ 아키텍처 개발[669]

구글 연구진은 기존 대규모언어모델(LLM)이 가진 구조적 한계를 극복하기 위해 새로운 신경망 아키텍처인 ‘타이탄(Titans)’을 개발하고, 2024년 12월 31일 관련 연구 결과를 공개하였다(Titans: Learning to Memorize at Test Time). 구글 연구진은 지난 10여 년 간 AI 분야에서 지속적으로 제기된 순환신경망(Recurrent Models)과 어텐션(Attention) 구조의 한계를 보완하기 위해, 두 구조의 장점을 결합한 새로운 신경망 아키텍처인 타이탄(Titans)’을 제안하였다. 기존 순환 모델은 정보를 고정 크기의 은닉 상태(hidden state)에 압축하여 장기 정보를 저장하는 데 한계가 있었고, 어텐션 기반 트랜스포머는 넓은 문맥을 정교하게 모델링할 수 있지만 계산 비용이 문맥 길이에 따라 제곱으로 증가해 실질적으로 긴 문맥을 처리할 수 없다는 구조적 제약을 갖고 있었다.

이를 해결하기 위해 연구진은 신규 신경 기반 장기기억 모듈(neural long-term memory module)을 도입하였다. 이 모듈은 과거 데이터를 장기적으로 저장하는 기능을 학습하며, 어텐션이 현재 문맥의 단기 의존성을 정확하게 처리하는 단기기억(short-term memory) 역할을 담당하도록 설계되었다. 신경 기억 모듈은 병렬화가 가능해 학습 속도가 빠르며, 추론 과정에서도 높은 효율성을 유지할 수 있다는 점이 특징이다.

[그림 6-12] 게이트로서의 메모리(MAG) 아키텍처



출처: Behrouz A 등[670]

669) Behrouz A, Zhong P, Mirrokni V. Titans: Learning to memorize at test time. arXiv preprint arXiv:2501.00663. 2024.12.

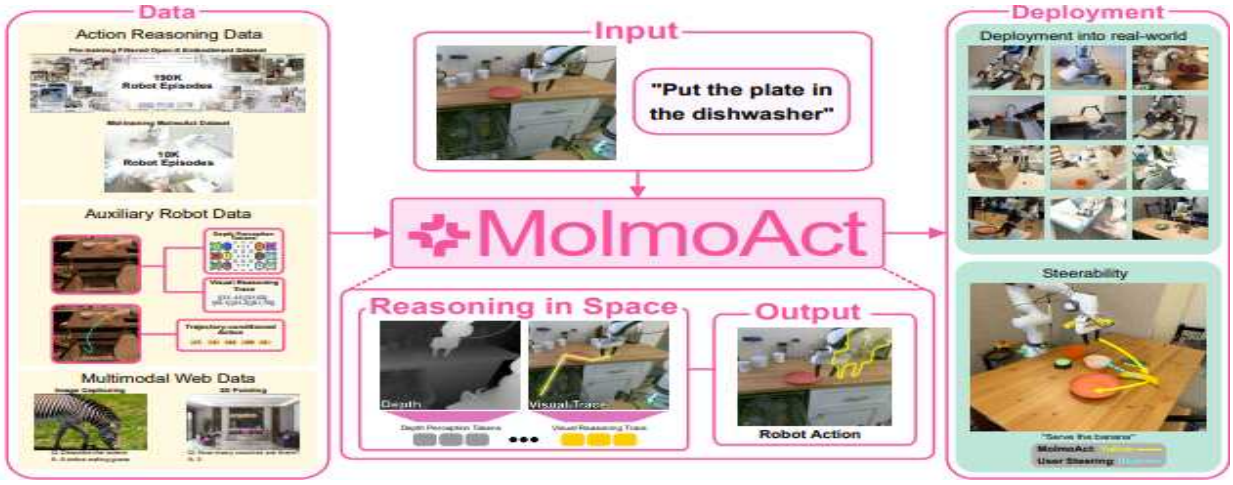
670) Behrouz A, Zhong P, Mirrokni V. Titans: Learning to memorize at test time. arXiv preprint arXiv:2501.00663. 2024.12.

이러한 메모리 구조를 바탕으로 연구진은 타이탄 아키텍처 세 가지 변형 모델을 제시하였으며, 언어모델링, 상식추론, 유전체 분석, 시계열 분석 등 다양한 영역에서 기존 트랜스포머와 최신 선형 순환 모델 대비 우수한 성능을 확인하였다. 특히, 200만 토큰(2M) 이상의 초장기 문맥에서도 높은 정확도를 유지하며, ‘needle-in-a-haystack’과 같은 장기기억 기반 검색·추론 과제에서 기존 모델보다 월등한 성능을 기록하였다.

(2) 앨런 AI 연구소, 3차원 공간에서 추론하는 ‘몰모액트’ 개발[671]

앨런 AI 연구소(AI2)는 2025년 8월 12일, 3차원 공간에서의 행동을 추론할 수 있는 동작 추론 모델(Action Reasoning Model, ARM) ‘몰모액트(MolmoAct)’를 허깅페이스에 오픈소스로 공개하였다. 몰모액트는 AI2의 비전-언어-행동(VLA) 모델인 ‘몰모(Molmo)’를 기반으로, 사람이 공간을 이용해 사고하듯 3차원 환경에서 물체 간 거리·위치 관계를 이해하고 행동을 계획·수행할 수 있도록 설계되었다. 이를 통해 텍스트 중심 LLM이 갖는 공간 이해 한계를 보완하고, 실제 로봇이 물리 공간에서 지시를 따르는 능력을 크게 강화하는 것을 목표로 한다.

[그림 6-13] ‘몰모액트’의 개요



출처: Allen AI Institute[672]

671) Allen AI Institute, MolmoAct: An Action Reasoning Model that reasons in 3D space, 2025.8.12.
672) Allen AI Institute, MolmoAct: An Action Reasoning Model that reasons in 3D space, 2025.8.12.

몰모액트는 깊이 정보와 기하 구조를 인코딩하는 1단계(환경 이해), 목표에 도달하기 위한 경로를 시각적으로 계획하는 2단계(경로 계획), 계획된 경로를 실제 로봇 팔·그리퍼 등 하드웨어 동작으로 변환하는 3단계(동작 실행)로 구성된 자기회귀 추론 사슬을 통해 로봇의 행동을 생성한다. 첫 번째 모델인 '몰모액트-7B'는 2,630만 개 샘플을 256개의 NVIDIA H100 GPU로 하루 만에 학습해 기존 VLA 모델보다 훨씬 효율적인 학습 비용을 달성했으며, 학습에 포함되지 않은 새로운 유형의 작업에서도 OOD(Out-of-Distribution) 작업 성공률 72.1%를 기록해, 구글 RT-2-X(64.3%), 마이크로소프트 Magma(62.6%)를 상회하는 성능을 보인 것으로 보고되고 있다.

3) 학계·기업

(1) AMD와 존스 홉킨스大 연구진, 과학 연구를 자율 수행하는 에이전트 개발[673]

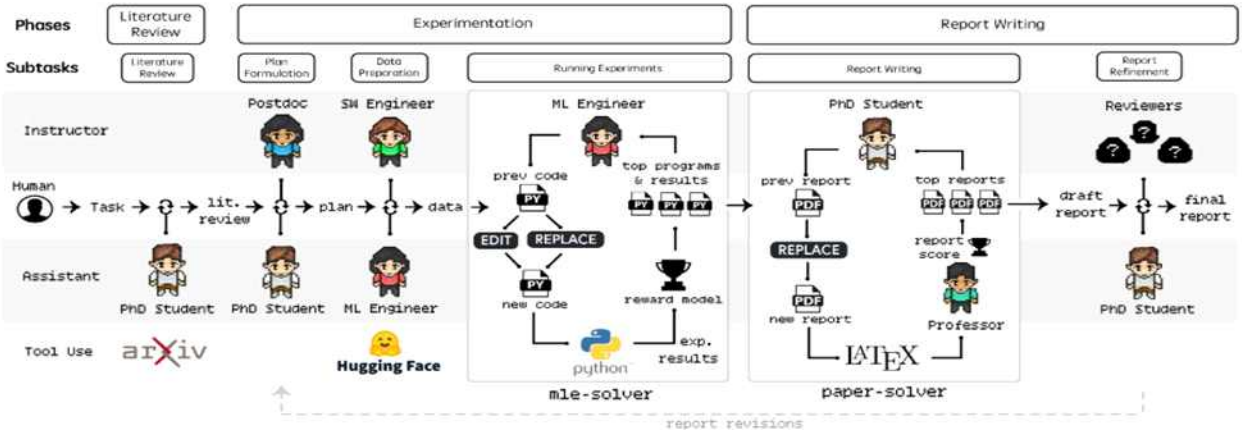
AMD와 존스홉킨스대학교 연구진은 2025년 1월 8일, 대규모언어모델(LLM)을 활용하여 과학 연구의 전 과정을 자동화·지원할 수 있는 자율 연구 에이전트 프레임워크 'Agent Laboratory'를 공개하고 관련 연구 결과를 발표하였다.

Agent Laboratory는 △문헌 검토, △실험 수행, △보고서 작성의 세 단계 모듈로 구성되며, 각 단계에서 LLM 기반 전문 에이전트가 상호 협력하는 구조를 갖는다. 이 과정에서 시스템은 관련 연구 논문 수집 및 분석, 실험 계획 수립, 데이터 전처리, 자동화된 실험 실행, 최종 보고서 작성까지 일관된 연구 흐름을 구현한다.

연구팀은 OpenAI의 o1-preview, o1-mini, GPT-4o 등 세 가지 모델을 적용해 Agent Laboratory 성능을 평가하였다. 그 결과, 연구 결과물의 품질과 보고서 작성 성능에서는 o1-preview와 o1-mini가 GPT-4o보다 우수한 성능을 보였으나, 시간 및 비용 효율성 측면에서는 GPT-4o가 가장 높은 효율을 나타낸 것으로 보고되었다.

673) Github, Agent Laboratory: Using LLM Agents as Research Assistants, 2025.1.8.

[그림 6-14] Agent Laboratory의 개요



출처: Github

4) 대학, 연구소, 기업 등의 트랜스포머 아키텍쳐 연구

2025년을 대학, 연구소, 글로벌 IT 기업을 중심으로 한 Transformer 아키텍처 개선 연구는 연산 효율성 제고, 컨텍스트 창 확장, 대체·하이브리드 아키텍처 개발, 그리고 모델 신뢰성 강화라는 방향으로 전개되고 있다. 기존 Transformer 구조의 성능 한계를 보완하고 대규모·장기 문맥 처리 요구에 대응하기 위한 구조적·시스템적 접근이 병행되고 있는 것이 특징이다.

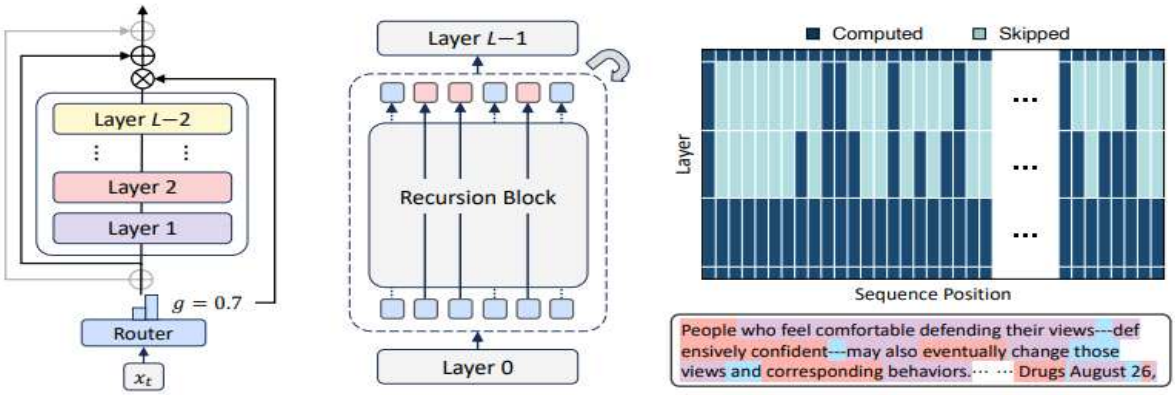
(1) 연산 효율성 및 속도 개선[674]

연산 효율성과 처리 속도 개선을 위한 연구가 지속적으로 고도화되고 있다. Transformer는 어텐션 연산과 대규모 파라미터로 인해 높은 계산 비용과 메모리 사용량이 발생하는 구조적 한계를 지닌다. 이를 완화하기 위해 Grouped-Query Attention(GQA)과 같은 기법이 제안되었으며, 이는 쿼리 벡터를 그룹 단위로 공유함으로써 긴 문맥 처리 시 메모리 사용량을 줄이고 추론 효율을 향상시키는 데 기여하고 있다. 또한 구글 딥마인드가 제안한 MoR(Mixture of Recursions)는 토큰별로 필요한 연산 깊이를 동적으로 조절하는 구조로, 의미적으로 단순한 토큰은 빠르게 처리하고 복잡한 토큰에만 연산을 집중함으로써 추론 속도를 크게 개선하는 접근으로 평가된다. 이와 함께 NPU, 전용 AI 가속기 등 하드웨어 환경에

674) Hugging Face, What changed in the Transformer architecture, 2025.3.8. (2025.11.17. 방문)
https://huggingface.co/blog/rishiraj/what-changed-in-the-transformer-architecture

최적화된 경량화 알고리즘과 희소 어텐션(Sparsity Attention)을 적용하려는 연구도 병행되고 있다.

[그림 6-15] MoR(Mixture-of-Recursions) 개요



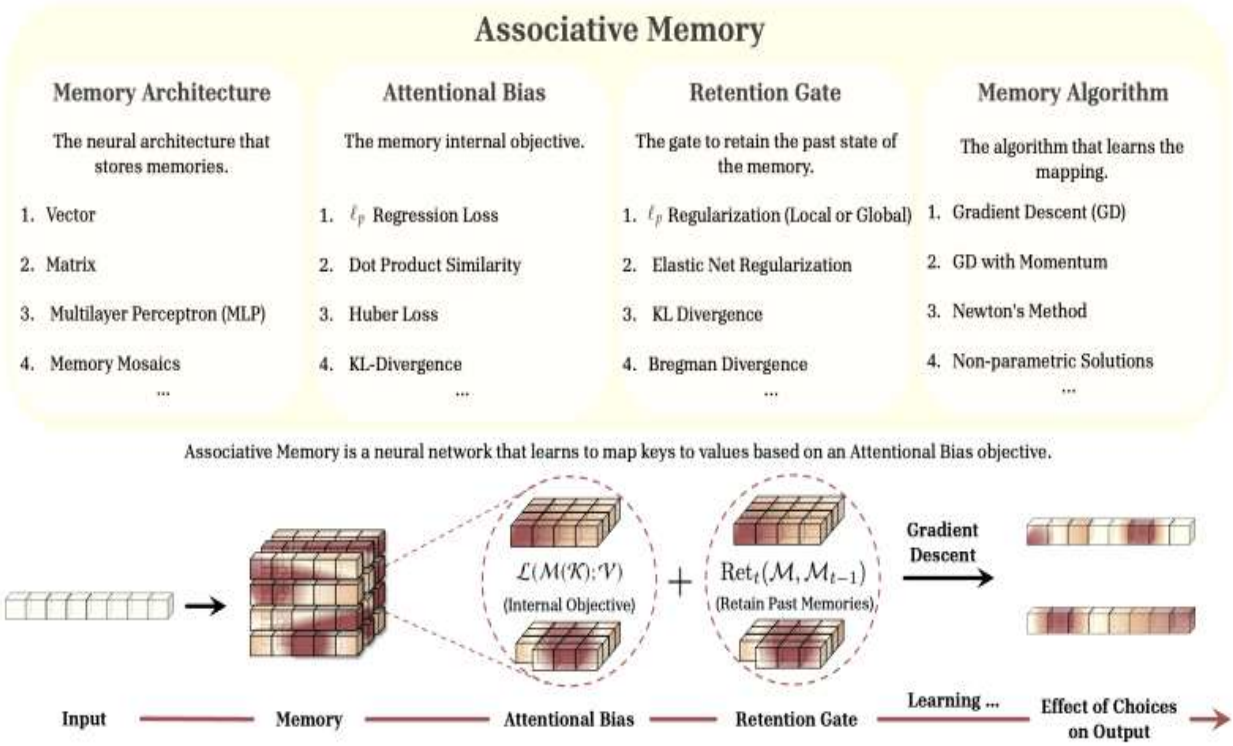
출처: arxiv.org/pdf/2507.10524

(2) 컨텍스트 창(Context Window)의 무한 확장[675]

컨텍스트 창을 획기적으로 확장하기 위한 장기 기억 메커니즘 연구가 중요한 흐름으로 부상하였다. 기존 Transformer는 컨텍스트 길이가 제한되어 장문 문서나 초장기 시계열 데이터 처리에 한계가 있었다. 이를 극복하기 위해 Infini-attention과 같은 접근법이 제안되었으며, 이는 스트리밍 방식의 어텐션과 압축된 메모리를 결합해 이론적으로 무한에 가까운 컨텍스트 처리를 가능하게 한다. 또한 구글의 Titans(2024.12), MIRAS(2025.4)와 같은 연구는 Transformer 구조에 장기 기억 모듈을 추가함으로써 방대한 문서 분석, 유전체 분석 등 초장기 문맥을 요구하는 영역에서의 활용 가능성을 확대하고 있다. 구글 연구진은 기존 LLM의 한계를 극복할 수 있는 새로운 신경망 아키텍처 '타이탄(Titans)'을 개발하고 2024년 12월 31일 관련 논문[676]을 공개했다.

675) Google Research, Titans + MIRAS: Helping AI have long-term memory, 2025.12.4
https://research.google/blog/titans-miras-helping-ai-have-long-term-memory
676) Google Research, Titans: Learning to Memorize at Test Time, 2024.12.31
https://arxiv.org/abs/2501.00663

[그림 6-16] MIRAS 아키텍쳐 개요



출처: Google Research(2025), https://arxiv.org/pdf/2504.13173

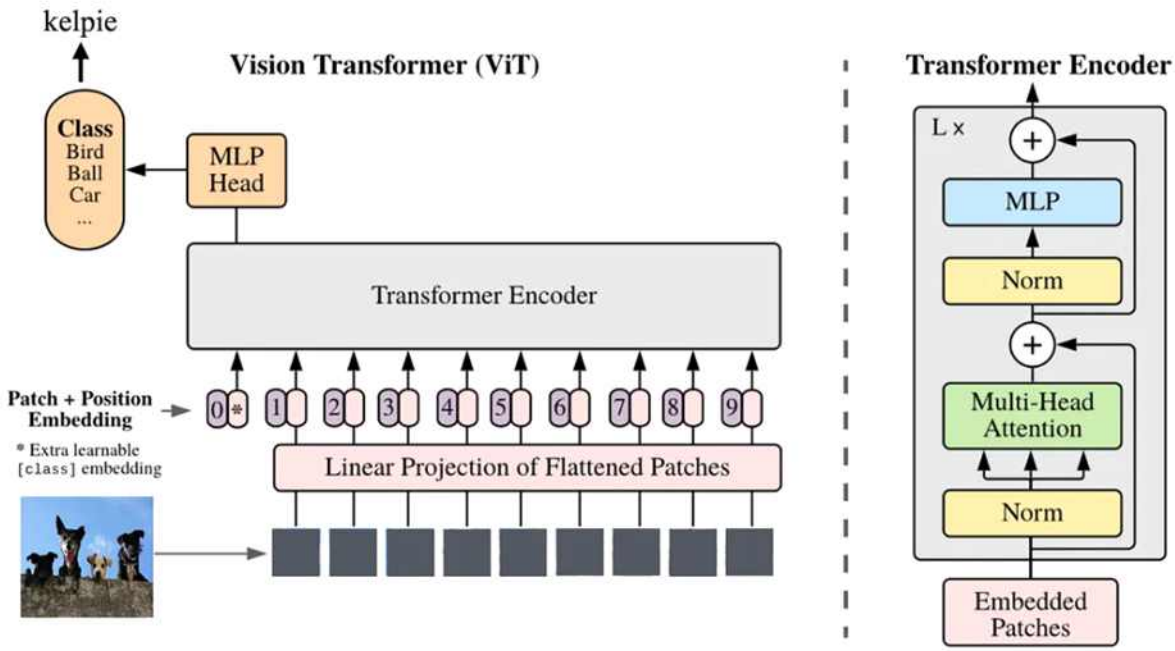
(3) 멀티모달 및 범용성 확대[677]

멀티모달 처리 능력과 범용성을 강화하기 위한 연구가 본격화되고 있다. 텍스트 중심의 Transformer를 넘어 이미지, 비디오, 시계열 데이터 등 다양한 입력을 통합적으로 처리하려는 시도가 주류로 자리 잡고 있다. 비전 트랜스포머(Vision Transforme:(ViT)[678]의 변형 모델과 VLM(Vision-to-Language Model) 계열 연구는 시각 정보와 언어 정보를 통합적으로 이해하는 방향으로 발전하고 있으며, 이는 자율주행, 의료 영상 분석, 로보틱스 등 다양한 산업 분야에서 활용 가능성을 확대하고 있다. 특히 NaViT(Normalized Vision Transformer)는 입력 이미지의 크기와 종횡비에 제약을 두지 않고 원본 해상도를 유지한 채 처리할 수 있는 구조로, 실제 응용 환경에서의 유연성을 크게 향상시킨 사례로 평가된다.

677) NetAppInstaclustr, Top 10 open source LLMs for 2025, (2025.11.17. 방문) https://www.instaclustr.com/education/open-source-ai/top-10-open-source-llms-for-2025/

678) Shaibal Saha, Lanyu Xu, Vision Transformers on the Edge: A Comprehensive Survey of Model Compression and Acceleration Strategies, 2025.2.26 https://arxiv.org/abs/2503.02891

[그림 6-17] 비전 트랜스포머 아키텍쳐(예)



출처: Medium.com[679]

(4) 트랜스포머 대체 및 하이브리드 아키텍처[680]

Transformer의 구조적 한계를 근본적으로 극복하기 위한 대체 및 하이브리드 아키텍처 연구가 활발히 진행되고 있다. Transformer의 핵심 제약 중 하나인 어텐션 연산의 이차 복잡도를 완화하기 위해 상태 공간 모델(SSM)과의 결합이 시도되고 있으며, Mamba-2와 같은 모델은 Transformer 수준의 성능을 유지하면서도 선형에 가까운 확장성을 제공하는 대안으로 주목받고 있다. 이와 함께 선형 어텐션(Linear Attention) 계열 연구도 진행되고 있으며, 기존의 제곱 복잡도를 선형 복잡도로 낮춘 구조를 통해 대규모 시퀀스 처리의 효율성을 개선하려는 접근이 이어지고 있다.

679) Mao Jia, Inside a Vision Transformer (ViT): How Image is Classified Step by Step, 2025.7.3. https://blog.gopenai.com/inside-a-vision-transformer-vit-how-image-is-classified-step-by-step-0af45d5acbca
680) Albert Gu, Tri Dao, Mamba: Linear-Time Sequence Modeling with Selective State Spaces, 2023.12.1. https://arxiv.org/abs/2312.00752

(5) 해석 가능성 및 신뢰성 연구[681]

AI 모델의 해석 가능성과 신뢰성을 강화하기 위한 연구도 중요한 연구 축으로 자리 잡고 있다. 대규모 언어모델이 블랙박스화됨에 따라, 모델 내부의 판단 근거를 이해하고 설명하려는 요구가 커지고 있다. 이에 따라 희소 회로(Sparse Circuits) 기술과 같이 모델 내부의 특정 신경망 경로와 활성 패턴을 추적·분석하여 의사결정 과정을 해석하려는 설명 가능한 AI 연구가 2025년 주요 연구 트렌드로 부상하고 있다. 희소 회로 연구는 신경망 내에서 특정 개념(예: 문법, 논리, 감정)을 담당하는 뉴런들의 연결망인 '회로'를 식별한다. Anthropic의 'Mapping the Mind of a Large Language Model'[682] 연구는 이러한 회로 분석의 대표적인 사례로 꼽힌다. 대규모 언어모델의 내부 작동 원리를 정밀하게 분석하기 위한 '현미경' 접근법을 바탕으로, 2025년 3월 Anthropic은 Claude 모델을 대상으로 두 편의 연구를 통해 모델 내부에서 해석 가능한 개념(특징)을 식별하고 이들이 계산 회로 형태로 연결되는 과정을 분석했으며, 입력 단어가 출력 단어로 변환되는 내부 경로의 일부를 실증적으로 제시하였다. 또한 Claude 3.5 Haiku를 대상으로 핵심 동작을 대표하는 10가지 과제에 대한 심층 분석을 수행해 프롬프트 응답 과정에서의 주요 내부 메커니즘과 활성 패턴을 체계적으로 제시함으로써, 대규모 언어모델의 추론이 구조화된 내부 표현과 연쇄적 계산에 기반함을 보여주었다.[683] 이러한 연구는 향후 공공·의료·국방 등 고신뢰성이 요구되는 분야에서 Transformer 기반 AI의 활용을 확대하는 데 중요한 기반이 될 전망이다.

681) EIMT, The Future of Generative AI: Trends to Watch in 2025 and Beyond, 2025.10.14. https://www.eimt.edu.eu/the-future-of-generative-ai-trends-to-watch-in-2025-and-beyond
682) Anthropic, Tracing the thoughts of a large language mode, 2025.3.27. https://www.anthropic.com/research/tracing-thoughts-language-model
683) Anthropic, Tracing the thoughts of a large language mode, 2025.3.27. https://www.anthropic.com/research/tracing-thoughts-language-model

2. 2025년 국내 연구기관 동향

1) 한국전자통신연구원(ETRI), AI 연구 및 표준 개발

(1) 상하이교통大 연구진, 소량의 고품질 데이터를 활용한 추론 AI 모델 개발[684]

한국전자통신연구원(ETRI)은 AI 분야에서 다양한 연구와 개발을 주도하고 있다. 한국전자통신연구원(ETRI)은 인공지능·고성능 컴퓨팅 연구의 핵심 조직으로 인공지능컴퓨팅연구소[685]를 운영하며, AI 기술혁신을 위한 기반 기술 개발을 주도하고 있다. 해당 연구소는 ETRI의 비전인 ‘디지털 혁신으로 행복한 미래세상을 만드는 기술 선구자’를 구현하기 위해 초고성능 하드웨어·소프트웨어 전략기술, AI·양자 기반의 차세대 컴퓨팅 아키텍처, 시스템 소프트웨어 등 컴퓨팅 전 주기 핵심기술을 연구한다. 특히 AI 슈퍼컴퓨터, 지능형 반도체, 양자컴퓨팅, 사이버보안 기술 등을 개발해 국가 차원의 AI 인프라 구축과 산업 현장 적용을 지원하고 있으며, 국내외 대학·기업과의 협력을 기반으로 개방형 기술혁신 생태계 조성에 기여하고 있다. 이를 통해 ETRI는 AI 연구의 실질적 성과가 사회문제 해결과 산업 경쟁력 강화로 이어지도록 ‘파괴적 혁신 컴퓨팅’ 기반을 마련하는 데 중점을 두고 있다.

또한 ETRI는 사내 인공지능창의연구소를 통해 인공지능 기술의 국가 경쟁력 강화를 위한 도전적·선도형 연구를 수행하고 있다. 동 연구소는 소재·부품부터 자율차·로봇과 같은 완성품에 이르는 폭넓은 분야를 아우르며, 단순 모방이 아닌 차세대 기술로 이어질 수 있는 창의·도전형 연구를 지향한다. 주요 연구 분야로는 사람처럼 지속적으로 보고·듣고·학습하는 차세대 인공지능 핵심기술, 자율·안전 운행이 가능한 지능형 모빌리티 및 로봇 기술, 소부장(소재·부품·장비) 경쟁력 확보를 위한 스마트 소재·차세대 반도체·지능형 부품 기술 등이 있다. 아울러 테라헤르츠, 사이버 브레인, 브레인 링크, 저탄소 집적기술 등 고위험·고난도 창의도전 연구를 통해 미래 유망 원천기술의 가능성을 탐색하고 있으며, 향후 기술이전, API·데이터 공개 등을 통해 국내 산업 생태계의 선순환을 촉진하고 글로벌 연구기관과의 협력을 기반으로 AI·지능형 모빌리티·로봇·소재부품 분야에서 차세대 기술 개발을 강화하고자 한다.

---

684) e Y, Huang Z, Xiao Y, Chern E, Xia S, Liu P. Limo: Less is more for reasoning. arXiv preprint arXiv:2502.03387. 2025.2.

685) ETRI, 인공지능안전연구소 (2025.11.17. 방문) https://www.etri.re.kr/kor/sub6/sub6_0101.etri

아울러 ETRI는 2024년 11월 인공지능안전연구소[686][687]를 신설하여 AI안전 성과 신뢰성 확보를 전담하는 조직 체계를 구축하였다. 인공지능안전연구소는 AI안전정책 및 대외협력실, AI안전 평가실, AI안전 연구실 등 3개 실로 구성되어 있으며, △AI 시스템의 안전성·신뢰성 평가, △AI 안전 관련 정책·제도 및 국제 동향 연구, △AI안전 기술 개발, △주요국 인공지능안전연구소와의 협력 네트워크 구축 등을 종합적으로 수행한다. 이를 통해 안전한 AI 개발·활용 문화를 확산하고, 국내 AI 기업이 글로벌 시장에서 요구되는 안전·신뢰 기준을 충족할 수 있도록 지원함으로써, 국가 차원의 AI 경쟁력 제고와 국제 협력 기반 강화에 기여하고 있다.

[그림 6-18] 인공지능안전연구소 조직도



출처: 인공지능안전연구소 홈페이지. 조직안내[688]

ETRI는 2025년 11월, 인공지능을 무선 전송·네트워크 제어·엣지 컴퓨팅 전반에 적용한 6G 지능형 무선 액세스(AI-RAN) 기술 개발을 완료하고, 5G 대비 최대 10배 수준의 전송 효율 향상을 기대할 수 있는 AI-네이티브 이동통신 기반을 제시하였다. AI-RAN은 채널 상태 분석을 기반으로 한 빔포밍·전력 제어, 기지국 간 협력 및 간섭 관리, 엣지 단 트래픽 예측·분산, 지연 최소화 등을 통해 초고밀도 환경에서도 안정적인 통신 품질을 유지할 수 있도록 설계되었으며, 특히 AI가 직접 무선 신호를 복원·보정하는 '뉴럴 리시버' 기술을

686) 인공지능안전연구소 (2025.11.17. 방문) https://www.aisi.re.kr/kor
687) ETRI, 인공지능안전연구소 (2025.11.17. 방문) https://www.etri.re.kr/kor/sub6/sub6_0101.etri
688) https://www.aisi.re.kr/kor/contents/20

통해 밀리미터파 환경에서 데이터 복원 정확도 약 18% 향상, 채널 예측 정확도 약 15% 향상, 데이터 손실률 30% 감소 등의 성과를 거두어 AI 기반 수신기가 기존 방식의 한계를 실증적으로 입증하였다. ETRI는 본 기술을 스스로 학습·진화하는 '셀프 이볼빙 RAN'으로 고도화하고, 과학기술정보통신부·IITP의 '6G 핵심기술개발사업'(서울대, 통신 3사, 주요 대학·기업 등 공동 참여)을 통해 우리나라가 3GPP 내 AI/ML 기반 무선 인터페이스 및 이동성 관리 표준화를 주도하는 등 6G AI 네트워크 분야의 국제 경쟁력 확보를 목표로 연구를 지속하고 있다.[689]

2) 한국전자통신연구원(ETRI), AI 연구 및 표준 개발

KISDI는 1985년 2월에 설립된 국책연구기관으로, 국내외 인공지능 기술, 산업 동향 및 정책 이슈를 포괄하여 국내·외 정보화 및 정보통신·방송분야의 정책·제도·산업 등에 관한 각종 정보를 수집·조사·연구한다. 특히 AI 분야의 이슈 페이퍼 및 학술 논문을 'KISDI AI Outlook'을 통해 게재하고 있다. 'KISDI AI Outlook'에 게재된 최근(2024~2025)의 AI 관련 2024년 주요 연구보고서들은 아래의 표와 같다.

689) 전자신문, ETRI, 6G 지능형 무선 액세스 기술 개발. AI 기반 이동통신 기반 마련, 2025.11.17. https://www.etnews.com/20251117000094

<표 6-5> KISDI의 인공지능 분야에서 발행된 주요 연구보고서 목록(2024~2025)

번호	보고서 제목	발행일
1	중국의 조용한 침공: AI, 반도체 붉은 공급망	2024.3.31
2	대량맞춤화를 실현하는 인공지능 기반 인간-로봇 협업	2023.3.31
3	생성 모델과 변분 오토인코더(Variational Autoencoder)	2024.6.28
4	인공지능 공정성 동향 분석	2024.6.28
5	인공지능 시스템 이용자 보호 관련 해외 규율 동향과 시사점	2024.9.30
6	딥페이크 관련 국내외 규제 동향 분석	2024.9.30
7	AI 커버곡을 둘러싼 법률 문제: 음성권의 보호 가능성을 중심으로	2024.12.30
8	글로벌 AI 패권 경쟁과 한국의 대응 전략	2024.12.30
9	인간-휴머노이드 AI간의 언어 형상화와 사용자 경험	2025.03.31
10	AI 기반 소프트웨어 정의 공장을 활용한 제조 혁신	2025.03.31
11	Physical AI 시대에 대응한 On-device AI 반도체 경쟁력 강화 방향	2025.06.30
12	인공지능 기술이 근로관계에 미치는 영향과 노동법적 대응 방안	2025.06.30
13	인공지능과 사이버보안: AI보안 거버넌스 체계 강화 방안을 중심으로	2025.09.30
14	AI 중견국의 전략	2025.09.30
15	글로벌 AI 안전 거버넌스의 다극화와 한국의 전략적 대응 방향	2025.09.30

출처: KISDI AI Outlook 웹페이지[690]

3) 소프트웨어정책연구소(SPRi)

소프트웨어정책연구소는 인공지능, 메타버스 등 새로운 소프트웨어 트랜드를 선도적으로 분석하고 이슈를 발굴하여 정책 대안 제시, 법제도 개선, 통계 생산 등 연구활동을 추진해 오고 있다. 소프트웨어정책연구소는 연구보고서, 이슈리포트, 산업정책동향 등을 발행하며, 최근(2024~2025)년도 AI 분야의 주요 연구보고서 목록은 아래의 표와 같다.

690) KISDI AI Outlook 웹페이지 (2025.11.17. 방문)

<표 6-6> SPRi의 인공지능 분야에서 발행된 주요 연구보고서 목록(2024~2025)

구분	보고서 제목	발행일
1	산업 수요 중심의 국가 AI R&D 전략성 강화를 통한 AI 확산 방안 연구	2024. 04.29
2	공공부문 AI 도입현황 연구	2024. 04.29
3	국외 AI 시장 및 창업지원 정책 동향 조사	2024. 04.29
4	국내 AI 창업기업 비즈니스 현황 분석	2024. 04.29
5	글로벌 AI 신뢰성 정책 동향 분석	2024. 04.29
6	메타버스산업 분류체계 연구	2024. 04.29
7	메타버스 성장통 극복 방안 연구	2024. 04.29
8	2023년 메타버스 시장 및 동향 분석	2024. 04.29
9	2023년 국내외 인공지능 산업 동향 연구	2024. 04.29
10	생성형 AI에 대응한 SW인재양성 정책 방향 연구	2024. 04.29
11	해외 AI안전연구소 추진 현황과 시사점	2024.07.23
12	유럽연합 인공지능법(EU AI Act)의 주요내용 및 시사점	2024.08.06
13	책임 있는 AI를 위한 기업의 노력과 시사점	2024.08.07
14	글로벌 초거대 AI 모델 현황 분석(2020~2023년)	2024.08.19
15	생성형 AI에 의한 소프트웨어 개발자 업무 영향 분석	2024.08.20
16	AI 위험 유형 및 사례 분석	2024.10.31
17	AI인프라에서 AI서비스로 : SW기업의 AI도입 모델과 신서비스 모델의 탐색	2024.12.19
18	AI 안전의 개념과 범위	2024.12.30
19	SW 개발자 채용시장의 변화와 생성형 AI의 영향	2025.01.07
20	미국의 AI안전·신뢰성 정책 추진 현황과 시사점	2025.01.20
21	글로벌 초거대 AI 모델 현황 분석(2024년 조사)	2025.02.14
22	특허지표를 활용한 주요국 AI 기술 경쟁력 분석	2025.04.01
23	AI Index 2025 주요내용 및 시사점	2025.04.15
24	인공지능 기술·산업 생태계 육성방안 연구	2025.04.30
25	AI 신뢰성 및 윤리 제도 연구	2025.04.30
26	AI 안전의 부상과 정책 대응 방안 연구	2025.04.30
27	2024년 공공부문 AI 도입현황 연구	2025.04.30
28	2024년 국내외 인공지능 산업 동향 연구	2025.04.30
29	SW개발자 채용 변화 전망과 생성형 AI	2025.04.30

출처: SPRi 홈페이지[691)]

691) SPRi 홈페이지
https://spri.kr/posts?data_page=1&code=data_all&study_type=&board_type= (2025.11.17. 방문)

4) 한국지능정보사회진흥원(NIA)

한국지능정보사회진흥원(NIA)은 국내외 인공지능 관련, 정책, 법제 연구와 사업 등을 수행하고 있다. 연구보고서 등 발간물은 원내에 인공지능정책실, 인공지능융합본부, 인공지능데이터본부에서 인공지능 관련 연구를 주로 추진한다.

NIA는 매년 국가지능정보화백서를 발행하고 있으며, 정보화정책저널(인공지능 포함) 지능사회 이슈분석 등 다양한 연구보고서를 발간하고 있다. 다음의 최신(2024~2024) 인공지능 관련 정보화정책저널의 목록은 아래의 표와 같다.

<표 6-7> NIA의 AI 관련 연구보고서 사례

구분	보고서 제목	게시일
1	ChatGPT는 우리에게 어떤 우려를 초래하는가?: 유튜브 영상 뉴스 댓글의 CTM(Correlated Topic Modeling) 분석을 중심으로	2024.03.31
2	가상현실 에이전트 외국어 교사를 활용한 외국어 학습의 몰입 융합 효과	2024.03.31
3	챗GPT 등장 이후 인공지능 환각 연구의 문헌 검토: 아카이브(arXiv)의 논문을 중심으로	2024.07.02
4	생성형 AI의 법적 문제와 규제 논의 동향	2024.09.30
5	기업의 인공지능 기술 도입에 영향을 미치는 요인 분석: 국내 기업 데이터를 이용한 실증연구	2024.09.30
6	생성형 AI 파워드 메타버스: 사용자 중심 경험과 사회·기술적 상호작용의 진화와 혁신	2025.03.31
7	인공지능을 활용한 메타버스 서비스에서의 개인정보보호 컴플라이언스 개발에 관한 연구	2025.03.31
8	공공행정에서의 인공지능 활용에 관한 실증 연구: 공공 AI를 위한 정책적 시사점)	2025.06.30
9	AI 활용 정부 서비스의 사례와 유형 분석: 새 정부 AI 정부 서비스 성공을 위한 전략 설정 제언	2025.09.30
10	인공지능 기반 서비스 이용에 있어서 개인의 정보평가 성향의 영향: 프라이버시 염려 및 보호 행동과의 조절된 매개효과 분석	2025.09.30
11	KoSBI 데이터셋을 활용한 LLMs의 사회적 편향성(Social Bias) 분류 성능 평가: GPT-3.5 Turbo와 GPT-4o의 비교를 중심으로	2025.09.30
12	토픽 모델링을 통한 심리상담 챗봇의 개입 유형별 사용자 경험 분석	2025.09.30

출처: NIA 웹페이지[692)]

692) NIA 홈페이지
https://www.nia.or.kr/site/nia_kor/ex/bbs/List.do?cbIdx=82618

제7장 인공지능 교육 및 인력 동향

2025년 전 세계 교육 업계와 AI로 인한 인간 노동, 일자리 및 고용 구조 변화 등에 대한 분석을 통해 다음과 같은 주요 트렌드를 도출하였다.

● AI가 노동시장 구조를 재편하면서 재교육·평생학습 수요 급증

UNCTAD는 AI로 인해 반복·규칙적 업무 중심 일자리는 감소하는 반면, 고숙련·창의·문제해결 중심 직무에서는 업무 재편과 생산성 향상을 통해 새로운 기회가 창출될 수 있다고 분석하였다. AI 노출도가 높은 직무일수록 임금·생산성 향상이 관측되지만, 직무 내용·요구 역량 변화 속도가 빨라지면서 재교육·평생학습 체계가 부족한 국가·계층에서는 고용 불안과 격차 확대 우려가 커지고 있어, 노동·교육 정책의 결합이 중요해지고 있다.

● 국가 단위의 생애주기 AI 교육·대규모 인재 양성 프로그램 확산

중국은 2025년 '디지털 교육 활성화에 관한 의견'을 통해 국가 스마트교육 플랫폼, 교육 특화 AI 모델·데이터셋, AI 기반 교육 혁신 및 국제화, 디지털 인프라·보안 체계 확충 등을 추진하는 등, 교육 시스템 전체를 AI 중심으로 재설계하는 국가 주도형 모델을 채택하였다. 한국은 초·중등·대학·평생교육 전 주기에 걸쳐 AI 기본교육을 강화하고, 17개 시도 AI 교육지원센터, AI 거점대학·연구단, 과학고·영재학교 AI·SW 특화 프로그램, 학·석·박 패스트트랙 등으로 지역·단계별 인재양성 체계를 구축 중이다.

● 학교·대학에서 생성형 AI 활용의 급격한 보편화와 준비 격차

Microsoft의 「AI in Education 2025」에 따르면 전 세계 교육기관의 86%가 이미 생성형 AI를 도입했으며, 교사와 학생 모두 과제 설계·수업 준비·학습 보조 등에 AI를 활용하고 있다. 그러나 상당수 교사는 AI를 수업에 통합할 준비가 되어 있지 않다고 응답해, '현장 활용 속도'와 교육과정·교원 역량·평가체계 등 '교육 시스템 정비' 간 괴리가 나타나고 있다.

● 기업 내 AI 역량 강화를 위한 조직 재편과 CAIO 확산

AWS 글로벌 생성 AI 도입 지수에 따르면 조사 대상 기업의 90%가 이미 생성형 AI 도구를 사용하고 있으며, 45%는 이를 향후 IT 예산의 최우선 투자 분야로 꼽고, 60%는 최고 AI 책임자(CAIO)를 지정하였

다. 그럼에도 인재 부족·비용·환각(hallucination)·데이터 보안·프라이버시 우려 등이 주요 장애 요인으로 나타나, 특히 중소기업·공공·비영리 영역에서 준비 수준 차이가 크다는 점이 지적되고 있다.

● 생성 AI 확산이 경력 초기 근로자에 미치는 불균형 고용효과와 '이중 구조' 인력전략 필요성 증가

스탠퍼드 연구에 따르면 미국 노동시장에서 전반적인 고용은 탄탄한 성장세를 보이지만, AI 노출도가 높은 직업에서 22~25세 젊은 근로자의 고용은 2022년 말~2025년 7월 사이 6% 감소한 반면, 숙련 근로자 고용은 6~9% 증가했다. AI 노출도가 가장 높은 직업군의 22~25세 고용은 노출도가 가장 낮은 직업군 대비 약 12% 감소해 통계적으로 유의미한 차이를 보인다. 이러한 결과가 생성 AI 확산이 임금보다 고용에 먼저, 특히 경력 초기 근로자에게 더 큰 충격을 주고 있으며, 자동화 중심 직업에서는 고용 감소가 나타나는 반면, AI가 인간 노동을 보완하는 직업에서는 고용이 유지된다는 점을 지적한다.

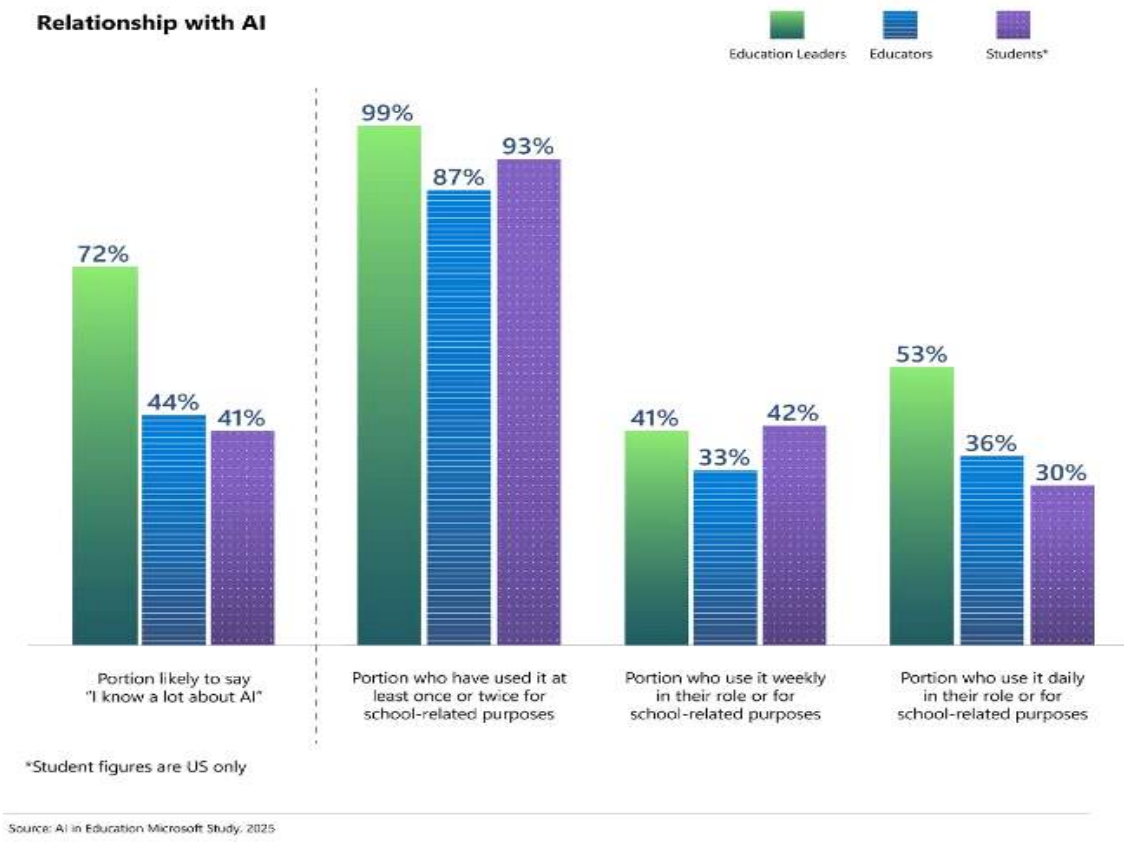
제1절 교육 분야

1. AI 교육 현황 및 추이[693]

1) 교육현장의 AI 도입 확산

Microsoft의 ‘AI in Education Report 2025’ 보고서는 미국, 브라질, 사우디아라비아, 영국, 일본, 호주의 교육자와 교육기관 최고 관리자를 대상으로 실시한 설문조사 결과를 종합해 분석하였다.[694] 현재 교육기관의 생성 AI 도입률은 86%로, 전체 산업 분야 중 가장 높은 수준을 기록했다. 특히 미국의 경우 학교에서 AI를 활용하는 학생이 전년 대비 26%p 늘었고, 교육자의 AI 활용은 21%p가 증가한 수치다. 한편, AI를 전혀 사용하지 않는 학생 비율은 20%p 감소하였다.

[그림 7-1] 교육 부문에서 생성 AI 사용



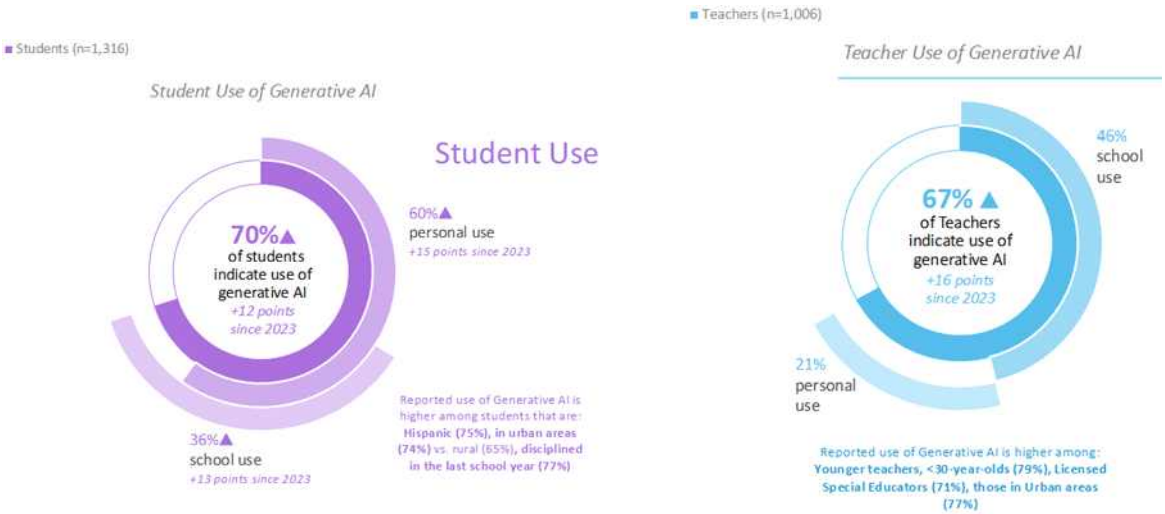
출처: Microsoft, AI in Education Report, 2025.08.20

693) Stanford HAI, AI Index Report 2024
694) Microsoft, AI in Education Report: Insights to support teaching and learning, 2025.08.20.

미국 비영리단체 민주주의기술센터(CDT, Center for Democracy & Technology)가 발표한 설문조사[695]에 따르면 생성 AI를 사용하는 교사는 2023년 51%에서 2024년 67%로, 학생은 2023년 58%에서 2024년에는 70%로 증가하였다. 이 같은 결과는 교육 현장에서 AI 확산이 많이 진행된 것을 보여준다.

2) 교육 현장에서 AI 활용

[그림 7-2] 학생과 교사의 생성 AI 사용



출처: Center for Democracy & Technology, Out of Step: Students, Teachers in Stride with EdTech Threats While Parents Are Left Behind, 2025.01.15.

학생들은 생성 AI를 학교 관련 용도(36%)보다 개인 용도(60%)로 더 많이 사용한다고 응답하였다.(CDT) 학생들의 학업 관련 AI 주 활용 목적은 과제 구상(37%), 정보 요약(33%), 신속한 정보 획득(33%), 피드백 수렴(32%)으로 시간 절약과 학습 성과 증대목적으로 활용되는 것을 알 수 있다.

반면, 교사들은 개인 용도(21%)보다 학교 관련 용도(46%)로 사용하는 비율이 더 높아 업무 효율성을 높이는 데 집중하는 경향을 보인다. MS 보고서에서는 수업 계획이나 자료 구상

695) Center for Democracy & Technology, Out of Step: Students, Teachers in Stride with EdTech Threats While Parents Are Left Behind, 2025.01.15.

(31%), 수업 자료나 과제 제작 또는 업데이트(29%), 학생을 위한 복잡한 주제 단순화(24%), 실시간 성과 데이터와 예측 분석(24%) 등에 주로 AI를 활용한다고 응답하였고, CDT 조사에서는 맞춤형 교육 커리큘럼 개발(25%), 시험 채점(21%), 기타 학교 관련 업무(수업 계획 수립, 이메일 작성, 과제 생성)(16%)에 생성 AI를 사용한다고 응답하였다. 교사들은 반복적인 행정 업무를 자동화하고, 학생들에게 더 나은 교육 콘텐츠를 제공하기 위해 생성 AI를 사용하는 것으로 나타났다.

교육기관 관리자들은 주로 과제 피드백과 개선 안내(36%), 운영 간소화(35%), 접근성 도구 제공(33%), 학생 성장 기회 파악(33%), 학생 지원 및 소통 능력 향상(30%)에 AI를 활용한다고 응답하였다. 교육 행정 및 학생 지원 시스템 개선을 목적으로 AI를 활용하는 것으로 보인다.

앤스로픽은 고등교육기관의 교육자 7만 4천 명이 클로드(Claude)에서 주고받은 익명화 대화를 바탕으로 대학 교육환경에서 교육자들의 AI 도입 현황을 분석한 보고서를 발표하였는데[696], 교육자들에게서 가장 두드러지는 AI 활용 분야는 교육과정 개발(57%)이며, 학술 연구(13%)와 학생 성취도 평가(7%)가 순으로 활용하였다. 또한 앤트로픽 결과에서는 교육자들의 클로드 사용 패턴이 업무의 성격에 따라 '증강(Augmentation)'과 '자동화(Automation)'라는 두 가지 뚜렷한 경향으로 나뉜다는 것을 볼 수 있다.

- 증강 경향이 높은 작업: 창의성과 전문성이 요구되는 작업에서는 AI를 협력적 파트너로 활용 (▲대학 교육 및 강의 자료 작성 (77.4%) ▲연구 보조금 제안서 작성 (70.0%))
- 자동화 경향이 높은 작업: 반복적이고 정형화된 업무에서는 AI에게 작업을 위임 (▲교육기관 재정 관리 (65.0%) ▲학생 기록 관리 및 성적 평가 (48.9%))

696) Anthropic, Anthropic Education Report: How educators use Claude, 2025.08.27.

3) AI 활용 학습 효과

AI 활용은 학습 효과 측면에서도 가시적인 성과를 보이고 있다. 호주 Macquarie 대학교의 실험에서 AI 챗봇을 활용한 학생들은 시험 성적이 평균 10% 향상되었으며[697], 나이지리아의 무작위 통제 실험에서도 Microsoft Copilot을 활용한 영어 학습집단이 0.31표준편차의 성취도 향상을 보였다. 특히 사회경제적 배경이 학습 성과에 미치는 영향은 통계적으로 유의미하지 않아, AI가 교육 형평성 개선에 기여할 가능성이 확인되었다.[698]

또한 Microsoft Research와 Cambridge University Press의 공동 연구에서는, AI 기반 독서보조도구를 활용한 학생들이 전통적인 필기 방식보다 내용 이해도와 몰입도가 높았으며, AI 단독 사용보다 전통적 학습과 병행할 때 효과가 극대화되는 것으로 나타났다. 이는 AI가 교육을 대체하기보다는 보완적 학습 파트너로 기능할 때 최적의 효과를 낸다는 점을 시사한다.

이 밖에도 교육 현장에서 다양한 방식으로 AI를 활용하고 있다. 미국 Fulton County 학군은 Copilot Chat을 교과 과정에 통합해 학생들의 창의적 사고와 자신감을 높였으며, 영국 Manchester 대학교와 북아일랜드 교육청은 Copilot을 활용해 수업자료 작성과 행정업무를 자동화하였다. 체코의 한 초등학교는 다문화 학생(15% 외국 출신)을 지원하기 위해 AI 번역도구를 활용하여 교사·학부모 간 의사소통 문제를 해결하였다.

697) Microsoft. Macquarie University students' exam scores up by nearly 10 per cent thanks to new AI-powered chatbot. 2025.3.24.
698) WORLD BANK. From Chalkboards to Chatbots: Evaluating the Impact of Generative AI on Learning Outcomes in Nigeria (English). 2025.5.19.

2. AI의 교육에 대한 영향

1) AI 리터러시 격차

교육분야의 생성 AI 사용 비율은 급증했지만, 전 세계 교육자의 45%와 미국 학생의 52%는 아무런 AI 교육을 받은 적이 없다고 응답해 도입속도와 리터러시 간 격차가 확인된다.[699]

OECD와 EU 집행위원회는 AI 시대의 초중등 학생들이 갖춰야 할 지식, 기술, 태도를 제시하는 AI 리터러시 프레임워크 초안을 발표했다.[700] 이 프레임워크는 OECD가 주관하는 국제학업성취도평가(PISA)에 처음으로 포함될 AI 리터러시 평가의 기반을 제공하는 한편, EU에서 양질의 포용적 디지털교육과 기술 증진 목표를 지원하게 될 전망이다.

AI 리터러시 프레임워크는 △AI와 상호작용(Engaging with AI) △AI를 활용한 창작(Creating with AI) △AI 관리(Managing AI) △AI 설계(Designing AI) 4축의 학습목표와 교실활용 시나리오를 제시하고 있으며, AI 리터러시를 교과목 전반에 통합하는 것을 목표로 한다.

2) 학교 AI 정책 부재와 청소년 안전 문제

또 하나의 불균형은 학교 정책의 불균형이다. 스탠포드 HAI 서밋에서는 교육 분야 AI 사용 지침을 마련한 주가 26개에 불과하다는 사실이 지적되었다. 캐서린 트루잇 전 노스캐롤라이나주 교육감은 이러한 주 차원 지침의 부재가 개별 교육구나 교사에게 복잡한 윤리적, 실용적 문제를 단독으로 해결하도록 부담을 전가하여 현장의 혼란과 위험을 가중시킨다고 비판했다. 앞서 언급한 CDT 설문조사 결과 중에는 학교 정책에 관한 응답도 포함되었다.[701] 학교에 생성 AI 정책이 있다고 응답한 교사 비율은 68%(허용 정책 36%, 금지 정책 32%)였으며, 교사의 21%는 학교에 AI 정책이 없다고, 11%는 잘 모르겠다고 응답하였지만, 교사 대상의 학교 AI 지침은 학생의 생성 AI 무단 사용이 의심되는 경우 취할 대응 방향 및 학생의 생성 AI 사용을 탐지하는 방법에 관한 내용이 주류를 차지했다.

699) Microsoft, AI in Education Report: Insights to support teaching and learning, 2025.08.20.
700) AILitFramework, New AI Literacy Framework to Equip Youth in an Age of AI, 2025.05.21.
701) Center for Democracy & Technology, Out of Step: Students, Teachers in Stride with EdTech Threats While Parents Are Left Behind, 2025.01.15.

또한, 학생의 생성 AI를 이용한 부정행위를 파악하고, 부정행위에 대한 지침도 부족하다. 교사의 60%는 학생의 생성 AI 사용을 탐지하기 위한 도구를 사용하거나 테스트 중이며, 이 중 39%는 정기적으로 사용한다고 응답하였으나, 학교 차원에서 생성 AI를 이용한 학생의 부정행위를 파악하기 위해 학생 활동을 모니터링한다는 응답은 37%에 불과했다. 교사의 35%는 학교에서 과제 작성용 생성 AI 사용이 필터링되거나 차단되어 있다고 답했으며, 학생의 39% 및 학부모의 41%도 학교에서 과제 작성용 생성 AI 사용이 필터링되거나 차단되어야 한다고 응답했다.

미국의 비영리단체 커먼센스 미디어(Common Sense Media)가 미국 전역의 13~17세 청소년 1,060명을 대상으로 실시한 설문조사 결과, 전체 응답자의 72%가 한 번 이상 AI 컴패니언을 사용한 경험이 있다고 응답했다.[702]

응답자의 33%는 역할놀이, 친구나 연인 관계 형성, 정서적 지지 등 사회적 상호작용과 대인관계 형성을 위한 수단으로 AI 컴패니언을 활용했으나, 46%는 AI 컴패니언을 단순한 도구나 프로그램으로 인식했다고 한다. 대화의 만족도 측면에서는 67%가 AI보다 실제 친구와의 대화가 더 만족스럽다고 답했으나, 31%는 AI 대화가 사람과 같거나 더 만족스럽다고 응답했다. 그러나 AI 컴패니언 사용자의 34%는 대화 중 불쾌감을 경험한 적이 있다고 답했고 33%는 중요한 문제를 실제 사람보다 AI와 먼저 상의한 경험이 있으며, 24%는 본인의 실명, 위치, 비밀 등 개인정보를 AI에 제공한 적이 있다고 응답하는 등 안전 우려가 커지고 있다.

이에 커먼센스 미디어는 설문조사 결과를 바탕으로 현재의 기술 수준과 안전조치로는 18세 미만 청소년의 AI 컴패니언 이용이 부적절하다고 판단하고, 기술 기업과 교육기관, 정책 당국에 연령확인・사용제한・위기개입 장치의 필요성에 대한 정책적 대응을 제안하였다.

- AI 컴패니언 개발 기업은 자가 신고 방식이 아닌 실효성 있는 연령 확인 수단을 도입하고 자살 충동 등 위기 상황에서 인간 전문가를 연결하는 개입 체계를 구축해야 하며, 과다 사용을 방지하기 위한 사용 제한 기능을 설정 필요

- 학교와 교육자들은 AI 컴패니언이 감정적 애착을 형성하는 방식을 설명하는 교육 커리큘

702) Common Sense Media, Talk, Trust, and Trade-Offs: How and Why Teens Use AI Companions, 2025.07.16.

럼을 개발하고 디지털 리터러시 프로그램에 AI 윤리 교육을 통합하며 AI 시스템에 개인정보 공유로 인한 개인정보 침해 위험에 교육을 진행

- 정책 당국은 AI 컴패니언 플랫폼에 대해 미성년자에게서 동의를 받아 데이터를 수집하는 행위를 금지하고 안전 기준과 위기 대응 체계 수립을 의무화할 필요

3) 국가 AI 교육 정책

미국과 중국은 정부가 주도하여 교육 부분 AI 활용도를 높이고, 국가 차원의 표준화 하려는 모습을 보였다. 두 국가는 매우 상이한 접근방식으로 교육 정책을 추진하고 있다.

미국의 트럼프 대통령은 2025년 4월 23일 교육과정에 AI를 통합하고 교사에게 AI 교육을 제공하며 차세대 AI 인력을 양성하기 위한 목적으로 '미국 청소년을 위한 AI 교육 발전' 행정명령[703]을 발표했다. 주요 내용은 ▲백악관 AI 태스크포스의 구성과 ▲AI 기반 교육 개선 ▲교사 대상 AI 교육 강화 ▲AI 관련 견습 제도 확대를 요구하는 것이다.

먼저 백악관 AI 교육 태스크포스는 행정명령에 명시된 정책을 포함하여 AI 교육과 관련된 연방정부의 노력을 조정할 예정이며, 백악관 과학기술정책실(OSTP) 실장을 의장으로 하고, 농무부, 노동부, 에너지부, 교육부의 장관, 국립과학재단(NSF) 총재, 국내 정책 담당 대통령 보좌관 등이 태스크포스 구성원으로 참여하게 된다. 또한, 행정명령 발표일로부터 90일 이내에 AI 분야에서 학생과 교사의 성과를 강조하고 AI 솔루션 개발을 장려하는 대통령 AI 챌린지 계획을 수립하여 12개월 이내에 챌린지를 개최 하기로 하였는데, 2025년 9월 9일에 발표된 내용에 의하면 챌린지 상금으로 Microsoft가 125 달러를 지원한다고 하였다.

다음으로 AI 기반 교육 개선은 AI 기업, 학계, 비영리단체 등과 민관협력을 통해 K-12(초중고) 학생들의 기초 AI 리터러시와 비판적 사고 능력 함양에 중점을 둔 온라인 교육 자원을 공동 개발하는 내용으로 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 향상

703) The White House, Advancing Artificial Intelligence Education for American Youth(EO 14277), 2025.04.23

할 방안을 모색할 예정이다.

세 번째로 교사 대상 AI 교육 강화의 경우 교육부 장관은 행정명령 발표일로부터 120일 이내에 교사 교육 관련 보조금 프로그램에서 AI 사용을 우선시하는 조치를 취하여 행정업무 감소, 교사 교육과 평가 개선, 기초 컴퓨터 과학 및 AI 관련 교사 전문성 개발 등을 교사들에게 지원하고, NSF 총재는 교육 분야에서 AI 활용 연구를 우선할 수 있도록 지원하며, 교사들의 AI 도구 활용을 지원하는 교사 연수 기회를 제공하는 내용이다.

마지막으로 AI 관련 견습 제도 확대는 행정명령 발표일로부터 120일 이내에 노동부 장관은 AI 관련 공인 견습 제도에 대한 산업 전반의 참여를 확대하고, AI를 개발· 활용하는 근로자에게 학습 기회를 제공하는 조치를 시행하기로 한 것으로, AI 관련 직종에서 공인 견습 프로그램을 확대하고 청소년 프로그램을 통한 AI 기술 교육을 촉진하며, 고등학생의 AI 교육과정 및 자격증 프로그램 수강 기회를 확대할 예정이다.

이에 대하여 백악관은 지난 2025년 9월 9일 역대 최대 규모의 민관협력 AI 교육 이니셔티브를[704] 발표하였는데 Google 등 50개 이상의 기업과 기관이 참여하여 전국의 학생, 교사, 학부모들에게 무료 AI 교육과 자원을 제공한다고 발표하였는데, 그 내용은 다음과 같다.

<표 7-1> 미국 청소년을 위한 AI 교육 자원 제공 이니셔티브 참여 기업 주요 내용

참여기업	내용
Google	• 미국 교육 및 직업 훈련 프로그램에 10억 달러 투자 • 전국 고교에 Gemini 2.5 Pro를 포함한 'Gemini for Education' 무료 접근 권한 제공 • 대학생을 위한 'AI for Education Accelerator'를 100개 대학에서 200개로 확대 • 향후 3년간 AI 교육과 디지털 웰빙 지원을 위해 1억 5천만 달러의 보조금을 약속 • 이 중 Code.org에 300만 달러, Flourish Fund에 200만 달러를 지원함
Microsoft	• 미국 전역의 K-12 학생들에게 OpenAI의 최신 AI 모델을 기반으로 구축된 AI 코파일럿 무료 접근권을 제공 • '대통령 AI 챌린지' 상금으로 125만 달러를 제공 • 학생과 교사 모두를 위한 LinkedIn Learning 라이선스를 제공해 체계적인 온라인 AI 과정 무료 수강 제공 • 이 과정은 LinkedIn 프로필에 인증서로 추가되어 학생과 교사들이 학습 성과를 입증하고 취업 기회와 연결될 수 있도록 지원

704) The White House, Major Organizations Commit to Supporting AI Education, 2025.09.09.

참여기업	내용
Meta	• AI 청소년 교육 공약을 위해 2천만 달러 이상을 투자 발표 • 피어슨과 협력해 전국 군인 가족을 위한 학군의 교사들을 위해 특별히 맞춤화된 AI 도구 제공 • 이 협력을 통해 AI가 교육자의 행정 부담을 어떻게 완화할 수 있는지에 대한 문제 해결 도구들을 개발할 예정 • 피어슨은 교사 25만 명에게 생성 AI 기술의 변혁적 잠재력을 교육하는 등 전문 개발 역량을 지원
Amazon	• 2028년까지 400만 명에게 AI 기술 교육을 지원 • 미국 교육자 1만 명이 AI 커리큘럼을 활용할 수 있도록 함 • 적격 기관에 최대 3천만 달러의 크레딧을 제공해 학생 및 교사에게 클라우드 및 AI 기술을 사용할 수 있도록 지원 • '대통령 AI 챌린지'도 지원 예정
NVIDIA	• 향후 5년간 K-12 AI 기술 및 교육 개발에 2천5백만 달러 상당의 투자를 약속 • 파트너들과 함께 보조금, 인프라 기부, 교육 프로그램 개발을 모색해 차세대가 AI 기술 역량을 갖추도록 할 계획 • NVIDIA가 매년 접근하는 35만 명의 미국 연구자 및 대학생 범위를 넘어 기존 교육 및 디지털 업스킬링 프로그램을 확장하는 계획
IBM	• 2028년까지 SkillsBuild 프로그램과 기타 과정을 통해 학습자 200만 명에게 AI 관련 기술 역량 강화 기회를 제공 • 이 계획은 커리큘럼 개발, 교사 교육, 학생 참여를 통한 고등학교 AI 교육 강화, 인력 개발 및 자격증을 통한 성인 인력 교육 발전, 학생 및 교육자 AI 중심 자원과 연구개발을 통한 민관 파트너십이 포함
OpenAI	• 미국 지역사회 및 경제계 리더들과 파트너십을 통해 AI 인재를 필요로 하는 고용주와 연결하는 플랫폼 구축함으로써 AI 기술 교육과 취업 기회를 직접적으로 제공 • 2030년까지 1천만 명의 근로자를 인증하고, 인증된 근로자들에게 AI 경제로 전환하는 데 필요한 네트워크를 제공할 계획
Code.org	• 향후 3년간 25개 주와 협력해 AI 경로, AI 표준, AI 교육법 제정 추진 • AI에 중점을 둔 새로운 무료 오픈소스 고등 과정을 개발하고 확장해 2028/2029학년도까지 연간 40만 명의 미국 고등학생에게 제공
APPLE	• 비영리 조직 및 교육기관과의 파트너십을 통해 새로운 도구와 자원으로 전국 교사 및 학생 대상 AI 교육 확대 • 커먼센스미디어와 함께 교육자들이 책임감 있는 AI를 가르칠 수 있도록 지원하는 새로운 커리큘럼과 교육 프로그램을 개발
Intel	• 미국 학생들을 미래 AI 잠재 인력이 될 수 있도록 전국적 혁신을 주도하기 위한 포괄적인 AI 교육 이니셔티브의 시작을 발표 • (K-12 AI 교육) 3-12학년을 위한 250시간의 학습 자원을 제공하는 'AI for Youth Program'을 지원하여 AI 준비 학교 모델을 전국적으로 확산시킬 계획
교육 전문 기업들	• 피어슨 에듀케이션: 미국 고등학교 교사 25만 명에게 '생성 AI 기초' 인증을 위한 무료 바우처를 제공 • 맥그로힐 에듀케이션: 2030년까지 학생 1,460만 명 교육을 목표로 하는 다면적 AI 접근법을 약속 • 센게이지 그룹: 2025년 가을 100개 이상의 새로운 AI 학습 콘텐츠 출시 예정

출처: The White House, Major Organizations Commit to Supporting AI Education, 2025.09.09.

중국은 중앙정부가 강력한 통제력을 바탕으로 교육 시스템 전반을 체계적으로 혁신하는 국가 주도형 모델을 채택하고 있다. 이는 국가 주도의 기술 주권 확보와 교육 시스템의 체계적 표준화를 목표로 하는 중앙집권적 접근 방식이다.

중국 교육부 등 9개 정부부처도 2025년 4월15일 AI와 같은 신기술을 교육 전반에 통합하여 교육 혁신을 촉진하기 위한 목적으로 '디지털 교육 활성화에 관한 의견(이하 의견)'을 발표하였다. 주요 내용은 △국가 스마트교육 플랫폼 구축 △AI 기반 교육 혁신 △교육 국제화 △디지털 교육 인프라 개선 △디지털 보안 체계 구축을 중점 추진이다. 먼저, 중앙과 지방의 모든 교육 자원을 통합·공유하는 국가 차원의 허브인 국가 스마트교육 공공서비스 플랫폼을 구축한다. 이를 통해 교육 데이터를 통합하고 국가 차원의 교육 빅데이터센터를 구축하여 데이터와 인터페이스 표준을 통일하는 교육 데이터 공유 네트워크를 구축하게된다. 다음으로는 교육 특화 AI 모델과 고품질 데이터셋을 구축하고 교육 분야의 AI 시범 사업을 체계적으로 추진하며 AI를 활용한 과학 연구의 패러다임 변화를 추구하는 AI 기반 교육 혁신으로 초중고 및 대학 교육과정에서 AI 전문 과목 개설을 장려하고 디지털 교과서 개발을 지원하며, 교사와 학생 대상의 디지털 리터러시 표준과 AI 활용 지침을 마련해 디지털 리터러시와 AI 활용 역량을 강화한다.

또한, 디지털 교육을 위한 국제협력 체계를 구축하고 글로벌 인재 양성을 위한 디지털 교육과정과 훈련기지를 구축하며, 교육 디지털화에 관한 국제 싱크탱크를 설립하는 등 국제적으로 디지털 교육자원의 공동 구축과 공유를 촉진하고 글로벌 인재를 양성하며 글로벌 디지털 교육 거버넌스에 적극 참여하고자 한다.

이 같은 내용을 실행할 수 있도록 디지털 교육 인프라도 개선한다. 학교의 네트워크 인프라 구축을 지원하고 디지털 교육 관련 표준 제도를 개선하며 교육 인프라 개선을 위한 투자 메커니즘을 구축하며, 디지털 교육 HW/SW 환경과 교육 데이터 관련 표준 규격을 제정하며, 정부와 사회, 기업이 참여하는 다각화된 투자 메커니즘을 구축해 교육 인프라 구축을 지원한다.

마지막으로 교육 플랫폼의 보안 역량을 강화하고 교육 데이터 관련 사이버보안과 개인정보 보호 체계를 구축하며, 알고리즘 안전성 평가 등 AI 관련 안전 보장 체계를 마련하는 것이 중국의 AI 교육 정책 주요 내용이다.

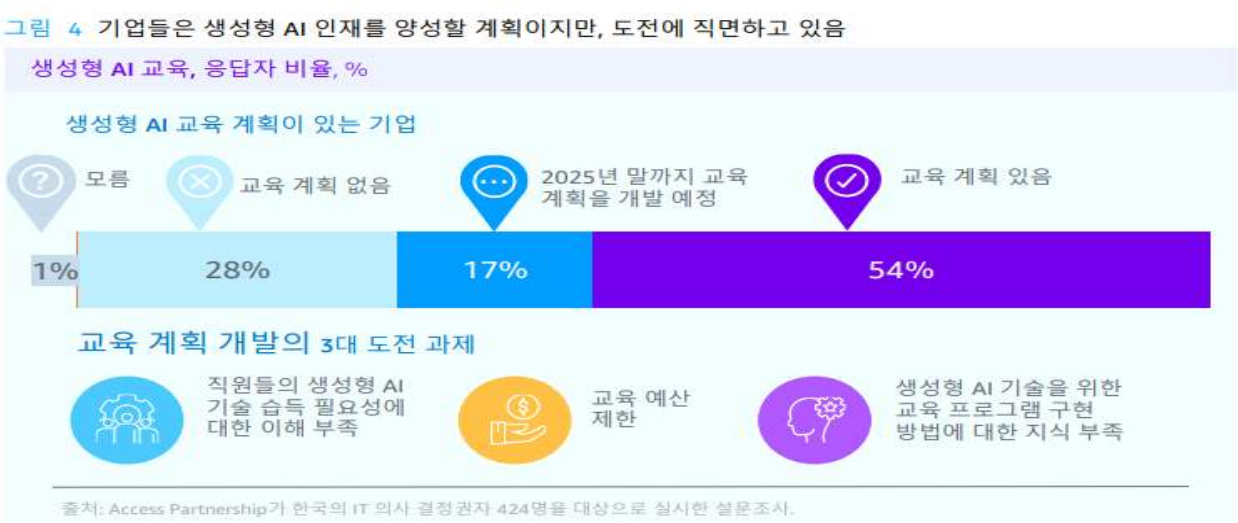
제2절 인력 및 고용 분야

1. 인력 및 고용 현황 및 추이

AWS가 9개국(독일, 미국, 브라질, 영국, 인도, 일본, 캐나다, 프랑스, 한국)의 ICT 의사결정권자 3,739명을 대상으로 진행한 설문조사를 바탕으로 2025년 5월 6일 발표한 ' 글로벌 생성 AI 도입 지수(Global Gen AI Adoption Index)' [705]에 따르면 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 답변하였고, 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 반영하고 있다.

기업들은 AI 인재 수요 증가에 대응해 사내 인력 양성과 외부 인재 영입을 병행하고 있다. 조사 기업의 56%가 생성 AI 인재 격차를 해소하기 위해 생성 AI 교육 계획을 수립했고, 19%는 2025년 말까지 수립 예정하여 사내 AI 인재를 양성하고 역량을 키울 것으로 보이며, 92%는 2025년에 생성 AI 전문 지식을 갖춘 인재를 채용할 계획이다. 또한 26%는 조직 내 신규 직무 중 절반 이상이 생성 AI 기술을 요구할 것으로 예상하고 있다.

[그림 7-3] 생성 AI 교육 계획



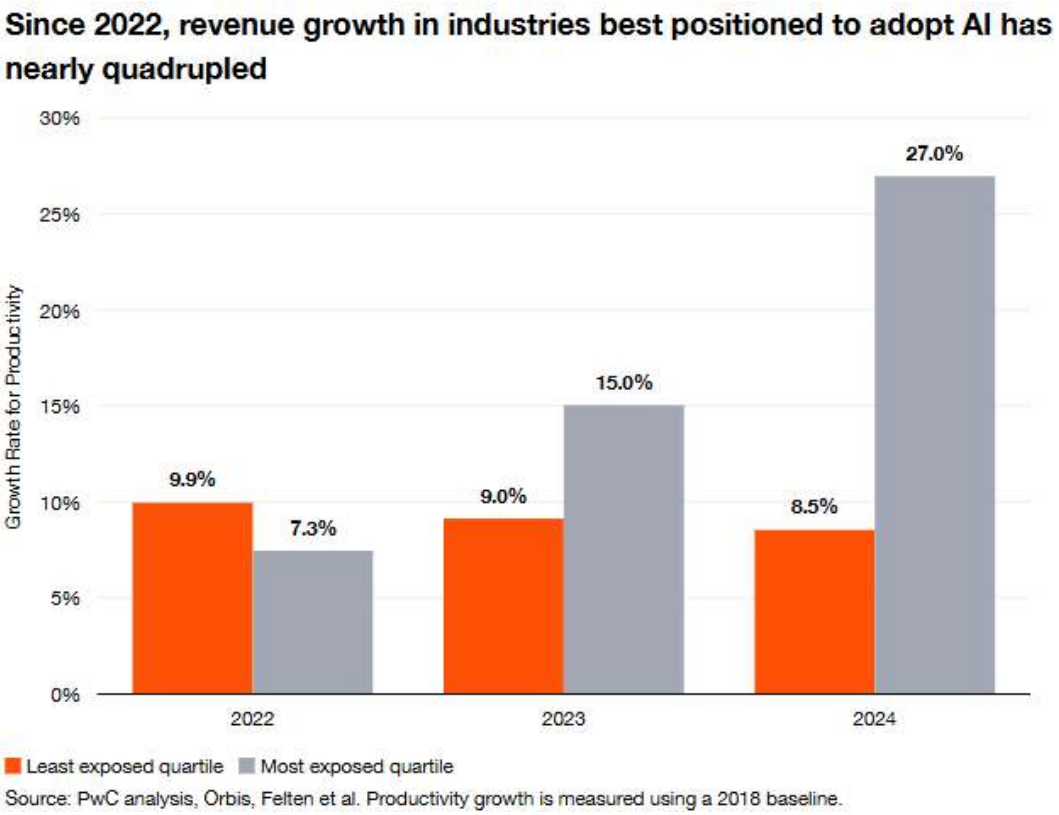
출처: 아이티데일리, 생성형 AI 도입 서두르는 기업들… 숙련된 인재 확보 '숙제' 2025.05.13.

705) Amazon, Generative AI Adoption Index, 2025.05.06.

한국의 경우[706] 424명의 IT 의사결정권자가 조사에 참여였는데, 54%가 생성 AI를 2025년 IT 예산의 최우선 과제로 꼽았고, 63%가 최고AI책임자를 임명했으며 추가로 26%가 2026년까지 임명할 계획이라고 응답하였다. 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 예정이다.

PwC는 AI가 일자리와 임금, 기술, 근로자 생산성에 미치는 영향을 분석한 '2025년 글로벌 AI 일자리 바로미터(2025 Global AI Jobs Barometer)'[707]를 발표하였는데, AI에 가장 많이 노출된 산업의 직원당 매출 증가율은 가장 적게 노출된 산업 대비 3배 높았으며, 임금은 두 배 더 빠르게 상승했다고 밝혔다.

[그림 7-4] AI 노출 대비 매출 증가율

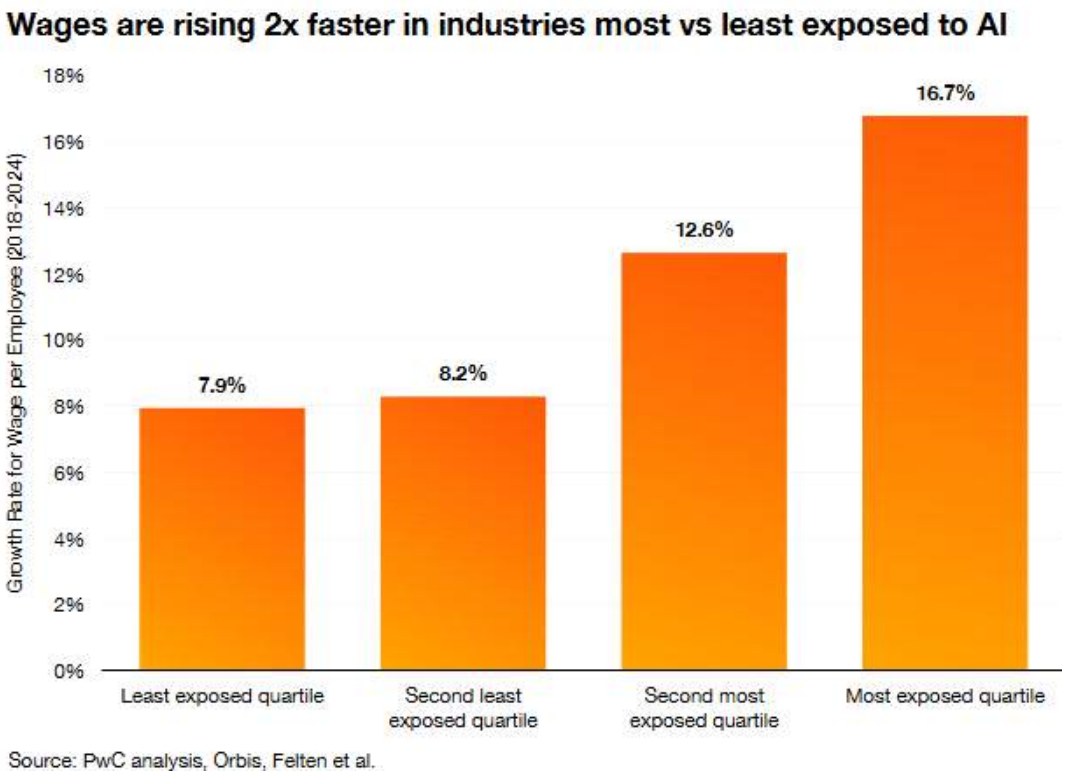


706) Amazon, AWS 연구: 생성형 AI 채택 지수-한국 기업에 대한 인사이트, 2025.05.07.
707) PwC, The Fearless Future: 2025 Global AI Jobs Barometer, 2025.06.03.

AI에 가장 적게 노출된 산업의 매출 증가율은 2022년 9.9%에서 2024년 8.5%로 소폭 하락했으나, AI에 가장 많이 노출된 산업의 매출 증가율은 7.3%에서 27.0%로 약 4배 증가했다.

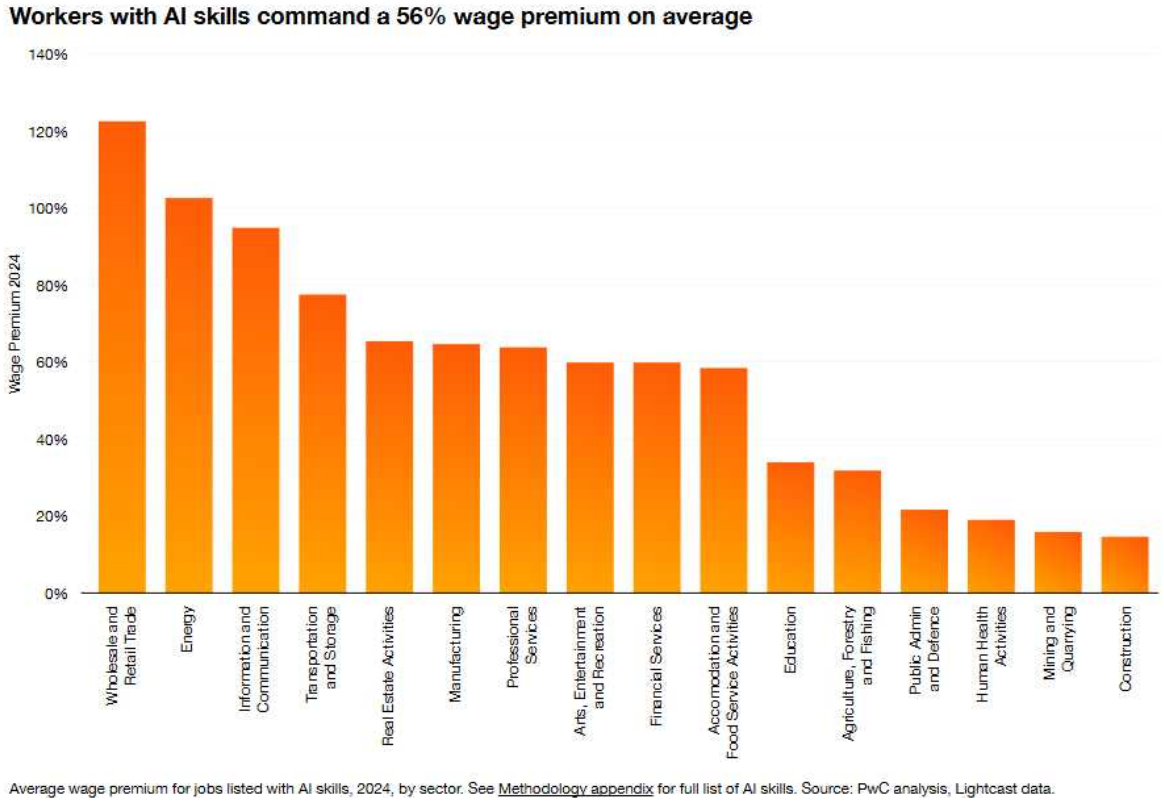
임금의 경우 AI에 가장 많이 노출된 산업은 가장 적게 노출된 산업 대비 임금이 두 배 더 빠르게 상승했으며, 고객 서비스 에이전트 같은 AI로 자동화될 수 있는 직업에서도 임금은 증가 추세를 볼 수 있다. 2018~2024년 사이 AI에 가장 적게 노출된 산업의 임금 증가율은 7.9%에 머물렀으나, AI에 가장 많이 노출된 산업의 임금 증가율은 16.7%를 기록했다.

[그림 7-5] AI 노출도가 가장 높은 산업과 가장 적은 산업 비교



머신러닝이나 프롬프트 엔지니어링과 같은 AI 기술을 보유한 근로자는 그렇지 않은 근로자 대비 임금이 평균 56% 높았으며, PwC가 분석한 모든 산업에서 AI 기술에 대하여 임금 프리미엄을 지급했고, 도소매업, 에너지, 정보통신, 교통·물류 순으로 높은 임금 프리미엄을 기록했다.

[그림 7-6] AI 기술 보유 근로자 수에 따른 임금 프리미엄

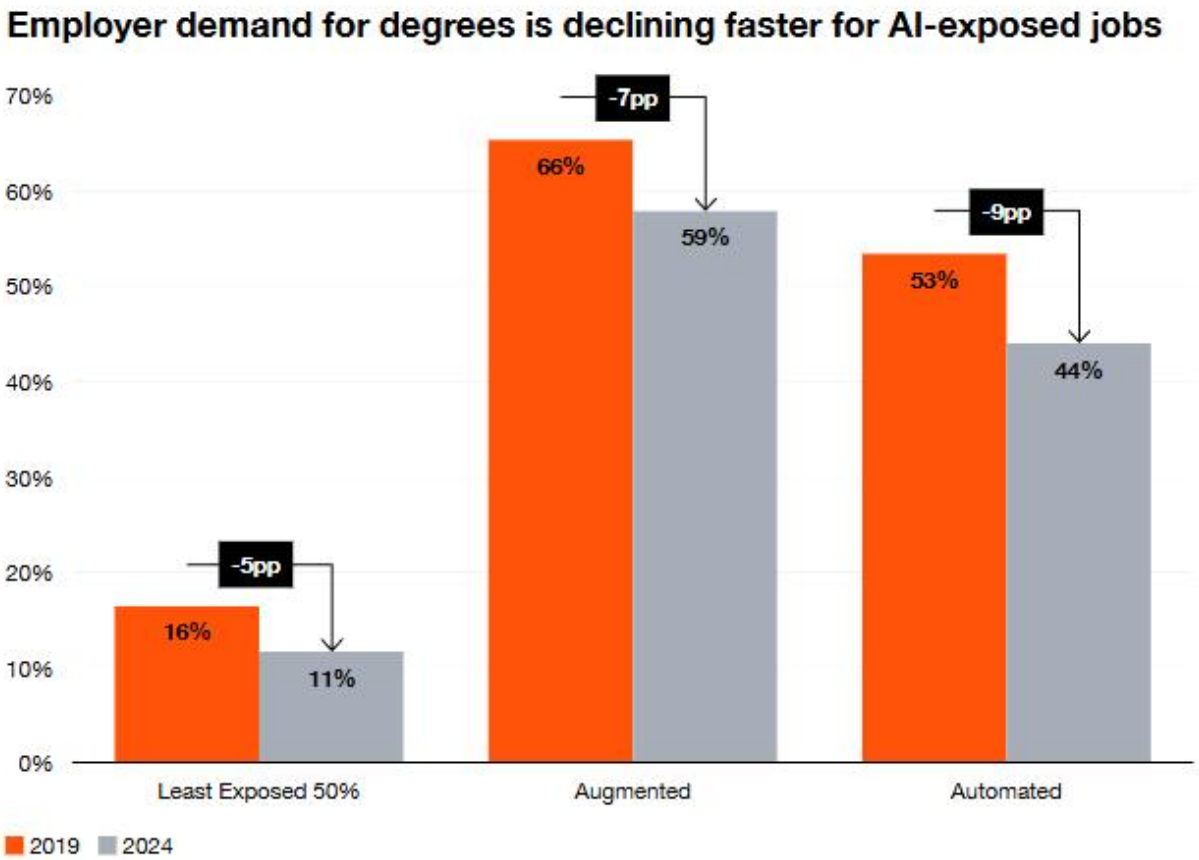


AI에 노출된 직업의 일자리 수는 전 세계적으로 사실상 모든 직업에서 증가 추세를 기록했으나, AI에 더 많이 노출된 직업의 일자리 수(예: 교육 전문가, 법률 전문가)는 2019~2024년 사이 38% 증가해 AI에 덜 노출된 직업의 일자리 수(예: 운전사)(65%)보다 증가 속도가 느린 것으로 확인되었다.

AI에 가장 많이 노출된 직업(예: 재무 분석가)에서는 고용주가 요구하는 기술이 가장 적게 노출된 직업(예: 물리 치료사)보다 66% 더 빠르게 변화하고 있으며, 학위 요건은 감소 추세이다.

AI에 노출된 직업은 인간의 업무를 개선하거나 지원하는 증강형 직업(예: 외과의사, 판사)과 AI가 자율적으로 완료할 수 있는 자동화형 직업(예: 소프트웨어 코더, 고객 서비스 담당자)으로 구분할 수 있다. AI에 노출된 직업에서 학위를 요구하는 비율은 AI 증강형 일자리에서는 2019년 66%에서 2024년 59%로, AI 자동화형 일자리에서는 2019년 53%에서 2024년 44%로 감소한 것을 확인할 수 있다.

[그림 7-7] AI에 노출된 직업에서 학위를 요구하는 비율



Source: PwC analysis, ILOSTAT, Felten et al. We calculate degree requirements as the percentage of jobs that explicitly state degree requirements divided by the total number of jobs that have education data. This analysis includes data from all countries in our study with sufficient quality data: Spain, New Zealand, Singapore, Netherlands, France, Canada, Australia, the US, Germany, the UK and Switzerland.

2. AI의 인력·고용에 대한 영향

1) 신규채용 시장의 변화

스탠포드 대학교 디지털경제연구소의 분석[708]에 따르면 미국 노동시장에서 AI 확산이 경력 초기 근로자들에게 불균형한 영향을 미치면서 AI 노출도가 높은 직업에서 경력 초기 근로자의 고용이 뚜렷하게 감소했음을 알 수 있다. 이러한 결과는 경기 상황 요인을 배제하거나 표본 구성을 달리했을 때도 일관되게 나타나, 생성 AI 확산이 경력 초기 근로자들에게 부정적 고용효과를 유발하고 있음을 입증했다고 할 수 있다.

연구진들은 미국 최대 급여 처리 서비스 기업 ADP의 급여 기록(2021년 1월~2025년 7월)을 활용해 생성 AI 도입이 노동시장에 미치는 영향을 조사하였는데, 그 결과 AI 확산이 미국 노동시장에서 경력 초기의 근로자들에게 불균형한 영향을 미치고 있다는 가설을 뒷받침하는 광범위한 증거를 제시하였다. 특히 소프트웨어 개발자나 고객 서비스 담당자와 같은 AI 노출도가 가장 높은 직업에 종사하는 경력 초기 근로자의 고용이 뚜렷하게 감소하는 경향을 볼 수 있는데, 2025년 7월까지 22~25세의 경력 초기 소프트웨어 개발자 고용은 2022년 말의 최고점 대비 약 20% 감소했으며, 반면 AI 노출도가 낮은 보건의료 보조 직업에서는 젊은 근로자 고용이 숙련 근로자보다 빠르게 성장하는 것을 볼 수 있다.

근로자 연령대로 살펴보면, 미국 노동시장에서 전반적인 고용의 탄탄한 성장세에도 불구하고 젊은 근로자의 고용은 2022년 말부터 정체되었으며, AI 노출도가 높은 직업군에서는 마이너스 성장을 기록한 것을 알 수 있는데, 앞서 언급한 AI 노출도가 가장 높은 직업에서 22~25세 근로자 고용은 2022년 말부터 2025년 7월 사이 6% 감소했으며, 이와 대조적으로 숙련 근로자의 고용은 6~9% 증가했음을 알 수 있다.

고용에 영향을 주는 경기 상황 요인과 기업 내 고용구조 차이를 배제하고 순수하게 AI 노출도가 고용 변화에 미치는 영향을 조사했을 때도, AI 노출도가 높은 직업에서 젊은 근로자의 고용 감소 효과는 확연했는데, 22~25세 근로자 중 AI 노출도가 가장 높은 직업군은 노출도가 가장 낮은 직업군 대비 고용이 약 12% 감소해 통계적으로 유의미한 차이를 기록했다.

---

708) Stanford University, Canaries in the Coal Mine? Six Facts about the Recent Employment Effects of Artificial Intelligence, 2025.08.26.

이러한 결과는 기술 직종이나 원격근무 가능 직종 제외, 시간제와 임시직 근로자 포함 등 대안적 표본 구성에서도 일관되게 나타나, 생성 AI 확산이 경력 초기 근로자에게 불균형적으로 부정적 고용효과를 유발하고 있음을 입증했다.

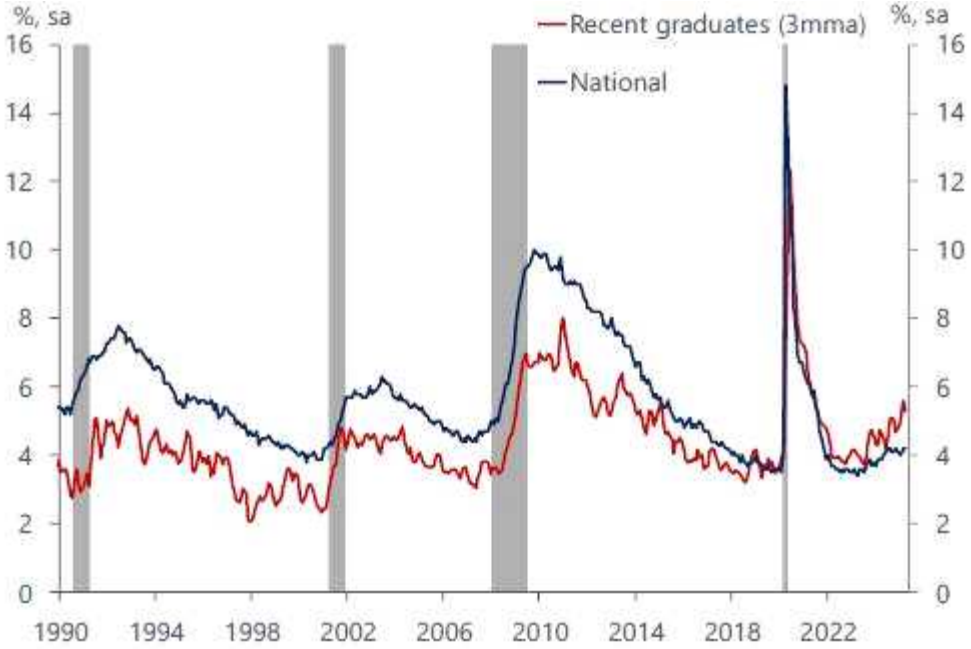
한편, AI가 노동시장에 미치는 영향은 임금보다 고용에서 더 뚜렷하게 나타나며, 고용 변화와 대조적으로 연령이나 직종별 AI 노출도에 따른 연봉 차이는 미미한 수준으로, 이는 단기적으로 임금 경직성이 존재하며, 초기에는 AI가 임금보다 고용에 더 큰 영향을 미칠 수 있음을 시사한다. 또한, AI로 자동화되는 직업에서는 경력 초기 근로자 고용이 감소했으나 AI 사용이 주로 인간 노동을 보완하는 직업에서는 고용이 유지되는 것으로 확인할 수 있다.

경제 분야 컨설팅업체 옥스퍼드 이코노믹스(Oxford Economics)또한 신규 채용에 대해 유사한 보고서[709]를 발표했다. 이에 따르면, 미국 노동 시장은 팬데믹 이후 강세를 보였으나 신규 대학 졸업자들의 실업률은 증가 추세임을 확인했다고 한다. 2023년 중반 이후 실업률 증가분의 85%는 일자리를 찾지 못한 신규 노동시장 진입자가 차지하며, 신규 대졸자들의 높은 실업률은 주로 기술 분야의 고용 둔화와 컴퓨터 과학 전공자의 증가에 의한 것으로 분석된다.

역사적으로 학사 이상 학력 소지자의 고용률은 특히 경기침체 시기에 저학력자보다 더 안정적인 경향을 보였으나, 최근 신규 대졸자의 실업률이 전국 실업률을 웃도는 현상은 AI로 인한 새로운 변화를 시사한다. 연구진은 신입 일자리가 줄어든 것을 주요 원인으로 설명하며, 기업 수요가 줄어든 분야에서 대졸자들이 최근 과잉 공급되었다고 지적했다.

709) Oxford Economics, Educated but unemployed, a rising reality for US college grads, 2025.05.27.

[그림 7-8] 신규 대졸자 실업률과 전국 평균 실업률 비교



지난 20년간 전문 서비스와 과학·기술 서비스는 대졸자의 고용 비중이 가장 많이 증가한 직종이나, 이 직종의 채용공고는 2021년 이후 40% 이상 급감했다. 연구진에 따르면 컴퓨터와 수리과학은 자동화의 영향이 큰 직종으로, 2022년 이후 해당 직무의 고용 감소는 신규 대졸자들이 차지하는 초급 수준의 일자리를 중심으로 AI에 의한 일자리 대체 효과를 시사한다. 연구진의 분석에 따르면 컴퓨터 과학과 수학 관련 직종에서 22-27세 사이의 대졸자 채용은 8% 감소하였으나, 다른 모든 직종에서는 대졸자 채용이 2% 증가한 것이 이를 뒷받침한다.

2) 근무 환경의 변화

기업에서 AI를 도입하여 활용하는 비율도 많이 늘었다. 세일즈포스가 2025년 6월 26일 발표한 '슬랙 인력 지수(Slack Workforce Index)' 조사 결과[710]에서는 전 세계 사무직 근로자들의 AI 도입률과 AI에 대한 신뢰가 대폭 증가한 것으로 확인할 수 있다. 2024년 11월 대비 AI를 도입했다는 응답이 50% 증가하였으며, 매일 AI를 사용하는 비율은 233% 증가했다. 근로자의 60%가 AI를 사용 중이며 23%는 AI 에이전트에 작업을 대신 맡겨 처리한 적이 있다고 응답하였다.

710) Salesforce, New Slack Research: Daily AI Usage Among Workers Surges 233%, 2025.06.26.

조사 결과 밀레니얼 세대(1981년~1997년 출생) 근로자들이 직장 내 AI 활용을 주도하는 세대로 부상했음을 보여준다. 밀레니얼 세대의 30%가 AI 에이전트를 완전히 이해한다고 답해, Z세대(1997년~2015년 출생)(22%)보다 높은 수치를 기록했으며, 밀레니얼 세대의 68%는 초안 작성, 요약, 아이디어 도출 등 전략적 작업에 주로 AI를 사용한다고 응답했다.

또한 매일 AI를 사용하는 근로자들은 생산성과 집중력, 직무 만족도에 대하여 "매우 좋음"이라고 응답하는 비율이 각각 41%, 41%, 38%로 非사용자(26%, 25%, 22%)를 능가하는 결과를 볼 수 있는데, AI나 AI 에이전트는 추가 자료조사의 필요성을 줄이고 작문과 의사소통을 지원하며 창의적 아이디어 도출을 위한 브레인스토밍을 도움으로써 생산성 향상에 기여하기 때문인 것으로 보인다. AI 사용자 중 96%가 자신의 역량이 부족한 업무 수행에 AI를 사용한 적이 있다고 답했으며, AI 에이전트에 기대하는 작업으로 프레젠테이션 작성(82%), 거래 관련 조사(80%), 브레인스토밍 지원(72%)을 제시했다.

근로자들은 AI가 다른 직원과의 관계를 대체하는 것이 아니라 오히려 강화한다고 인식하였다. 매일 AI를 사용하는 근로자들은 非사용자 대비 동료와 더 연결되어 있다고 느낄 가능성이 246% 더 높았고, 직장 내 소속감을 느끼는 비율도 62% 높은 것으로 확인할 수 있다. Z세대 응답자의 34%는 AI 덕분에 직장에서 더욱 긴밀히 연결되었다고 답했으며, 29%는 팀원과의 협업과 같은 직장 내 참여도가 더 높아졌다고 응답했다.

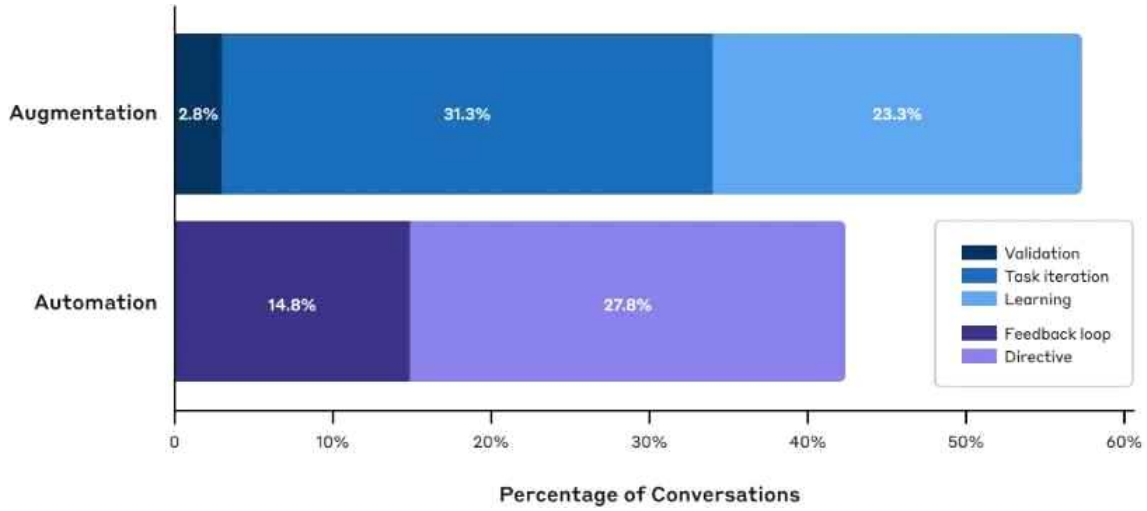
앤스로픽은 2025년 2월 10일에 발표한 '앤스로픽 경제 지수(Anthropic Economic Index)' 보고서[711]에서 AI가 노동시장과 경제에 미치는 영향을 이해하기 위해 '클로드(Claude.ai)'에서 이루어진 수백만 건의 익명화된 대화를 기반으로 경제 전반에서 AI가 실제 업무에 어떻게 통합되고 있는지를 분석하였다. 보고서에 따르면 소프트웨어 개발 및 기술문서 작성에 AI가 가장 많이 사용되며, 전체 직업의 약 36%가 업무의 25% 이상에서 AI를 일정 수준 사용된다고 한다. 클로드에 입력된 질문의 37.2%가 소프트웨어 수정, 코드 디버깅, 네트워크 문제 해결 등 소프트웨어 엔지니어링 직무와 연관된 "컴퓨터 및 수학" 범주에 해당하며, 두 번째로는 "예술, 디자인, 스포츠, 엔터테인먼트, 미디어"(질문의 10.3%) 직무로 나타났다. 신체 노동 중심의 "농업, 어업, 임업" 범주에 해당하는 질문은 0.1%로 최하위를 기록했다.

711) Anthropic, The Anthropic Economic Index, 2025.02.10.

전체 직업 중 약 4%만 업무의 75% 이상에서 AI를 사용했으며, 저임금이나 초고임금 일자리보다 컴퓨터 프로그래머와 데이터 과학자 등 중간~고임금 범위의 직업이 AI를 가장 많이 사용했다.

AI를 업무 자동화(Automation: AI가 문서 서식 지정과 같은 작업을 직접 수행)와 증강(Augmentation: AI가 사용자와 협력하여 작업을 수행) 중 주로 어디에 사용하는지 분석한 결과는, 업무의 57%가 증강에 해당하고 자동화 비중은 43%로 확인되었다.

[그림 7-9] 클로드 대화에서 업무 증강 vs 업무 자동화 비율



제8장 결론

2025년은 생성 AI와 에이전트형 AI가 실험 단계에서 벗어나 산업·공공·교육·일상 전반에 깊이 스며들기 시작한 전환점의 해였다. 본 보고서는 전 세계 AI 발전 수준, 시장과 투자 구조, 주요국 정책과 국제 거버넌스, 기업과 기술·연구, 교육·인력·고용 동향을 종합적으로 분석하여 2025년 글로벌 AI 생태계의 구조적 변화를 살펴보았다. AI는 이제 개별 기술이 아니라 데이터·컴퓨팅·인력·제도·윤리·국제협력이 얽힌 복합 인프라로 자리 잡고 있으며, 그 기회와 위험이 동시에 증폭되고 있다. 이에 따라 2025년 글로벌 AI 분야 동향을 통해 도출되는 핵심 시사점과 정책 제언을 제시하고자 한다.

1. 2025년 AI 주요 동향

1) 생성 AI의 가치 실현을 위한 도전

AI는 '보편적 업무 도구' 단계로 진입했지만, 가치 실현 측면에서는 여전히 '스케일링 격차'가 뚜렷하다.

스탠퍼드 HAI의 2025년 AI 인덱스 보고서에 따르면 2024년 기준 전 세계 조직의 78%가 이미 AI를 활용하고 있으며, 이는 2023년 55%에서 급격히 높아진 수치다. 맥킨지의 「The state of AI in 2025」 조사에서도 응답 기업의 88%가 최소 한 개 이상의 비즈니스 기능에서 AI를 사용하고 있다고 보고된다. 특히 IT 서비스 헬프데스크, 지식 관리, 고객 응대 등 정보 집약적 업무 영역에서 생성 AI와 에이전트형 AI가 빠르게 도입되고 있으며, 많은 조직이 다수의 도메인 특화 에이전트를 동시에 운영하는 단계에 진입하고 있다. 그럼에도 불구하고, AI 활용이 조직 전반으로 확산된 '스케일 단계'에 도달한 기업 비중은 20%대 초반에 그치며, 에이전트형 AI를 핵심 업무 프로세스에 깊이 통합한 기업 역시 소수에 불과하다.

이는 AI 기술 자체보다는 데이터 품질과 거버넌스, 레거시 시스템, 인력 역량, 조직 문화와 같은 비기술적 요인이 AI 효과 확산의 병목으로 작용하고 있음을 시사한다. 글로벌 차원에서 AI는 "남들이 모두 쓰는 기술"이 되었지만, "실질적인 경쟁우위와 생산성 향상으로 연결하는 능력"에는 국가·기업 간 큰 편차가 존재하는 것이다.

2) AI 에이전트와 피지컬 AI

생성 AI와 에이전트형 AI, 피지컬 AI가 소프트웨어 산업 구조와 가치 사슬을 재편하고 있다.

ABI 리서치와 시장조사 기관들의 분석에 따르면 전 세계 AI 소프트웨어 시장은 2024년 약 1,220억 달러에서 2030년 4,674억 달러 수준으로 연평균 25% 이상 성장할 것으로 전망된다. 특히 생성 AI 분야는 같은 기간 전체 시장보다 높은 성장률을 기록하며, 기존 SW 시장과 구분하기 어려울 정도로 범용 인프라로 편입되고 있다. AI 에이전트 시장만 보더라도 2024년 51억 달러에서 2030년 471억 달러로, 연평균 40%가 넘는 고성장이 예상된다. 골드만삭스는 에이전트형 AI 확산으로 2030년 글로벌 소프트웨어 총시장 규모가 기존 전망치 대비 최소 20% 이상 확대될 것으로 보고, 전체 SW 산업 매출의 60% 이상이 '지능형 AI 레이어'에서 창출될 것이라고 전망하였다. 구글 클라우드의 'ROI of AI' 보고서와 맥킨지 조사 결과를 합쳐 보면, 기업들은 고객 서비스, 마케팅, 보안, IT 운영, 연구·개발 등 전 가치사슬에 걸쳐 에이전트를 도입하고 있으며, 다수의 기업이 10개 이상의 에이전트를 동시에 운영하고 있다. 이 과정에서 AI는 단순 보조 도구를 넘어, 업무 흐름 전체를 설계·조정하는 '오케스트레이터'로 기능하고 있으며, 플랫폼·클라우드·데이터·반도체·서비스가 결합된 통합 생태계로 시장 구조가 빠르게 재편되고 있다.

2025년 인공지능 산업과 정책 환경의 또 다른 중요한 변화는 피지컬 AI, 즉 실세계와 직접 상호작용하는 인공지능 기술이 주요국의 중장기 전략 의제로 본격 편입되기 시작했다는 점이다. 그간 생성 AI와 초거대 언어모델(LLM)이 디지털 영역의 정보 처리와 콘텐츠 생성 중심으로 발전해 왔다면, 2025년을 기점으로 AI는 로봇, 자율주행, 스마트 제조, 물류, 에너지, 돌봄 등 물리적 환경에서 인지·판단·행동을 수행하는 단계로 확장되고 있다. 특히 중국은 정부 업무보고와 산업정책을 통해 '체화형 AI'(Embodied AI)를 미래 핵심 산업으로 명시하며 휴머노이드 로봇과 지능형 로봇을 차세대 성장 동력으로 격상하였고, 미국·EU·일본 역시 기존의 로봇·자율시스템·스마트팩토리 정책을 초거대 모델 및 에이전트 기술과 결합하는 방향으로 재정렬하고 있다. 이러한 흐름은 AI 기술 경쟁이 단순한 알고리즘 성능이나 디지털 서비스 영역을 넘어, 제조·물류·모빌리티·도시 인프라 등 실물경제 전반의 생산성과 안전성, 국가 산업 경쟁력을 좌우하는 구조적 변수로 이동하고 있음을 시사한다. 즉, 2025년은 생성 AI 중심의 1단계 전환을 넘어, 에이전트형 AI와 피지컬 AI가 결합된 'AI의 실체화(physicalization)'가 본격화되는 전환기로 평가할 수 있다.

3) AI와 글로벌 불균형 심화

AI 투자와 인프라, 기술력은 소수 국가·기업에 집중되며 글로벌 불균형을 심화시키고 있다.

HAI의 분석에 따르면 2024년 전 세계 AI 관련 총투자액은 2,523억 달러로 2023년 대비 25.5% 증가했다. 민간투자는 전년 대비 44.5% 늘어나며 가장 가파른 성장세를 보였고, M&A도 12.1% 증가했다. 투자 분야별로는 AI 인프라·연구·거버넌스(373억 달러), 데이터 관리·처리(166억 달러), 의료·헬스케어(110억 달러)에 투자가 집중되었는데, 이는 OpenAI, Anthropic, xAI 등 소수 선도 기업에 대한 대규모 자본 유입과 맞물려 있다. CB 인사이트가 집계한 2025년 3분기까지의 투자 동향을 보면, 경기 둔화에도 불구하고 분기별 AI 투자액은 4분기 연속 450억 달러를 상회하고 있으며, 10억 달러 이상 초대형 라운드는 대부분 프런티어 모델 기업과 관련 인프라 기업에 집중되고 있다.

한편, UNCTAD는 2025년 기술·혁신 보고서에서 2033년까지 전 세계 AI 시장이 약 4조 8천억 달러 규모로 성장할 수 있다고 전망하면서도, 현재 AI의 경제적 혜택이 소수 국가와 빅테크 기업에 집중되고 있음을 경고한다. 또한 전 세계 일자리의 최대 40%가 AI의 영향을 받을 수 있지만, 많은 개발도상국은 아직 국가 AI 전략조차 마련하지 못해 향후 ‘AI 격차’와 대표성 부족 문제가 심화될 수 있다고 지적한다.

이러한 흐름은 AI가 국가 간·기업 간 ‘초격차’를 만들어내는 핵심 요인으로 작동하고 있으며, 데이터·연산·인재가 결합된 AI 인프라 영역에서 선도국과 추격국의 간극이 구조적으로 고착될 위험을 시사한다.

4) AI 윤리, 거버넌스 및 안전성 강화

책임 있는 AI 거버넌스는 급속히 제도화되는 동시에, 국가·지역별 불균형과 규범 경쟁이 가속화되고 있다.

AI 인덱스 2025에 따르면 2016년 이후 39개국에서 최소 1건 이상의 AI 관련 법률이 제정되었고, 2024년까지 통과된 AI 관련 법률은 총 204건에 이른다. 연간 통과 건수는 2016년 1건에서 2024년 40건으로 증가해, AI 거버넌스가 더 이상 선언적 논의 단계가 아니라 구체적 규율

체계로 전환되고 있음을 보여준다.

EU는 2024년 AI 법(AI Act)을 최종 승인하여 2025년부터 단계적으로 시행을 앞두고 있으며, 위험 기반 규율 체계와 투명성·인권 보호 의무 등을 통해 세계 최초의 포괄적 AI 규제 프레임워크를 구축했다. UN 차원에서는 고위급 AI 자문단 보고서 「Governing AI for Humanity」를 바탕으로 AI 국제과학패널과 글로벌 AI 거버넌스 대화가 출범했고, AI가 인권·개발·평화에 미치는 영향을 다층적으로 다루는 글로벌 논의 플랫폼이 확대되고 있다.

OECD, UNESCO, 국제인권기구 등도 신뢰할 수 있는 AI, 교육·노동에서의 책임 있는 활용, 인권·프라이버시 보호 기준을 제시하는 가이드라인을 잇달아 발표하였다. 한편, HAI는 책임 있는 AI(RAI) 영역에서 규정·원칙은 빠르게 늘고 있으나, 개발사 간 책임성 벤치마크와 평가 기준이 통일되어 있지 않아 모델 간 비교와 규제 집행에 어려움이 있다고 지적한다. 즉, 규범과 법제는 빠르게 늘어나고 있지만, 실제 구현과 평가, 국제 공조는 아직 초기 단계이며, 미국·EU·중국 등 주요 권역 간 규제 철학의 차이가 글로벌 AI 생태계의 규범 경쟁으로 이어지고 있는 것이다.

5) AI 역량과 준비 수준의 격차 심화

AI는 교육·노동·사회 구조를 재편하는 핵심 인프라로 부상했지만, AI 역량과 준비 수준의 격차가 크게 벌어지고 있다.

Microsoft의 AI in Education 보고서에 따르면 전 세계 교육기관의 86%가 이미 생성 AI를 도입했으며, 학교·대학 현장에서 교사와 학생 모두 과제 설계, 수업 준비, 학습 보조 등에 AI를 활용하고 있다. 그러나 교사의 상당수는 AI를 수업에 통합할 준비가 되어 있지 않다고 응답해, ‘활용 속도’와 ‘교육 시스템 정비’ 간 괴리가 나타나고 있다. AWS의 글로벌 생성 AI 도입 지수는 조사 대상 기업의 90%가 이미 생성 AI 도구를 사용하고 있으며, 45%는 생성 AI를 향후 IT 예산의 최우선 투자 분야로 꼽고, 60%는 최고 AI 책임자(CAIO)를 두고 있다고 보고한다.

그럼에도 인재 부족과 비용, 환각(hallucination) 문제, 데이터 보안과 프라이버시 우려가 주

요 장애 요인으로 지적되며, 특히 중소기업·공공·비영리 영역에서 준비 수준의 편차가 크다. 한편, UNCTAD는 AI로 인해 반복·규칙적인 업무 중심의 일자리는 감소할 수 있으나, 고숙련·창의·문제해결 중심 직무에서는 업무 재편과 생산성 향상을 통해 새로운 기회가 창출될 수 있다고 분석한다.

AI 노출도가 높은 직무일수록 임금 성장과 생산성 향상이 관측되지만, 동시에 직무 내용과 요구 역량의 변화 속도가 빨라지면서 재교육과 평생학습 체계가 갖추어지지 않은 국가와 계층에서 고용 불안과 격차 확대 우려가 커지고 있다. 이러한 흐름 속에서 AAAI 2025 등 국제 학술대회는 추론·사실성·평가·윤리·에이전트·로보틱스 등 분야를 중점 의제로 다루며, 기술 연구의 중심축이 단순 성능 경쟁에서 '인간과 공존하는 안전하고 책임 있는 AI'로 이동하고 있음을 보여준다.

요약하자면, 2025년 글로벌 AI 동향은 ▲폭넓은 보급과 제한적인 스케일링 ▲생성형·에이전트형 AI 중심의 시장 구조 재편 ▲투자와 인프라의 초집중과 불균형 심화 ▲책임 있는 AI 거버넌스의 급속한 제도화와 규범 경쟁 ▲교육·노동·사회 전반의 대전환과 역량 격차 확대가 부각되었다. 이러한 변화는 우리나라가 향후 AI 정책을 설계할 때, 단순한 기술 추격을 넘어 산업·사회·거버넌스 전 영역을 포괄하는 '전략적 선택'을 요구하고 있다.

2. 정책제언

앞 장들에서 살펴본 글로벌 동향과 본 연구의 분석 결과를 종합하면, 우리나라 AI 정책은 더 이상 "AI 기술 자체의 확보"에 머물러서는 안 되며, "AI를 기반으로 한 국가 전반의 구조 전환"을 어떻게 설계할 것인가에 초점을 옮겨야 한다. 또한 2025년은 인공지능기본법 제정과 AI 서울 정상회의, 국내 AI안전연구소 설립 등 우리나라 AI 정책 체계가 새로운 전기를 맞은 시기라는 점에서, 향후 5년을 내다본 전략적 재정렬이 필요한 시점이기도 하다. 이에 본 연구는 다음과 같은 방향에서 우리나라 AI 정책의 고도화를 제언한다.

1) 초거대·에이전트형 AI 시대에 대응하는 '전 주기 AI 인프라 전략' 구축

AI 인덱스 2025와 본 보고서의 투자 분석에서 확인되듯이, 글로벌 AI 투자는 AI 인프라·연

구·거버넌스와 데이터 관리 영역에 집중되고 있으며, 이 분야에 대한 장기·대규모 투자가 국가·기업 간 격차를 결정하는 핵심 요인이 되고 있다.

우리나라는 반도체·네트워크·모바일 등 기반 산업 경쟁력을 보유하고 있음에도, 초거대 모델 학습에 필요한 연산 자원, 고품질·고가치 데이터, 실험·검증 인프라에 있어서는 미국·중국·EU에 비해 제약이 크다. 향후 정책은 단순히 데이터센터와 GPU 공급을 늘리는 수준을 넘어, ▲국가 차원의 AI 컴퓨팅 허브와 공공·민간이 공동 활용하는 클라우드-온프레미스 하이브리드 구조 ▲산업·의료·교육·공공 등 계층별·도메인별 고가치 데이터 스페이스 구축 ▲연구개발·검증·규제 샌드박스를 아우르는 ‘AI 실험 인프라’ 조성 ▲AI 반도체·시스템 소프트웨어·모델·서비스를 묶은 전략투자 프로그램 등으로 이어져야 한다. 또한 이러한 인프라를 소수 대기업에만 집중하기보다, 중소·스타트업·연구기관·지역 혁신주체가 접근할 수 있는 공공성 있는 플랫폼으로 설계하여, 인프라 격차가 곧 혁신 격차로 이어지지 않도록 하는 것이 중요하다.

2) ‘파일럿 중심’에서 ‘스케일업 중심’으로 전환하는 AI 활용 정책 마련

글로벌 조사에서 확인되듯이, 다수의 국가와 기업은 AI를 여러 부문에서 도입하고 있음에도 불구하고, 실제로 조직 전체의 업무 구조와 성과 지표를 바꾸는 수준으로 활용하는 사례는 제한적이다.

우리나라 역시 공공·제조·금융·의료·교육 등 다양한 분야에서 시범사업과 PoC가 축적되어 왔지만, 표준화된 참조 아키텍처와 성과 관리 지표 부재, 데이터와 레거시 시스템의 분절, 규제·책임 범위의 불명확성 등으로 인해 ‘파일럿의 벽’을 넘지 못하는 사례가 적지 않다. 따라서 향후 정책은 ▲산업별·업무별 고가치 AI 활용 시나리오를 발굴하고, 이를 기반으로 한 ‘스케일링 로드맵’과 참조 아키텍처를 제시하며 ▲에이전트형 AI를 전제로 한 업무 프로세스 재설계(BPR)와 조직문화 변화까지 포함하는 종합 지원 프로그램을 마련하고 ▲성과 창출 사례를 확산·재사용할 수 있도록 공공·민간이 함께 공유하는 ‘AI 활용 레퍼런스 라이브러리’를 구축하는 방향으로 진화해야 한다. 특히 중소기업의 경우, 자체적으로 데이터·인력·시스템을 갖추기 어렵다는 점을 감안해, 클라우드 기반 AI 서비스, 공동 데이터 플랫폼, 표준화된 에이전트 템플릿 등 “완성형 패키지”를 제공하는 정책 수단이 효과적일 것이다.

3) 국가 AI 전략의 핵심 축으로 피지컬 AI 생태계 강화

거대 모델과 데이터·컴퓨팅 인프라 중심의 기존 전략을 한 단계 확장하여, 피지컬 AI를 국가 AI 전략의 핵심 축 중 하나로 명확히 설정할 필요가 있다.

우리나라는 반도체, 제조, 자동차, 조선, 로봇, 가전 등 세계적인 경쟁력을 보유한 실물 산업 기반을 갖추고 있어, 피지컬 AI는 단순한 기술 추격이 아닌 산업 주도형 AI 경쟁력 확보의 현실적인 돌파구가 될 수 있다. 이를 위해 첫째, 휴머노이드 로봇, 자율주행·무인이동체, 스마트팩토리, 물류·건설 로봇 등 분야를 아우르는 피지컬 AI 중점 전략 분야를 명시적으로 정의하고, 초거대 모델·에이전트 기술과의 결합을 전제로 한 범부처 R&D·실증 프로그램을 체계화할 필요가 있다.

피지컬 AI는 소프트웨어뿐 아니라 센서, 모터, 배터리, 엣지 컴퓨팅, AI 반도체, 시뮬레이션(디지털 트윈)까지 결합되는 복합 시스템인 만큼, AI 반도체·로봇 부품·제조 장비 산업과 연계한 정책 패키지를 통해 국내 공급망 경쟁력을 함께 강화해야 한다. 피지컬 AI의 안전성과 신뢰성 확보를 위해 자율주행, 로봇, 산업설비 등 실제 환경에서의 테스트·검증을 가능하게 하는 실증 특례, 규제 샌드박스, 안전 인증 체계를 선제적으로 정비함으로써, 기술 상용화의 불확실성을 낮출 필요가 있다.

AI 인재 정책 측면에서도 소프트웨어 중심 인력 양성에서 나아가, 기계·전기·제어·로봇 공학과 AI를 융합한 피지컬 AI 융합 인재 양성 체계를 구축함으로써 장기적인 기술 자립과 산업 경쟁력을 뒷받침해야 할 것이다. 이러한 전략은 우리나라가 초거대 AI 기술 경쟁을 넘어, AI가 실제 산업과 사회 문제 해결에 기여하는 ‘현장 중심 AI 강국’으로 도약하기 위한 핵심 정책 방향이 될 것이다.

4) 한국형 책임 있는 AI 거버넌스의 정교화

EU AI 법과 UN 거버넌스 논의를 참고한 ‘한국형 책임 있는 AI 거버넌스’를 정교화해야 한다.

2025년을 전후로 EU, 미국, 영국, 일본 등은 AI 안전·신뢰성·인권 보호를 위한 제도와 전문기관을 본격 가동하고 있다. EU AI 법의 위험 기반 규율, UN의 AI 국제과학패널과 글로벌

AI 거버넌스 대화, OECD·UNESCO의 책임 있는 AI 원칙 등은 앞으로 국제 표준과 규범 형성에 중요한 기준점이 될 것이다.

우리나라 역시 인공지능기본법 제정과 AI 안전연구소 설립, 딥페이크·개인정보 보호 관련 법제 정비 등을 통해 규범 인프라를 갖춰가고 있으나, 아직 전체적인 거버넌스 체계와 실행 메커니즘은 정립 과정에 있다. 향후 정책은 ▲위험 수준에 따른 차등 규율과 혁신 공간을 조화시키는 한국형 위험 기반 규제 모델 수립 ▲AI 시스템의 설계·학습·배포·운영 전 주기에 걸친 투명성·설명가능성·인권영향평가 체계 마련 ▲국가 AI 안전·시험·인증 인프라 구축과 국제 상호인정 체계 참여 ▲데이터 보호, 알고리즘 공정성, 딥페이크·허위정보 대응을 포괄하는 통합적인 규범 설계를 통해 '혁신 친화적이면서도 권리를 보호하는' 거버넌스를 구축해야 한다. 아울러, 법제와 규제기관 중심 접근을 넘어, 기업·연구기관·시민사회가 자율규범과 모범사례를 축적할 수 있도록 지원하고, 공공영역에서 선도적으로 책임 있는 AI 활용 기준을 제시하는 것도 중요하다.

5) AI 인력·교육 체계 재편

AI 인력·교육 체계를 'AI 활용 역량 보편화'와 '고급 인재 전략 육성'이라는 이중 구조로 재편해야 한다.

세계적으로 K-12 단계에서 컴퓨터과학 교육을 제공하거나 계획 중인 국가는 전체의 3분의 2를 넘어서고 있으나, 교사와 학생의 준비 수준은 국가·지역별로 크게 차이가 나는 것으로 보고된다.

우리나라 역시 초·중등과정의 디지털·AI 교육 강화, 대학·대학원의 AI 융합전공 확대, 재직자 교육 프로그램 등 다양한 정책을 추진하고 있으나, 현장의 수업 여건과 교원 역량, 지역·계층 간 접근성 측면에서는 여전히 격차가 존재한다. 향후 정책은 ▲모든 국민이 기본적인 AI 리터러시를 갖추도록 하는 '디지털·AI 기본권 교육' 체계 구축 ▲교원 양성과 연수 단계에서 AI 활용 역량을 필수화하고, 교실·강의실에서 AI를 실질적 교육 도구로 활용할 수 있도록 실습 중심 지원을 확대하며 ▲직업훈련·평생교육 체계 전반에 AI 기반 직무 전환·업스킬링 과정을 통합하는 방향으로 나아가야 한다. 동시에, 초거대 모델·AI 반도체·로보틱스·에이전트·안전·윤리 등 전략 분야의 최고급 인재를 육성하기 위해, 대학·연구기관·기업이 연계된 장기 펠로우십, 글로벌 공동연구, 창업 연계 프로그램 등을 대폭 확충할 필요

가 있다. 이는 AI 기술 추격을 넘어, 새로운 연구 패러다임과 응용 분야를 선도하기 위한 필수 조건이다.

6) 산업·사회·국제협력 정책 통합 거버넌스 마련

AI를 국가 전략산업이자 사회 인프라로 인식하고, 산업정책·사회정책·국제협력정책을 통합하는 거버넌스가 요구된다.

UNCTAD가 경고하듯, AI는 세계 경제 성장의 새로운 엔진이 될 수 있는 동시에, 국가 간·계층 간 격차를 확대할 위험을 안고 있다.

우리나라의 경우 제조·수출·디지털 인프라 측면에서 AI 활용 잠재력이 높지만, 인구구조 변화와 지역·산업 간 양극화, 중소기업의 디지털 격차 등 구조적 제약도 동시에 존재한다. 따라서 AI 정책을 산업·기술 정책에만 한정하지 말고, 노동·복지·교육·지역균형발전·문화·언론 등 사회 전반의 제도와 연계하는 통합 전략이 필요하다. 구체적으로는 ▲AI 도입으로 인한 일자리 구조 변화에 대응하는 직업훈련·전직 지원과 사회안전망 결합 ▲AI를 활용한 돌봄·보건·환경·교통·재난 대응 등 공공서비스 혁신과 디지털 포용 정책의 동시 추진 ▲지역 기반의 AI 혁신클러스터와 생활 밀착형 AI 서비스 확산을 통해 수도권·대기업 중심의 혁신 집중 완화 ▲AI 기술과 콘텐츠가 민주주의·담론·언론 생태계에 미치는 영향을 고려한 정보 신뢰성·미디어 리터러시 강화 정책 등을 종합적으로 설계해야 한다. 또한 AI 서울정상회의를 계기로 형성된 국제 네트워크를 활용하여, 기후·보건·재난·개발협력 등 글로벌 공공재 영역에서 AI를 활용한 공동 프로젝트를 선도함으로써, 우리나라가 책임 있는 AI 활용의 모범사례를 제시하는 것도 중요하다.

6) 중견국으로서 글로벌 AI 거버넌스와 표준 설정에서 '연결자(hub)' 역할 수행

글로벌 AI 거버넌스 논의는 미국·EU·중국 등 거대 권역과 G7 중심으로 전개되고 있으며, 다수 개발도상국과 중견국은 아직 논의 과정에서 충분한 대표성을 확보하지 못하고 있다.

우리나라는 디지털 인프라와 민주주의, 산업 경쟁력, 개발협력 경험을 두루 갖춘 국가로서, 선도국과 개발도상국을 이어주는 다리 역할을 수행할 수 있는 위치에 있다. 이를 위해 ▲UN

AI 국제과학패널, OECD·G7 AI 논의, 다양한 다자·복수국 협의체에 적극 참여하고, 아시아·태평양 지역을 대표하는 사례와 목소리를 제시하며 ▲AI 안전·윤리·교육·포용적 개발 등 특정 주제를 중심으로 한 국제 협력 이니셔티브를 주도하고 ▲국내에서 축적된 공공데이터·디지털정부·전자정부 경험을 AI 시대에 맞게 재설계하여 글로벌 디지털 공공재 형태로 공유하는 전략이 요구된다. 이러한 노력을 통해 우리나라는 단순한 '규범 수용자'를 넘어, '규범 설계자이자 조정자'로서 역할을 확대할 수 있을 것이다.

3. 기대효과

본 보고서에서 분석한 2025년 국내외 AI 산업·정책·기술·인력·교육 동향과 위의 제언들은 어디까지나 출발점이다. AI 기술과 시장, 거버넌스 환경은 향후 몇 년간 지금보다 더 빠른 속도로 변화할 가능성이 높다. 따라서 정부와 산업계, 학계, 시민사회는 정기적인 동향 분석과 정책 점검을 통해 전략을 유연하게 조정할 수 있는 '학습하는 거버넌스' 체계를 구축해야 한다. 「2025년 국내외 인공지능 산업 동향 조사」가 제시한 분석과 결론이 향후 인공지능기본법 하위 법령과 제도 정비, 국가 AI 전략 고도화, 산업·교육 정책 설계, 기업과 연구기관의 전략 수립에 폭넓게 활용되어, 우리나라가 혁신성과 책임성을 겸비한 AI 선도국가로 도약하는 데 일조할 수 있기를 기대한다.

[참고 문헌]

1. AISI, AIセーフティに関する評価観点ガイドの公開, 2024.9.18.
2. Allen Institute for AI, OLMo: Open Language Model, 2024.02.01.
3. Anthropic, Claude 3.5 Sonnet, 2024.06.21.
4. Anthropic, Sleeper Agents: Training Deceptive LLMs that Persist Through Safety Training, 2024.01.15.
5. AP News, It’s not just Elon Musk: ChatGPT-maker OpenAI confronting a mountain of legal challenges, 2024.03.07.
6. Apple, Introducing Apple Intelligence, the personal intelligence system that puts powerful generative models at the core of iPhone, iPad, and Mac, 2024.06.10.
7. Arxiv, AgentCourt: Simulating Court with Adversarial Evolvable Lawyer Agents, 2024.08.15.
8. Autorite de la Concurrence, Generative artificial intelligence: the Autorite issues its opinion on the competitive functioning of the sector, 2024.06.28.
9. Baidu, Baidu Presents a Suite of Toolkits and Models to Supercharge AI Creativity at Create 2024, 2024.04.16.
10. Bmbf, Zwischenbericht: Ein Jahr KI-Strategie, 2019.11.15.
11. BMW Group, Successful test of humanoid robots at BMW Group Plant Spartanburg, 2024.08.06.
12. BMWK, Strategie Künstliche Intelligenz der Bundesregierung Fortschreibung 2020. 2020.12.02.
13. California Legislative, SB-1047 Safe and Secure Innovation for Frontier Artificial Intelligence Models Act.(2023-2024), 2024.9.3.
14. CB Insights, Game Changers 2025, 2024.9.
15. CIO Korea, 미국, 한국 등 28개국, ‘안전하고 책임있는 AI 개발’ 강조한 블레츨리 공동 선언 발표, 2023.11.02.
16. Cisco, Cisco’s 2024 AI Readiness Index: Urgency Rises, Readiness Falls, 2024.11.19.
17. Cisco, Cisco’s 2024 AI Readiness Index-South Korea, 2024.11.19.
18. CMA, AI Foundation Models: udate paper, 2024.04.11.
19. Constellation Research, Google Cloud rolls out new Gemini models, AI agents, customer engagement suite, 2024.9.24.
20. EC, EU and Canada boost their strategic digital partnership to address new challenges, 2024.2.1.
21. EDPB, Statement 3/2024 on data protection authorities’ role in the Artificial Intelligence Act framework, 2024.07.17.
22. European Commission, Coordinated Plan on Artificial Intelligence 2021 Review, 2021.4.21.
23. European Commission, European research development and deployment of AI,

2024.10.11.
24. European Commission, First Draft of the General-Purpose AI Code of Practice published, written by independent experts, 2024.11.14.
25. European Commission, Guidelines on the responsible use of generative AI in research developed by the European Research Area Forum, 2024.03.20.
26. European Data Protection Supervisor, EDPS Guidelines on generative AI: embracing opportunities, protecting people, 2024.6.03.
27. European Ombudsman, How the European Commission decides on and uses artificial intelligence, 2024.03.15.
28. Fedscoop, OpenAI, Anthropic enter AI agreements with US AI Safety Institute, 2024.08.29.
29. Figure AI, Figure unveils Figure 02, its second-generation humanoid, setting new standards in AI and robotics, 2024.08.06.
30. Fonearena, McAfee Project Mockingbird that can detect Deepfake audio demoed, 2024.01.12.
31. Forrester, Predictions 2025: An AI Reality Check Paves The Path For Long-Term Success, 2024.10.22.
32. Gartner, Intelligent Agents in AI Really Can Work Alone. Here’s How, 2024.10.1.
33. Generative AI Misuse: A Taxonomy of Tactics and Insights from Real-World Data.
34. Gigazine, OpenAI announces it has broken down GPT-4’s thoughts into 16 million interpretable patterns, 2024.06.07.
35. Global Times, OpenAI’s cut of China’s access to its API service would ’push domestic developers to catch up vigorously’, 2024.07.10.
36. Google Deepmind, AlphaGeometry: An Olympiad-level AI system for geometry, 2024.01.17.
37. Google Deepmind, Generating audio for video, 2024.06.17.
38. Google Deepmind, Introducing the Frontier Safety Framework, 2024.05.17.
39. Google Deepmind, Mapping the misuse of generative AI, 2024.08.02.
40. Google DeepMind, Shaping the future of advanced robotics, 2024.01.04.
41. Google Research, Introducing ASPIRE for selective prediction in LLMs, 2024.01.18.
42. Google, AlphaFold 3 predicts the structure and interactions of all of life’s molecules, 2024.05.08.
43. Google, Bard becomes Gemini: Try Ultra 1.0 and a new mobile app today, 2024.02.08.
44. Google, Gemini 1.5 Pro 2M context window, code execution capabilities, and Gemma 2 are available today, 2024.06.27.
45. Google, Gemini breaks new ground with a faster model, longer context, AI agents and more, 2024.05.14.
46. Google, Gemma 2 is now available to researchers and developers, 2024.06.27.

47. Google, How we’ re increasing transparency for gen AI content with the C2PA, 2024.09.17.
48. Google, How we’re addressing explicit fake content in Search, 2024.07.31.
49. Google, New generative media models and tools, built with and for creators, 2024.05.14.
50. Google, New in Gemini: Custom Gems and improved image generation with Imagen 3, 2024.08.28.
51. Google, Our next-generation model: Gemini 1.5, 2024.02.15.
52. Google, Smaller, Safer, More Transparent: Advancing Responsible AI with Gemma, 2024.07.31.
53. Gouvernement.fr , 25 recommandations pour l’IA en France, 2024.03.13.
54. Gov.uk, Assuring a Responsible Future for AI, 2024.11.06.
55. Gov.uk, National AI Strategy, 2022.12.18.
56. Gov.uk, Research on parent and pupil attitudes towards the use of AI in education, 2024.08.28.
57. Gov.uk, Teachers to get more trustworthy AI tech, helping them mark homework and save time, 2024.08.28.
58. GOV.UK, UK-Canada cooperation in AI compute: memorandum of understanding, 2024.1.13.
59. Government of Canada, Artificial Intelligence and Data Act, 2023.9.27.
60. Government of Canada, Canada launches Canadian Artificial Intelligence Safety Institute, 2024.11.12.
61. Government of Canada, Generative AI in your daily work, 2024.10.15.
62. Government of Canada, Guide on the Scope of the Directive on Automated Decision-Making, 2024.6.27.
63. Government of Canada, Guide on the use of generative artificial intelligence, 2024.10.15.
64. Government of Canada, Pan-Canadian Artificial Intelligence Strategy, 2022.7.20.
65. Government of Canada, Voluntary Code of Conduct on the Responsible Development and Management of Advanced Generative AI Systems, 2023.5.28.
66. IMDA, Singapore launches AI Verify Foundation to shape the future of international AI standards through collaboration, 2023.06.07.
67. IMF, Gen-AI: Artificial Intelligence and the Future of Work, 2024.01.14.
68. info.gouv.fr, 25 recommandations pour l’IA en France, 2024.03.13.
69. Kakao, 카카오, ‘if(kakaoAI) 2024’ 에서 그룹 AI 비전 공개…AI 메이트 ‘카나나’ 도 첫 선, 2024.10.22.
70. McAfee, McAfee Unveils Advanced Deepfake Audio Detection Technology at CES 2024 to Defend Against Rise in AI-Generated Scams and Disinformation, 2024.01.08.

71. Meta Training and Inference Accelerator.
72. Meta, Introducing Llama 3.1: Our most capable models to date, 204.07.23.
73. Meta, Introducing Meta Llama 3: The most capable openly available LLM to date, 2024.04.18.
74. Meta, Labeling AI-Generated Images on Facebook, Instagram and Threads, 2024.02.06.
75. Meta, Meet Your New Assistant: Meta AI, Built With Llama 3, 2024.04.18.
76. Meta, With 10x growth since 2023, Llama is the leading engine of AI innovation, 2024.08.29.
77. Microsoft, Discover the New Multi-Lingual, High-Quality Phi-3.5 SLMs, 2024.08.22.
78. Microsoft, Microsoft announces the largest investment to date in France to accelerate the adoption of AI, skilling and innovation, 2024.05.13.
79. Microsoft, VASA-1: Lifelike Audio-Driven Talking Faces Generated in Real Time, 2024.04.16.
80. Ministère de l'économie, La stratégie nationale pour l' intelligence artificielle. 2024.05.22.
81. Ministry of Electronics & IT, 6th meeting of the GPAI Ministerial Council held on 3rd July 2024 at New Delhi, 2024.07.03.
82. Ministry of Foreign Affairs of Japan, Prime Minister Kishida' s attendance at the Side Event on Generative AI at the OECD Ministerial Council Meeting, 2024.05.02.
83. MIT News, New algorithm discovers language just by watching videos, 2024.6.11.
84. National Artificial Intelligence Research Resource Task Force, Strengthening and Democratizing the U.S. Artificial Intelligence Innovation Ecosystem: An Implementation Plan for a National Artificial Intelligence Research Resource. 2023. 1.
85. Nature, AI models collapse when trained on recursively generated data, 2024.07.24.
86. NCSL, Artificial Intelligence 2024 Legislation, 2024.9.
87. NIST, Artificial Intelligence Risk Management Framework, AI RMF 1.0. 2023. 5.
88. NIST, U.S. AI Safety Institute Signs Agreements Regarding AI Safety Research, Testing and Evaluation With Anthropic and OpenAI, 2024.08.29.
89. NSF, Democratizing the future of AI R&D: NSF to launch National AI Research Resource pilot. 2024.1.24.
90. Nvidia, NVIDIA Blackwell Platform Arrives to Power a New Era of Com
91. OECD, AI, data governance and privacy: Synergies and areas of international co-operation, 2024.06.26.
92. OECD, Artificial intelligence and the changing demand for skills in the labour market, 2024.04.10.
93. OECD, Artificial intelligence and wage inequality, 2024.04.10.
94. OECD, Explanatory memorandum on the updated OECD definition of an AI

system, 2024.03.05.
95. OECD, Generative artificial intelligence in finance, 2023.12.15.
96. OECD, 'OECD Artificial Intelligence Review of Germany, 2024.06.11.
97. OECD, Recommendation of the Council on Artificial Intelligence, 2022. 5.
98. OECD, The potential impact of Artificial Intelligence on equity and inclusion in education, 2024.08.
99. OECD, Using AI in the workplace: Opportunities, risks and policy responses, 2024.03.15.
100. OMB, Advancing Governance, Innovation, and Risk Management for Agency Use of Artificial Intelligence, 2024.03.28.
101. OMB, Memorandum For The Heads Of Executive Departments And Agencies, 2024.03.28
102. OpenAI, C2PA in DALL·E 3, 2024.02.06.
103. OpenAI, Extracting Concepts from GPT-4, 2024.06.06.
104. OpenAI, GPT-4o mini: advancing cost-efficient intelligence, 2024.07.18.
105. OpenAI, Hello GPT-4o, 2024.05.13.
106. OpenAI, How OpenAI is approaching 2024 worldwide elections, 2024.01.15.
107. OpenAI, Introducing OpenAI o1-preview, 2024.09.12.
108. OpenAI, Introducing the GPT Store, 2024.01.10.
109. OpenAI, Navigating the Challenges and Opportunities of Synthetic Voices, 2024.03.29.
110. OpenAI, OpenAI acquires Rockset, 2024.06.21.
111. OpenAI, SearchGPT Prototype, 2024.07.25.
112. OpenAI, Superalignment Fast Grants, 2023.12.14.
113. OpenAI, Weak-to-strong generalization, 2023.12.14.
114. Parliament of Canada, Bill C-27, An Act to enact the Consumer Privacy Protection Act, the Personal Information and Data Protection Tribunal Act and the Artificial Intelligence and Data Act and to make consequential and related amendments to other Acts, 2023.4.24.
115. PDPC, Singapore's Approach to AI Governance, 2024.11.17. .
116. Prime Minister of Canada, Securing Canada's AI advantage, 2024.4.7.
117. Reuters, Meta to deploy in-house custom chips this year to power AI drive, 2024.02.02.
118. Sakana.ai, The AI Scientist: Towards Fully Automated Open-Ended Scientific Discovery, 2024.08.13.
119. Senate Dems, Majority Leader Schumer Delivers Remarks To Launch SAFE Innovation Framework For Artificial Intelligence At CSIS. 2023.6.21.
120. SiliconAngle, Baidu unveils Ernie 4.0 Turbo in a 'significant upgrade' to its AI chatbot, 2024.06.28.

121. SiliconAngle, French open-source AI model startup Mistral AI raises $640M at $6B valuation, 2024.06.11.
122. SiliconAngle, Meta faces AI accuracy issues as tech industry tackles hallucinations, deepfakes, 2024.07.31.
123. Smartnation, Factsheet - Artificial Intelligence (AI) initiatives launched to uplift Singapore's economic potential, 2024.03.01.
124. Smartnation, National AI Strategy, 2019.11.30.
125. Smartnation, Singapore National AI Strategy 2.0 (NAIS 2.0), 2023.12.05.
126. Stability.ai, Introducing Stable LM 2 1.6B, 2024.01.19.
127. Stability.AI, Stable Diffusion 3, 2024.02.22.
128. Stanford HAI, AI Index Report 2024.
129. TechCrunch, Google launches Gemini Ultra, its most powerful LLM yet, 2024.02.08.
130. TechCrunch, OpenAI buys Rockset to bolster its enterprise AI, 2024.06.21.
131. TechCrunch, OpenAI created a team to control 'superintelligent' AI — then let it wither, source says, 2024.05.18.
132. TechRadar, Jarvis AI is real - Google accidentally leaks its AI agent that browses the web for you, 2024.11.6.
133. The federal government, AI - a brand for Germany, 2018.11.15.
134. The New York Times, California Legislature Approves Bill Proposing Sweeping A.I. Restrictions, 2024.08.28.
135. The New York Times, U.S. Clears Way for Antitrust Inquiries of Nvidia, Microsoft and OpenAI, 2024.06.05.
136. The Verge, Gemini 1.5 is Google's next-gen AI model — and it's already almost ready, 2024.02.16.
137. The Verge, Google says Gemini AI is making its robots smarter, 2024.07.11.
138. The Verge, Nvidia reveals Blackwell B200 GPU, the 'world's most powerful chip' for AI, 2024.03.19.
139. The Verge, OpenAI announces SearchGPT, its AI-powered search engine, 2024.07.26.
140. The Wall Street Journal, There's a Tool to Catch Students Cheating With ChatGPT. OpenAI Hasn't Released It., 2024.08.04.
141. The White House, FACT SHEET: Biden-Harris Administration Outlines Coordinated Approach to Harness Power of AI for U.S. National Security, 2024.10.24.
142. Tortois Intelligence. The Global Artificial Intelligence Index, 2024.09.19.
143. UK AI Safety Institute, AI Safety Institute releases new AI safety evaluations platform, 2024.05.10.
144. UK AI Safety Institute, Collaboration on the safety of AI: UK-US memorandum of understanding, 2024.04.02.

145. UK AI Safety Institute, International Scientific Report on the Safety of Advanced AI, 2024.05.17.
146. UK Cabinet Office, Generative AI Framework for HMG, 2024.01.18.
147. UNESCO, Foundation Models such as ChatGPT through the prism of the UNESCO Recommendation on the Ethics of Artificial Intelligence, 2023.6.
148. UNESCO, Generation AI: Navigating the opportunities and risks of artificial intelligence in education, 2024.07.22.
149. United Nations AI Advisory Body, Governing AI for Humanity - Final Report, 2024.09.19.
150. United Nations(General Assembly), Seizing the opportunities of safe, secure and trustworthy artificial intelligence systems for sustainable development, 2024.3.11.
151. Venturebeat, Apple researchers achieve breakthroughs in multimodal AI as company ramps up investments, 2024.03.15.
152. VentureBeat, Apple researchers develop AI that can ‘see’ and understand screen context, 2024.04.01.
153. Venturebeat, DeepMind researchers discover impressive learning capabilities in long-context LLMs, 2024.04.24.
154. VentureBeat, Elon Musk’s xAI defies ‘woke’ censorship with controversial Grok 2 AI release, 2024.08.14.
155. Venturebeat, iOS gets an AI upgrade: Inside Apple’s new ‘Intelligence’ system, 2024.07.29.
156. Venturebeat, Meta’s Self-Taught Evaluator enables LLMs to create their own training data, 2024.08.19.
157. Venturebeat, Stable Diffusion 3.0 debuts new diffusion transformation architecture to reinvent text-to-image gen AI, 2024.02.22.
158. whitehouse, Fact Sheet: Biden-Harris Administration Announces Key AI Actions Following President Biden’s Landmark Executive Order. 2024.1.19.
159. whitehouse, FACT SHEET: Biden-Harris Administration Announces New AI Actions and Receives Additional Major Voluntary Commitment on AI. 2024.7.26.
160. Whitehouse, Fact Sheet: Key AI Accomplishments in the Year Since the Biden-WJ Harris Administration’s Landmark Executive Order, 2024.10.30.
161. Wired, Why Elon Musk Had to Open Source Grok, His Answer to ChatGPT, 2024.03.11.
162. xAI, Announcing Grok-1.5, 2024.03.28.
163. xAI, Grok-1.5 Vision Preview, 2024.04.12.
164. xAI, Grok-2 Beta Release, 2024.08.13.
165. xAI, Open Release of Grok-1, 2024.03.17.
166. ZDNet, Apple claims its on-device AI system ReaLM ’substantially outperforms’ GPT-4, 2024.04.02.

167. 한국경제, SKT “기가와트급 데이터센터 조성…’AI 인프라 슈퍼 하이웨이’ 만든다“, 2024.11.04.
168. 한국과학기술기획평가원(KISTEP), 국가전략기술 R&D(연구・개발) 인력 실태조사, 2024.5.
169. 한국교육학술정보원(KERIS), 생성형 AI를 활용한 교수학습 운영 가이드, 2024.5.
170. 한국교육학술정보원(KERIS), 생성형 AI를 활용한 교수학습 프레임워크 개발 연구, 2024.1.
171. 한국전자통신연구원, ETRI, AI안전 및 신뢰성지원 국제표준 제정, 2024.07.16.
172. 人民邮电报, 工信部拟筹建人工智能标准化技术委员会, 2024.07.05.
173. 全国网络安全标准化技术委员会, 《人工智能安全治理框架》1.0版, 2024.0.09.
174. 内閣府, AI に関する暫定的な論点整理, 2023.5.26.
175. 内閣府, AIセーフティ・インスティテュートの設立について, 2024.2.14.
176. 内閣府, 統合イノベーション戦略 2023, 2023.6.9.
177. 内閣府, 統合イノベーション戦略 2024, 2023.6.4.
178. 内閣府, 統合イノベーション戦略 2024, 2024.6.4.
179. 国家互联网信息办公室关于, 人工智能生成合成内容标识办法（征求意见稿, 2024.09.14.
180. 工业和信息化部, 国家人工智能产业综合标准化体系建设指南（2024版）, 2024.06.05.
181. 广东省人民政府, 关于人工智能赋能千行百业若干措施的通知, 2024.06.06.
182. 成都市出台了《成都市人工智能产业高质量发展三年行动计划（2024—2026年）》, 2024.06.20.
183. 湖南省工业和信息化厅, 《湖南省人工智能产业发展三年行动计划(2024-2026年)》, 2024.09.04.
184. 科技日报, 培育新质生产力在行动丨培育千亿级人工智能产业创新集群 陕西发布行动计划推进五大工程, 2024.04.08.
185. 経団連, AI政策・制度検討の動向, 2024.9.26.
186. 経済産業省, AI事業者ガイドライン（第1.0版）」を取りまとめました, 2024.4.19.
187. 経済産業省; 総務省, AI 事業者ガイドライン（第 1.0 版）, 2024.4.19.
188. 防衛省, 防衛省ＡＩ活用推進基本方針, 2024.7.
189. 首相官邸, ＡＩ戦略会議・ＡＩ制度研究会合同会議, 2024.8.2.

[첨부 1] 2025년 주요국 AI 관련 분야별 정책 동향 요약

<표> 2025년도 주요국의 AI 관련 분야별 정책 동향 요약

국가	정책/법제	기업/산업	기술/연구	인력/교육
미국	• 미국의 AI 리더십의 장벽 제거(EO 14179) - 과도한 규제 완화와 AI 혁신 가속화로 미국의 글로벌 AI 패권 유지 전략 제시 • 미국 AI 실행계획 - 규제 장벽 제거, AI 혁신 가속, 초거대 모델·컴퓨팅 인프라 확충, 오픈소스·민간 주도 생태계 강화, 국가안보·글로벌 AI 리더십 확보를 핵심 축으로 한 범정부 실행 로드맵	• 데이터센터 인프라 연방 허가절차 가속화(EO 14318) - 연방 토지 임대 및 특별지원으로 대규모 프런티어 AI데이터센터 신속 구축 추진	• AI 표준혁신센터(CAISI) 설립 - NIST 산하 AI 안전연구소를 표준·혁신 중심으로 재편하여 AI 기술 표준화와 혁신 지원 강화	• AI 교육 행정명령(EducateAI) - K-12 AI 교육 커리큘럼 개발 지원 및 재직자 대상 AI 기술 재교육 프로그램 확대 등을 통해 AI 인재풀 확대
EU	• AI 법(AI Act) - 세계 최초 포괄적 AI 규제법으로 위험 기반 접근을 통해 신뢰성·안전성 확보를 추구함	• AI 대륙 실행계획 - 컴퓨팅 인프라 구축, SME 지원 'AI 팩토리' 등 5대 전략으로 유럽 AI 산업 경쟁력 및 기술주권 강화 추진	• AI 팩토리·기가 팩토리 추진 - 스타트업·중소기업·연구자 지원 HPC 허브인 'AI 팩토리' 13개 구축 및 초거대 모델 개발 위한 'AI 기가팩토리' 최대 5개 건설 계획	• AI 스킬 아카데미 출범 - AI 분야 고급 기술교육 및 재교육 프로그램을 제공하여 EU 내 AI 인재 육성가속화
중국	• 'AI 플러스' 행동계획 - 2027년까지 과학기술·산업·민생 등 6대 분야에 AI 심층 융합을 실현하고 지능형 경제로 도약하는 로드맵 제시	• 컴퓨팅 자원 상호연결 행동계획 - 국가 초거대 컴퓨팅 인프라의 효율적 연계를 추진하여 AI 산업 기반 컴퓨팅 역량 강화	• AI 과학기술 윤리 관리서비스(시범) - 대학·연구기관 등의 AI 연구활동에 윤리심사 절차를 도입해 AI 연구 윤리준수 체계화	-

국가	정책/법제	기업/산업	기술/연구	인력/교육
일본	• AI 추진법(AI 관련 기술 R&D 및 활용 촉진법) - 2025년 제정된 일본 최초 AI 기본법으로 AI 개발 촉진 및 안전 확보 위한 정부 권한·거버넌스 강화	• AI 이용·개발 계약 체크리스트 - 기업 간 AI 개발·이용 계약 시 발생할 수 있는 이익·위험 배분 등 주요 사항을 정리한 지침을 마련하여 안전한 AI 활용 촉진	-	-
영국	• AI 기회 실행계획 - 2025년 1월 발표된 국가 AI 전략 이행계획으로 컴퓨팅 인프라 확보, 공공부문 AI 신속 도입, 핵심 분야 AI 챔피언 육성 등 50개 권고안 제시	• AI 성장특구(AI Growth Zone) - 지역별 AI 산업 클러스터 조성 추진으로 민간의 AI 혁신 투자와 실증을 촉진 (2025년 첫 후보지 선정)	• AI 안전연구소 역량확대 - 첨단 AI 안전 연구기관에 예산 확보 및 법정기구 지위 부여를 통해 프론티어 AI 위험 대응과 기술표준 개발 주도	• AI 인재전략 - AI 전공 졸업생 확대, 해외 인재 유치 장학제도신설, 다양성 제고 등으로 AI 인력풀 강화하고 기술 혁신 인재 확보
캐나다	• 연방 공공서비스 AI 전략 - 2025~27년 시행되는 캐나다 최초의 공공부문 AI 종합전략으로 중앙 AI 역량 구축, 정책·법률 정비, 인재 양성, 투명성 제고 등 4대 목표 수립	-	• 주권 AI 컴퓨팅 전략 - 2024년부터 5년간 20억 CAD 투자로 첨단 AI 컴퓨팅 인프라에 대한 국내 접근성 확보 및 AI 연구·산업 경쟁력 강화	• 청정에너지 분야 AI 인재·연구 투자 - 2025년 9월 캐나다 정부가 청정에너지 AI 인력 양성및 연구역량 확충에 2,870만 달러 투자 발표
프랑스	• CNIL 2025-2028 전략계획 - AI 시스템 라이프사이클별 준수 기준및 감시 기술 아키텍처 개발을 추진하고, 2025년 하반기 AI 책임 프레임워크에 대한 권고안 마련 예정	-	• 국립 AI 평가·안전 연구소(INESIA) - 2025년 설립된 AI 신뢰성·안전성 평가전문 연구기관으로, 산하 기관 연합을 통해 AI 모델 성능 검증 및 정책 자문 수행	• 해외 AI 연구자 유치 프로그램 - INRIA 주관으로 해외 AI 인재 채용 시 연구비·정착 지원(가족 동반 이주 지원 등)을 제공하여 세계 최고급 AI 인재 장기 유치추진

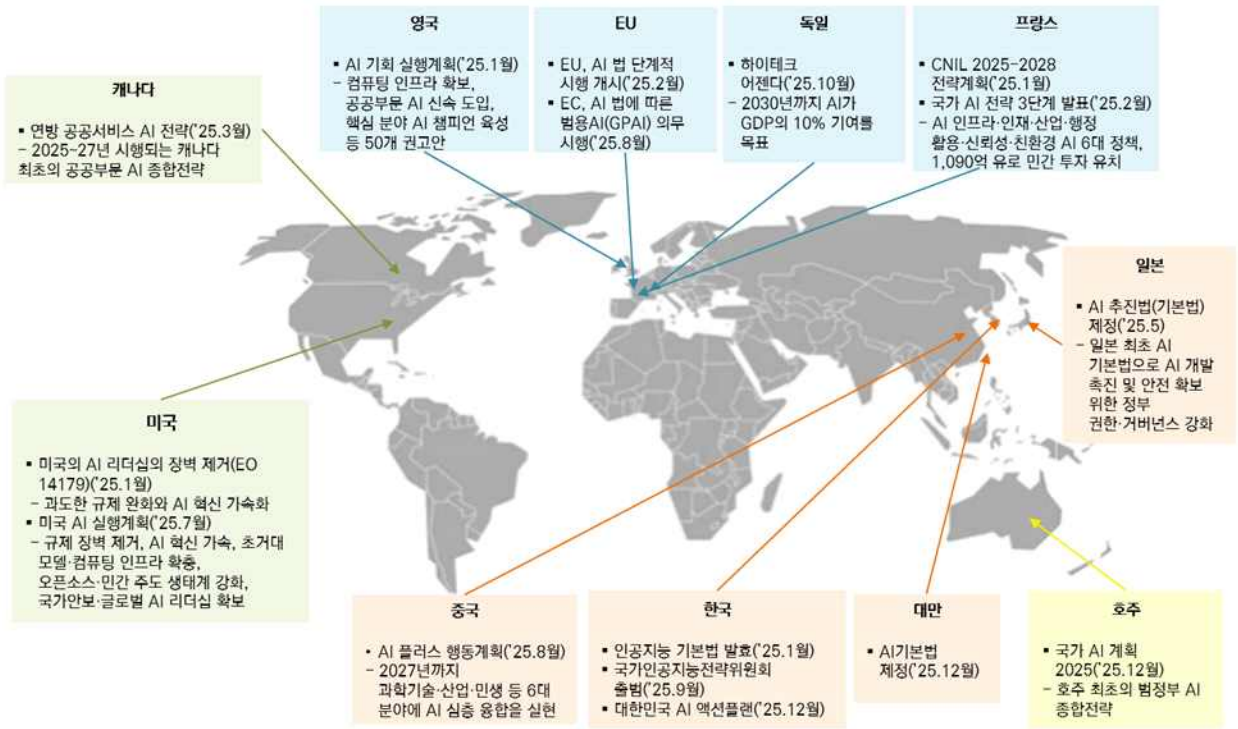
국가	정책/법제	기업/산업	기술/연구	인력/교육
독일	• 하이테크 어젠다(Hightech Agenda 2025) - 2030년까지 AI가 GDP의 10% 기여를 목표로 하는 AI 가속화 정책발표(AI 혁신 Offensive)	• 디지털국가현대화부(BMDS) 신설 - 연방 차원의 AI·디지털 컨트롤타워를 설치해 분산된 기능을 통합하고 AI 정책의 일관된 추진 및 조율 강화	• AI 기가팩토리 건설 계획 - EU 초대형 HPC 센터 입찰 참여를 통해 독일 내 AI 슈퍼컴 센터(기가팩토리) 구축 추진(도이치텔레콤과 협력)	-
싱가포르	• 국가 AI 전략 2.0 - Tusted AI를 핵심 가치로 AI 거버넌스를 강화하고 산업 전반의 AI 활용 촉진을 도모하는 차세대 국가 AI 계획 추진	• 글로벌 AI 신뢰보증 샌드박스 - 싱가포르 IMDA와 AI Verify 재단이 기업의 생성형 AI 모델을 안전하게 테스트·검증할 수 있는 글로벌 협력 이니셔티브 출범	• 하이브리드 양자-고전 컴퓨팅 이니셔티브 - 약 S$2,450만을 투자해 양자컴퓨팅과 HPC/AI 인프라 통합 기술(HQCC 1.0)을 개발하고 금융·물류 등 분야 적용 촉진	-
이스라엘	• 국가 AI 본부 설립 - 2025년 10월 총리실 산하에 국가 AI 전략 이행을 총괄할 AI 본부 설립을 승인하여 AI를 경제·안보·사회 전반 핵심역량으로 전환하는 거버넌스 구축	-	-	-
인도	• 선진 인도를 위한 AI 로드맵('24.3월) - 2025년 「India AI Mission」 성과를 토대로 AI 인프라 확충, 데이터 플랫폼 구축 등 AI 생태계 전반의 발전을 추구하는 정책 방향 제시	-	• 토종 AI 인프라 및 모델 육성 - 38,000개 GPU 기반 초거대 AI 연구 인프라구축 등 ₹10,300억 투자로 자국산 파운데이션 모델개발과 기술 자립 가속	• AI 교육 CoE 설립 - 2025년 예산에 교육 분야 AI 센터에 50억 루피 배정, 총 4개의 AI 센터통해 인도 내 AI 인재 양성과 생태계 구축 가속

국가	정책/법제	기업/산업	기술/연구	인력/교육
UAE	• AI 기반 규제 지능 생태계 - 입법 과정 전반에 AI 기술을 통합하여 전세계 법률·정책 데이터를 실시간 분석하는 스마트 입법 시스템구축 및 규제의 선제적 대응 도모	-	-	-
대만	• AI 기본법 제정 - 2025년 8월 제정된 AI 개발 및 응용 촉진을 위한 기본법으로 우호적 AI 환경 조성과 대만의 AI 허브화 기반 마련	• 산업혁신조례 개정 - 2025년 4월 투자 세액공제 대상에 'AI'를 추가하고 핵심기술 해외투자 심사를 강화하여 AI 도입 촉진및 기술유출 방지 도모	• AI 혁신투자 강화계획 - 2024년 수립된 프로그램으로 향후 10년간 NT$100억을 투자해 기술력 검증 완료된 AI 스타트업의 스케일업및 신속한 상용화 지원	• AI 인재 육성 계획 - 2025년부터 4년간 AI 전문인력 20만 명양성을 목표로 NT$10억 투자 (세제 혜택, 산학협력, 경진대회 등으로 청년층 AI 역량 강화)
호주	• 국가 AI 계획 2025 - 2025년 12월 발표된 최초의 범정부 AI 종합전략으로 AI 선도기술 투자, 사회 전반 AI 혜택 확산, 신뢰할 수 있는 AI 구축 등 3대 목표 수립	• 첨단 AI 인프라 투자 촉진 - 고성능 데이터센터 등 AI 인프라에 대한 민간 투자 유치를 장려하고 AI 활용으로 산업 경쟁력과 생산성 향상 도모	• 대규모 데이터 개방 - 정부·민간 보유 데이터셋의 공유를 촉진하여 AI 모델 개발과 혁신을 가속하고 기술 발전을 지원하는 환경 구축	• AI 인력 재스킬링 - AI로 인한 일자리 변화에 대비해 근로자 재교육을 강화하고 근로자 역량 향상 및 생산성 이익의 공정 분배를 유도
한국	• 인공지능 기본법 - 2024년 12월 국회 통과, 2025년 1월 시행된 국내 AI 분야 거버넌스 기본법으로 AI 산업 육성, 위험 예방 등 제도 기반을 확립	• 국가 AI 역량 강화 방안 - 2025년 2월 발표된 AI G3 도약 전략으로 AI 컴퓨팅 인프라 확충, 차세대 AI 모델 개발, AI+X 가속화 등 3대 전략과 6대 과제로 산업 경쟁력 강화추진	• World Best LLM 프로젝트 - 세계 최고 수준의 초거대 언어모델개발을 목표로 정예 연구팀을 선발·지원하고 글로벌 AI 챌린지 개최 등을 추진	• AI 인재양성 종합대책 - 초중등 AI 교육시간 확대('25년 730교→'28년 2,000교), 대학 AI 교육과정 신설·확산, 재직자 AID 집중과정 확대 등 생애주기별 AI 인력 양성 강화

[첨부 2] 2025년 주요국 AI 정책 전 세계 맵(World Map)

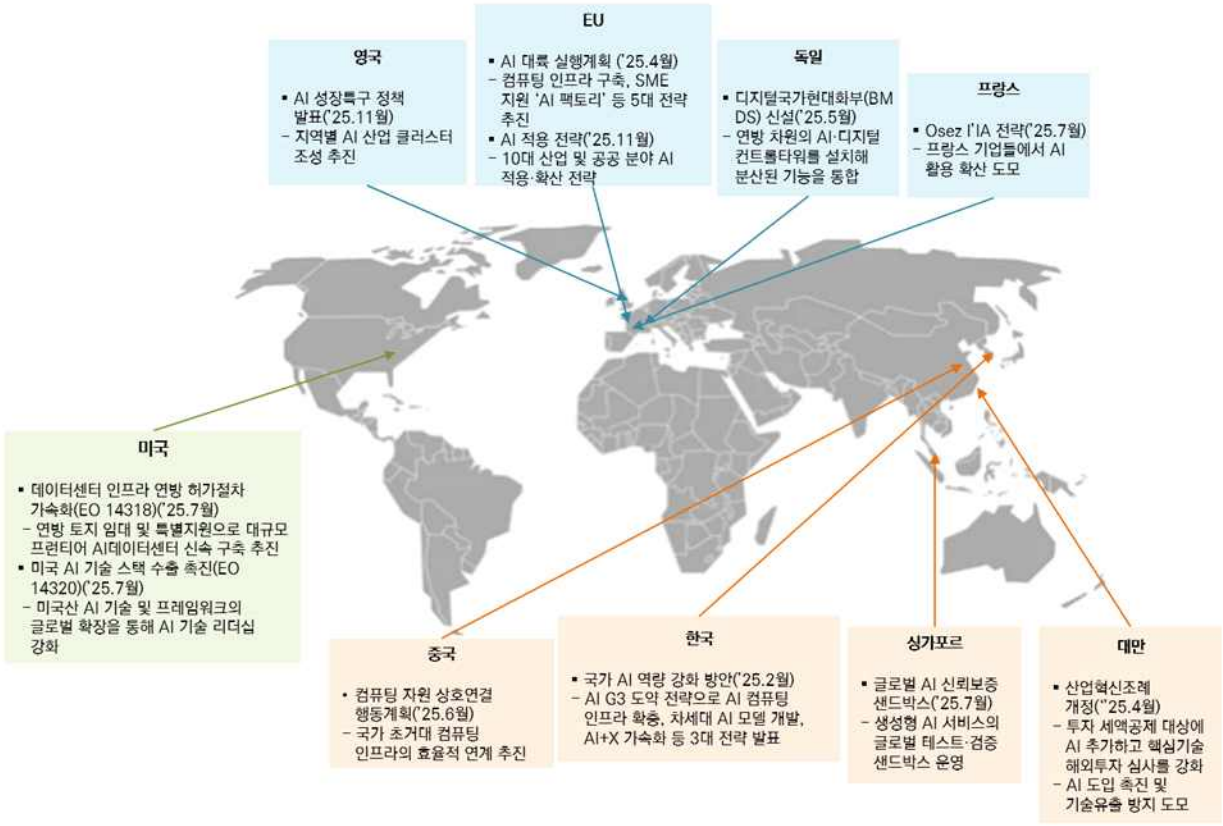
① 정책/법제

[그림] 2025년 주요국 AI 정책(정책/법제 분야) 전 세계 맵(World Map)



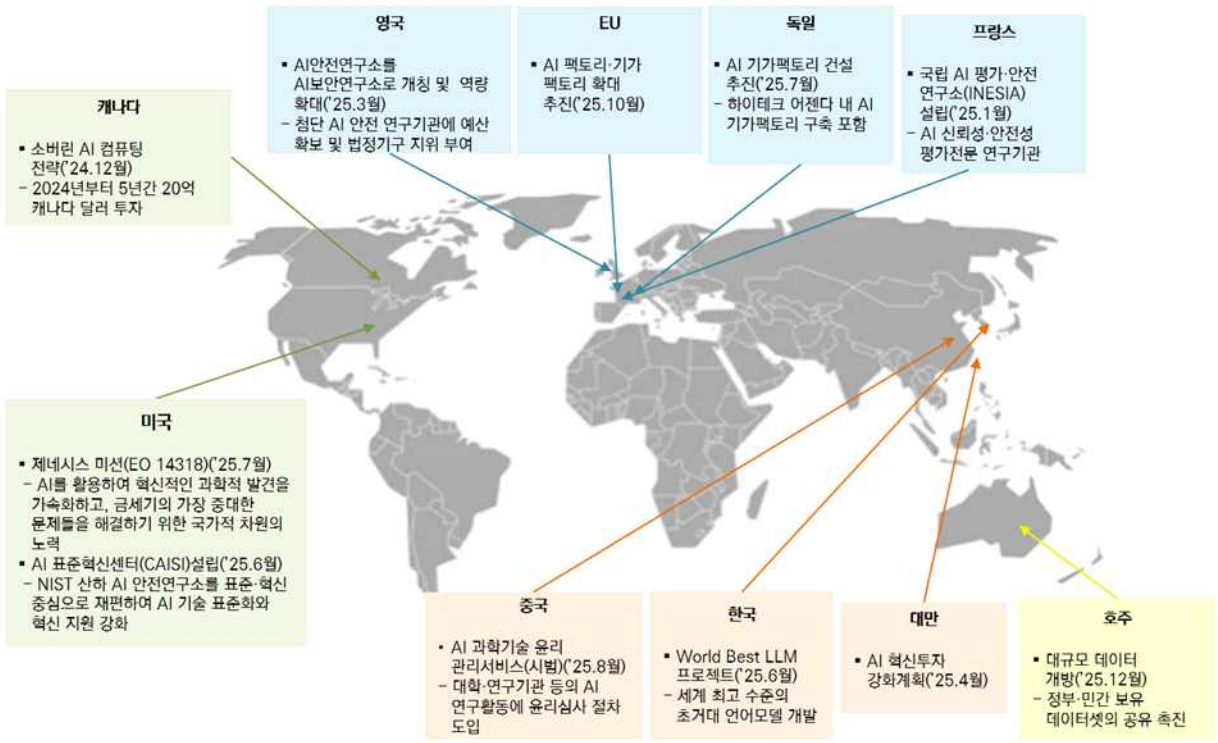
② 기업/산업

[그림] 2025년 주요국 AI 정책(기업/산업 분야) 전 세계 맵(World Map)



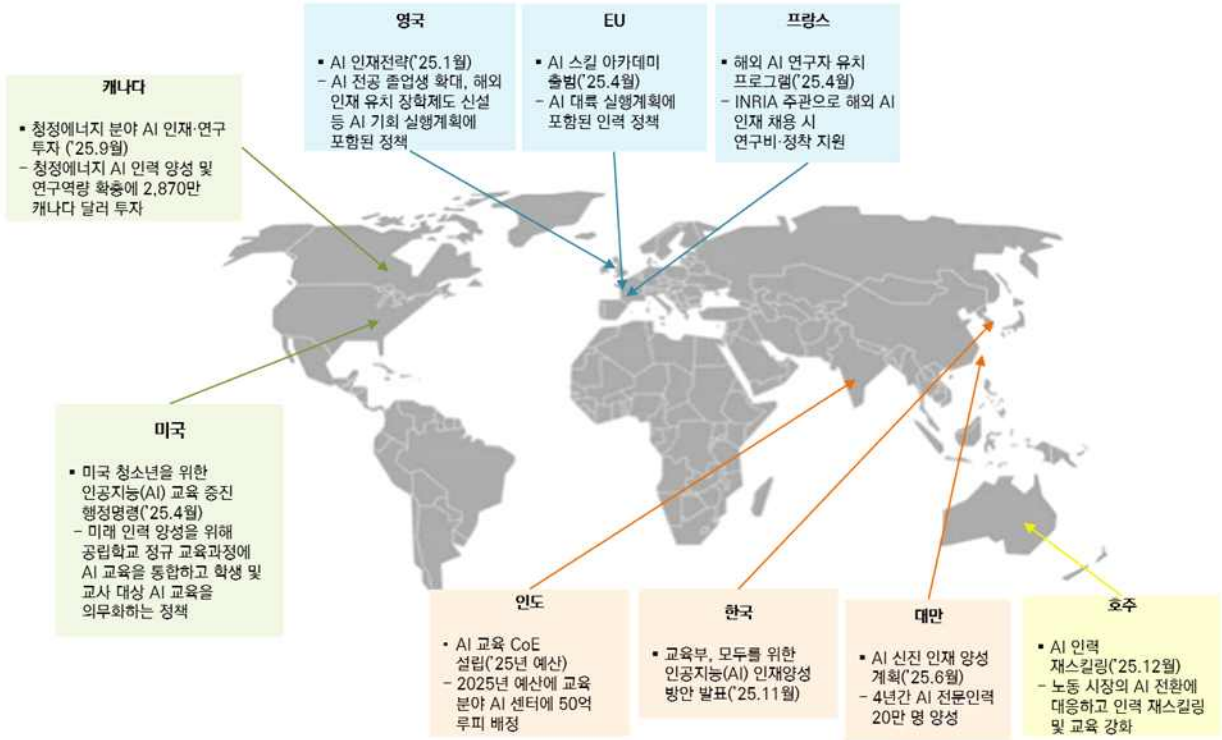
③ 기술/연구

[그림] 2025년 주요국 AI 정책(기술/연구 분야) 전 세계 맵(World Map)



④ 인력/교육

[그림] 2025년 주요국 AI 정책(인력/교육 분야) 전 세계 맵(World Map)



[첨부 3] 국내외 주요 기업 동향

① 국내 AI 기업 동향

<표> 국내 AI 기업 동향

가치사슬	기업명		주요 활동 (기술·모델 개발, 투자유치(인수합병), 사업전략(협업제휴), 제품/서비스 출시 등)
애플리케이션 · 서비스	콘텐츠 제작	뤼튼테크놀로지	- (서비스) 대화형 챗봇 '챗 뤼튼' 출시('23.3), AI 챗봇(나만의 AI)에 AI 검색 기능 업데이트2('23.8), AI 캐릭터 채팅 서비스 '크랙' 출시('25.4), GPT-5 서비스 신규 업데이트('25.8) - (플랫폼) AI 플랫폼 '뤼튼 2.0' 출시('23.4), '뤼튼 3.0' 출시('25.4) - (투자) 프리시리즈 B 투자 250억 원 유치('24.6), 830억 원 추가 투자 유치('25.3) - (활동) AX전환사업 자회사 '뤼튼AX' 설립('25.9)
		이스트소프트	- (서비스) 구독형 AI 휴먼 영상 제작 'AI 스튜디오 페르소'와 대화형 AI 휴먼 '페르소 라이브' 출시('24.5), 보고서 생성과 추론 특화 모델 '앨런 LLM' 출시('25.6), AI 검색엔진 '앨런' 구독 서비스 출시('25.7) - (제휴) 일본 LLM 응용 서비스 기업 스파이럴A와 'AI 휴먼 공동 사업' MOU 체결('24.2), 네이버클라우드와 대학 AI 교육 플랫폼 구축('24.9), 콘텐츠 AI 기업 포바이포와 AI 기술과 비즈니스 협력 체결('25.5)
		스캐터랩	- (서비스) 엔터테인먼트·콘텐츠·커머스 시장용 sLLM '핑퐁-1' 공개('23.11), AI 스토리 플랫폼 '제타' 출시('24.4), 제타 회원 200만 돌파('25.4)
		라이너	- (서비스) 생성 AI 검색 서비스 출시('23.6), 학술 모드 AI 검색 베타 서비스 출시('24.7), 국내 첫 '딥 리서치' 출시('25.3), 신규 리서치 AI 에이전트 '피어 리뷰'와 '설문 시뮬레이터' 공개('25.9), '논문 퀵리뷰' 출시('25.12) - (기술) 카이스트와 AI 에이전트 신뢰도 측정하는 벤치마크 연구('24.6) - (투자) 270억 원 규모 시리즈 B2 투자 유치('24.10) - (제휴) 비씨카드와 AI 검색 강화 제휴('25.7)
		네오사피엔스	- (투자) 165억 규모 프리 IPO 투자 유치('25.12) - (제휴) 교육 스타트업 단꿈아이와 제휴해 실시간 AI 튜터 공개('25.9)
		라이언로켓	- (서비스) AI 웹툰 도구 '젠버스 알파' 출시('25.1), AI 기반 캐릭터 시각화 서비스 '젠버스 라이트' 출시('25.5)
		딥브레인AI	- (플랫폼) 딥페이크 자동 탐지 솔루션 공개('24.11), AI 영상 합성 플랫폼 'AI 스튜디오 4.0' 공개('25.6) - (서비스) AI 스튜디오 신기능으로 이미지 생성 기능과 광고 제작 자동화 기능 출시('25.9) - (제휴) 서울사이버대학과 'AI 교수' 개발 MOU 체결('24.9), 일본 지상파 뉴스에 AI 아나운서 제공('24.11)
	모빌리티	포티투닷	- (기술) 베트남에 R&D 센터 설립('25.7) - (서비스) 카타르에서 로보택시 실증('25.2) - (투자) 현대차·기아, 포티투닷에 5000억 '전략적 출자('25.8)
		Autonomous A2Z	- (서비스) '새벽동행 자율주행버스' 서비스('24.11), 자율주행 브랜드 'AZTON' 론칭('25.7) - (제휴) 싱가포르 모빌리티 플랫폼 '그랩'과 제휴해 동남아 시장 자율주행차 운영 공동 연구('25.3), UAE 스페이스테크 기업 '스페이스42'와 현지 합작법인 설립('25.11)
		스트라드비젼	- (제품) AI 기반 영상인식 SW 'SVNet' 탑재 누적 차량대수 400만 대 돌파('25.8) - (서비스) 데이터 관리 워크플로우를 고도화한 'SVDataFlow' 파이프라인 발표('25.9) - (제휴) AI 시스템온칩(SoC) 공급업체 '악세라'와 자율주행 혁신 가속화 제휴 체결('25.4), AMD와 AI 기반 인식 기술 고도화 협력('25.11)

가치사슬	기업명		주요 활동 (기술·모델 개발, 투자유치(인수합병), 사업전략(협업제휴), 제품/서비스 출시 등)
	제조	포스코DX	- (서비스) 환경부와 협력해 AI 기반 동물 찻길 사고 예방 및 모니터링 시스템 시범 운영('25.4) - (제휴) AWS와 AI 엔지니어 에이전트 개발 MOU('25.7), 금호타이어와 스마트 타이어 공동 개발 제휴('25.9) - (투자) 미국 산업용 휴머노이드 로봇 기업 '페르소나 AI'에 포스코기술투자와 함께 300만 달러 투자('25.12)
		마키나락스	- (플랫폼) AI EXPO KOREA 2025에서 경량화 AI 플랫폼 런웨이 공개('25.5) - (사업/제휴) HD현대와 용접 로봇용 AI 솔루션 개발 MOU('25.6), 독자 AI 모델 업스테이지 컨소시엄 참여('25.8), 제조 AX 얼라이언스 참여('25.9), 이삭엔지니어링과 업무협약('25.11)
		인터엑스	- (모델) 자율제조 특화 AI '젠AI' 공개('24.10), 사출성형 제조 특화 AI '젠엑스' 공개('25.5) - (서비스) 제조 문서 분석 솔루션 'Document.AI' AWS 마켓플레이스 출시('25.8), 산업 AI EXPO에서 통합 AI 솔루션 공개('25.12)
	데이터/인프라	슈퍼브AI	- (모델) 산업 특화 비전 파운데이션 모델 '제로' 공개('25.6) - (서비스) 영상 관제 에이전트 '슈퍼브 영상관제 솔루션' 공개('25.3) - (협업/제휴) 독자 AI 모델 LG AI연구원 컨소시엄 참여('25.8), K-휴머노이드 연합 참여('25.9) - (투자) 140억 규모의 프리 IPO 투자 유치('25.12)
		래블업	- (플랫폼) 생성 AI 개발 및 훈련 플랫폼 '백엔드닷에이아이 v25.3' 업데이트 및 클라우드 솔루션 '백엔드닷에이아이 컨티넘' 출시('25.3), 생성 AI 개발 플랫폼 'AIAI:DOL(AI Deployable Omnimedia Lab)' 베타 출시('25.9)
		프렌들리AI	- (서비스) AI 개발용 GPU 최적화 서비스 '서버리스 엔드포인트' 출시('24.1), 기업 맞춤형 AI 개발 서비스 '프렌들리 파인튜닝' 출시('24.10) - (플랫폼) 딥시크 AI 모델 사용할 수 있는 서버리스 API 출시('25.2), 자체 추론 플랫폼에 LG AI연구원 '엑사원 4.0' 서비스 연동('25.7), AI 모델 비교 플랫폼 'WBA' 출시('25.8) - (투자) 275억원 규모 시드 익스텐션 라운드 투자 유치('25.8) - (제휴) 허깅페이스와 제휴해 생성형 AI 모델 서빙 솔루션 '프렌들리 데디케이트 엔드포인트' 제공('25.1), 엔비디아 최신 AI 모델 '네모트론 3'을 프렌들리AI플랫폼에서 배포('25.12)
		노타	- (플랫폼) 온디바이스 AI 개발 플랫폼 '넷츠프레소' 출시('22.4) - (서비스) 실시간 영상관제 솔루션 'NVA' 정식 출시('25.9), LLM 최적화 서비스 출시('25.9), 소아·청소년 호흡음 분석 AI 진단 솔루션 개발('25.11) - (제휴/협업) 온디바이스 생성 AI로 퀄컴 AI 허브 지원('25.2), 두바이 교통국과 생성형 AI 기반 지능형 교통체계(ITS) 솔루션 공급 계약 체결('25.4), 딥엑스와 온디바이스 AI 시장 제휴('25.5), 윈드리버와 온디바이스 생성 AI 제휴('25.5), 르네사스와 초소형 엣지 AI 기반 운전자 모니터링 솔루션 개발('25.7), 'NVA'의 동남아 시장진출 목적으로 킬사글로벌과 협업('25.9), 텔레칩스의 '돌핀 5'에서 안면인식 AI모델 경량화 및 최적화 기술 적용 협력('25.11), 퓨리오사AI와 NPU 기반 AI 기술협력을 위한 업무협약 체력('25.12) - (투자) 300억 원 규모 시리즈C 투자 유치('24.7)

가치사슬	기업명		주요 활동 (기술·모델 개발, 투자유치(인수합병), 사업전략(협업제휴), 제품/서비스 출시 등)
	엔터프라이즈	삼성SDS	- (서비스) 생성 AI 서비스 '패브릭스'와 '브리티 코파일럿' 출시('24.5), 아시아 최대 보험 컨퍼런스 PIC에서 AI 동시통역 서비스 제공('25.9) - (사업) 에이전트 중심의 AI 풀스택 전략 고도화 발표('25.9)
		데이원컴퍼니	- (서비스) 몰입형 AI 학습 콘텐츠 'AI 월드' 출시('25.9) - (협업) 독자 AI 모델 업스테이지 컨소시엄 참여('25.8) - (사업) 미국 법인 설립('25.4), 음악 콘텐츠 스타트업 'NOA' 인수('25.7)
		오케스트로	- (서비스) 생성 AI 솔루션 '클라리넷' GS 인증 1등급 획득('25.8), AI GovTech인 생성형 AI 솔루션 'G-AIDSP', 가상화 솔루션 '콘트라베이스', 클라우드 네이티브 운영관리 플랫폼 '비올라' 공개('25.12) - (사업/제휴) 생성 AI 스타트업 라이프로그 인수('24.10), 독자 AI 모델 업스테이지 컨소시엄 참여('25.8), SR과 철도 운영 분야 AI 도입 제휴('25.9), 마이그레이션·재해복구 전문기업 '데이타커맨드' 인수('25.10), 연세대 AI혁신연구원과 AI기술 공동연구협약 체결('25.12)
		롯데이노베이트	- (서비스) 기업용 AI 플랫폼 '아이멤버'에 추론 기능 탑재('25,3), 사내용 AI 플랫폼 '아이멤버' 대외서비스 출시('25.4), 기업용 AI 플랫폼 '아이멤버 3.0' 공개('25.7), 제주도 관광형 자율주행 서비스 '일출봉Go!' 운행 ('25.9), 롯데백화점 AI 컨시어지 서비스 '더스틴' 구축('25.12)
		콕스웨이브	- (서비스) AI 챗봇 구축・운영・분석 지원하는 AI 엔터프라이즈 솔루션 출시('24.2), AI 튜터 플랫폼 '에듀탭' 출시('25.9.) - (제휴) 인도 AI 기업 펀다멘토와 콜센터 AI 전환 솔루션 개발 위한 전략적 파트너십 체결('24.7), 인도 AI 기업 레난과 AI 공동 개발 MOU('25.4), 기업용 AI 기술 고도화 위해 투플랫폼과 전략적 협업('25.6), 방산기업 퀀텀에어로와 국방 AI 역량 강화를 위한 기술 개발 업무협약 체결('25.10) - (투자) 45억 원 규모의 시드 투자 유치('24.1)
		몰로코	- (제휴) SKT와 협력해 AI 광고 플랫폼 '어썸 2.0 '출시('24.6), 무신사와 AI 추천기술 고도화 MOU('25.1) - (사업) AI 광고 솔루션으로 놀유니버스/당근 통합 마케팅 지원('25.7)
	생산성	한컴	- (서비스) 문서 요약과 작성, 번역 등 지원하는 '한컴독스 AI' 출시('24.6), 자회사 씽크프리, 글로벌 기업용 AI 검색 서비스 '리파인더 AI' 베타 출시('24.8), AI 문서작성도구 '한컴어시스턴트'와 AI 질의응답 솔루션 '한컴피디아' 출시('24.12), AI 한컴어시스턴트 유료 구독 서비스 출시('25.9) - (기술) AI 학습용 PDF 데이터 추출 엔진 오픈소스 공개('25.9), AI 하이브리드 협업 기술 ITU 국제표준(ITU-T F.746.19)으로 승인('25.11) - (사업/제휴) BHSN과 리걸 AI 도입 MOU('25.7), 독자 AI 모델 LG AI연구원 컨소시엄 참여('25.8), 텐센트 클라우드와 글로벌 AI 사업 협력을 위한 업무협약을 체결('25.12)
		플리토	- (서비스) AI 이미지 번역 서비스 출시('25.4), B2C AI 통번역 솔루션 '챗 트랜스레이션' 출시('25.12) - (사업/제휴) 부산외대와 다국어 데이터 구축 MOU 체결('25.6), 퓨리오사AI와 NPU 기반 통번역 고도화 업무협약 체결('25.7), APEC CEO 서밋 코리아 2025 퓨처-테크 포럼 AI에 AI 통번역 기술을 공식 지원('25.10)
	법률	로앤컴퍼니	- (서비스) MS 애저 기반 법률 AI 어시스턴트 '슈퍼로이어' 공개('25.1) - (협업/제휴) 독자 AI 모델 업스테이지 컨소시엄 참여('25.8), SKT 주도 AI 연합체 'K-AI 얼라이언스' 가입('25.9) - (투자) 500억 원 규모 시리즈C2 투자 유치('25.7)
	의료	뷰노	- (사업) 골연령 분석 AI 솔루션 'BoneAge' 사업 일체를 마이허브에 양도('25.9) - (기술) 급성심근경색 조기탐지 AI 모델 연구성과 국제학술지 게재('25.5), 콩팥기능 저하 선별 AI 의료기기 식약처 허가 획득('25.6), 유럽심장학회에서 AI 심전도 분석 연구 4건 발표('25.8) - (협업/제휴) 아주대의료원과 의료 AI 공동연구 MOU('25.5), 독자 AI 모델 업스테이지 컨소시엄 참여('25.8)
		루닛	- (서비스) 유방암 진단 AI 솔루션 '루닛 인사이트 MMG' 및 '루닛 인사이트 DBT' 도입 미국 의료기관 200곳 돌파('25.5) - (기술) 북미영상의학회에 '루닛 인사이트'를 활용한 14편의 연구초록 발표('25.12), AI 바이오마커 플랫폼 '루닛스코프'를 적용해 HER2를 진단한 연구결과 국제 학술지 게재('25.12) - (제휴) 애질런트와 AI 기반 동반 진단 솔루션 공동 개발('25.9), 글로벌 임상수시험수탁기관 셀카르타와 AI 바이오마커 플랫폼 '루닛 스코프'의 임상시험 활용을 위한 전략적 파트너십('25.9)

가치사슬	기업명	주요 활동 (기술·모델 개발, 투자유치(인수합병), 사업전략(협업제휴), 제품/서비스 출시 등)
AI 모델·플랫폼	네이버	- (기반모델) 초거대 언어모델 '하이퍼클로바 X' 공개('23.8), 하이퍼클로바X 신규 모델 '대시' 공개('24.4), 하이퍼클로바X에 이미지 처리 기능 업데이트('24.8), 하이퍼클로바X 씽킹 공개('25.6) - (서비스) AI 쇼핑앱 '네이버 플러스 스토어' 출시('25.3), 네이버웍스에 'AI 스튜디오' 출시('25.12) - (기술) 네이버 ASF(AI Safety Framework) 발표('24.6), 하이퍼클로바 X 경량모델 3종 (3B/1.5B/0.5B) 오픈소스 공개('25.4) - (제휴) 서울대와 소버린 AI·인재 양성 위해 공동연구 및 업무협력('25.8)
	카카오	- (기반모델) 카나나 플래그, 카나나 에센스, 카나나 나노 전체 라인업 구축('25.2), 카나나 언어모델 4종(8B, 2.1B 등) 오픈소스 공개('25.5), AI 안전성 검증 모델 '카나나 세이프가드' 오픈소스 공개('25.5), 멀티모달 AI '카나나-o' 공개('25.5), 경량 멀티모달 AI '카나나-1.5-v-3b'와 MoE 모델 '카나나-1.5-15.7b-a3b' 오픈소스 공개('25.7), 허깅페이스에 '카나나-2' 오픈소스 공개('25.12) - (서비스) AI 비서 '카나나' 공개('24.10), 카나나 베타 서비스 시작('25.5), MCP 실험공간 'PlayMCP' 베타 공개('25.8), 카카오톡과 AI 결합 발표('25.9), 오픈AI와 공동 프로덕트 출시('25.10)
	KT	- (기반모델) 거대 AI '믿:음' 발표('23.10), 믿:음 2.0 오픈소스 공개('25.7) - (제휴) MS와 AI·클라우드·정보기술·IT 분야 협력을 위한 전략적 파트너십 체결('24.6), 데이터브릭스와 생성 AI 및 데이터 서비스 개발 협력('24.11), 팔란티어와 AI 협력 체결('25.3), 베트남 비엣텔 그룹과 AI 전환 협력 체결('25.5), 서울대-KAIST와 AICT(AI+ICT) 공동연구 협력 확대('25.9), 크루AI와 전략적 파트너십 체결('25.10)
	SKT	- (기반모델) 자체 개발 '에이닷 엑스 3.1'와 Qwen 2.5 기반의 한국어 특화 LLM '에이닷 엑스 4.0' 공개('25.7) - (플랫폼) 생성 AI 플랫폼 '엔터프라이즈 AI 마켓' 출시('24.1), 에이닷에서 글로벌 LLM을 종합한 멀티 LLM 에이전트 서비스 제공('24.8) - (서비스) AI 개인비서 '에이닷' 출시('23.9), 글로벌향 AI 에이전트 '에스터' 공개('24.11), AI 에이전트 '에이닷 비즈' 베타테스트 개시('25.2), AI 기반 설문조사 결과 제공하는 'AI 서베이' 출시('25.4), 에이닷 4.0 출시('25.8), 티맵에 에이닷 적용('25.9), '에이닷'에 'A.X 4.0'과 오픈AI 'GPT-5' 시리즈 신규 도입 및 문자 보안 기능 추가('25.9) - (협업/제휴) Humane/Perplexcity AI와 상호투자, 공동마케팅, 글로벌향 AI 에이전트 서비스 고도화 등 협력('24.09), 생성 AI 연구 위한 MIT '생성 AI 컨소시엄' 참여('25.2), 크래프톤과 언어모델 공동 개발('25.7), 독자 AI 파운데이션 모델 프로젝트 주관 사업자 선정('25.8) - (투자) Anthropic에 1억 달러 투자('23.8), AI 검색엔진 기업 Perplexity AI에 1천만 달러 투자('24.6), 미국 AI 스타트업 TogetherAI에 전략적 투자('25.3) - (사업) 사내 'AI 거버넌스 포털' 오픈('25.9), AI 혁신을 위한 AI CIC(사내회사) 신설('25.9)
	LGU+	- (서비스) 보이스피싱 실시간 탐지 가능한 AI 통화 비서 '익시오' 출시('24.11), 보이스피싱 차단 '안티딥보이스' 기술을 익시오에 탑재('25.6), AI 기반 교사 행정 관리 서비스 'U+슈퍼스쿨' 출시('25.12), OpenAI의 기술 활용해 성형 AI 기반 구독형 콜봇 서비스인 'Agentic 콜봇' 출시('25.12) - (제휴) 뤼튼과 '익시오'에 콘텐츠 제공 MOU 체결('25.5), 오픈AI와 에이전틱 AI 컨택센터 개발 MOU('25.7) - (기술) 익시오 보이스피싱 탐지기능 고도화 위한 ICT 규제 샌드박스 실증특례 승인('25.9)
	LG AI연구원	- (기반모델) 엑사원 공개('21.12), 엑사원 2.0 공개('23.7.), 엑사원 3.0 오픈소스 공개('24.8), 엑사원 3.5 오픈소스 공개('24.12), 추론 AI '엑사원 딥' 오픈소스 공개('25.3), 의료 AI 모델 '엑사원 패스 2.0'과 엑사원 4.0 공개('25.7) - (서비스) 데이터셋 저작권 문제 파악하는 에이전트 '넥서스' 출시('25.2) - (사업/제휴) 미국 잭슨랩과 알츠하이머와 암 연구 AI 기반 공동 연구개발('25.8), 독자 AI 모델 주관 사업자 선정('25.8)

가치사슬	기업명	주요 활동 (기술·모델 개발, 투자유치(인수합병), 사업전략(협업제휴), 제품/서비스 출시 등)
	솔트룩스	- (하드웨어) 생성 AI 개발용 하드웨어 '루시아 온' 출시('24.3) - (기반모델) 자체 개발 LLM '루시아2' 공개('24.5), '루시아 2.5' 글로벌 출시('25.4), 추론 AI 모델 '루시아 3.0' 공개('25.5) - (서비스) 기업용 검색 솔루션 '구버 엔터프라이즈' 출시('24.9), AI 검색 서비스 '구버' 프리뷰 버전 공개('24.11), '구버' 업데이트 버전 공개('25.5), AI 에이전트 '구버' 정식 출시('25.7) - (제휴) 콘텐츠 기업 NEW와 AI 콘텐츠 서비스 개발 MOU('25.2)
	마음AI	- (기반모델) 기업용 거대언어모델 'MAAL1' 공개('24.3), 온프레미스 LLM 제품 'MAAL'을 라마 3.1 기반으로 업그레이드('24.8), 딥시크 기반 기업용 AI 모델 '말 알바트로스' 출시('25.2), 로봇 파운데이션 모델(RFM) 상용화('25.7) - (플랫폼) 로봇 전문 플랫폼 '피지컬 AI 몰' 오픈('25.8) - (기술) 전장 특화 음성 AI 기술 'SUDA' 공개('25.4) - (인수) 멀티모달 기반모델 특화 기업 맨드언맨드 인수('24.5)
	업스테이지	- (기반모델) '솔라 프로' 프리뷰 공개('24.9). 광학문자인식(OCR) 모델 'Document Pharse' 공개('24.10), '솔라 프로' 정식 출시('24.12), 차세대 LLM '솔라 프로2' 출시('25.7) - (플랫폼) LLM 평가 플랫폼 '이벨버스' 공개('24.4) - (서비스) '솔라' 기반 번역 서비스 베타 공개('24.5), AI 기반 한글 교열 서비스 '에디트업' 출시('25.5) - (투자) 1천억 원 규모 시리즈 B 투자 유치('24.4), 620억 원 규모 투자 유치('25.8) - (사업) 독자 AI 모델 주관사업자 선정('25.8)
	포티투마루	- (기반모델) AI 엑스포 2024에서 산업 특화 LLM 모델 'LLM42' 공개('24.5), LLM42로 AI-MASTER 신뢰성 인증 획득('25.7) - (투자) LG유플러스로부터 100억 투자 유치('24.1) - (제휴) LG유플러스, 샌즈랩과 사이버보안 강화 LLM 기술 공동 개발 MOU('24.3), 오라클과 클라우드 인프라 및 LLM 사업 협력('24.7), 한패스와 AI 기반 외국인 생활 정보 서비스 개발 MOU('25.7)
	트웰브랩스	- (기반모델) 영상언어 AI 'Pegasus-1' 및 멀티모달 영상 이해 모델 'Marengo 2.6' 업데이트('24.3), 'Pegasus-1.2' 업데이트 버전 공개('25.2) - (투자) 엔비디아 등으로부터 700억 규모 시리즈A 투자 유치('24.6), SKT로부터 3,000만 달러 전략적 투자 유치('24.12) - (서비스) 아마존 베드록에서 'Marengo'와 'Pegasus' 정식 공급 개시('25.7), 영상 이해 MCP 서버 출시('25.9), 플러그인 AI 'DocuAgent42' v1.1 출시('25.12) - (제휴) LG CNS와 영상 이해 기술 비즈니스 고도화 MOU('25.7)
	코난테크놀로지	- (기반모델) 자체 개발 '코난 LLM' 출시('23.8), 사용자 목소리를 다양한 언어로 변환하는 음성합성 모델 공개('24.4), 일반 모드와 추론 모드 통합 LLM 'ENT-11' 공개('25.3) - (플랫폼) AI 기반 검색 플랫폼 'RAG-X' 공개('25.4), 의료 생성 AI 플랫폼 'HAI(Hallym Artificial Intelligence)' 구축('25.7), APEC 2025 포럼에서 AI 회의 통역 시스템 도입('25.8), 사내 구축형 범용 생성 AI 플랫폼 'KEMI' 한국남부발전에 실무 적용('25.8) - (서비스) 기업향 검색증강 생성 솔루션 '코난 RAG-X' 공개('25.4), 추론통합모델 코난 LLM 'ENT-11' 기반 법률 특화 서비스 공개('25.9) - (제휴) AI 반도체 스타트업 리벨리온과 생성 AI 키오스크 개발 협력 체결('24.8), 한림대학교 의료원과 의료 AI 플랫폼 'HAI' 개발('25.7)

가치사슬	기업명	주요 활동 (기술·모델 개발, 투자유치(인수합병), 사업전략(협업제휴), 제품/서비스 출시 등)
	모레	- (기반모델) 자체 개발 LLM 'MoMo-70B' 허깅페이스 LLM 평가에서 세계 1위('24.1), 라마 기반 한국어 LLM 'Motif' 오픈소스 공개('24.12), 자회사 모티프 sLLM '모티프 2.6B' 오픈소스 공개('25.6) - (제휴) AMD와 전략적 파트너십 맺고 2,200만 달러 규모 투자 유치('23.10), Tenstorrent와 공동으로 효율적인 AI 데이터 센터 솔루션 공개('24.11) - (기술) 미국 AI 인프라 서밋에서 분산 추론 시스템 발표('25.9)
	매스프레소	- (기반모델) 업스테이지와 수학 특화 AI 모델 'MathGPT' 개발('24.1) - (제휴) NE능률과 저작물 이용 계약 체결('24.1), KT 통해 AI 학습 플랫폼 '콴다' 구독 서비스 출시('25.2) - (서비스) 미국향 대학교육 서비스 '프렙파이(Prep.Pie)' 출시(24.3), 2025년 하반기 AI 학습 에이전트 '콴다 3.0' 출시 계획 발표('25.7)
	올거나이즈	- (기반모델) 금융 전문 AI 모델. '알리 파이낸스 LLM' 출시('23.9) - (서비스) MCP 빌더 출시('25.4) - (플랫폼) LLM 에이전트 역량 평가 플랫폼 '올인원 벤치마크' 출시('25.2) - (사업/제휴) 독자 AI 모델 업스테이지 컨소시엄 참여('25.8), AI 솔루션 알리를 KT 클라우드 플랫폼에 통합해 금융 AI 시장 공략 제휴('25.8)
	엔씨소프트(NC AI)	- (기반모델) 오픈소스 기반 LLM '라마 바르코 LLM' 공개('24.9), VARCO-VISION 2.0 멀티모달 AI 모델 오픈소스 공개('25.7), 초경량 멀티모달 AI 모델 '바르코 비전 2.0' 공개('25.7) - (사업/제휴) 독자 AI 모델 주관사업자 선정('25.8), MBC와 미디어 콘텐츠 분야 AI 전환 MOU('25.8) - (기술) 음성·언어 기술 학회 INTERSPEECH 2025에서 몬스터 사운드 생성·변환 AI 기술 공개('25.8), SIGGRAPH2025에서 '바르코 애니메이션'과 '바르코 싱크패이스' 기술 공개('25.7)
클라우드	KT 클라우드	- (플랫폼) AI 전환 플랫폼 'AI 파운드리' 첫 솔루션으로 'RAG 스위트' 출시('25.9) - (인프라) 국내 AI 반도체 스타트업 '리벨리온'의 아톰 적용한 NPU 인프라 구축, AI 전용 데이터센터 백석 AIDC 개관('24.8), 10MW 규모 경북 AI 데이터센터 준공('25.5) - (서비스) 엔비디아 GPU 기반 AI 학습 전용 인프라 서비스 'AI 트레인' 출시('24.5) - (제휴) 스테이지, 디노티시아, 폴라리스오피스, 리벨리온과 'AI 파운드리' 사업 전략적 제휴 체결('25.4), AI 파운드리 파트너십 확대로 올거나이즈, 인핸스, 세이지, 스피링크, 몽고DB와 MOU('25.8)
	네이버클라우드	- (플랫폼) 클로바 스튜디오를 통해 기업 고객 대상 하이퍼클로바X 모델 제공('23.10) - (사업/제휴) 인텔과 AI 공동연구센터 설립('24.4), 리벨리온과 사우디 AI 사업 협력 MOU('25.2), 태국 AI 기업 '시암 AI 클라우드'와 LLM 및 AI 에이전트 개발 MOU 체결('25.5), nvidia 등과 컨소시엄으로 모로코에 차세대 AI 데이터센터 구축 계획 발표('25.6), '독자 AI 파운데이션 모델' 프로젝트 주관 사업자 선정('25.8), 씽크풀과 금융 AI 챗봇 개발 협력('25.8) - (서비스) AI 영상 분석 서비스 'MAIU' 공개('25.5)
	NHN클라우드	- (플랫폼) AI 개발 플랫폼 'AI 이지메이커'에 모델 평가 자동화 기능 도입('25.7) - (인프라) 광주 국가 AI 데이터센터 운영 개시('24.3), 정부 AI 인프라 사업 1조 수주('25.8) - (제휴) 머신러닝 기업 베슬에이아이 코리아와 AI 인프라 및 MLOps 사업 협력 MOU 체결('24.6), 디노티시아와 AI 클라우드 사업 MOU 체결('25.3)

가치사슬	기업명	주요 활동 (기술·모델 개발, 투자유치(인수합병), 사업전략(협업제휴), 제품/서비스 출시 등)
하드웨어	삼성전자	- (제품) AI 탑재한 갤럭시 S24 시리즈 공개('24.1), 갤럭시 S25 공개('25.1), AI PC '갤럭시 북5 프로' 국내 출시('25.1), AI 강화한 갤럭시 Z 플립7과 폴드7 공개('25.7), 갤럭시 AI 탑재 '갤럭시 S25 FE·탭 S11 시리즈·버즈3 FE' 출시('25.9), AI 탑재 '갤럭시 A17 LTE' 출시('25.11), 제미나이 탑재한 '비스포크 AI 냉장고' 신모델 공개('25.12) - (기술) '칩렛' 패키징 적용 4나노 AI 칩 개발('25.6), AI 업무 생산성 측정 지표 '트루벤치(TRUEBench)' 공개('25.9) - (제휴) 리벨리온과 파운드리 4나노공정 이용 AI반도체 '리벨' 공동 개발 진행('23.10~), 네이버클라우드와 하이퍼클로바 구동을 위한 AI반도체 솔루션 개발 진행('22.12.~), OpenAI와 차세대 AI 데이터센터 구축을 위한 LOI체결('25.10), 엔비디아와 '반도체 AI 팩토리' 구축 발표('25.10) - (투자) 미국 AI 반도체 스타트업 Enfabrica에 투자('24.11), 영상 분석 AI 기술 기업 'Memories.AI'에 투자('25.7)
	SK하이닉스	- (제품) 초고성능 AI 메모리 'HBM3E' 양산 시작('24.3), 온디바이스 AI 용 모바일 낸드 솔루션 'ZUFS 4.0' 개발('24.5), AI데이터센터용 고용량 SSD 'PS1012 U.2' 개발('24.12), 서버용 저전력 D램 모듈 '소캠(SoCAMM)' 2025년 내 공급 개시 계획 발표('25.7), HBM4 양산 체제 구축('25.9), AI폰용 낸드 솔루션 제품인 'ZUFS 4.1' 공급 개시('25.9) - (제휴) SK텔레콤, 펭귄 솔루션스와 AI 데이터센터 사업 협력('25.1), 네이버클라우드와 AI 서비스 성능과 효율 혁신 협력('25.9)
	리벨리온	- (제품) 신규 AI 칩 '리벨쿼드' 발표('25.8) - (제휴) Arm 및 삼성전자와 제휴해 AI CPU 칩렛 플랫폼 개발 추진('24.10), AI 교육 솔루션 기업 엘리스그룹과 국산 AI 반도체 생태계 활성화 MOU 체결('25.1), 시스템 반도체 설계 전문 기업 코아시아세미와 차세대 AI 칩렛 공동개발 협력('25.7) 반도체 기업 Marvell과 맞춤형 AI 인프라 공동 개발('25.7), 말레이시아 주정부와 AI 반도체 협력('25.10) - (사업/투자) 리벨리온-사피온 합병법인 공식 출범('24.12), 1,650억 원 규모의 시리즈 B 투자 유치('24.1), 사우디 법인 설립('25.8), 3,400억 원 투자 유치('25.9), 시리즈C 라운드에서 약 3,400억 원 규모 투자('25.9), 시리즈C에 실리콘밸리 벤처캐피탈 '킨드레드벤처스', '탑티어 캐피탈 파트너스' 추가 유치('25.11)
	퓨리오사AI	- (제품) AI 반도체 워보이(WARBOY) 개발 완료 후 삼성전자 파운드리 통해 양산 개시('23.4), 2세대 AI 추론용 특화 고성능 RNGD(레니게이드) 공개('24.8), LG 엑사원에 레니게이드 전면 도입('25.7) - (제휴) 네이버클라우드와 협력해 GPU-NPU 팜 구축('23.12), RNGD를 MS 애저 마켓플레이스에 출시('25.4), 레니게이드를 LG 엑사원에 전면 도입 협업('25.7), 오픈AI 코리아 개소식에서 RNGD 기반 gpt-oss-120B 챗봇 시연('25.9) - (투자) 케이스톤파트너스, 200억 원 투자('25.6), 1,700억 원 규모 투자 유치('25.7),
	딥엑스	- (제품) 온디바이스 AI용 NPU 공개('24.5), 온디바이스 AI 1세대 제품 DX-M1 양산 돌입('24.8), 삼성전자 2나노 공정 활용 차세대 NPU 'DX-M2' 생산('25.4), 'Yotta 2025' 전시회에서 암페어, 네트워크 옵틱스와 AI 영상관제 시스템 공개('25.9) - (사업/투자) 사모펀드 기관 중심 1,100억 원 규모의 투자 유치('24.5), 대만 지사 설립('29.9) - (제휴) 대만 서버 솔루션 기업 AIC와 엣지 AI 서버 개발 협력('25.5), 노타와 온디바이스 AI 협력('25.5), 한컴아카데미와 엣지 AI 반도체 유통 협력('25.6), 삼성전자, 가온칩스와 차세대 AI 반도체 'DX-M2' 개발 위한 2나노(㎚) 공정 계약 체결('25.8), 바이두와 온디바이스 AI 반도체 공급 MOU('25.8)
	넥스트칩	- (제품) 현대자동차 출시 예정 소형 모빌리티 플랫폼 '모베드'에 카메라 모듈칩 탑재('25.7), 방산용 드론 적용 비전센서칩 개발('25.8) - (제휴) 스텔란티스 자회사 에이아이모티브와 자율주행 기술 개발 협력('25.7)
	텔레칩스	- (제품) 모빌리티용 고성능 네트워크 프로세서 'AXON' 출시('25.2), 차량용 SIP(System-in-Package) 모듈 공개('25.3), 차량용 AI 가속기 'A2X' 출시('25.3) - (플랫폼) 차량용 SoC 기반 오픈 하드웨어 플랫폼 'TOPS G' 시리즈 출시('25.7), 오픈소스 하드웨어 플랫폼 'TOPST' 출시('25.9) - (제휴) Arm과 협력해 차세대 인포테인먼트 SoC 'Dolphin7' 개발 추진('25.4)
	모빌린트	- (제품) 론 전문 다온아이앤씨의 군집 드론 솔루션에 AI 반도체 1,000대 공급('25.3), AI 가속기 모듈 'MLA100' 출시('25.4). NPU 기반 독립형 AI PC MLX-A1 연세대 공급('25.9) - (제휴) AWS IoT 그린그래스' 플랫폼에 고성능 NPU를 연동 협력('25.5), 코오롱베니트 AI 얼라이언스 합류('25.7), 대만 임베디드 플랫폼 에티나와 AI 가속기 카드와 엣지 AI 솔루션 MOU('25.9)

② 해외 AI 기업 동향

<표> 해외 AI 기업 동향

가치사슬	기업명		주요 활동 (기술·모델 개발, 투자유치(인수합병), 사업전략(협업제휴), 제품/서비스 출시 등)
애플리케이션 · 서비스	콘텐츠 제작	Adobe	- (기반모델) 이미지 생성 AI 'Firefly Image 3' 공개('24.4), AI 사진 편집 모델 'Magic Fixup' 오픈소스 공개('24.8), 'FireFly Video Model' 공개('24.10), 비디오 생성 AI 'Firefly Video' 출시('25.2), 맞춤형 AI 모델 개발용 'Adobe Firefly Foundry' 공개('25.10) - (서비스) PDF 편집기에 AI 어시스턴트 출시('24.2), 생성 AI 음악 편집 도구 프리뷰 버전 공개('24.2), 기업용 광고 생성 AI '젠스튜디오' 출시('24.2), Adobe Firefly 서비스 신기능 출시('25.5), Firefly에 음향 효과 생성 기능 추가 및 모바일 앱 출시('25.7), Adobe LLM Optimizer 출시('25.10) - (플랫폼) Adobe Experience 플랫폼에서 10개 에이전트와 오케스트레이션 툴 출시('25.3), 고객 경험과 마케팅 특화 기업용 AI 에이전트 출시(Audience Agent, Journey Agent 등)('25.9)
		Luma AI	- (기반모델) 동영상 생성 AI 'Dream Machine' 공개('24.6), 'Dream Machine 1.6' 공개('24.9), 동영상 AI 모델 'Ray2' 출시('25.1), 'Ray3 Modify' 출시('25.12) - (서비스) 'Dream Machine' 프레임 편집 기능 'Reframe' 공개('25.5), 비디오 편집 툴 'Modify Tool' 출시 ('25.6), 'Dream Machine' 편집용 기능 'Modify with Instructions' 출시('25.8)
		Udio	- (서비스) 음악 생성 AI 앱 'Udio' 공개('24.4), iOS 앱 출시('25.5), AI 음악 편집 툴 Sessions' 출시('25.6) - (제휴) Universal Music Group과 제휴('25.10), Warner Music Group과 제휴('25.11)
		Stability AI	- (기반모델) Stable Diffusion 2.0 ('22.11), Stable Diffusion XL('23.7), Stable Cascade('24.2), Stable Diffusion 3.0 ('24.2), Stable Diffusion 3.5('24.10), Stable Video 4D 2.0 업그레이드 ('25.5), Stable Audio 2.5('25.9) - (서비스) 텍스트 기반 애니메이션 생성 도구 출시('23.5) 이미지를 3D 동영상으로 변환하는 'Stable Video 3D' 발표('24.3), 멀티뷰 확산 모델 'Stable Virtual Camera' 출시('25.3) - (제휴) Amazon Bedrock에서 Stable Image Services 제공('25.9), Universal Music Group과 제휴 ('25.10), Warner Music Group과 제휴('25.11)
		Midjourney	- (기반모델) 이미지 생성 모델 V6('23.12), V7('25.4), 비디오 모델 V1 출시('25.6) - (서비스) 이미지 일관성을 위한 'Character Reference' 기능('24.3), 통합 웹이미지 편집기 출시 ('24.8), 이미지 편집 기능 출시('24.10), V1 'Style Explore' 신기능 출시('25.9)
		Runway	- (기반모델) 동영상 생성 AI Gen-3 Alpha 출시('24.6), Gen-3 Alpha Turbo 출시('24.8), Gen-4('25.4), Gen-4 Image API 출시('25.5), 동영상 생성과 편집 AI 'Runway Aleph' 출시 ('25.7), Gen-4.5('25.12) - (서비스) 영상 속 인물 캐릭터화 AI 도구 'Act-One' 출시('24.10)
		Canva	- (서비스) 이미지 생성 서비스 'Dream Lab' 출시('24.10), ChatGPT Deep Research와 통합 및 자체 MCP 서버 출시('25.6), 대화형 창작 파트너 'Canva AI' 글로벌 확대 발표('25.9), AI 신뢰성 도구 모음 'Canva Shield' 출시('25.11) - (플랫폼) 올인원 비주얼 커뮤니케이션 플랫폼 'Visual Suite 2.0' 출시('25.4), 콘텐츠 창작 플랫폼 'Creative Operating System' 출시('25.10)
		Eleven Labs	- (기반모델) 32개 언어 지원하는 텍스트-음성 변환 모델 Turbo v2.5('24.7), 음성-텍스트 모델 'Scribe v1' 출시('25.2), TTS 모델 'Eleven v3 Alpha'('25.6), 음성-텍스트 모델 'Scribe v2 Reatltime'('25.11) - (서비스) 음향 효과 생성 'Sound Effects 출시('24.6), 배경 소음 제거하는 'AI Voice Isolator' 출시 ('24.7), 텍스트-오디오 변환 모바일 앱 'Reader' 출시('24.8), 음성 생성 서비스 'Voice Design' 출시 (24.10), 대화형 AI 'Conversational AI 2.0'('25.5), 음악 생성 서비스 'Eleven Music'('25.8), 콘텐츠 편집 관리형 서비스 'Productions' 출시('25.9), ElevenLabs Agents('25.9), 'ElevenLabs Image & Video' 출시('25.11) - (투자) 1억 8천만 달러 시리즈 C 투자 유치('25.1)

가치사슬	기업명		주요 활동 (기술·모델 개발, 투자유치(인수합병), 사업전략(협업제휴), 제품/서비스 출시 등)
		Pika Labs	- (기반모델) 동영상 생성 AI 모델 'Pika 1.0' 출시('23.11), 'Pika 1.5' 출시('24.10), Pika 2.0('24.12) - (서비스) Pika 1.0에 음성 기능 및 음향효과 자동 생성 기능 추가('24.3), 영상에 인물이나 사물 추가하는 'Pikaadditions' 출시('25.2), iOS에서 'Social AI Video' 앱 출시('25.8) - (투자) 8천만 달러 규모의 시리즈 B 투자 유치('24.6)
		SUNO	- (서비스) 음악 생성 AI 'v3' 공개('24.3), 'V3.5' ('24.5), v4.5('25.5), 고급 음악 편집 기능 업데이트 ('25.6), 오디오 생성 워크스테이션 'Suno Studio' 출시('25.9) - (인수/제휴) 브라우저 기반 디지털 오디오 워크스테이션 WavTool 인수('25.6), Warner Music Group과 제휴('25.11) - (투자) Menlo Ventures 주도 시리즈 C 투자에서 2억 5천만 달러 유치('25.11)
		Moonvalley	- (기반모델) 윤리적 소스로 훈련한 비디오 생성 AI 모델 'Marey' 출시('25.3), 영화 제작 툴 'Marey Realism v1.5' 출시('25.8) - (투자) 비공개 투자자로부터 5,300만 달러 규모 투자 유치('25.5)
		Synthesia	- (기반모델) AI 아바타용 비디오 및 음성 엔진 'Expres-2' 공개('25.9) - (서비스) AI Video Assistant 출시('24.1), AI 비디오 커뮤니케이션 플랫폼 'Synthesia 2.0'('24.6), Synthesia에 Veo 3 기반 비디오 생성 기능 'Generative Assets' 추가('25.7), Synthesia 3.0 출시 ('25.10) - (투자) NEA 주도 시리즈 D 투자에서 1억 8,000만 달러 투자 유치('25.1)
	검색엔진	Perplexity	- (서비스) 긴 문서 생성 AI 서비스 'Page' 출시('24.05) 다단계 심층 검색 지원하는 'Pro Search' 출시('24.07), 유료 사용자 대상 쇼핑 기능 추가('24.11), Deep Research 기능 출시('25.2), 생산성 도구 'Perplexity Labs' 출시('25.5), AI 웹 브라우저 'Comet' 출시('25.7), 정부용 퍼플렉시티 출시('25.9), 이메일 어시스턴트 및 검색 API 출시('25.9), AI 웹 브라우저 'Comet' 무료 공개('25.10), AI 기반 특허 검색 서비스'Perplexity Patents' 출시('25.10), Comet에서 이용가능한 'Comet Assistant'업데이트('25.10), 안드로이드용 'Comet' 앱 출시('25.11) - (투자) 엔비디아, 제프 베이조스 등 투자자로부터 7,360만 달러 유치('24.1), 1억 2천만 달러 추가 조달('25.2), 1억 달러 투자 유치('25.7) - (제휴) 언론사와 수익 공유 프로그램 공개('24.7)
		MS	- (서비스) 검색엔진 Bing에 GPT-4 결합('24.5), Bing Generative Search 베타 공개('24.7), GPT-4 Turbo 통합('25.1), Bing Generative Search 출시('25.2), Bing Chat Enterprise('25.2) OpenAI Sora 기반 Bing Video Creator 출시('25.6), Bing Image Creator에 DALL-E3와 GPT-4o 추가('25.8), MS 365 Copilot에 Researcher 에이전트와 Cluade 모델 추가('25.9)
		Glean	- (서비스) 기업용 검색 서비스 'Glean Search'('21.9), 기업용 에이전트 플랫폼 Glean Agents('25.5), 기업용 AI 에이전트 'Glean Assistant' 출시('25.9) - (투자) 시리즈F에서 1억 5,000만 달러 투자 유치('25.6)
	챗봇	Character.ai	- (서비스) 다수의 AI 캐릭터가 참가하는 그룹 채팅 기능 출시('23.10), AI 캐릭터와 음성 통화 기능 출시('24.6), 비디오 생성 서비스 'Avatar FX' 출시('25.4), AI기반 소셜 피드 출시('25.8) - (인수) 구글과 기술 라이선스 및 핵심 인력 영입으로 사실상 인수 계약 체결('24.8)
		Quora	- (서비스) AI 채팅 플랫폼 'Poe' 출시('23.2), AI 챗봇 수익 공유 프로그램 출시('24.4), 앱 제작도구 'Preview' 기능 출시('24.7), AI 기반 앱 생성 도구 'Poe App' 출시('25.2) - (투자) VC '앤드리슨 호로위츠로'부터 7,500만 달러 투자 유치('24.1)

가치사슬	기업명		주요 활동 (기술·모델 개발, 투자유치(인수합병), 사업전략(협업제휴), 제품/서비스 출시 등)
	생산성 · 엔터프라이즈	Google	- (서비스) Google Workspace 용 생성 AI 도구 'Duet AI' 출시('23.3), Google Workspace용 동영상 제작 도구 'Google Vids' 공개('24.4), 기업 고객에 Gemini와 Notebook LM Plus 기능 추가('25.1), Google Workspace에 Gemini 기반 생성 AI 기능 업데이트('25.5), 구글 AI 울트라 기업 고객에 이미지 생성 모델 'Whisk' 및 영상 생성 모델 'Flow' 생성 한도 확대, 구글 독스에서 문서-오디오 변환 기능 추가('25.8) - (사업) 데이터센터 및 에너지 인프라 기업 인터섹트 파워 인수('25.12)
		Writer	- (기반모델) 적응형 추론 모델 'Palmyra X5' 출시('25.4), 소형 고성능 모델'Palmyra-mini' 공개('25.9) - (서비스) 기업용 AI 앱 제작 도구 'AI Studio' 출시('24.6), 기업용 AI 에이전트 'Super Agent' 공개('25.7) - (플랫폼) 기업 AI 에이전트 구축 및 운영을 위한 AI HQ 플랫폼 출시('25.4) - (투자) 'Premji Invest'등 다수의 기업으로부터 2억 달러 규모 시리즈 C 투자 유치('24.11)
		DeepL	- (서비스) AI 글쓰기 도구 'DeepL Write Pro' 출시('24.4), 기업용 번역 솔루션 'DeepL Pro' 165개국 신규 진출('24.6), 음성 번역 솔루션 'DeepL Voice' 출시('24.11), AI-사용자 상호작용을 위한 'Clarify' 기능 추가('25.3), 지원 언어 36개 언어로 확대('25.6), 'DeepL Agent' 정식 출시, 사용자의 언어 요구 사항을 자동으로 적용하는 'Customization Hub' 도입('25.11) - (플랫폼) 상용 API 솔루션 플랫폼 'DeepL Marketplace'('25.9) - (협업) Zoom에 'DeepL Voice' 통합('25.7) - (투자) VC 'Index Ventures' 주도로 3억 달러 규모 투자 유치('24.5)
		Anysphere	- (서비스) AI 코딩 도구 'Cursor' 출시('23), Cursor 1.0('25.6), Cursor 웹 앱 출시('25.7), Composer 모델 출시('25.10) - (투자) VC 'Thrive Capital' 주도로 9억 달러 규모의 투자 유치('25.6), NVIDIA 등으로부터 23억 달러 규모 시리즈 D 투자 유치('25.11) - (인수) 코드리뷰 자동화 스타트업 'Graphite' 인수('25.11)
		Notion	- (서비스) 'Notion AI' 정식 출시('23.2), 맞춤형 이메일 서비스 'Notion Mail' 출시('25.4), 생성 AI 기반 'Notion AI for Work' 출시('25.5), Notion AI Agents를 도입한 'Notion 3.0' 출시('25.9)
		WindSurf	- (서비스) 코딩 에이전트 'Codiumate' 공개('25.4), 코딩 모델 'SWE-1' 시리즈 공개('25.5), Cognition Devin 기능 통합 등 신기능 추가 업데이트 발표('25.8), 'SWE-1.5' 발표('25.10) - (피인수) Cognition에 인수('25.7)
		Cognition	- (서비스) AI 코딩 에이전트 'Devin' 출시('24.12), 'Devin 2.0'출시('25.4) - (투자/인수) Windsurf 인수('25.7), 4억 달러 규모 투자 유치('25.9)
		Harvey	- (서비스) 법률 특화 생성 AI 서비스 출시('23), 오픈AI 기술 기반 'Deep Research for Legal' 출시('25.6), Harvey 모바일 앱 출시('25.9), MS Office 365와 통합 서비스 개시('25.10) - (투자) 3억 달러 시리즈 D 투자 유치('25.2), 3억 달러 시리즈 E 투자 유치('25.6), 1억 6천만 달러 시리즈 F 투자 유치('25.12) - (제휴) LexisNexis와 제휴해 법률 질의응답 서비스 'Ask LexisNexis' 공개('25.8)
		Grammarly (Superhuman)	- (서비스) AI ROI 측정 도구 출시('24.10), AI 기반 글쓰기 플랫폼 Docs와 AI 에이전트 출시('25.8), 다국어(스페인어, 포르투갈어, 프랑스어, 이탈리아어, 독일어) 글쓰기 지원('25.9) - (투자) AI 생산성 플랫폼 확장 위한 10억 달러 투자 유치('25.5) - (사업) 사명을 'Superhuman'으로 변경('25.10) - (인수) 생산성 스타트업 Coda 인수('24.12), AI 이메일 기업 Superhuman 인수('25.7)

가치사슬	기업명	주요 활동 (기술·모델 개발, 투자유치(인수합병), 사업전략(협업제휴), 제품/서비스 출시 등)
AI모델 · 플랫폼	OpenAI	- (기반모델) GPT-4('23.3), GPT-4 Turbo('23.11), GPT-4o('24.5), GPT-4o Mini('24.7), o1('24.9), ChatGPT Pro('24.12), o3-mini('25.1), GPT-4.5('25.2), 챗GPT-4o 이미지 생성 모델('25.3), o3, o4-mini, GPT-4.1('25.4), o3-Pro('25.6), 음성모델 gpt-realtime, 개방형 가중치 모델 gpt-oss, GPT-5('25.8), Sora 2('25.10), GPT-5.1, GPT-5.1-Codex-Max('25.11), GPT-5.2, GPT-5.2-Codex, ChatGPT Images 1.5('25.12) - (서비스) ChatGPT Enterprise('23.8), Dall-E 3('23.9), ChatGPT 'Memory 기능 추가('24.2), 맞춤형 챗봇 GPTs 공개('23.11), 챗GPT 검색('24.10), AI 에이전트 'Operator'('25.1), Deep Research('25.2), 코딩 에이전트 'Codex', 챗GPT 쇼핑 기능 추가('25.5), 챗GPT Study Mode와 에이전트('25.7), GPT-5-Codex('25.9), 맞춤형 브리핑 제공 'ChatGPT Pulse'와 결제 기능 'Instant Checkout'('25.9) - (플랫폼) AI 에이전트 구축 플랫폼 'Response API' 출시('25.3), AI 웹브라우저 'ChatGPT Atlas' 공개('25.10), 'ChatGPT 그룹 채팅', 'ChatGPT 쇼핑 어시스턴트' 공개('25.11) - (인수) 데이터 분석 기업 Rockset 인수('24.6), AI 하드웨어 스타트업 io 인수('25.5) - (기술) 목소리 복제하는 'Voice Engine' 공개('24.3), GPT-4가 생성한 코드 오류 식별하는 'CriticGPT' 연구('24.7), AI 에이전트 조율하는 'Swarm' 프레임워크('24.10), 에이전트 검색 벤치마크 BrowseComp 공개('25.4), 의료 성능 평가 위한 HealthBench 도입('25.5), AI 모델 환각과 스키밍 연구 및 실제 작업에서 LLM 성능 벤치마크 GDPval 개발('25.9) - (제휴/투자) 미국 AI 안전 연구소와 모델 평가 합의('24.8), 66억 달러 투자 유치('24.10), 언론사·미디어 Conde Nast, Time, News Corp., FT, Axel Sprinter, FT 등과 콘텐츠 계약 체결, AI 인프라 구축 위한 '스타게이트' 설립('25.1), 400억 달러 투자 유치('25.4) 영국 정부와 AI 협력 MOU('25.7), NVIDIA와 AI 인프라 제휴 및 1,000억 달러 투자 유치 ('25.9)
	Anthropic	- (기반모델) Claude 2('23.7), Claude 3('24.3), Claude 3.5 Sonnet('24.6), Calude 3.5 Sonnet 업데이트 및 Claude 3.5 Haiku 출시('24.10), Claude Sonnet 3.7('25.2), Claude Opus 4와 Claude Sonnet 4('25.5), Claude Opus 4.1('25.8), Claude Sonnet 4.5('25.9), Claude Opus 4.5('25.11) - (서비스) Claude 생성물을 실시간 편집 및 수정하는 'Artifacts' 기능 출시('24.8), 기업용 Claude 구독 요금제 출시('24.9), 개발자용 프롬프트 도구 출시('24.11), Claude 웹 검색 추가(25.3), Claude Code('25.5), Claude 기반 금융 분석 솔루션 출시('25.7), 기업 구독자용 메모리 기능 추가('25.9) - (기술) AI-데이터 연결 표준 MCP 오픈소스 공개('24.11), LLM 사고 추적 도구 공개('25.6), LLM 성격 변화에 관한 '페르소나 벡터' 연구('25.8) - (제휴) 미국 AI 안전 연구소와 모델 평가 합의('24.8), Palantir 및 AWS와 제휴해 미국 국방/정보 부처에 모델 제공('24.11), 클라우드 기업 CoreWeave와 19억 달러 규모 클라우드 계약 체결('25.3) - (투자) 아마존으로부터 40억 달러 투자 유치('24.11), 구글로부터 10억 달러 유치('25.1), 35억 달러 투자 유치('25.2), 130억 달러 투자 유치('25.9), 미국 AI 인프라에 500억 달러 투자 발표('25.11)
	Google (Deepmind)	- (기반모델) PaLM('23.3), Bard('23.3), PaLM2(23.5), 음악 생성 AI Lyria('23.11), Gemini 1.0 및 Imagen 2('23.12), Gemma 및 Gemini 1.5 Pro('24.2), Gemini 1.5 Flash 및 Gemma 2('24.6), Gemma 2 2B('24.7), Imagen 3('24.8), Gemini 2.0('25.2), Gemma 3('25.3), Veo 3, Imagen 4, Gemma 3n('25.5), Gemini 2.5 Pro('25.3), Gemini 2.5 Flash('25.6), Gemma 3 270M, Genie 3('25.8), VaultGemma, Gemini Robotics 1.5, Gemini 2.5-Flash-Lite('25.9), Gemini 3, Gemini 3-Flash, Gemini 3 Pro Image('25.12) - (서비스) AI 음성비서 'Gemini Live' 출시('24.8), Gemini에 Imagen 3 통합('24.8), 영화 제작 도구 Flow('25.5), 코딩 에이전트 AlphaEvole('25.5), 개발자용 에이전트 'Gemini CLI'('25.6), Gemini에 이미지 편집 기능 추가('25.8) - (기술) 동영상 생성 AI 'Lumiere' ('24.1), AI 신뢰도 향상 'ASPIRE' 시스템 ('24.1), 로봇 제스처 생성 AI 'GemEM'('24.2), 반도체 칩 설계 AaphaChip('24.10), AI 에이전트 프레임워크 'Talker-Reasoner('24.10), AI 에이전트 통신 프로토콜 A2A('25.4), Gemini Robotics On-Device('25.6), AlphaGenome('25.6), 의료 AI 모델 'MedGemma', AlphaEarth('25.7) - (인수) Manus 인수('25.12)

가치사슬	기업명	주요 활동 (기술·모델 개발, 투자유치(인수합병), 사업전략(협업제휴), 제품/서비스 출시 등)
	Meta	- (기반모델) Llama 1('23.2), Llama 2('23.7), Code Llama 70B('24.1). Llama 3('24.2), Llama 3.1('24.7), 멀티모달 모델 Llama 3.2('24.9), 동영상 생성 모델 'Meta Movie Gen('24.10), Llama 3.3('24.12), Llama 4('25.4) - (제품) AI 안경 Meta Ray-Ban Display와 Oakley Meta Vanguard 출시, Meta AI를 탑재한 AR 안경 Orion 공개('25.9) - (서비스) 자사 서비스에 Meta AI 탑재('24.4), AI 캐릭터 생성 'AI Studio'('24.7), Meta AI 단독 앱 ('25.4), 개발자의 'Llama' 모델 활용 지원을 위한 'Llama API'('25.4), Meta AI에 AI 동영상 피드 'Vibes'('25.9) - (기술) 영상에서 누락된 부분 예측해 학습하는 V-JEPA 개발('24.2), 합성 데이터로 훈련한 3D 콘텐츠 생성 모델 'Vfusion3D' 공개('24.8), AI 모델 평가용 'Self-Taught Evaluator' 개발('24.10), V-JEPA 2 공개('25.6), DINOv3('25.8) - (인수) Scale AI 지분 49% 인수('25.6) AI 음성 스타트업 WavForms 인수('25.8)
	Microsoft	- (기반모델) Phi-1('23.6), Phi-2('23.12), Orca 2('23.11), Orca-Math('24.3), Phi-3('24.4), Phi-3.5('24.8), Ph-4('24.12), Phi-4 mini 및 Phi-4 Multimodal('25.2), , 온디바이스 모델 Mu('25.6), 자체 AI 모델 MAI-Voice-1와 MAI-1-preview 공개('25.8), MAI-Image-1 공식출시 ('25.11) - (서비스) Windows 11에 Copilot 도입('23.9), 기업용 MS 365 Copilot('23.11), 개인용 Copilot Pro('24.1), Copilot for Security('24.3), Copilot에 음성대화와 시각 인식 기능 추가 ('24.10), 신규 에이전트 공개 ('23.11), AI 에이전트 생성 프로그램 'Copilot Studio' 출시('24.11), MS 365 Copilot 업데이트('25.4), Bing Video Creator('25.6) , Edge 브라우저에 Copilot Mode 도입('25.7) - (투자) UAE AI 기업 G42에 15억 달러 투자('24.4) - (기술) 의료 진단 AI 도구 'MAI-DxO' 개발('25.7)
	Amazon	- (기반모델) 'Titan Image Generator V2' 출시('24.8), Nova 음성 및 동영상 생성 모델('25.4) - (서비스) 기업용 AI 챗봇 'Q' 출시('23.11.), 상품 추천하는 AI 챗봇 'Rufus' 출시('24.2), 기업용 앱 생성 AI 도구 'App Studio' 프리뷰 공개('24.7), 판매자 지원 AI 도구 'Amelia' 공개('24.9), AI 기반 쇼핑 가이드 출시('24.10), AI 음성비서 'Alex+'('25.2), AI 에이전트 'Nova Act'('25.4), 쇼핑에 AI 음성 요약 기능 도입('25.5), 판매자 지원 AI 에이전트와 도구 출시('25.9) - (투자/인수) HUMAIN과 사우디아라비아 AI 인프라와 AI 개발에 50억 달러 투자('25.5), AI 웨어러블 기업 Bee 인수('25.7)
	Apple	- (기반모델) 이미지-텍스트 멀티모달 모델 'MM1'('24.3) - (플랫폼) 'Apple Intelligence' 출시('24.10), Apple Intelligence 신기능 업데이트('25.6), 에어팟 실시간 번역 및 Apple Intelligence 언어 지원 확장 iOS 26.1 베타 공개('25.9) - (기술) 모바일 화면의 UI를 이해하는 멀티모달 모델 개발('24.4), 멀티모달 모델 훈련용 프레임워크 '4M' 데모 공개('24.6), 단백질 구조 예측 AI 모델 'SimpleFold' 개발('25.9) - (제휴) 오픈AI와 파트너십 발표('24.6)
	xAI	- (기반모델) 'Grok' 공개('23.11), Grok 오픈소스화 및 Grok-1.5 공개('24.3), 'Grok-1.5 Vision' 프리뷰 공개(24.4), 이미지 생성 통합한 'Grok-2' 공개('24.8), Grok-3('25.2), Grok-4('25.7), Grok Code Fast 1('25.8), Grok-4 Fast('25.9), Grok-4.1('25.11), Grok-4.1 Fast('25.11) - (서비스) 미국 정부용 'Grok for Government' 출시('25.7), 미 전쟁부에 GenAI.Mil 제품군의 일부로 납품('25.12) - (투자) 60억 달러 시리즈 B 투자 유치('24.5), 50억 달러 투자 유치('24.11), 60억 달러 유치('24.12)

가치사슬	기업명	주요 활동 (기술·모델 개발, 투자유치(인수합병), 사업전략(협업제휴), 제품/서비스 출시 등)
	NVIDIA	- (기반모델) Hugging Face 및 ServiceNow와 코드 생성 AI 'Starcoder 2' 공개('24.2), 텍스트에서 3D 생성하는 Latte3 발표('24.3), 'Nemotron-4 340B' 공개('24.6), 'Llama-3.1-Nemotron-70B-Instruct' 공개('24.10), 텍스트-오디오 생성 AI 모델 '푸가토' 개발('24.12), Llama-3.1 Nemotron Ultra('25.4), Cosmos Physical 모델과 다국어 음성 모델 'NVIDIA Canary-1b-v2'('25.8), 'Nemotron 3' 발표('25.11) - (플랫폼) 기업용 AI 및 데이터 분석 플랫폼 'NVIDIA AI Enterprise('21.8), 맞춤형 AI 모델 구축 및 배포 플랫폼 'NVIDIA AI Foundry('24.7) - (기술) 휴머노이드 로봇용 범용 기반모델 '프로젝트 그루트' 공개('24.3), 휴머노이드 개발을 위한 개방형 파운데이션 모델 'GRooT N1' 와 로봇 개발을 위한 오픈소스 물리엔진 Newton 발표('25.3), 로봇 학습을 위한 'Isaac Sim'과 'Isaac Lab' 프리뷰 공개('25.6) - (협업) Microsoft, Anthropic과 전략적 파트너십 발표('25.11)
	Cohere(캐나다)	- (기반모델) 다국어 LLM 출시('24.2), 기업용 신규 LLM 'Command-R+' 출시 ('24.4), 다국어 LLM 'Aya Expanse' 출시('24.10), 비전 모델 'Aya Vision', 기업용 LLM 'Commad A'('25.3), 멀티모달 검색 모델 'Embed 4'('25.4), 번역 특화 'Command A Translate'와 추론 AI 모델 'Command A Reasoning' ('25.8), 검색 결과 순위 재조정 모델 'Rerank 4'('25.12) - (서비스) 기업용 AI 챗봇 '코랄(Coral)' 출시('23.7), 기업용 AI 플랫폼 'North' ('25.1) - (투자) 5억 달러 규모의 시리즈 D 투자 유치('24.7), 1억 달러 투자 유치('25.9)
	IBM	- (기반모델) 기업용 LLM 'Granite' 오픈소스 공개('24.5), 'Granite 3.0' 공개('24.10), 'Granite 3.1'('24.12), 추론 특화 'Granite 3.2'('25.2), Granite 4.0 Tiny Preview('25.5), 'Granite 4.0' 정식 공개('25.10) - (서비스/플랫폼) 'watsonx Orchestrate' 정식 출시, IBM Consulting에서 독점 사용하는 AI 기반 플랫폼 'Quantum Safe Migration Orchestrator'('25.12) - (기술) NASA와 제휴해 태양 데이터 관측 오픈소스 모델 'Surya' 공개('25.8)
	SalesForce	- (기반모델) 기업용 기반모델 'xGen-Sales' 및 'xLAM' 공개('24.9) - (서비스) 완전 자율형 AI 에이전트 'Einstein Service Agent 공개('24.7), - (플랫폼) 로우코드 기반 생성 AI CRM 플랫폼 'Einstein 1 Studio' 공개('24.3), 로우코드 기반 자율형 AI 에이전트 'Agentforce' 공개('24.9) 에이전트 마켓플레이스 'Agent Exchange'('25.3), AI 에이전트 플랫폼 'Agentforce 3'('25.6), 정부용 AgentForce('25.8)AI 에이전트 조율 'MuleSoft Agent Fabric' 출시('25.9). - (인수) AI 스타트업 Zoomin 인수('24.9), 클라우드 데이터 관리기업 Informatica와 AI 에이전트 기업 Convergence.ai 인수('25.5), 자연어-SQL 플랫폼인 Waii와 AI 솔루션 기업 Regrello 인수('25.8)
	Hugging Face	- (기반모델) 1,000여 명 연구자 커뮤니티가 참여한 빅사이언스 프로젝트의 공동개발로 파라미터 1,760억 개 이상의 오픈소스 LLM 공개('22.7), 온디바이스AI용 'SmolLM2' 공개('24.11), 로봇용 시각-언어-행동(VLA) 모델 'SmolVLA'('25.6) - (서비스) 15조 개 토큰의 대규모 오픈소스 데이터셋 공개('24.6), AI 배포 서비스 HUGS 발표('24.10), AI 기반 웹앱 개발 도구 'OpenAI-Gradio' 공개('24.10), 오픈소스 Deep Research 공개(25.2) - (제품) 휴머노이드 로봇 HopeJR과 Reachy Mini 공개*'25.5), 'Reachy Mini' 정식 출시('25.7) - (기술) Computer Use용 GUI 에이전트 'Smol2Operator' 공개('25.9)
	Databricks	- (기반모델) 매개변수 1,320억 개의 'DBRX' 오픈소스 공개('24.3) - (플랫폼) 기업용 AI 에이전트 개발 자동화 솔루션 'Agent Bricks' 출시('25.6) -(제휴) AWS와 생성 AI 기능 제공 제휴 체결('24.10), Palantir와 AI 서비스 제공 제휴, Anthropic과 AI 에이전트 시장 확대 제휴('25.3), Google Cloud와 제휴로 'Data Intelligence Platform'에 Gemini 모델 기본 제공('25.6), OpenAI와 'Agent Bricks'에 오픈AI 모델 기본 제공 제휴('25.9) - (투자/인수) 자체 AI 모델 구축 지원하는 'Mosaic ML' 인수('24.3), 100억 달러 규모 시리즈 J 투자 유치('24.12)
	Snowflake	- (기반모델) 기업용 오픈소스 LLM 'Arctic' 공개('24.4) - (플랫폼) LLM 기반 앱 구축 지원하는 완전 관리형 서비스 'SnowFlake Cortex' 운영, 지능형 에이전트 'Snowflake Intelligence'와 데이터 사이언티스트 지원 'Data Science Agent' 출시('25.6), 스타트업용 AI 앱 개발 플랫폼 'Snowflake for Startup' 공개('25.9)

가치사슬	기업명	주요 활동 (기술·모델 개발, 투자유치(인수합병), 사업전략(협업제휴), 제품/서비스 출시 등)
	Mistral AI (프랑스)	- (기반모델) 'Mistral Large' 공개('24.2) 'Mixtral 8x22B' ('24.4), 코드 생성 모델 'Codestral' ('24.5), 'Mistral Nemo' 및 'Mistral Large 2' ('24.7), 'Pixtral 12B' ('24.9), 온디바이스용 'Les Ministro' ('24.10) 'Pixtral Large'('24.11), 'Mistral Small'('25.1), 코딩 특화 Devstral('25.5), 음성 AI 모델 'Voxstral'('25.7), 'Codestral 25.08'('25.7) - (서비스) 챗봇 서비스 Le Chat 출시('24.2), 유해 콘텐츠 탐지용 'Moderation API' 공개('24.11), 'Le Chat'에 웹 검색, 캔버스, 에이전트, 이미지 생성 기능 추가('24.11), 'Le Chat'의 모바일 앱과 기업용 버전 출시('25.2), 'Mistral Code'('25.6), 'Le chat'에 딥리서치와 음성 입력 기능 추가('24.7), 'Le chat에 메모리 기능 추가('25.9) - (투자) Microsoft로부터 1,500만 유로 투자 유치('24.2) 6,400만 달러 투자 유치('24.6) - (제휴/투자) Google과 클라우드 파트너십('23.12), UAE G42와 차세대 AI 플랫폼 구축 협력('25.5), 17억 유로 투자 유치('25.9)
	G42 (UAE)	- (기반모델) 자회사 Inception, JAIS 공개('23.8), JAIS 70B 공개('24.8), MBZUAI와 공동 개발한 'K2 Think' 공개('25.9) - (투자) Microsoft, 15억 달러 투자로 G42 일부 지분 확보 및 이사회 참여('24.4) - (제휴) 오픈AI와 AI 인프라 구축 협력('25.5)
	AI21 Labs (이스라엘)	- (기반모델) 'Jamba 1.0' ('24.3), 'Jamba 1.5 Mini와 Jamba 1.5 Large' ('24.8), Jamba 1.6('25.6) - (플랫폼) AI 플래닝 및 오케스트레이션 플랫폼 'Maestro' 출시('25.3) - (투자) Goolge, Nvidia, Intel 등으로부터 2억 8천만 달러 유치('23.11), 3억 달러 투자 유치('25.5)
	Zhipu.AI (중국)	- (기반모델) ChatGLM2-6B('23), GLM4-9B('24.6), 멀티모달 GLM-4V-Plus('24.8), 추론 AI 모델 'GLM-Z-1-32B-0414('25.4), GLM-4.5와 GLM-4.5-Air, CogVideoX-3 비디오 생성 모델 출시 ('25.7), GLM-4.5V('25.8), GLM-4.6V('25.9), GLM-4.7V('25.12) - (서비스) AI 에이전트 'AutoGLM Rumination'('25.3) - (투자) Tencent와 Alibaba로부터 3억 4,200만 달러 투자 유치('23.10), 10억 위안 투자 유치('25.7)
	Alibaba(중국)	- (기반모델) LLM 'Qwen-7B'와'Qwen-7B-Chat'('23.8), 동영상 생성 AI 모델 'EMO' ('24.2), 'Tongyi Qianwen 2.5' ('24.5), 수학 특화 AI 모델 'Qwen2-Math'('24.8), 20분 이상 영상 분석 가능한 'Qwen-VL' ('24.9), Qwen2.5-Coder ('24.11), 추론 특화 'Marco-o1' ('24.11), Qwen-2.5 Max('25.1), 비디오 생성 AI 'Wan 2.1'('25.2), 추론 특화 'QWQ-32B'('25.3), 추론 특화 'Qwen 3'('25.5), 추론 특화 Qwen3-235B-A22B-Thinking-2507'('25.7),영상 생성 AI 'Wan2.2'('25.7), Qwen3-Coder와 Qwen Image('25.8), Qwen3-Max-Preview와 Qwen3-Omni, Qwen3-Next, Wan2.5 프리뷰 ('25.9) - (플랫폼) AI 모델 오픈소스 플랫폼 'ModelScope' 공개('22.11)
	Tencent (중국)	- (기반모델) 기업용 LLM 'Hunyuan' 공개('23.9), 'Hunyuan-Turbo' 공개('24.9), Hunyuan-Large 공개('24.11), 'HunyuanVideo'('24.12), 'Hunyuan Turbo S'('25.2), 추론 AI 모델 'Hunyuan-T1과 HunyuanVideo-I2V('25.3), HunyuanVideo-Avatar('25.5), Hunyuan3D-2('25.6), Hunyuan Instruct 7B/4B/1.8 B/0.5 B('25.8), 번역 AI 모델 'Hunyuan-MT', Hunyan Image 3.0, Hunyuan 3D 3.0('25.9) - (서비스) AI 챗봇 'Yuanbao' 출시('24.5), Yuanbao 업데이트('25.3)
	Baidu (중국)	- (기반모델) 'Ernie 3.5' 공개('23.6) 'Ernie 4.0' 공개('23.10), 'Ernie 4.0 Turbo 공개('24.6), 'Ernie 4.5'와 'Ernie X1'('25.3), 'Ernie 4.5 Turbo'와 'Ernie X1 Turbo'('25.4), Ernie X1.1('25.9) - (서비스) AI 챗봇 'Ernie Bot' 출시('23.3), 'Ernie Bot' 사용자 3억 돌파('24.6), AI 코딩 어시스턴트 Comate 3.5('25.9) - (플랫폼) 딥러닝 오픈소스 프레임워크 'PaddlePaddle' 운영, AI 기반으로 검색 플랫폼 대규모 업데이트 ('25.7),PaddlePaddle v3.2 업그레이드('25.9)
	Sensetime (중국)	- (기반모델) 'SenseNova 4.0 공개('24.2), 'SenseNova 5.0' 공개('24.4), 실시간 멀티모달 모델 'SenseNova 5.5' 공개('24.7), SenseNova V6('25.4), SenseNova V6.5('25.8) - (서비스) AI 챗봇 'SenseChat' 공개 ('23.4), AI 아바타 동영상 생성 'Vimi' 공개 및 Code Raccoon 업데이트('24.7), AI 챗봇 Sensechat에 실시간 오디오 및 비디오 상호작용 등 신기능 추가('25.6)

가치사슬	기업명	주요 활동 (기술·모델 개발, 투자유치(인수합병), 사업전략(협업제휴), 제품/서비스 출시 등)
	ByteDance (중국)	- (기반모델) 텍스트-비디오 생성 모델 'MagicVideo-V2' 공개('24.1), 로봇 AI 초거대 모델 'GR-2'('24.10), 이미지 기반 전신 영상 생성하는 'OmniHuman-1'('25.2) 추론 AI 모델 'Dubao 1.5'('25.4), 로봇 AI 'GR-3 VLA' 모델과 범용 양팔 이동 로봇 'Bytemini' 발표('25.7), 오픈소스 모델 'Seed-OSS-36B'('25.8), 이미지 생성 AI 'Seedream 4.0'('25.9) - (서비스) AI 챗봇 'Dubao' 출시('23.8) 동영상 생성 AI 도구 'Jimeng AI' 앱스토어 출시('24.8), Dubao에 음악 생성 기능 추가('24.8), Dubao에 영상 생성 기능 추가('24.11) - (플랫폼) AI 챗봇 개발 플랫폼 'Coze' 출시('24.2), AI 에이전트 플랫폼 'Coze Space' 공개('25.4)
	MInimax (중국)	- (기반모델) LLM 'ABAB 6.5' 출시('24.4). 동영상 생성 모델 'ABAB-Video-01' 출시('24.9), 텍스트 모델 MiniMax-Text-01, 이미지와 텍스트 모델 MiniMax-VL-01('25.1), 추론 AI 모델 'MiniMax-M1'('25.6) - (플랫폼) 소비자향 AI 플랫폼 'Hailuo AI' 공개('24.5), Hauluo AI에 이미지-비디오 기능 출시('24.10)
	DeepSeek (중국)	- (기반모델) LLM 'V3'('24.12), 추론 AI 모델 'R1'('25.1), 이미지 AI 모델 'Janus Pro'('25.1), V3 업데이트('25.3), R1-0528 업데이트('25.5), V3.1('25.8), V3.1 업데이트('25.9), V3.2-Exp('25.9)
	Manus AI (싱가포르)	- (플랫폼) AI 에이전트 플랫폼 'Manus' 공개('25.3), Manus 구독 요금제와 모바일 앱 출시('25.4), 'Manus 1.5 ', 'Manus 1.5 lite '('25.10) 'Manus 1.6 ', 'Manus 1.6 Max'('25.12) - (서비스) 심층연구 기능 'Wide Research' 공개('25.7) - (인수) Meta에 인수('25.12)
	Moonshot AI (중국)	- (기반모델) LLM 'Kimi K1.5' 공개('25.1), Kimi-VL과 Kimi-Audio('25.4), Kimi-Dev와 Kimi K1.5 기반 에이전트 모델 'Kimi-Reseacher' 공개('25.6), LLM 'Kimi K2' 오픈소스 공개('25.7), Kimi K2-Instruct-0905('25.9), 'Kimi K2 Thinking'('25.11)
	Allen AI Institute	- (기반모델) 오픈소스 LLM 'OLMo' 공개('24.2), OLMo 1.7-7B 공개('24.4), Molmo 4종 공개 ('24.9), Tülu 3('25.1), 3D 공간 행동추론 모델 'MolmoAct'('25.8), 개방형 음성인식 모델 'OLMoASR'('25.8), 지구 관측 기반 모델 'OlmoEarth' 제품군 공개('25.11), 'Molmo 2'출시('25.12) - (기술) AI의 과학 논문 작업 성능 평가하는 SciArena 플랫폼 개발('25.7), AI 모델 훈련 프레임워크 'FlexOlmo' 공개('25.7), 과학연구 AI 에이전트 프레임워크 'Asta' 공개('25.8), 'Asta DataVoyager' 정식 출시('25.10)
	World Labs	- (사업) '24.4월 설립 후 1억 달러 유치('24.9) - (기반모델) 단일 이미지에서 3D 월드 생성하는 월드 모델 초기 프리뷰 공개('24.12), 3D 월드 모델 'Marble' 프리뷰 공개('25.9)
	Sakana AI(일본)	- (기반모델) 일본어 LLM 'EvoLLM-JP', 'EvoVLM-JP' 공개('24.3) - (기술) AI 모델을 병합해 새로운 모델을 생성하는 방식 연구('24.3), 과학 연구 자동화하는 AI Scientist 개발('24.8), 효율적 추론 모델 개발을 위한 '강화학습교사' 기법 개발('25.6), LLM 간 협력으로 추론 능력 높이는 'AB-MCTS' 기법 개발('25.7), 과학연구 프로그램 진화 프레임워크 'ShinkaEvolve'('25.9) - (투자) 1억 달러 규모의 시리즈 A 투자 유치('24.9), 200억 엔 규모의 시리즈 B 투자 유치('24.11)

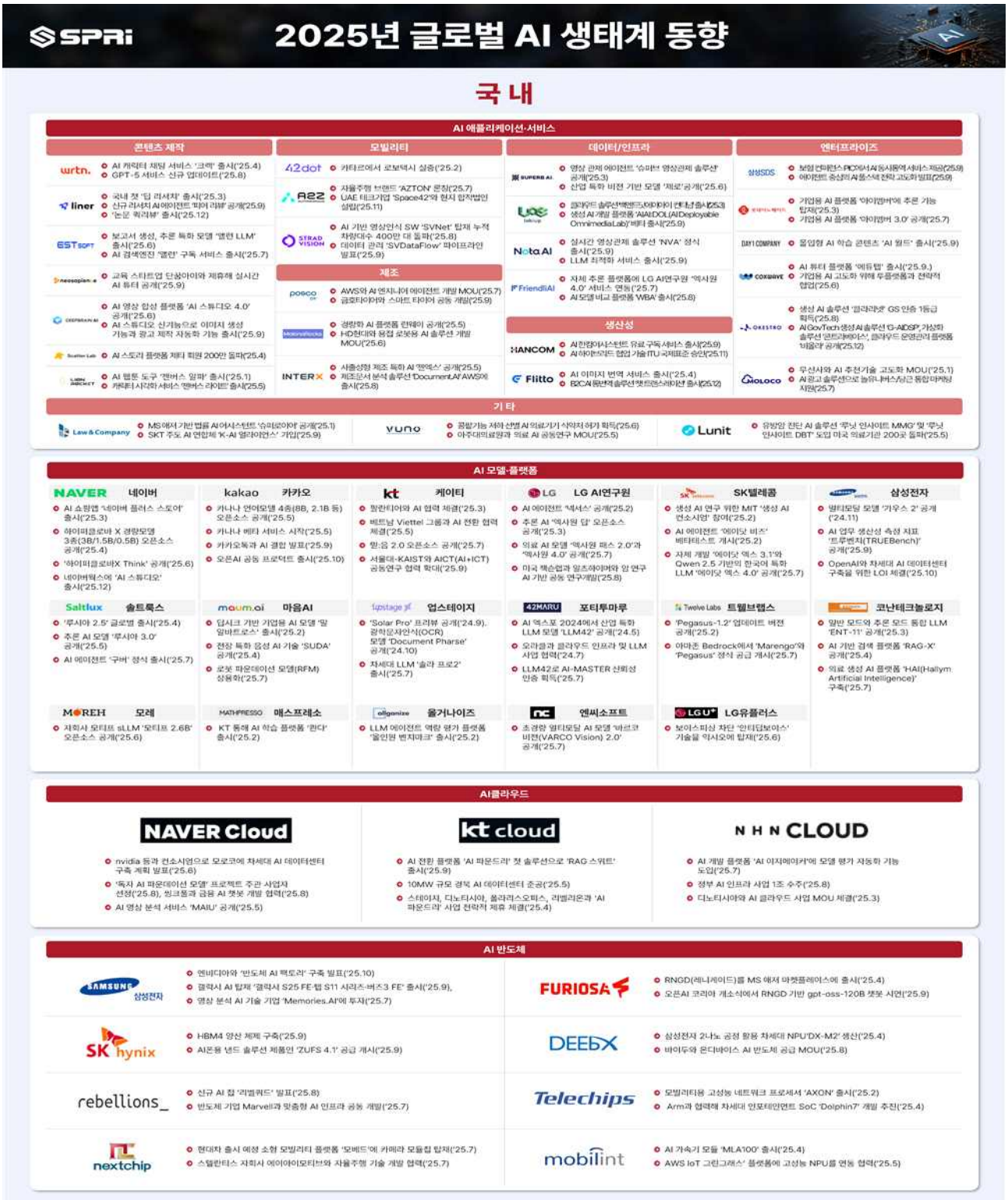
가치사슬	기업명	주요 활동 (기술·모델 개발, 투자유치(인수합병), 사업전략(협업제휴), 제품/서비스 출시 등)
클라우드	AWS	- (서비스) 기업의 AI 생성 서비스를 지원하는 'Amazon BedRocks' 공개('23.4), 머신러닝 모델 구축, 훈련, 배포를 지원하는 완전 관리형 서비스 Amazon SageMaker 운영, 기업용 에이전트 구축 플랫폼 'Amazon Bedrock AgentCore' 및 AI 기반 통합개발환경 'Kiro' 출시('25.7), AI 에이전트 개발 지원 Nova Act 확장 프로그램 출시('25.9) - (기업 투자) Anthropic에 40억 달러 투자('23.9), 생성 AI 스타트업 양성에 2억 3천만 달러 투자('24.6), Anthropic에 40억 달러 추가 투자('24.11) - (인프라) 독일 클라우드 인프라 확장에 100억 유로 투자('24.6), 호주 데이터센터 확장에 200억 호주 달러 투자, SK와 함께 울산 데이터센터 투자(25.6)
	Oracle	- (서비스) 'Oracle AI Agent Studio for Fusion Applications' 출시('25.3), '오라클 AI 데이터베이스 26ai', 'AI 데이터 플랫폼' 출시, 'Oracle Autonomous AI Lakehouse' 발표('25.10) - (사업) G42와 AI 기반 글로벌 의료 서비스 플랫폼 출시를 위한 전략적 파트너십('25.5), AMD와 협력하여 'AMD Instinct™' MI355X GPU 지원 계획을 발표('25.6)
	MS Azure	- (인프라) MS-Azure Open AI API 서비스 제공 참여, Hugging Face의 Azure 활용, Meta의 LLaMA2 활용('23.7), AI 데이터센터에 연간 800억 투자 계획 발표('25.1), 위스콘신에 두 번째 AI 데이터센터 건설 계획('25.9), 영국 AI 인프라에 300억 달러 투자('25.9), 자체 개발 차세대 CPU 'Azure Cobalt 200'를 발표하고 MS 데이트 센터에서 가동('25.11) - (플랫폼) 생성 AI 앱 개발 플랫폼 Azure AI Studio를 확장한 AI 개발 플랫폼 'Azure AI Foundry' 출시('24.11), 'Azure AI Foundry'에 신규 모델(xAI Grok3) 및 에이전트 서비스 추가 (25.5), 클라우드 솔루션과 AI 앱, 에이전트용 'Microsoft Marketplace' 공개('25.9), 'Foundry'에 전문 보안 에이전트 기능 'Azure AI Content Safety 에이전트' 추가('25.10)
	Google Cloud	- (인프라) 미국 주요 주에서 2년간 데이터센터와 AI 인프라에 250억 달러 투자계획 발표('25.7), 버지니아주 AI 인프라에 90억 달러 투자 및 오클라호마 AI 인프라에 90억 달러 투자('25.8), 자체 설계 TPU '아이언우드(Ironwood)' 및 CPU '악시온(Axion)'로 AI 인프라 대폭 업그레이드('25.11) - (제휴) Hugging Face와 생성 AI 개발 지원 협력('24.1), Nvidia와 제휴해 Blackwell AI 컴퓨팅 플랫폼 도입 및 Google Cloud에 Nvidia DGX Cloud 서비스 적용('24.3), Oracle과 멀티 클라우드 파트너십 발표('24.6), AI 에이전트 파트너 프로그램 공개('24.11), OpenAI에 서비스 공급 계약 체결('25.6) - (투자/인수) 클라우드 보안 플랫폼 기업 Wiz 인수('25.3), AI 스타트업 지원 위한 신규 투자 프로그램 'AI Futures Fund 출시('25.5)
	IBM	- (제휴) Intel과 제휴해 IBM 클라우드상에서 AI 가속기 'Gaudi 3' 제공('25.5) - (인수) 데이터 및 AI 컨설팅 기업 Hakoda 인수('25.4) - (서비스) IBM cloud에서 비즈니스 인텔리전스 지원 'IBM watsonx BI' 출시('25.8), IBM Watson X Orchestrate에 에이전트 관찰과 거버넌스 기능 도입('25.8)
	Alibaba Cloud (중국)	- (인프라) 3년간 클라우드와 AI 인프라에 3천 800억 위안(약 75조 원) 투자 계획 발표('25.2), 한국에 두 번째 데이터센터 개소('25.6), 브라질 프랑스 네덜란드에 데이터센터 설립 및 2026년 멕시코, 일본, 한국, 말레이시아에 추가 데이터 센터 설립 계획('25.9) - (협업) 고객, 개발자, 파트너 대상 생성 AI 교육과정 제공(24.5), SAP와 제휴해 클라우드 인프라에서 SAP의 기업용 소프트웨어 공급(25.5), 싱가포르에 AI 글로벌 역량 센터 출범('25.7) - (서비스) AI 플랫폼(PAI)에 Deepseek 모델 추가('25.2), PAI에 고성능 AI 추론 기능 강화('25.7) - (투자) 중국 AI 스타트업 Moonshot AI에 8억 달러 투자('24.5), 글로벌 파트너 생태계 강화와 AI 혁신 가속화 위한 6천만 달러 투자 계획('25.7)
하드웨어	NVIDIA	- (제품) 신규 AI 칩 'Blackwell' 공개('24.3), 개인용 AI 슈퍼컴퓨터 'Project Digit('25.1), 미국산 AI 슈퍼컴퓨터 생산 계획 발표('25.4), 2025년 하반기 'Blackwell Ultra' 출시 및 2026년 'Vera Rubin' 출시 계획 발표('25.3), 'Blackwell Ultra' 출시, Blackwell 기반 로봇 컴퓨터 'Jetson Thor' 출시('25.8), 대규모 컨텍스트 추론 특화 GPU 'Rubin CPX' 발표('25.9) - (제휴) Softbank 지원해 일본 최대 AI 슈퍼컴퓨터 구축 추진('24.11), 유럽 기업 및 정부와 협력해 유럽 AI 인프라 구축 계획('25.6) - (투자/제휴) 미국에서 4년간 최대 5,000억 달러 AI 인프라 투자 계획 발표('25.4), OpenAI에 1,000억 달러 투자, 영국 AI 스타트업 생태계에 20억 파운드 투자, Intel과 AI 인프라와 개인용 컴퓨터 개발 제휴 ('25.9)

가치사슬	기업명	주요 활동 (기술·모델 개발, 투자유치(인수합병), 사업전략(협업제휴), 제품/서비스 출시 등)
	Intel	- (제품) AI 칩 'Gaudi 3' 공개('24.4), 차세대 AI 칩 'Lunar lake' 공개('24.9), AI 성능 높이는 CPU 신제품 'Xeon 6' 출시('25.5), Dell AI 플랫폼에 Gaudi 3 탑재('25.5), 'Lunar lake' 후속작 'Panther Lake' 출시 예고('25.10) - (인력) AI PC Acceleration Program 일환 AI PC Developer 프로그램 신설('24.2) - (제휴/투자) LG이노텍과 제휴해 AI 기반 스마트팩토리 구축 지원('25.8), NVIDIA와 맞춤형 데이터센터와 PC 제품 공동 개발 및 50억 달러 투자 유치('25.9)
	Qualcomm	- (제품) Snapdragon 7 Gen 4 모바일 플랫폼 공개('25.5), 차세대 PC 프로세서 'Snapdragon X2 Elite Extreme'과 모바일 CPU 'Snapdragon 8 Elite Gen 5' 공개('25.9), - (제휴/인수) 엣지 기기용 AI 개발 플랫폼 Edge Impulse 인수('25.3), 베트남 스타트업 VinAI의 생성 AI 사업부 인수('25.4), HUMAIN과 AI 데이터센터 협력('25.5), Harman과 자동차 분야 생성 AI 협력, Google Cloud와 자동차 산업의 에이전틱AI 협력('25.9), Advantech과 IoT용 엣지 AI 협력('25.5)
	AMD	- (제품) 대규모 AI 개발 특화 GPU 'M1300X' 공개('23.6), AI 칩 'Instinct MI300' 시리즈('23.12), 3세대 라이젠 AI 프로세서 등 신제품 공개('24.6), 신규 AI 칩 'MI325X' ('24.10), 차세대 AI 칩 MI400 ('25.6) - (플랫폼) Nvidia CUDA에 대항하는 'ROCm 7.0' 출시('25.10) - (인수) AI SW 기업 Nod.ai 인수 ('23.10), 핀란드 AI 스타트업 Silo AI 인수('24.7), AI 서버 기업 ZT System 인수('24.8) - (제휴) Cohere와 글로벌 협력 확대('25.9)
	Groq	- (제품) AI 추론 칩 LPU 공개('24.2) - (제휴) 삼성전자와 차세대 AI 칩 생산 위한 파트너십 체결('23.8), 데이터센터 기업 Equinix와 협력해 핀란드에 첫 유럽 데이터센터 개설('25.7) - (인수) 비즈니스 AI 솔루션 기업 Definitive Intelligence 인수('24.3) - (투자) 삼성전자 주도의 시리즈D 투자 라운드에서 6억 4천만 달러 유치(24.8), 사우디아라비아로부터 15억 달러 투자 유치('25.2), 7억 5천만 달러 투자 유치('25.9)
	Tenstorrent	- (제품) AI 추론 특화 PCIe 확장 카드 'Grayskull' e75와 e50 발표('24.3) - (제휴) 삼성전자와 차세대 AI 칩 생산 제휴 체결('23.10), UAE AI 스타트업 AIREV와 생성 AI 스택 개발 제휴('25.6) - (투자) 시리즈 D 투자 6억 9,300만 달러 유치('24.12)
	Meta	- 자체개발 AI 가속기 'MTIA v1' 공개('23.5) 'MTIA v2' 공개('24.4), MTIA 칩 시험 생산 개시('25.3)
	OpenAI	- (제휴) Broadcom 등 AI 칩 설계 기업과 협력 논의('24.7) 자체 AI 칩 개발 위해 '26년 가동하는 TSMC 최첨단 공정 'A16' 예약('24.9), Broadcom 및 TSMC와 자체 AI 칩 개발 추진('24.11), 2026년 TSMC에서 자체 AI 칩 양산 계획 수립('25.2), Broadcom과 자체 AI 칩 설계 개발 완료 및 2026년 양산 계획('25.9)
	Microsoft	- 자체개발 AI 가속기 'Maia 100'와 CPU 'Cobal 100' 공개('23.11), AI PC 'Copilot Plus PC' 공개 ('24.5), 데이터센터용 AI 칩 2종 출시('24.11), 자체 AI 칩 'Maia'의 양산 일정 2026년으로 연기 발표 ('25.6)
	Google	- (제품) ARM 기반 맞춤형 CPU 'Axion' 공개('24.4), 텐서플로우 연산에 최적화된 TPU(Tensor Processing Unit) v6 공개 및 TPU 신제품 'v5p' 정식 버전 출시 ('24.5), 추론 최적화 차세대 TPU 'IronWood' ('25.4) - (서비스) TPU 모니터링 라이브러리 공개('25.7)
	Cerebras	- (제품) 세계 최고 속도의 AI 칩 'WSE-3' 공개 ('24.3) 자사 추론 클라우드 플랫폼에서 고속 추론 모델 'Qwen3-235B' 공개('25.7) - (제휴) 생성 AI용 컴퓨팅 인프라 제공 위해 Dell과 협력('24.6), NinjaTech와 협력해 AI 기반 개발 솔루션 'Fast Dep Coder' 출시 지원('25.8) - (인프라) 추론 전문 데이터센터 미국과 유럽에 6개 구축 발표('25.3)
	SambaNova	- (제품) 저비용 구동 가능한 프로세스 'SN40L 출시('23.9), AWS Marketplace에서 AI 플랫폼 공개 ('25.5) - (서비스/플랫폼) 데이터센터용 턴키 AI 추론 솔루션 출시('25.7)

[첨부 4] 2025년 글로벌 AI 생태계 동향

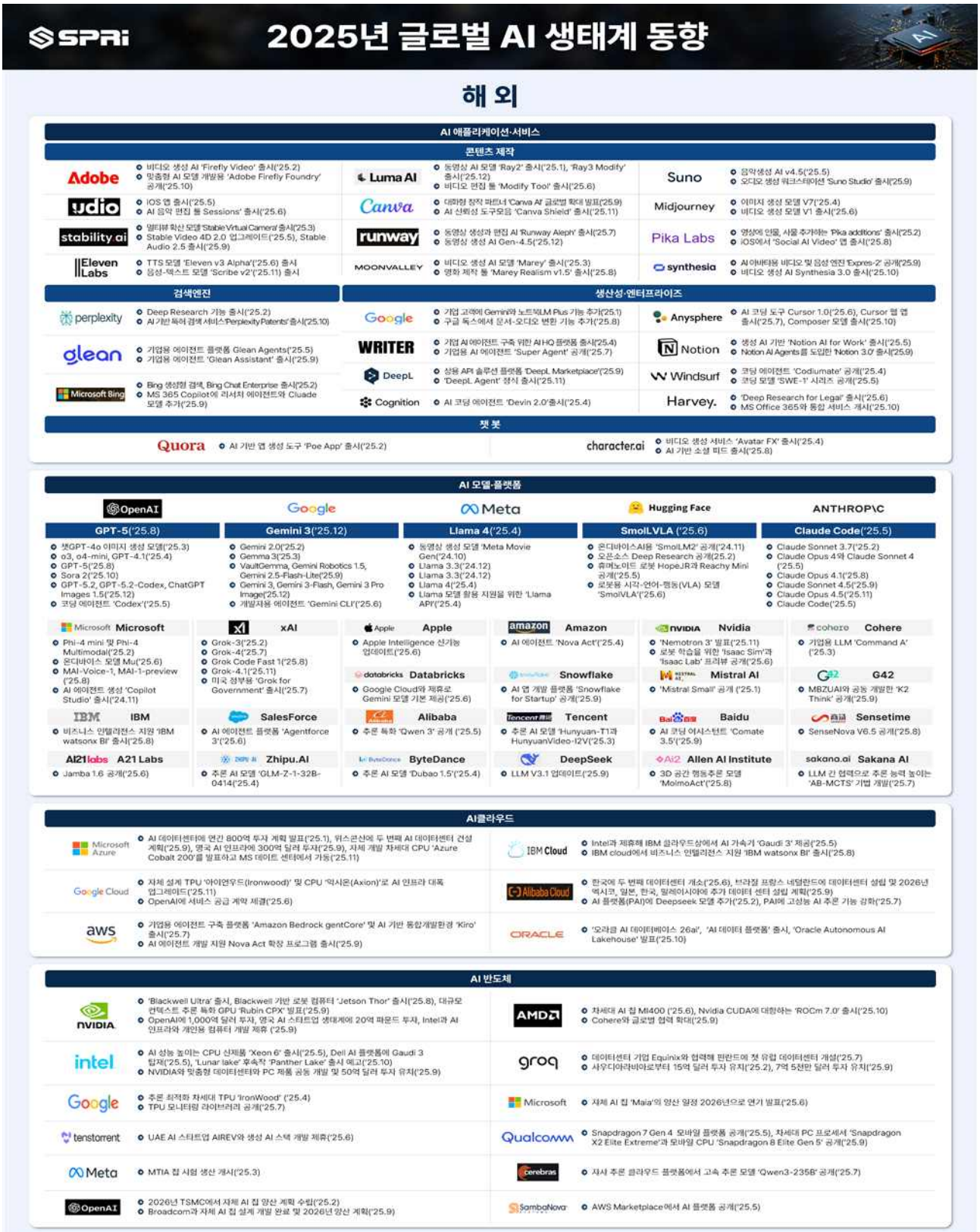
① 국내 AI 기업 동향

[그림] 2025년 글로벌 AI 생태계 동향(국내 AI 기업)



② 해외 AI 기업 동향

[그림] 2025년 글로벌 AI 생태계 동향(해외 AI 기업)



[첨부 5] AI 산업 동향 전문가 간담회 개최 결과

① 제1차 전문가 간담회 개최 결과

〈표〉 제1차 중소기업 AI 활용 간담회 개최 결과 주요 내용

구 분	주 요 내 용
목적	• 중소기업 대표이사 또는 담당 임원을 초청해 중소기업의 AI 활용 현황 및 장애 요인, 정책 지원 방향에 대해 의견 청취
일시 및 장소	• 일시: 2025년 4월 21일(월) 오전 10시~12시 • 장소: B-TIME (강남역 5번 출구 인근 비트빌딩 B1)
참석자	• 총 10명(전문가 4명 + 주관기관 3명 + 수행사 3명) - 전문가(4명): 중소기업 종사자 (A: 금융 및 자동차 분야 SW 대표, B: 바이오 분야 연구이사, C: 운수업 분야 대표, D: 모바일 라이브 스트리밍 분야 SW 대표) - 주관기관(3명): SPRi 이해수 선임연구원, 유재흥 책임연구원, 강호준 선임연구원 - 수행사(3명): 넥스텔리전스 김동진 연구위원, 조현도 수석연구원, 박찬미 주임연구원
주요 안건	• AI 실제 도입 현황, 성과 및 애로사항 • AI 활용 계기 및 협업·도입 방식 • AI 활용을 위한 정책적·제도적 지원
회의 결과 (전문가 주요 의견)	**□ AI 실제 도입 현황, 성과 및 애로사항** o A • 사업 개요: 금융 및 자동차 분야 소프트웨어 테스트 자동화 전문 기업 • AI 활용: 테스트 케이스 생성 및 결과 분석에 생성형 AI 도입, 퍼블릭AI와 CND 방식 협업 • 어려움: AI 인력 자체 채용은 어려워 제휴 방식 선택. 유료 AI 도입에 대한 정부 지원 필요 제안 • 성과: 주요 고객사 대상 추가 계약 유도 o B • 사업 개요: 건강기능식품 원료 개발 제조업체 • AI 활용: 개인적으로 마케팅 기획, 법률 검토 등에서 ChatGPT 활용 • 어려움: 회사 차원의 AI 활용은 없음. 조직 차원의 수요나 인식 부족 • 제안: "AI + 업무" 융합형 실무 교육 프로그램이 더 실효성 있을 것 o C • 사업 개요: 운수업용 ERP·무인 관제 솔루션 운영 • AI 활용: - ERP 내 반복 업무 자동화, 특이사항 감지 - CCTV+AI 통한 무인 상황 인식 시스템 - 챗봇 기반 경영 대시보드 및 AI 영상 콘텐츠 제작 • 어려움: 전담 AI 인력 확보가 어려워 프리랜서 및 지인 네트워크 활용 • 성과: 서울 시내 12% 법인택시 고객사 확보. 인력비 절감 및 효율 향상 o D • 사업 개요: 모바일 라이브 스트리밍 앱 운영

구 분	주 요 내 용
	• AI 활용: - 방송 중 배경 교체, 트래킹, TTS·STT 적용 - 고객 응대 자동화 등 다방면 활용 • 어려움: 자체 모델 개발은 인력·비용 부담 큼. 대부분은 기존 오픈모델 활용 • 제안: 각 산업별 사례와 노하우가 공유될 수 있도록 정부 차원의 정리 필요 **□ AI 활용 계기 및 협업·도입 방식** o A • 고객 요청, 내부 필요성 인식, 본인 주도 세 요인이 복합적 • 퍼블릭AI와의 협업은 전략 임원 간 네트워크 기반 • 자체 AI 역량 확보보다는 TX(테스트 경험) 축적과 제휴를 핵심 전략으로 설정 o B • 전통 제조업 특성상 회사 차원의 도입 유인은 적음 • 개인의 실무에서 AI 필요성을 체감하고 활용 o C • 운수업 경험에서 ERP 필요성을 느끼고 창업 • 내부는 핵심 백엔드만 유지하고, AI 기능은 외주와 프리랜서 적극 활용 • AI 기반 챗봇·영상 툴 활용 증가 o D • GCP 등 클라우드 플랫폼 네트워크 활용해 최신 AI 트렌드 파악 • 처음에는 오픈모델 활용 → 자체 모델 개발 필요성 인지 → 리소스 제약으로 구현 어려움 **□ AI 활용을 위한 정책적·제도적 지원** o A • 산업 전문가의 실제 경험을 바탕으로 한 '현장형 발표회' 제안 • 업무+AI 융합 전문 인력 육성이 중요 • 정부가 산업별 '멀티 인재' 양성을 간접 지원할 필요 o B • 전통 제조업에 적합한 AI 교육 프로그램 필요 • 신입 등에게 AI 활용 교육이 실질적 도움 될 수 있음 o C • 정부 정책과 기술의 간극 존재, 중소기업 입장에선 매칭·서류 부담이 커서 진입 장벽 큼 • 중간 거버넌스 체계를 통한 중소기업-대기업 협업 중재 필요 o D • 정부 R&D 사업 등에서 실효성 떨어졌던 경험 있음 (지식 전가, 결과물 공유 불균형) • 각 산업별 AI 활용 사례가 정리돼 공유되는 체계가 필요

② 제2차 전문가 간담회 개최 결과

<표> 제2차 피지컬 AI 전문가 간담회 개최 결과 주요 내용

구 분	주 요 내 용
목적	• 피지컬 AI 관련 기업 대표·담당 임원, 전문가를 초청해 정책 지원 방향에 대해 의견 청취
일시 및 장소	• 일시: 2025년 9월 29일(월) 오후 14시~16시 • 장소: 토즈 강남컨퍼런스점
참석자	• 총 10명(산업계 전문가 3명 + 연구계 전문가 5명) - 전문가(5명): 중소기업 종사자 (A: AMR 피지컬 AI 이사, B: Vision AI 및 MLOps 전문 기업 대표, C: 지능형 로봇 솔루션 전문기업 본부장, D: 로봇학과 교수, E: 로봇 분야 연구원 본부장) - 주관기관(3명): SPRi 이해수 선임연구원, 유재흥 책임연구원, 강호준 선임연구원 - 수행사(2명): 넥스텔리전스 김동진 연구위원, 박찬미 주임연구원
주요 안건	• 피지컬 AI 관련 최신 동향 • 규제 및 제도 관련 애로사항 • 글로벌 경쟁 환경(미국·중국 대비)과 기회
회의 결과 (전문가 주요 의견)	□ **피지컬 AI 관련 최신 동향** o A • 현재 회사는 AMR(자율이동로봇) 중심 사업 전개 • 청소로봇이 주 사업모델이었으나, 현재는 물류, 병원에서의 자율 이송이 주력 • 이동·관제를 하고 있지만 조작기술은 부족, 글로벌 파운데이션 모델 의존도가 높음 • 클라우드 기반 지원 중단 가능성, 비식별 데이터 문제 등 불확실성 존재 o B • 비전AI 회사로, 물리 환경 인식, 로봇 협업 지원 • 피지컬 AI 명칭을 앞세운 유사 기업 난립 → 시장 혼탁 우려 • 국내 SaaS 도입 문화가 부족(한국은 온프레미스 위주) → 산업 발전 저해 우려 o C • 회사는 비전 AI를 통한 객체 인식·제어 기술 보유 • 피지컬 AI 개념 자체가 불분명, 용어 중심의 정책 혼선 많음 • 여러 토론회, 간담회에 많이 참석하는 편인데 남는 것이 많지 않음 • 현장에서는 자율주행, 관제, 디지털 트윈 등 점진적 기술 수요가 현실적 o D • 실제 수요는 제조 AI이지만, ‘피지컬 AI’ 용어가 확산되며 정책 키워드화 • 조작기술은 빅테크(NVIDIA, Google 등) 중심으로 급속 발전, 국내 적용은 미비 • 제조업에서는 LLM보다는 현장 노하우 전승 문제 때문에 RAG가 더 유용 • 물류분야는 다품종 환경이라 조작 기술 수요가 크고, 이 때문에 피지컬 AI가 주목 받음 o E • 피지컬 AI 정의 혼선, 부처마다 다른 해석 존재 • 연구 현장에서는 인력·데이터 부족이 가장 큰 제약 • 소프트웨어 단가가 지나치게 낮게 책정돼 AI 투입도 어려움 • 현장 협업과 테스트베드, 현실적인 정책 지원 필요

구 분	주 요 내 용
	□ 규제 및 제도 관련 애로사항 o A • 규제 관련한 이야기는 학습용 데이터 비식별 처리 문제 같은 주제 같은데, 현장 이슈는 아님 • 다만 안전 규제가 모호함. 예를 들어 충돌 안전 규제(속도·충격량 기준)의 현장 적용 불합리 • 글로벌 수출 고려 시 국내보다 엄격한 기준을 지켜야 하므로 비용 부담 • 실효성 있는 테스트베드 부재 실제 운영 데이터를 볼 수 없는 상황 o B • 피지컬 AI는 태동되는 분야라 아직 규제보다는 문화, 사회적 합의가 문제 • 예를 들어 비식별 데이터 처리 문제, 현장 카메라 설치 시 노동자 반발이 큰 현실 • 안전 목적이라면 일정 수준 강제할 수 있는 법적 장치 필요 o C • 안전 규격(ISO 10218-1,2 등) 개정 예정 → AMR, 드론, 휴머노이드까지 확대 가능 • 개인정보 규제로 데이터 수집 어려움. 중국과 달리 국내는 공공 데이터 수집이 제한적 • 데이터는 기업이 보유하고 활용하지 못하는 경우 많아, 정부가 구매·공유하는 구조 필요 • o D • 휴머노이드 보급이 늘면 배터리·충돌 안전 규제가 핵심 쟁점이 될 것 • 중국산 QDD 모터가 글로벌 휴머노이드 확산을 견인 중, 충돌 제어 부재 문제 지적 • 새로운 안전 규제와 제도적 공백 발생 가능성이 높음 o E • 내재적 규제 문제(데이터, 인력, 전력 등) 해결이 우선 • 지식재산권 보호 미흡 → 핵심 기술이 해외로 유출되는 위험 우려 • 배터리·구동기 관련 규제가 산업 확산의 병목이 될 수 있음 □ 글로벌 경쟁 환경(미국·중국 대비)과 기회 o A • 중국산 청소·서비스 로봇에 시장 잠식, 성능도 우위 • 국내는 안전·보안을 차별점으로 삼을 필요가 있음 • 부품·데이터 활용에서 중국과의 협업 범위도 정책적으로 정해야 함 o B • 중국은 데이터 탑다운 정책으로 글로벌 벤치마크 모델을 배출 • 한국은 데이터 구축은 했지만 글로벌 표준 모델로 연결된 사례가 없음 • 독자 모델 개발도 중요하지만, 글로벌 확산 전략이 병행돼야 함 o C • 미국·중국을 그대로 추격하는 건 불가능 • ‘현실화된 피지컬 AI’에서 틈새 경쟁력(예: 자연어 기반 조작, 인간 친화적 관제) 확보 필요 • 대규모 경쟁보다 고급화된 틈새 공략이 잠재력이 있다고 생각 • 미국 인력 유입(예: H-1B 비자 문제 등)을 한국 창업 기회로 활용 가능

구 분	주 요 내 용
	o D • LLM·거대 모델 경쟁은 불리하나, 틈새형·특화형 모델 개발은 기회가 있음 • “바나나 꼭지” 사례처럼 현실 문제 해결형 모델을 지원해야 산업화 가능 • 대규모 투자 경쟁보다 소형·현장 맞춤형 모델 지원이 유효 o E • 인력·연구 환경 개선이 전제돼야 글로벌 경쟁에서 틈새 공략 가능 • 교육·연구 체질 개선 없이는 장기적 경쟁력 확보 어려움

주 의

1. 이 보고서는 소프트웨어정책연구소에서 수행한 연구보고서입니다.
2. 이 보고서의 내용을 발표할 때에는 반드시 소프트웨어정책연구소에서 수행한 연구결과임을 밝혀야 합니다.

비매품/무료
95060
9 791191 525748
ISBN 979-11-91525-74-8 (PDF)

[소프트웨어정책연구소]에 의해 작성된 **[SPRI 보고서]**는 공공저작물 자유이용허락 표시기준 제 4유형**(출처표시-상업적이용금지-변경금지)**에 따라 이용할 수 있습니다.
(출처를 밝히면 자유로운 이용이 가능하지만, 영리목적으로 이용할 수 없고, 변경 없이 그대로 이용해야 합니다.)