

AI Scientist 기술 개발 동향

Technology Development Trends in AI Scientist Systems

김영길 (Y.K. Kim, kimyk@etri.re.kr)
이용주 (Y.J. Lee, yongju@etri.re.kr)
임수중 (S.J. Im, isj@etri.re.kr)
송화진 (H.J. Song, songhj@etri.re.kr)
정문영 (M.Y. Chung, mchung@etri.re.kr)
권오욱 (O.W. Kwon, ohwoog@etri.re.kr)

언어지능연구실 책임연구원
시각지능연구실 책임연구원/실장
언어지능연구실 책임연구원/실장
복합지능연구실 책임연구원/실장
데이터지능연구실 책임연구원/실장
지능정보연구본부 책임연구원/본부장

ABSTRACT

Advances in foundation models, scientific agents, multimodal reasoning, automated experimentation, and laboratory robotics enable artificial intelligence (AI) systems to participate in the scientific discovery cycle. However, capabilities remain fragmented, and no common framework unifies reasoning, experimentation, validation, and collaboration. This article defines an AI Scientist as an agentic system that acquires knowledge, prioritizes hypotheses, conducts digital or physical experiments, analyzes results, and refines research plans. It proposes capability levels spanning AI Co-Scientists, autonomous AI Scientists, and AI research organizations; reviews international and Korean systems across scientific domains; and presents a human-AI collaborative reference architecture. The analysis shows strong performance in literature synthesis, hypothesis generation, code-based experimentation, and closed-loop laboratory automation, but limits in long-horizon coherence, reproducibility, and physical-world reliability. These findings support priorities in human-AI collaboration, evidence-centered evaluation, standardized tool interfaces, and staged domain validation.

KEYWORDS AI Co-Scientist, AI for Science, AI Scientist, Autonomous Scientific Discovery, Self-Driving Laboratory

* DOI: <https://doi.org/10.22648/ETRI.2026.J.410507>

* 이 논문은 한국전자통신연구원 연구운영비지원사업(기본사업) 「과학특화 멀티모달 파운데이션 모델 개발 및 적용」 연구개발과제 (26CB1200)의 결과물입니다.



본 저작물은 공공누리 제4유형
출처표시+상업적이용금지+변경금지 조건에 따라 이용할 수 있습니다.

© 2026 한국전자통신연구원

I. 서론

과학기술 분야에서도 AI를 접목한 R&D 패러다임의 대전환이 일어나고 있으며, AI for Science(AI4S) 시대가 도래하고 있다. 과거 ‘도구로서의 AI’를 넘어, 스스로 가설을 세우고 실험하는 ‘연구 주체로서의 AI’가 과학적 발견의 속도를 가속화하면서 연구 패러다임이 변하고 있다. AI 글로벌 경쟁력은 물론 과학혁신 가속화를 위해 미·중·한·EU를 중심으로 정부 주도로 AI 기반 과학기술 패러다임 대전환을 추진하고 있으며, '25년 11월 미 정부는 AI 기반 국가적 과학혁신 Genesis Mission을 공식 실행하는 행정명령에 서명하고 과학기술 패러다임 대전환을 선언한 바 있다.

최근 단백질 구조 예측, 신소재 생성, 물성 예측 등 다양한 과학 도메인에서 생성되는 데이터의 양과 복잡성은 인간 과학자의 인지적 한계와 전통적인 R&D 방법론의 임계점을 넘어서고 있다. 거대한 탐색 공간 속에서 유의미한 가설을 세우고 이를 실험을 통해 검증하는 일련의 과정은 막대한 시간과 비용을 요구하며, 이는 국가 R&D 생산성 정체의 주요 원인으로 지적된다.

이러한 한계를 극복하기 위해 인공지능을 과학 연구의 단순한 도구로 활용하는 단계를 넘어, 스스로 지식을 학습하고 가설을 생성·검증하는 ‘AI for Science’가 새로운 과학기술의 패러다임으로 부상하고 있다. 특히 기하급수적으로 증가하는 과학 논문과 실험 데이터 속에서 인간 연구자가 미처 발견하지 못한 숨겨진 상관관계를 찾아내고, 초고속으로 R&D 루프를 회전시킬 수 있는 자율형 연구 패러다임으로의 전환은 이제 선택이 아닌 국가 기술 주권 확보를 위한 필연적 과제이다.

본고에서는 AI 과학자 기술의 전 세계적인 개발 트렌드를 분석하고, AI 과학자 기술 개발의 방향성

을 제안하기 위해 다음과 같이 구성되었다. II 장에서는 AI 과학자 개념 및 레벨을 정의한다. III 장에서는 국내외 기술 개발 동향으로 글로벌 빅테크 및 선도 연구소들의 기술 개발 동향과 실제 도메인 적용 사례를 살펴본다. IV 장에서는 AI 연구동료 구성도 및 이를 뒷받침하는 핵심 기술을 살펴본다. V 장에서는 AI 과학자 기술 발전 방향 및 국내 연구 개발에 대한 시사점과 결론을 도출하고자 한다.

II. AI Scientist 개념 및 레벨 정의

1. AI Scientist의 정의 및 범위

본고에서 다루는 AI 과학자(AI Scientist)는 기존의 데이터 분석용 AI나 특정 실험 장비의 자동화 시스템과 명확히 구별된다. AI 과학자에 대한 정의가 다양하지만, 과학적 전문지식을 이해하는 과학특화 파운데이션 모델을 기반으로, 스스로 미해결 문제를 탐색하여 가설을 수립하고, 디지털 또는 물리적 실험을 설계·수행하며, 결과를 분석·검증하고, 연구 산출물을 작성한 후 그 결과로부터 피드백을 받아 가설을 수정·보완하는 자율적·반복적 연구 루프를 수행할 수 있는 AI 에이전틱 시스템을 의미한다고 할 수 있다.

이 정의에서 AI 과학자가 하나의 거대 모델을 의미하지 않으며, 실제 시스템은 범용 또는 과학특화 파운데이션 모델, 문헌 및 데이터 검색, 지식 그래프, 코드 실행기, 시뮬레이션, 과학 도구 소프트웨어, 실험장비, 장·단기 메모리, 평가 모델, 인간 협업 인터페이스 등이 결합된 복합 시스템으로 구현된다.

AI 과학자가 수행해야 하는 과학적 발견 과정은 ‘연구목표 설정 → 과학지식 탐색 → 연구문제 구체화 → 가설 생성 → 가설 평가·선정 → 실험 설계 → 실험·시뮬레이션 실행 → 결과 분석 → 가설 검

증 → 연구결과 작성·공유 → 후속 가설 생성의 순환 구조'로 표현할 수 있다. 따라서, AI 과학자는 적어도 다음과 같은 능력을 갖추어야 한다.

첫째, 연구목표를 해석하고 이를 검증 가능한 과학적 질문과 하위 과제로 분해해야 한다. 둘째, 선행 연구, 과학 지식 및 데이터를 근거로 복수의 가설을 생성하고 신규성, 타당성, 검증 가능성 및 기대 효과 등을 평가해야 한다. 셋째, 가설을 검증하기 위한 실험이나 시뮬레이션을 설계하고 필요한 도구를 선택·사용해야 한다. 디지털 환경에서의 시뮬레이션을 넘어 자율 실험실(Self-Driving Lab)과 연계되어 실제 물리적 세계의 실험까지 자율적으로 주도해야 한다. 넷째, 실험 결과의 통계적·과학적 유효성 등 결과를 분석하고 근거를 기반으로 가설의 지지 또는 기각 여부를 판단하고 검증해야 한다. 다섯째, 연구 과정과 결과를 추적 가능한 형태로 기록하고 다음 연구 주기를 계획할 수 있어야 한다.

2. AI 과학자의 레벨 정의

AI 과학자 기술은 문헌 검색과 가설 생성에서부

터 시뮬레이션 실행, 물리 실험, 자율 프로젝트 운영에 이르기까지 그 적용 범위와 자율성 수준이 매우 다양하다. 그러나, 현재는 단순한 연구지원 도구, AI 연구동료, 자율실험 시스템, 다중 에이전트 연구조직이 모두 'AI Scientist'라는 동일한 명칭으로 혼용되고 있어 기술 수준과 발전 방향을 객관적으로 비교하기 어렵다.

따라서 연구 과정에서 AI가 담당하는 범위와 인간의 개입 정도를 기준으로 체계적인 레벨을 정의할 필요가 있다. 표 1은 AI 연구동료를 C1~C3, AI 과학자를 S1~S2, AI 연구조직을 L1~L2로 구분하고, 가설 생성에서 디지털·물리 실험의 폐쇄루프, 자율 프로젝트 및 연구조직 운영으로 발전하는 구조를 제시한다.

현재의 주요 시스템을 본 분류에 적용하면 문헌 및 가설 중심 시스템은 C1, 데이터 및 코드 기반 전주기 시스템은 C2, 로봇 실험실은 잘 정의된 실험공간 안에서 S2 수준에 해당한다. L1~L2는 장기간 독립 연구 수행이 가능한 AI 과학자들의 집단지성을 이용할 수 있는 연구 실험실, 연구소 등 연구조직 단계로 볼 수 있다. Sakana AI의 The AI Scientist는 S1,

표 1 AI 연구 범위와 인간 역할을 고려한 AI 과학자 레벨 정의

구분	레벨	이름	정의	인간 역할
AI 연구동료 (Co-Scientist)	C1	가설 생성형	근거 기반의 검증 가능한 가설과 연구후보 제안	문제 설정, 후보 선택, 실험 수행
	C2	In-Silico 실행형	문헌 검색-가설생성-인실리코 검증	목표 설정, 실행 감독, 결과 판단
	C3	인간협력 실험검증형	wet-lab 포함 실험 계획 및 검증 제안-인간 실험 협업	실험 실행, 안전 관리, 결과 승인
AI 과학자 (Scientist)	S1	시뮬레이션 폐쇄루프	가설-실험설계-시뮬레이션 실험-검증 연결된 폐쇄루프 자율 수행	목표 설정, 최종 실험 및 검증
	S2	물리실험 폐쇄루프	자율실험실 연동한 실제 실험 및 검증 연결된 폐쇄루프 자율 수행	목표 설정, 안전·품질 관리, 장비 운영 보조, 예외 개입
AI 연구조직 (Laboratory)	L1	자율 프로젝트 수행	미션 프로젝트 설정하여, 여러 독립 연구 장기간 수행, 실패복구 및 재현성, 스케일업 등 관리	프로젝트 디렉터·책임자
	L2	범용 자율 연구조직	가치 및 예산을 반영한 연구주제 발굴부터 안전 관리까지 통합 자율 운영	연구소장 거버넌스

Google의 Co-Scientist는 C1 단계로 볼 수 있다[2-5].

국내에서는 K-문샷 프로젝트의 하나로 '35년 연구 소장급 AI 과학자 개발로 노벨상급 성과 창출을 추진 중이다. 국가전략 과학 난제 해결을 위해 연구 전 주기 자율 수행 AI 과학자 플랫폼을 확보하고, 도메인별 과학적 발견 검증 및 연구 생산성 혁신을 실증하는 것을 목표로 한다. 주요 마일스톤으로는 '28년 인간-AI 협업형 AI 연구동료 시스템(C3 레벨) 구현, '31년 인간 감독형 자율 과학 시스템(S2 레벨) 구현, '35년 전 주기 자율 연구 연구소장급 AI 과학자(L2 레벨) 구현을 추진한다.

III. 국내외 기술 개발 동향

‘AI for Science’는 초기에는 단백질 구조 예측, 신소재 탐색, 기후 예측 등 개별 과학 문제 해결을 위한 도메인 특화 AI 모델 개발을 중심으로 발전해 왔다. 최근에는 이러한 도메인 특화 모델과 다양한 과학 도구를 통합하여, 문헌 탐색, 가설 생성, 실험 설계·수행, 도구 활용 및 결과 검증까지 지원하는 AI Co-Scientist와 AI Scientist로 발전하고 있다.

이에 따라 AI for Science의 연구 범위는 도메인 특화 기반모델 개발을 넘어, 인간과 협업하는 AI 연구동료, 물리적 실험과 연계된 자율실험 시스템, AI 과학자의 성능·신뢰성·안전성을 검증하는 평가 체계로 확대되고 있다. 다음에서는 이러한 기술 발전 흐름을 바탕으로 국내외 주요 기술 동향을 살펴본다.

1. 국외 기술 동향

첫째, 신약 개발, 신소재 탐색, 화학 반응 시뮬레이션, 기후 예측 등 다양한 과학 분야에서 과학적 발견을 가속하기 위한 도구 개발이 활발히 추진되고

있다. 특히, 분야별 특성에 최적화된 도메인 특화 AI 과학 도구가 글로벌 빅테크 및 연구기관을 중심으로 상용 플랫폼으로 출시되고 있다.

Google DeepMind는 특정 과학 분야의 난제를 해결하기 위한 전문 AI 모델인 알파(Alpha) 시리즈를 통해 과학 플랫폼을 구축하고 있다. 알파폴드는 단백질 구조 예측 문제를 해결하여 구조 생물학의 지형을 바꾸었다. '25년 기준, 190개국 300만 명 이상의 연구자가 알파폴드를 사용하고 있으며, 알파폴드3는 단백질뿐만 아니라 RNA, DNA 등 다양한 생체 분자와 이들의 상호작용을 예측할 수 있도록 하여, 신약 개발 및 항생제 내성 연구를 획기적으로 가속화하고 있다[1]. 재료 탐색을 위한 그래프 네트워크 AI 모델인 GNoME는 220만 개 이상의 안정 후보 구조를 발견하였고[2], 전 세계 연구자들이 양자컴퓨터용 초전도체나 차세대 배터리 소재를 개발하는데 활용하고 있다. 기후 변화 대응과 기상 재해 예측을 위해 웨더넥스트와 알파어스를 개발하였다. 기존 모델보다 8배 빠른 속도로 온도, 기압, 풍속 등 기상 요소를 예측하고, 수학 분야에서는 알파프루프와 알파지오메트리가 국제 수학 올림피아드 수준의 문제를 해결하며 뛰어난 논리적 추론 능력을 입증하였다[3].

NVIDIA는 가상 시뮬레이션 환경 기반의 재료 및 화학 시뮬레이션 도구로 ALCHEMI와 BioNeMo를 개발하여 신약 개발 및 재료 탐색을 위한 시뮬레이션 환경과 생성형 AI 프레임워크를 통해 과학적 연구 가속화를 위한 인프라를 제공하고 있다. 재료 및 화학 시뮬레이션을 위한 AI 추론을 최적화한 ALCHEMI NIM 마이크로 서비스를 제공하여 연구자들이 가상 AI 실험실에서 화합물의 특성과 재료의 안정성을 시뮬레이션 기반으로 평가할 수 있도록 지원한다[4]. 생명과학과 신약 개발에 특화된 생성형 AI 모델의 학습, 미세조정, 추론 및 배포를 지

원하는 프레임워크로 BioNeMo NIM 마이크로서비스도 제공하고 있다[5].

둘째, 단순 과학 데이터 분석 도구에서 인간과 협업하는 AI 연구동료 및 가설 생성과 실험 자동화를 수행하는 자율적 연구 에이전트로 진화 중이며, 과학 분야의 검증 가능한 가설을 생성하고 실험할 후보를 좁혀 주는 단계에 도달하고 있다.

Sakana AI는 인간의 개입 없이 자율적으로 과학적 발견을 수행하는 자율 AI 과학자 프레임워크를 최초로 제안하였다[6,7]. '24년 8월 LLM을 활용하여 과학적 발견의 전 과정(아이디어 생성, 코드 작성, 실험 수행, 결과 시각화, 논문 작성, 피어 리뷰)을 자동화하는 프레임워크를 제안하였고, 시스템이 생성한 논문 가운데 한 편이 ICLR 워크숍 심사를 통과한 사례를 보고했다[6]. 초기 모델(v1)은 사람이 작성한 코드 템플릿에 크게 의존했던 반면, '25년 4월 공개한 v2에서는 코드 템플릿에 대한 의존성을 완전히 제거하였다[7].

Google Research는 과학 문헌을 읽고, 데이터를 추론하며, 새로운 실험을 제안하고 인간과 협업하는 AI 연구동료(AI Co-Scientist) 플랫폼을 제안하였다. Gemini 기반의 멀티 에이전트 시스템으로 인간의 과학적 탐구 과정을 모방하여 새로운 가설을 생성하고 타당성을 검증하여 연구 계획을 제시하고 인간과 협업하도록 설계되었다[8].

OpenAI의 GPT-Rosalind와 Anthropic의 Claude Science 등 최신 AI 모델들은 신약 개발 분야에서 혁신적인 가능성을 입증하고 있다. 이들 모델은 가설 생성 능력의 향상에 그치지 않고, 문헌 조사 및 연구 데이터 분석, 논문 작성, 신약 후보 물질 선정 등 연구의 전 과정을 아우르는 종단형(End-to-End) 연구 보조 기능을 수행하고 있다[9,10]. 이는 단일 업무 지원을 넘어 연구 전반의 효율성을 도모하는 연구 동료로서 AI의 확장된 역할을 보여준다.

Shanghai AI Lab은 가설 수립부터 실험적 검증에 이르는 전 과정에 걸쳐 연구자를 지원하기 위해 설계된 과학적 발견 플랫폼인 Intern-Discovery를 공개했다. 특히 InternAgent-1.5는 장기 과학연구 과제를 수행하기 위한 통합 에이전트 프레임워크로 연구계획 생성, 검증, 개선 및 장기기억 기능을 포함한다. 문헌조사 중심의 답리서치뿐만 아니라 계산과학, 솔루션 최적화, 실험실 연구를 연결하며, 에이전트가 실패 경험과 중간 결과를 장기적으로 축적해 연구전략을 수정하도록 설계되었다[11].

셋째, 신소재 · 화학 · 촉매 등 일부 과학 분야에서 로봇 기반 자율실험실(Self-Driving Laboratory)과 결합하여, 실험 계획-물리적 실행-자동 분석-후속 실험 설계를 반복하는 폐쇄루프(S2) 레벨로 발전하고 있다. 이 과정에서 오케스트레이터는 고수준의 가설로부터 AI 또는 최적화 알고리즘이 제안한 실험 조건을 장비가 실행할 수 있는 제어 명령과 워크플로로 변환하고, 로봇 암, 고체 · 액체 분주기 등의 동작을 조정하며, 측정 결과를 다음 의사결정 단계로 전달하는 역할을 수행한다. 완전 자율형 물리실험 폐쇄루프는 아직 모든 과학 분야에 보편화된 단계는 아니며, 장비 간 상호운용성, 데이터 품질, 예외 상황 복구, 실험 안전성 및 인간 연구자의 검증과 감독이 여전히 중요한 기술적 과제로 남아 있다.

'26년 2월 OpenAI는 GPT-5와 Ginkgo Bioworks의 자동화 클라우드 실험실을 연결하여 무세포 단백질 합성 조건을 폐쇄루프 방식으로 최적화하여 단백질 생산 비용을 40% 절감했고[12], 중국과학원(CAS)은 방대한 과학 논문 데이터를 분석하고 실험 도구를 오케스트레이션하는 자율형 AI 모델 Science-One을 기반으로, 문헌 마이닝부터 실험 합성까지 전 과정을 자동화하는 폐쇄루프형 신재료 발견 연구 시스템을 발표하였다[13].

미국 로렌스 버클리 국립연구소(LBNL)의 A-Lab

은 계산 재료 데이터베이스, 문헌에서 추출한 합성 지식, 머신러닝, 능동학습 및 로봇 실험장비를 통합한 무기재료 자율실험실이다. 17일 동안 연속 운영하여 57개 목표 물질 중 36개를 합성했으며, 실패 데이터를 분석해 후속 합성전략을 개선했다. AI 재료 예측과 실제 물리 실험을 폐쇄루프로 연결한 대표 사례이다[14].

넷째, AI 연구동료 에이전트의 과학 연구 전 주기 벤치마크가 구축되고 공개되고 있다. AI 과학자 시스템을 객관적이고 체계적으로 평가하는 벤치마크(Benchmark) 프레임워크는 최근 기술 동향 분석에서

가장 중요하게 다뤄지는 연구 주제 중 하나이다. 기존 모델 평가 방식(예: MMLU, GSM8K 등 단답형/객관식 문제 풀이)으로는 자율적으로 긴 시간 동안 여러 과학 도구를 쓰며 연구 흐름을 완결해야 하는 에이전트로서의 능력을 측정할 수 없기 때문이다.

AstaBench는 문헌 탐색, 연구질문 형성, 데이터 분석, 실험계획 및 결과 해석 등 과학연구 전 주기를 평가하기 위한 2,400개 이상의 문제로 구성된 벤치마크이다. 검색 환경과 평가조건을 통제된 상태에서 다양한 AI 과학자를 비교하며, 현재 AI가 개별 태스크에서는 성과를 내지만 전 주기 과학 연구를 안

표 2 주요 AI 과학자의 적용 분야, 레벨 분류 및 핵심 기술

주요 AI Co-Scientist/Scientist (개발기관)	주요 적용 분야	레벨 분류	핵심 기술
AI Co-Scientist (Google Research)	신약 재창출, 항생제 내성·유전자 전달 기전	C1	Gemini 기반 멀티 에이전트, 생성-토론-진화, 토너먼트식 가설 순위화, 추론 시점 연산 확장 기술
Kosmos (Edison Scientific)	대사체학, 재료과학, 유전체학	S1	병렬 에이전트 실행, 구조화된 월드모델, 장기 문맥 관리, 근거·코드 추적 기술
Robin (FutureHouse)	치료표적·후보 발굴, 세포실험 설계	C3	전문 에이전트 오케스트레이션, LLM 심사위원 기반 후보 토너먼트, Lab-in-the-Loop 폐쇄루프
LabOS (Stanford·Princeton)	암 면역치료 표적, 줄기세포 공학	C3	자기진화형 멀티에이전트, 동적 도구 생성, 3계층 연구 메모리, XR 기반 실시간 인간-AI 협업, Dry Lab-Wet Lab 폐쇄루프
The AI Scientist v2 (Sakana AI)	딥러닝, 모델 일반화, 정규화, 컴퓨터 비전	S1	에이전틱 트리 탐색, 실험관리 에이전트, 코드 작성·실행, 데이터 분석·시각화, 논문 작성·AI 리뷰 기술
Intern-Discovery (Shanghai AI Lab)	단백질 공학, 신약·분자설계, 기후·태풍, 물질과학	S2	동적 지식 흐름, 그래프 강화 몬테카를로 탐색, 구조화된 장기 인지 메모리와 경험 진화, 실험장비의 폐쇄루프 연결
AlphaEvolve (Google DeepMind)	수학, 데이터센터·칩·AI 학습 최적화	S1	Gemini 코딩 에이전트, 진화적 탐색, 프로그램 변이·선택, 다중 자동 평가 기술
Claw AI Lab (NTU 등)	AI·컴퓨터과학 연구 자동화	S1	5계층 연구 아키텍처, 역할별 다중 에이전트 연구팀, 실행 하네스, 롤백·재개 기술
Coscientist (CMU)	유기합성, 촉매 반응, 반응조건 최적화, 로봇 화학실험	S2	다중 LLM 기반 중앙 계획 에이전트, 코드 생성·실행 및 오류 수정, 로봇 실험장비 API 연계, 폐쇄루프 실험 및 데이터 기반 최적화
A-Lab (LBNL·UC Berkeley)	배터리, 태양전지, 기능성 무기재료	S2	로봇 고체합성·특성분석, 능동학습, 열역학 모델, 워크플로 오케스트레이션 기술
AILA (IIT Delhi 등)	재료 표면 분석, 나노기계 특성 평가	S2	API 장비 제어, 코드 실행·디버깅, 이미지·측정값 분석, 실시간 파라미터 최적화

출처 참고문헌 [7,8,11,14,17-22]을 바탕으로 재구성

정적으로 수행하기에는 여전히 큰 격차가 있음을 보여준다[15].

HeurekaBench는 AI 연구동료를 어떻게 제대로 평가할 것인가, 그 기준과 프레임워크를 제시하고 있다. 실제 실험 데이터가 주어지고 이 데이터로 논문 한 섹션을 작성하게 한 후, 주어진 데이터로 적절한 분석을 했는지, 결과가 정합적인지, 결과 해석이 데이터에 근거했는지 등을 종합 평가한다[16].

표 2는 최근의 주요 AI 과학자들의 레벨 분류, 적용 분야, 핵심 기술을 나타내며, 레벨 분류는 '26년 7월까지 공개된 논문과 기술 자료에서 입증된 기능을 기준으로 하였다.

2. 국내 기술 동향

국내에서는 출연연, 대학 및 기업을 중심으로 공공 연구정보 인프라, 과학기술 특화 파운데이션 모델, 도메인 특화 과학적 발견 플랫폼 등이 연구 개발되고 있지만, 연구 전 주기를 통합하는 AI 기반 연구 플랫폼 구축은 시작 단계에 있다.

ETRI는 전략연구사업으로 '26년부터 과학기술용 특화 멀티모달 파운데이션 모델 개발에 착수하였으며, '27년 KISTI와 함께 과학기술 다학제 AI 연구동료 플랫폼 구축 사업을 수행할 예정이다. KISTI는 약 1억 8천만 건에 달하는 국가 과학기술 데이터를 통합 제공하는 ScienceON 플랫폼을 운영하고 있으며, 이를 기반으로 구축된 ScienceON AI를 통해 논문 요약, 번역, 용어 해설부터 다중 논문 비교 분석 및 AI 질의응답 서비스 등 연구자의 선행연구 조사를 체계적으로 지원하는 지능형 연구 인프라를 제공한다[23]. KAERI는 '24년 원자력 특화 거대언어모델 AtomicGPT를 공개하고[24], 원자력 논문, 전문용어 사전, 규제문서 및 연구보고서 등을 학습시켜 원자력 문서 분석, 연구 정보 검색, 요약·생성

기능에 적용하였다.

대학에서는 'AI for Science' 기반의 가설 자동 생성, 신소재 설계, 질병 예측 등 기초·응용과학 전반에 걸친 기초 연구를 수행하고 있다. 특히, KAIST, UNIST, 서울대, POSTECH 등에서는 기존의 단순 AI 과학 데이터 분석 알고리즘 개발을 넘어, AI 에이전트의 가설 수립, 실험 계획, 물리적 로봇 실험 장비를 연동하는 폐쇄루프 자율실험실 구축 및 실증이 진행되고 있다.

기업에서는 신약 및 신소재 분야에서 후보 물질의 탐색·설계를 지원하는 AI 역량이 고도화되고 있으나, 실험 설계-수행-검증-재학습으로 연결되는 완전한 폐쇄루프 자율연구 구현은 초기 단계이다. LG AI 연구원은 재료개발 플랫폼 EXAONE Discovery를 활용해 하루 만에 42만 개 이상의 화합물을 검토하고 데이터센터용 액침냉각 유체 후보를 개발한 사례를 제시했다[25]. 스탠다임은 타깃 발굴, 화합물 디자인, 약물성 및 특허 가능성 예측 모듈을 유기적으로 연계해 신약 개발 초기 단계의 연구 의사결정을 다각도로 지원하는 모듈형 에이전트 구조를 확보하였다. 아스테로모프는 인간의 개입 없이 스스로 가설을 세우고, 실험을 설계 및 분석하여 새로운 과학적 발견을 도출해 내는 AI 과학자 모델을 개발 중이다. 라이너는 수집·분석된 학술 정보를 바탕으로 구조화된 보고서 작성을 지원하는 에이전틱 AI 지원 범위를 확대하고 있으며, 연구 기획부터 논문 작성·검토에 이르는 연구 전 주기 지원 체계를 구축 중이다.

IV. AI 연구동료 구성도 및 핵심 기술

현재의 AI는 방대한 문헌·데이터·가설 후보를 인간보다 빠르게 탐색하고 시뮬레이션과 분석을 대규모로 수행할 수 있지만, 연구목표의 가치 판단, 과

학적 타당성 검토, 예외 상황 대응, 실험 안전 및 결과에 대한 책임에서는 여전히 인간 연구자의 전문성과 통찰이 필수적이다. 따라서 단기적으로는 인간이 연구목표 설정과 핵심 의사결정을 담당하고 AI가 문헌 탐색, 가설 도출, 실험 설계, 데이터 분석 및 대안 탐색을 지원하는 인간-AI 협업형 AI 연구동료를 우선 개발하여 연구 생산성을 높이고, 실제 연구 현장에서 신뢰성·재현성·안전성을 검증할 필요가 있다.

이러한 협업 과정에서 축적되는 고품질 연구 데이터, 인간의 판단·수정 이력, 실험 프로토콜, 성공·실패 사례 및 연구 전 주기 운영 경험은 향후 AI가 가설 설정부터 실험 실행, 결과 분석, 재계획을 자율적으로 반복하는 완전 자율형 AI 과학자로 진화하는 데 필요한 기술적 토대가 될 수 있다. 즉, 인간-AI 협업형 AI 연구동료는 현실적인 단기 연구성과를 창출하는 동시에 완전 자율화에 따르는 오류·안전·책임 위험을 단계적으로 낮추는 필수적

인 중간 단계이며, 이를 기반으로 자율성의 범위와 인간의 감독 수준을 점진적으로 확대하는 것이 기술적 타당성, 현장 수용성 및 연구윤리를 함께 확보할 수 있다.

AI 연구동료는 개별 기술의 단순 조합이나 언어 모델을 중심으로 여러 기능을 단순 연결한 시스템이 아니라, 과학 데이터 및 과학 도구의 공유와 활용이 가능하고 연구 상태와 관련 증거를 공유하면서 계획-실행-검증을 자율적으로 반복하는 플랫폼 구조로 설계되어야 한다. 그림 1은 ETRI와 KISTI가 공동으로 설계하고 있는 인간-AI 협업형 전 주기 AI 연구동료 플랫폼의 구성도를 나타낸다. 본 플랫폼은 개별 학문의 경계를 넘어 방대한 글로벌 과학기술 데이터를 통합 분석하고, 전문 연구원 수준의 창의적 추론과 자율 검증 루프를 통해 과학적 발견을 가속화하는 AI 연구동료 구축을 목표로 한다.

본 플랫폼의 핵심 기술로는 과학특화 멀티모달 파운데이션 모델, 다학제 과학적 발견을 위한 가설

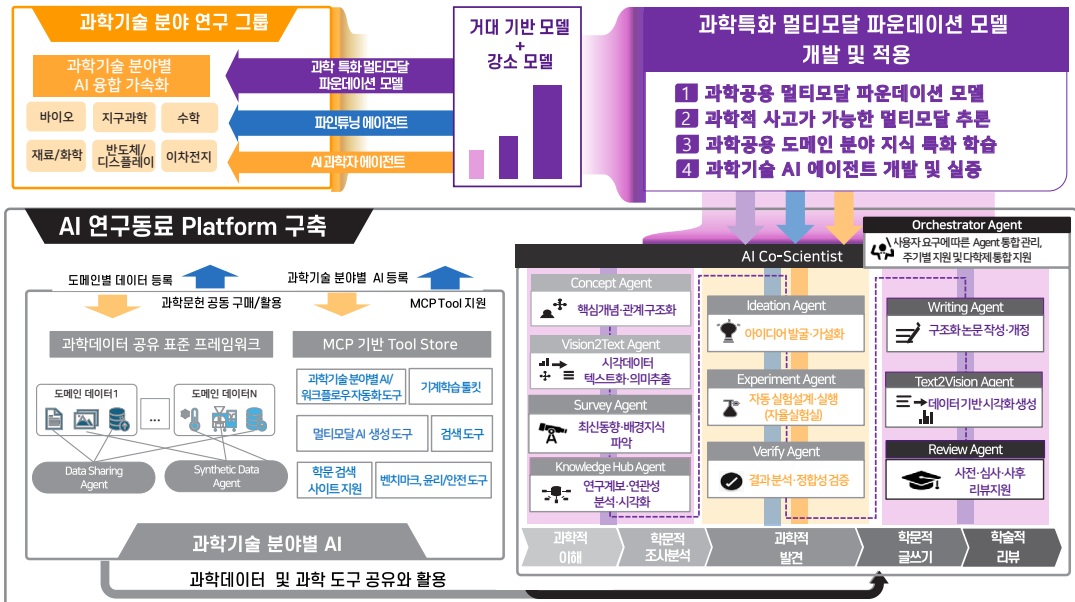


그림 1 전 주기 AI 연구동료(Co-Scientist) 플랫폼 구성도

생성 및 리뷰 기술, 전문지식 기반 과학적 발견(가설 선정-실험-검증) 기술, 과학 데이터 및 도구 활용 플랫폼 기술, 다학제 에이전트 오케스트레이션 기술을 포함한다.

표 3은 AI 연구동료 핵심 기술 및 주요 연구 내용을 나타낸다. 이 중에서 다학제 과학적 발견을 위한 가설 생성 및 리뷰 기술은 이질적인 과학 도메인(예: 물리학-생물학, 화학-경제학) 간의 다학제 지식 그래프를 구축하고, 인간이 발견하기 어려운 데이터 간의 빈틈을 찾아 창의적 가설을 생성하고 리뷰한다. 전문지식 기반 과학적 발견(가설 선정-실험-검증) 기술은 생성된 가설을 검증하기 위해 독립적 연구 방법론을 수행하는 AI로 가설 선정, 실험 설계, 자율 실험(Dry/Wet) 수행, 결과 분석, 실패 시 가설 수정 루프를 수행한다. 다학제 에이전트 오케스트레이션

기술은 인간 연구자가 연구의 방향을 조정하거나 의사결정에 참여할 수 있는 협업 프레임워크와 다중 AI 에이전트 간의 협업을 조율한다.

그리고, ‘AI 과학자의 자율 발견 · 검증 핵심 기술은 다학제 과학지식 허브 → 연구 공백 분석 → 가설 생성 → 가설 랭킹 · 진화 → 실험설계 · 도구 호출 → 프로비넌스 기반 검증 → 메타리뷰 · 논문화 → 연구자 피드백 기반 고도화의 구조’로 구현될 수 있다.

V. 발전 전망 및 결론

1. AI 과학자의 기술 발전 방향

AI 과학자 기술은 단일 거대언어모델 중심의 연구보조 도구에서 벗어나, 과학 지식 · 데이터 · 추론 · 시뮬레이션 · 실험장비 · 다중 에이전트를 통합

표 3 AI Co-Scientist 핵심 기술 및 주요 연구 내용

핵심 기술	주요 연구 내용	설명
1. 다학제 과학적 발견을 위한 가설 생성 및 리뷰 기술	다학제 과학지식 허브 구축	과학 문헌 멀티모달 심층 이해에 기반한 다학제 연구 지식 생성 및 맞춤형 검색이 가능한 지식 허브 개발
	논문 작성	다학제 연구결과 해석 기반 자동 논문 생성 및 근거 제시 기술 개발
	메타 리뷰	과학적 발견에 대한 인과관계, 설명/해석 방법론 연구, 학술적 리뷰 및 논문 개선 기술 개발
	다학제 과학적 가설 생성	딥러시저형 멀티모달 데이터 기반 융합적 과학 가설 디스커버리 및 생성 기술 개발
2. 전문지식 기반 과학적 발견 (가설 선정-실험-검증) 기술	가설 선정	과학적 성과 창출과 실험 검증 가능성 기반 유효 가설 선정 기술 개발
	가설 실험	가설 타당성 검증과 실험 변수/조건 최적화를 위한 인간-AI 협업 가상 실험 환경 구축, 실험 설계 및 수행 기술 개발
	가설 검증	실험 결과 분석, 프로비넌스 기반 가설 유효성, 신뢰성 제공 검증 추론 기술 개발
	전문지식 기반 과학적 발견	가설 생성-실험-검증 단계의 자율적 반복을 통한 전문지식 기반 과학적 발견 절차 기술 개발
3. 과학 데이터 및 도구 활용 플랫폼 기술	MCP Tool Store 구축	과학기술 연구 전 주기 활용에 필요한 논문 검색, 과학전문 AI 모델, 시뮬레이션 실험-도구 등 외부도구 호출 및 활용 기술 개발
	과학 데이터 공유	과학 도메인별 데이터 사일로 문제 해결을 위한 과학데이터 공유 표준 및 과학 데이터 안전 공유 프레임워크 개발
	과학 데이터 합성	실험환경 특성을 모사한 고충실도 과학데이터 합성/증강 기술 개발
4. 다학제 에이전트 오케스트레이션 기술	오케스트레이션	과학 연구 전 주기 관리를 위한 AI 연구동료 협업 프로세스 생성·수정이 가능한 오케스트레이션 기술 개발
	과학 AI 연구동료 실증	인간-AI 협업 물리적 연구 실험실 실증: 과학 분야 적용(신소재 및 바이오 우선)

하는 전 주기 자율 연구 시스템으로 발전하고 있다. 단기적으로는 인간 연구자가 목표와 최종 판단을 담당하고 AI가 대규모 문헌 탐색과 가설 · 실험 후보 생성을 지원하는 인간-AI 협업형 AI 연구동료가 확산될 것으로 예상되며, 중장기적으로는 평가 기준이 명확하고 실험 자동화 수준이 높은 제한된 과학 분야부터 반자율 · 완전자율형 AI 과학자로 발전할 전망이다[6-8,14,17].

1.1 단일 모델에서 복합 과학 AI 시스템으로 전환

향후 AI 과학자의 경쟁력은 기반 파운데이션 모델의 크기만으로 결정되지 않는다. 과학특화 모델, 지식베이스, 인과추론, 형식검증, 시뮬레이터, 최적화 알고리즘 및 로봇을 적절히 조합하는 시스템 설계 능력이 중요하다. 언어모델은 문제 해석과 개방형 아이디어 생성에 강점이 있지만 수치 계산, 화학 구조 검증, 수식 증명 및 통계 분석에서는 전문 도구가 더 신뢰할 수 있다. 따라서, LLM이 모든 답을 직접 생성하기보다 적절한 전문 모델과 도구를 선택하고 결과를 통합하는 구조가 바람직하다.

1.2 정적 RAG에서 동적 과학 지식-세계모델로 발전

현재의 RAG는 질문 시점에 관련 문서를 검색하지만, 장시간 연구에서는 어떤 가설을 검토했고 어떤 실험이 실패했으며 무엇이 미해결 상태인지 지속적으로 관리해야 한다. 향후 시스템은 논문, 데이터, 가설, 실험 및 결과를 동적으로 연결한 과학 세계모델을 구축할 것이다. 세계모델은 사실의 집합뿐만 아니라 불확실성, 경쟁 가설, 인과관계, 시간 및 출처를 포함해야 한다. 실험 결과가 추가될 때마다 관련 가설의 신뢰도와 후속 실험 우선순위를 갱신해야 한다.

1.3 생성 중심에서 증거·검증 중심으로 전환

초기 AI 과학자 연구는 얼마나 많은 아이디어를 생성하는가에 초점을 두었으나, 잘못된 아이디어를 얼마나 효과적으로 제거하고 올바른 결론을 재현하는가가 더 중요하다. 향후에는 ‘먼저 생성하고 나중에 설명하는’ 방식에서 ‘주장마다 검증 가능한 근거를 확보한 후 결론을 도출하는’ 방식으로 전환해야 한다. 자동 논문 작성보다 주장-근거 추적, 재현 패키지 생성, 독립 검증 및 불확실성 표현이 핵심 성능이 될 것이다.

1.4 단발성 태스크에서 장시간 임무 수행 자율 연구로 확장

AI 과학자가 실제 연구 생산성을 높이려면 수분 단위의 질의응답을 넘어 수일 또는 수개월에 해당하는 연구 과정을 관리해야 한다. 이를 위해 계층적 계획, 연구 체크포인트, 상태 복구, 비용 · 시간 관리, 실패 학습 및 지속적 메모리가 필요하다. 장시간 실행은 단순히 더 많은 모델 호출을 의미하지 않는다. 탐색 가치가 낮은 가설을 중단하고, 유망한 경로에 자원을 집중하며, 성과가 없는 반복을 감지하는 연구 포트폴리오 관리 능력이 요구된다.

1.5 디지털 연구에서 물리 실험으로 확장

높은 자율성은 먼저 코드, 수학, 시뮬레이션과 같이 결과가 자동 검증되는 영역에서 발전할 가능성이 크다. 이후 화학 · 소재 · 바이오 분야의 표준화된 실험 프로토콜과 자동화 장비로 확대될 것이다. 물리 실험으로의 확장에는 장비 표준화, 디지털 트윈, 로봇 조작, 센서 융합, 실시간 오류 감지 및 안전 정책이 필수적이다. 모든 실험을 자동화하기보다 반복적 · 고처리량 실험부터 AI가 담당하고, 복잡한 예외와 가치판단은 인간이 담당하는 혼합 구조가 현실적이다.

1.6 개별 AI 과학자에서 연구 조직 AI로 확장

단일 AI 모델이 개별 논문 작성이나 단순 실험 설계를 담당하던 단계에서 벗어나, 다양한 역할을 맡은 AI들이 협업하는 집단지성형 연구 생태계로 진화할 것으로 전망된다.

개별 AI 과학자에서 연구조직 AI로의 확장은 단순한 성능 향상을 넘어 연구 패러다임의 근본적 전환을 의미한다. 연구조직 AI는 실제 인간 연구소의 조직 구조(연구소장, 이론 연구원, 데이터 엔지니어, 개발 연구원, 연구 보조원 등)를 멀티 에이전트 기반 아키텍처로 모방한다. 역할의 분업화는 물론 서로 다른 가설을 동시에 검증하여 수개월이 걸릴 탐색 과정을 수 시간 단위로 단축시킬 수 있고, 복수 프로젝트를 병렬 수행하는 압도적인 연구 생산성을 확보하게 된다.

1.7 인간-AI 협업형 연구 체계의 확산

단기적으로 가장 높은 실용성을 가질 형태는 인간을 배제한 완전자율 시스템보다 인간-AI 협업형 AI 연구동료이다. AI는 인간이 처리하기 어려운 규모의 문헌과 후보 공간을 탐색하고, 인간은 과학적 중요성, 연구 방향, 암묵지 및 윤리적 책임을 제공한다. 향후 한 명의 연구자가 문헌조사, 이론, 실험설계, 데이터 분석, 통계 및 리뷰를 담당하는 복수의 AI 전문 에이전트를 관리하는 형태가 확산될 수 있다. 이때 연구자의 역할은 세부 작업 수행자에서 연구목표 설정자, AI 연구팀 관리자 및 최종 과학적 판단자로 이동할 것이다.

그러나 인간이 단순히 AI 결과를 형식적으로 승인하는 구조가 되어서는 안 된다. 자동화 편향을 줄이기 위해 AI가 결론뿐만 아니라 대안, 불확실성, 실패 가능성 및 반대 증거를 함께 제시해야 한다. 인간이 주요 가정과 증거를 검토할 수 있는 설명 인터페이스가 필요하다.

2. 국내 연구 개발에 대한 시사점

국내에서 범용 기반 모델만으로 글로벌 선도 AI 과학자와 경쟁하기는 어렵다. 국내 연구기관이 보유한 과학기술 문헌, 연구 데이터, 전문모델, 컴퓨팅 및 실험 인프라를 통합하고, 이를 신뢰 가능한 연구 페쇄루프로 구현하는 차별화 전략이 필요하다.

첫째, 과학 지식 허브를 구축해야 한다. 국내외 논문, 연구보고서, 특허, 데이터, 코드, 연구시설, 온톨로지 및 실패 실험을 연결하고, 주장과 근거의 출처를 추적할 수 있어야 한다. 특히 국내 공공 연구성과와 한국어 과학기술 문헌은 글로벌 범용모델이 충분히 학습하지 못한 차별적 자산이다.

둘째, 모델 독립적인 과학 에이전트 플랫폼을 개발해야 한다. 특정 상용 LLM에 종속되지 않고 과제와 보안 수준에 따라 국산·공개·상용 모델을 선택할 수 있어야 한다. 문헌 검색, 데이터 분석, 시뮬레이션, 전문 AI 모델 및 실험 장비를 표준 인터페이스로 등록·조합하는 과학 도구 저장소가 필요하다.

셋째, 과학 임무 오케스트레이션 기술을 확보해야 한다. 연구 전 주기의 작업 분해, 전문 에이전트 구성, 장시간 상태관리, 자원 배분, 오류 복구 및 인간 승인 절차를 지원해야 한다. 분야별 연구자가 코드를 직접 작성하지 않고도 연구 프로세스를 구성·수정할 수 있는 워크플로 설계환경이 요구된다.

넷째, 검증과 재현성을 우선하는 플랫폼을 구축해야 한다. 모든 결과에 데이터, 코드, 모델, 문헌 및 실행로그가 연결되어야 하며, 별도의 검증 에이전트가 결과를 독립적으로 재실행해야 한다. 생성된 논문의 문장 품질보다 재현 패키지의 완전성을 핵심 성과지표로 설정할 필요가 있다.

다섯째, 국가 AI 과학자 벤치마크를 구축해야 한다. 단순 과학 질의응답이 아니라 국내 연구자가 실제 수행하는 문헌조사, 가설 생성, 데이터 분석, 시

물레이션, 실험설계 및 재현 과제를 포함해야 한다. 시간분리형 발견 평가, 전문가 블라인드 평가 및 실제 실험실 검증을 결합하는 것이 바람직하다.

여섯째, 신소재·바이오 등 특화된 전략 분야를 중심으로 과학적 발견 전 주기 실증을 추진해야 한다. 초기에는 디지털 시물레이션과 자동 평가가 가능한 문제를 우선 선정하고, 이후 인간 실험실 연계형 Lab-in-the-Loop와 자율실험실로 확장하는 단계적으로 발전시키는 접근이 필요하다.

일곱째, 과학 연구행위의 위험 수준에 따라 AI의 자율성과 권한을 차등 부여하는 위험도 기반 자율성 관리체계를 마련해야 한다. 문헌 검색, 데이터 분석, 샌드박스 환경에서의 코드 실행 등 저위험 업무에는 비교적 높은 수준의 자율성을 허용할 수 있으나, 외부 시스템이나 데이터의 변경, 대규모·고비용 계산자원의 사용, 위험물질을 취급하는 실험, 생물학적 실험과 같이 안전·윤리적 영향이 큰 행위에는 인간의 사전 승인과 별도의 접근권한을 요구하도록 설계해야 한다.

3. 결론

AI 과학자는 특정 과학 문제의 예측을 수행하는 개별 모델을 넘어, 과학 문헌과 데이터를 이해하고 가설을 생성하며 실험을 설계·수행하고, 그 결과를 분석·검증하는 등 과학적 발견 전 주기를 자율적으로 수행하는 기술로 발전하고 있다.

그러나 현재까지 보고된 성과는 대부분 제한된 연구 분야와 통제된 실험 환경에서 검증된 결과라는 한계가 있다. 문헌 기반으로 생성된 가설의 실질적 신규성과 과학적 타당성, 장시간 연구 과정에서 발생하는 오류의 누적과 전파, 생성 코드와 통계 분석의 신뢰성, 물리적 실험에서의 예외상황 대응, 제3자에 의한 독립적 재현 가능성, 연구윤리 및 안전

성 확보 등은 여전히 해결해야 할 핵심 과제이다. 따라서 높은 수준의 언어 생성 품질이나 자동화된 리뷰 점수만으로 AI의 과학적 발견 역량을 평가해서는 안 되며, 명확한 근거 제시와 실제 실험, 독립적 재현을 통한 외부 검증을 중심으로 성능을 평가할 필요가 있다.

단기적으로는 인간 연구자의 전문적 판단과 AI의 대규모 탐색·분석 능력을 결합한 AI 연구동료가 연구 생산성을 크게 높일 것으로 전망된다. 특히 반자율형 AI 과학자는 코드, 수학, 시물레이션처럼 결과를 객관적으로 평가할 수 있는 분야와 실험 자동화 수준이 높은 연구환경을 중심으로 우선 확산될 가능성이 크다. 궁극적으로 AI 과학자의 경쟁력은 LLM 자체의 규모나 언어 생성 능력보다 고품질 과학 지식의 확보, 과학 도구 및 실험 인프라와의 연계, 장시간 연구 과정에서의 상태와 논리적 일관성 유지, 독립적 검증과 재현성 확보, 연구 프로비넌스 관리, 안전·윤리 거버넌스를 하나의 폐쇄루프 체계로 통합하는 시스템 공학 역량에 의해 좌우될 것이다. 신뢰할 수 있는 AI 과학자의 발전은 인간 연구자를 대체하기보다 연구자가 더욱 광범위한 가설 공간을 탐색하고, 더 많은 실험을 정밀하게 수행하며, 복잡한 과학적 난제를 더욱 신속하게 해결할 수 있도록 지원하는 새로운 연구 패러다임을 형성할 것으로 기대된다.

용어해설

AI 과학자(AI Scientist) 과학특화 파운데이션 모델을 기반으로, 스스로 미해결 문제를 탐색하여 가설을 수립하고, 디지털 또는 물리적 실험을 설계·수행하며, 결과를 분석·검증하고, 연구 산출물을 작성한 후 그 결과로부터 피드백을 받아 가설을 수정·보완하는 자율적·반복적 연구 루프를 수행할 수 있는 AI 에이전트 시스템

AI 연구동료(AI Co-Scientist) 인간 연구자와 협력하여 가설 수립, 실험 설계·수행, 결과 분석·검증, 연구 산출물 작성과 피드백 기반 가설 수정·보완 등 과학 연구의 전 과정을 지원하는 인공지능 시스템

자율실험실(Self-Driving Lab) AI가 실험 목표에 따라 실험 조건을 설계하고, 로봇·자동화 장비를 이용해 실험을 수행하며, 측정 결과를 분석한 뒤 다음 실험을 스스로 결정하는 폐쇄루프형 연구 시스템

참고문헌

- [1] J. Abramson et al., "Accurate Structure Prediction of Biomolecular Interactions with AlphaFold 3," *Nature*, vol. 630, 2024, pp. 493-500.
- [2] A. Merchant et al., "Scaling Deep Learning for Materials Discovery," *Nature*, vol. 624, 2023, pp. 80-85.
- [3] T.H. Trinh et al., "Solving Olympiad Geometry without Human Demonstrations," *Nature*, vol. 625, 2024, pp. 476-482.
- [4] NVIDIA, "NVIDIA ALCHEMI," NVIDIA Developer Documentation, [Accessed July 2026]
- [5] P. St. John et al., "BioNemo Framework: A Modular, High-Performance Library for AI Model Development in Drug Discovery," *arXiv preprint*, 2024. doi: 10.48550/arXiv:2411.10548v5
- [6] C. Lu et al., "The AI Scientist: Towards Fully Automated Open-Ended Scientific Discovery," *arXiv preprint*, 2024. doi: 10.48550/arXiv:2408.06292
- [7] Y. Yamada et al., "The AI Scientist-v2: Workshop-Level Automated Scientific Discovery via Agentic Tree Search," *arXiv preprint*, 2025. doi: 10.48550/arXiv:2504.08066
- [8] J. Gottweis et al., "Towards an AI Co-Scientist," *arXiv preprint*, 2025. doi: 10.48550/arXiv:2502.18864
- [9] OpenAI, "Introducing GPT-Rosalind for Life Sciences Research," OpenAI Research Release, 2026.
- [10] Anthropic, "Claude Science, an AI Workbench for Scientists, Is Now Available," Anthropic News, 2026.
- [11] S. Feng et al., "InternAgent-1.5: A Unified Agentic Framework for Long-Horizon Autonomous Scientific Discovery," *arXiv preprint*, 2026. doi: 10.48550/arXiv:2602.08990
- [12] OpenAI, "GPT-5 Lowers the Cost of Cell-Free Protein Synthesis," OpenAI Research Publication, 2026.
- [13] Chinese Academy of Sciences, "China Introduces Upgraded AI Foundation Model to Accelerate Scientific Discovery," *China Daily*, 2026.
- [14] N.J. Szymanski et al., "An Autonomous Laboratory for the Accelerated Synthesis of Inorganic Materials," *Nature*, vol. 624, 2023, pp. 86-91.
- [15] J. Bragg et al., "AstaBench: Rigorous Benchmarking of AI Agents with a Scientific Research Suite," *arXiv preprint*, 2025. doi: 10.48550/arXiv:2510.21652
- [16] S.S. Panigrahi et al., "HeurekaBench: A Benchmarking Framework for AI Co-Scientist," in *Proc. Int. Conf. Learn. Represent.*, (Rio de Janeiro, Brazil), Apr. 2026.
- [17] L. Mitchener et al., "Kosmos: An AI Scientist for Autonomous Discovery," *arXiv preprint*, 2025. doi: 10.48550/arXiv:2511.02824
- [18] A.E. Ghareeb et al., "Robin: A Multi-Agent System for Automating Scientific Discovery," *arXiv preprint*, 2025. doi: 10.48550/arXiv:2505.13400
- [19] L. Cong et al., "LabOS: The AI-XR Co-Scientist That Sees and Works with Humans," *arXiv preprint*, 2025. doi: 10.48550/arXiv:2510.14861
- [20] A. Novikov et al., "AlphaEvolve: A Coding Agent for Scientific and Algorithmic Discovery," *arXiv preprint*, 2025. doi: 10.48550/arXiv:2506.13131
- [21] Y. Fei et al., "Agentic LLM Reasoning in a Self-Driving Laboratory for Air-Sensitive Lithium Halide Spinel Conductors," *arXiv preprint*, 2026. doi: 10.48550/arXiv:2604.11957
- [22] I. Mandal et al., "Autonomous Microscopy Experiments through Large Language Model Agents," *arXiv preprint*, 2025. doi: 10.48550/arXiv:2501.10385
- [23] 한국과학기술정보연구원, "대화형 AI 통합 검색 서비스, ScienceON AI(SAI) 출시," KISTI 보도자료, 2024.
- [24] 한국원자력연구원, "원자력 특화 거대언어모델 'AtomicGPT' 공개," KAERI 보도자료, 2024.
- [25] LG AI연구원, "LG Showcases EXAONE's Real-World Industry Impact and Latest AI Research at ICML 2026," LG AI Research News, 2026.