

2 AI Scientist, R&D 병목을 해소하는 실행 인프라로 부상

➔ 연구 판단 병목에 따른 사람-AI 공동연구 확산

- R&D 병목, 실험 전 탐색·선별 단계로 이동
 - 과학 연구의 핵심 병목은 실험 장비의 처리량을 넘어, 급증하는 논문·특허·실험데이터와 후보물질 가운데 무엇을 먼저 검토하고 검증할지 정하는 판단 단계로 이동
 - 특히 신약·소재·화학은 분자·조성·실험 조건의 조합이 실제 실험 가능한 범위를 크게 넘어서면서, 방대한 탐색 공간에서 검증 가치가 높은 대상을 가려내는 일이 연구 자원 배분의 핵심으로 부상
 - 기존 자동화가 실험 실행과 데이터 수집의 속도를 높이는 데 집중했다면, 최근에는 자료 탐색·후보 선별·가설 비교·검증 우선순위 설정까지 지원하는 방향으로 연구 자동화의 요구 범위가 확대
 - 이에 제한된 인력·시간·비용을 검증 가치가 높은 실험에 배분하고, 검증 가능한 연구 경로를 설계하는 역량이 R&D 성과를 좌우하는 핵심 요인으로 부상
- AI Scientist, 사람-AI 공동연구 체계로 부상
 - AI Scientist는 문헌 탐색·데이터 분석·가설 생성·실험 설계·결과 검토 등 분절된 연구 단계를 하나의 목표 아래 연결해 수행하는 연구 특화 AI 시스템
 - AI Scientist의 핵심은 과학자 대체가 아니라, 사람이 감당하기 어려운 대규모 탐색과 반복 분석을 AI가 맡고 연구자는 문제 설정과 과학적 판단·검증에 집중하는 역할 분담에 있음
 - 실제 적용 시 문헌·데이터 분석, 후보물질 선별, 실험 조건 탐색 등 처리 범위와 결과를 비교적 명확하게 검증할 수 있는 연구 구간을 중심으로 확산
 - 현재의 성과는 연구 전 과정의 완전한 자율화보다, 시간과 비용이 집중되던 특정 병목을 줄여 연구자가 검증할 대상과 경로를 좁히는 수준에 해당
 - 이처럼 성과가 개별 모델의 성능만이 아니라 데이터·모델·연산·실험·검증의 연결 수준에 좌우되면서, R&D 경쟁력의 기반도 연구자 개인의 역량을 넘어 연구 실행 인프라의 구축 역량까지 확대

➔ AI Scientist, 흩어진 연구 단계를 하나의 실행 흐름으로 연결

● (개념) 개별 AI 도구를 넘어 연구 오케스트레이터로 진화

- AI Scientist는 개별 연구 과업을 수행하는 단일 AI 도구를 넘어, 문헌, 데이터, 모델, 실험 도구를 하나의 목표 아래 연결해 연구 흐름을 조율하는 시스템으로 발전
- 구체적으로는 연구 질문이나 목표를 입력받아 자료·데이터 탐색, 가설 수립, 분석, 실험 방향 설정 등의 과업을 연계해 수행하는 연구 특화 AI 에이전트와 자율 연구 시스템을 지칭
- AI Scientist는 미리 설정된 절차를 반복하는 기존 자동화와 달리, 각 단계의 결과를 평가해 다음 과업을 선택하고 필요에 따라 기존 계획을 수정하는 일부 자율성을 보유했다는 점에서 구분
- 자율성의 적용 범위에 따라 AI Scientist는 특정 과제의 후속 행동을 결정하는 전문 에이전트부터, 로봇·실험 장비를 제어하며 조건 설정과 검증을 반복하는 자율 연구 시스템까지 다양하게 구현

● (연구 절차) 선형 연구에서 순환형 연구로 전환

- 기존 연구에서는 문헌 검색, 데이터 분석, 실험 설계, 논문 작성 등 단계별 도구를 연구자가 직접 연결해 왔다면, AI Scientist는 이 과정을 하나의 연속된 연구 흐름으로 통합
- 각 단계에서 생성된 자료와 분석 결과를 관련 맥락과 함께 후속 과업으로 자동 전달하고, 실험·분석 결과를 기존 가설과 계획에 다시 반영하는 내부 피드백 구조를 형성
- 이에 연구 과정은 각 단계가 정해진 순서대로 이어지는 선형 절차에서, 중간 결과를 앞선 가설과 계획에 반영해 후속 과업을 조정하는 순환형 연구 절차로 변화

● (작동 구조) 계획·실행·평가·기억을 반복하는 에이전트 루프

- AI Scientist는 연구 목표를 세부 과업으로 분해하고 필요한 도구를 실행한 뒤, 그 결과를 평가해 다음 행동과 계획을 조정하는 에이전트 루프를 통해 연구를 단계적으로 수행
- (계획) 연구자가 제시한 상위 목표를 문헌 탐색, 데이터 분석, 실험 설계 등 실행 가능한 하위 과업으로 분해하고, 과업 간 선후관계와 구체적인 수행 순서를 설정



- (실행) 수립한 계획에 따라 논문과 데이터베이스를 검색하고 코드, 시뮬레이션 도구를 실행해 근거와 분석 결과를 확보하며, 실험 장비와 연계된 경우 설정 조건에 따라 물리적 작업까지 수행
- (관찰 및 평가) 실행 과정에서 생성된 오류 메시지와 분석값·실험값을 제약조건, 기존 근거, 검증 지표와 대조하고, 기준을 충족하지 못하면 계획이나 도구 선택을 수정해 재실행
- (기억) 활용한 자료와 도구, 실행 조건, 성공 및 실패 결과를 저장해 동일한 오류의 반복을 줄이고, 이후 과업 수행과 재계획에 필요한 맥락을 지속적으로 유지

⇒ 검증 가능성에 따른 AI Scientist 자율성 차등화

- 검증 가능한 과업부터 자율 수행 범위 확대
 - AI Scientist의 자율 수행 범위는 연결 가능한 과업의 수보다, 각 단계의 결과를 얼마나 빠르고 명확하게 검증하고 그 결과를 다음 행동에 반영할 수 있는지에 따라 결정
 - 결과를 즉시 재실행·비교·검증할 수 있는 과업에서는 AI가 오류를 수정하며 다음 과업까지 수행하나, 판단 기준이 불명확하거나 검증 비용이 큰 영역에서는 사람의 개입이 필요
 - 이에 AI 자율성은 코드·계산·데이터 처리처럼 컴퓨터 환경에서 검증 가능한 작업에서 가장 넓게 형성되고, 가치 판단·물리 실험·안전 관리가 필요한 단계로 갈수록 축소
- (문헌·데이터) AI가 먼저 맡는 컴퓨터 기반 연구 실무
 - 논문·특허 검색, 분석 코드 작성, 데이터 정제, 통계분석은 결과가 디지털 형태로 남고 동일 조건에서 재실행하거나 기존 자료와 대조할 수 있어, AI 자율 수행이 가장 먼저 확대되는 영역
 - 코드·계산은 실행 결과와 오류를 재현할 수 있고 문헌 검색은 원문·출처와 대조할 수 있어, AI가 문제를 점검하고 방법을 수정한 뒤 재수행하는 자기 교정을 반복 가능
 - 다만 이러한 자기 교정은 기술적 오류와 출처의 일치 여부를 확인하는 데 그치므로, 문헌의 누락이나 분석 결과의 과학적 의미처럼 별도 해석이 필요한 영역에서는 자율 수행 범위가 제한

- (가설 수립) 후보 생성은 AI가 지원, 최종 판단은 연구자 몫
 - 컴퓨터 기반 작업을 넘어 가설 생성과 연구 방향 설정으로 범위가 확대되면, 결과의 타당성을 즉시 확인하기 어려워지면서 AI가 독자적으로 수행할 수 있는 판단의 범위도 축소
 - AI는 기존 문헌에서 연구 공백을 찾고 다수의 가설을 생성할 수 있으나, 신규성, 학술적 가치, 실현 가능성은 연구 맥락과 자원 여건까지 고려해야 하므로 객관적 지표만으로 판단하기 곤란
 - 이로 인해 가설과 아이디어의 생성 부담은 줄어들지만, 제한된 자원을 투입할 후보를 선별하고 연구 우선순위를 결정하는 판단은 여전히 연구자의 역할로 잔존
- (물리 실험) 물리적 위험과 안전 기준에 따른 자율 실험 범위 제한
 - AI Scientist가 가설 수립을 넘어 장비를 직접 제어해 실제 물리적 실험까지 수행하면, 자율 수행 범위는 장비를 가동하고 측정 결과에 따라 다음 실험을 결정하는 단계로 확대
 - 해당 단계에서는 AI가 설정한 조건에 따라 장비가 실험을 수행하고, 그 결과를 다시 분석해 다음 조건을 정함으로써 연구자가 매 실험에 개입하지 않아도 유망한 조건 중심의 반복 수행이 가능
 - 다만 장비 제어 오류나 사전에 승인되지 않은 기계 동작은 안전사고로 이어질 수 있어, 물리 실험에서는 결과 검증 가능성에 더해 안전 기준이 자율 수행의 범위를 추가로 제한
- 사람의 최종 책임을 전제로 한 제한적 자율성
 - AI의 과업 수행 능력이 향상되더라도 연구 결과의 진위와 실험 안전에 대한 최종 책임은 연구자와 기관에 있으므로, 성능 향상만을 근거로 최종 판단 권한까지 이전하기는 곤란
 - 특히 연구 방향과 결과의 타당성, 실험의 지속 여부를 결정하는 판단은 과학적 설명과 책임을 수반하므로, AI의 성과와 관계없이 사람의 최종 검토와 승인이 필요
 - 따라서 AI Scientist의 자율성은 사람의 개입이 사라진 상태가 아니라, 최종 책임 주체를 명확히 한 상태에서 정해진 범위의 실행은 AI에 맡기고 핵심 판단과 책임은 사람이 유지하는 방식으로 이해

➔ AI Scientist, 연구 병목 완화 사례 확산

- 완전 자율화보다 병목 완화가 먼저 현실화
 - AI Scientist 계열 시스템은 연구 전 과정을 대체하기보다, 시간과 비용이 많이 드는 자료 탐색, 분석 실행, 후보 선별, 반복 실험을 중심으로 활용 범위가 확대



- 2026년에는 주요 시스템의 성과가 동료 심사 논문으로 검증되고, AI가 발굴한 후보가 후속 실험과 임상 3상, 제품화로 이어지면서 개념 검증을 넘어 제한적인 현장 적용 가능성을 제시
- 구체적인 활용은 분절된 디지털 연구 업무를 하나의 흐름으로 연결하거나, 방대한 탐색 공간에서 유망 후보를 선별하고, 실제 장비를 활용해 물리 실험을 반복 수행하는 방식으로 구분 가능
- 이들 사례는 AI가 연구자를 얼마나 대체했는가보다, 연구 과정의 어떤 병목을 완화했으며 후속 검증과 사람의 판단이 어느 지점에서 필요한지를 보여주는 데 의미

(1) 흠어진 디지털 연구 업무를 하나의 흐름으로 연결

- (사카나 AI) 아이디어 도출부터 논문 심사까지 연결하는 종단간 자동화
 - 사카나 AI의 AI Scientist는 연구 아이디어 도출부터 코드 작성·실행, 실험 결과 정리, 논문 작성과 AI 심사까지 디지털 연구의 주요 단계를 연계한 종단간 연구 에이전트
 - 연구 목표를 입력하면 기존 문헌을 검토해 아이디어의 신규성을 점검하고, 코드를 작성·실행한 결과를 논문 형식으로 정리한 뒤 별도의 AI 심사자를 통해 완성도를 평가
 - 이를 통해 생성한 논문 1편이 ICLR 2025 워크숍의 블라인드 심사에서 정식 채택 기준을 충족하여, AI가 디지털 연구 절차를 일정 수준까지 자동 연결할 가능성을 제시(Nature, '26.3)
 - 다만 해당 논문에서 존재하지 않는 문헌을 인용한 오류가 확인되면서, 형식적 완성도와 학술 심사 통과만으로는 과학적 타당성을 보장하기 어렵다는 한계도 노출
- (바이옴니) 전문 도구와 데이터베이스를 연동하는 생의학 연구 에이전트
 - 사카나 AI가 디지털 연구 절차를 종단간 연결했다면, 바이옴니(Biomni)는 생의학 연구에 분산된 자료 탐색과 분석 실행을 하나의 흐름으로 연결하는데 초점을 둔 연구 에이전트
 - 연구자가 일상 언어로 과제를 입력하면 필요한 전문 도구와 데이터를 스스로 선택하고, 여러 단계의 생의학 분석 절차를 구성해 순차적으로 수행
 - 실제로 연구자가 60시간 이상 수행하던 450여 개 파일의 정제·시각화·패턴 분석을 40분 만에 완료해, 반복적인 데이터 처리에 투입되는 시간과 업무 부담을 대폭 단축(Science, '26.7)

- 이미 학계와 산업계의 1만여 개 연구실에서 활용되며, AI가 반복적인 분석 실무를 맡고 연구자는 문제 설정과 결과 해석에 집중하는 역할 분담이 연구 현장으로 확산될 가능성을 제시

(2) 방대한 가설 후보군에서 실제 검증 대상 선별

- (구글 코사이언티스트) 여러 에이전트의 상호 검증으로 유망 가설 선별
 - 구글 Co-Scientist는 여러 에이전트가 가설의 생성, 비판, 순위 평가, 발전을 분담하여 상호 토론하며, 방대한 가설 가운데 실험으로 검증할 유망 후보를 좁히는 다중 에이전트 시스템
 - 제미나이 모델을 기반으로 가설을 생성한 뒤, 동료 심사 방식의 비판과 후보 간 우선순위 평가를 거치고 문헌·데이터와 대조해 근거가 부족한 주장을 걸러 내면서 유망 가설을 구체화
 - '26년 5월 구글은 Co-Scientist가 도출한 급성 골수성 백혈병, 간 섬유화, 항생제 내성 관련 가설을 각각 독립 실험으로 검증해, AI가 선별한 후보의 실험적 유효성을 입증(Nature, '26.5)
 - 해당 사례는 AI가 단독으로 결론을 내리기보다, 여러 에이전트가 후보를 반복적으로 평가하며 연구자가 최종 검증할 대상을 단계적으로 좁히는 접근으로 이해할 필요
- (인실리코 메디슨) AI 후보 탐색을 신약 임상 단계로 연결
 - 인실리코 메디슨은 AI를 활용해 질병 표적을 발굴하고 후보물질을 설계한 뒤, 이를 전임상과 사람 대상 임상 개발로 연결한 AI 신약 개발 사례
 - 노화 관련 생물학 데이터를 분석해 폐 섬유증에 관여하는 핵심 표적 단백질 TNIK을 찾아내고, 해당 단백질 구조에 결합하는 치료 후보물질을 생성형 AI로 설계
 - '26년 7월에는 AI로 발굴하고 설계한 TNIK 표적 치료제 렌토서팁(Rentosertib)의 임상 3상을 환자 320명 규모로 시작하며, 해당 후보를 후기 임상 단계까지 발전시켰다(Insilico, '26.7)
 - 다만 이는 AI가 신약 개발 전 과정을 자동으로 완성한 사례라기보다, 표적 탐색과 후보 설계에 드는 시간을 줄이고 발굴한 후보를 후기 임상으로 연결할 가능성을 보여준 성과로 이해할 필요



● (LG AI 연구원) 대규모 후보 탐색을 신소재 개발로 연결

- LG AI 연구원의 EXAONE Discovery는 비정형 논문에서 분자구조와 소재 정보를 추출하고, 연구자의 질의에 적합한 후보물질을 제안하는 소재 발굴 특화 AI Co-Scientist 플랫폼
- 실제로 42만여 개 후보물질을 하루 만에 유망 후보군으로 좁혀 탈모 치료용 신소재 램시딜을 발굴하고 제품 개발까지 연결하며, 소재 발굴 분야의 역량을 입증(LG AI 연구원, '26.7)
- 이후 동일한 후보 탐색 체계를 GS칼텍스와의 데이터센터용 액침 냉각유 소재 개발에도 적용해, 특정 제품을 넘어 성격이 다른 소재 분야로 확장할 수 있음을 실증

(3) 실험실로 확장되는 AI Scientist

● (RoboChem-Flex) 진입장벽을 낮춘 저비용 자율실험실

- 암스테르담대의 RoboChem-Flex는 화학반응 자동화 장치에 AI 기반 조건 최적화를 결합해, 목표한 결과를 얻기 위한 반응 조건을 스스로 탐색하는 저비용 자율실험실
- AI가 시약 비율·반응 시간·광량 등의 조건을 제안하면 장치가 해당 조건으로 실험을 수행하고, 측정된 수율과 순도를 다시 AI에 전달해 조건을 수정하는 과정을 반복
- 이를 통해 광촉매·생촉매 등 6개 반응에서 수율을 높이는 조건 조합을 연구자 개입을 최소화한 상태에서 자율적으로 탐색했으며, 최소 구성 기준 약 5천 달러에 구현(Nature, '26.4)
- 측정 단계에 연구자의 개입이 필요하다는 한계는 있으나, 고가 장비를 갖춘 대형 기관에 집중됐던 자율 실험을 소규모 연구실로 확산할 가능성을 보여준 사례

● (AP-Lab) 실험실 성과를 양산으로 연결한 산업형 자율실험실

- 중국 선전 연구진의 AP-Lab은 소재 발굴에 머물던 기존 자율실험실과 달리, 발굴한 소재를 실제 공장 생산으로 연결하도록 설계한 산업형 자율실험실
- AI가 진단용 자성나노입자의 제조 조건을 제안하면 로봇이 이를 실험하고, 성능 결과를 다시 AI에 반영해 더 나은 조건을 찾아가는 과정을 반복함으로써 소재 개발을 자동화
- 이때 실험실 성능만 높이는 데 그치지 않고, 바이러스 검출 성능과 실제 공장 생산 데이터를 함께 반영해 대량생산에도 적용할 수 있는 조건을 탐색

- 이를 통해 진단 소재 개발 기간을 기존 4~6개월에서 3주로 단축했으며, 도출한 조건을 100만 회 검사분 규모의 생산에 그대로 적용해 상용 제품에 준하는 성능을 확인(Advanced Science, '26.2)
- 이는 자율 실험의 성과를 실험실 내부의 소재 발굴에 머물게 하지 않고, 개발부터 양산까지의 기간을 단축해 제품화 속도를 높일 수 있음을 보여준 사례

➔ R&D 플랫폼 경쟁, 모델 성능을 넘어 연구 실행 환경으로 확장

● 연구 기록과 실행 절차가 축적되는 환경 확보 경쟁 확대

- 앞선 활용 사례에서는 AI Scientist의 성과가 개별 모델의 성능만으로 결정되는 것이 아니라, 문헌·데이터·도구·실험이 하나의 흐름으로 연결될 때 구체화된다는 공통점이 확인
- 이에 따라 경쟁의 초점도 더 높은 성능의 모델 개발을 넘어, 연구자가 과제를 수행하고 결과를 축적하는 연구실행 환경을 누가 제공하느냐로 이동
- 실제로 '26년 상반기 주요 빅테크와 AI 기업이 연구용 작업공간과 플랫폼을 잇달아 출시하면서, AI Scientist를 활용한 연구 자동화가 시험적 도구를 넘어 상용 서비스로 구체화
- 경쟁은 연구자가 처음 접속하는 작업공간, 조직의 R&D 절차를 관리하는 운영 플랫폼, 여러 환경에서 호출되는 전문 계산 기반, 이들 계층을 자체적으로 연결하는 생태계 통합으로 전개

(1) (연구자 점점) 대화형 AI의 연구자 기본 작업 공간화

● (OpenAI) 무료 개방 기반의 연구자 점점 확대

- 연구자는 문헌 검색, 집필, 코드 실행, 데이터 분석마다 서로 다른 도구를 사용해 왔으나, 대화형 AI 기업이 이를 하나의 환경에 통합하면서 연구자 기본 작업 공간을 둘러싼 경쟁이 확대
- OpenAI는 '26년 1월 GPT-5.2를 내장한 무료 논문 작업공간 Prism을 출시하고, 계정만 있으면 제한 없이 이용하도록 개방해 폭넓은 연구자 접점을 확보(OpenAI, '26.1)
- 이어 '26년 3월 코딩 에이전트 Codex를 통합해 Prism에서 코드를 직접 실행할 수 있도록 하면서, 논문 집필 공간을 데이터 분석과 계산 실행까지 아우르는 연구 플랫폼으로 확장
- 이는 무료 개방으로 이용자 규모를 먼저 키우는 전략으로, 확보한 이용자층을 발판 삼아 순차적으로 기능을 확장하며 연구 작업의 기본 도구로 자리 잡으려는 접근



- (Anthropic) 민감 데이터 보호와 감사 가능성을 결합한 연구 워크벤치
 - Anthropic은 '26년 6월 연구용 워크벤치 Claude Science를 공개하고, 문헌 분석·시각화·원고 작성 등 분산된 연구 작업을 하나의 환경에서 수행하도록 구성(Anthropic, '26.6)
 - 민감 데이터는 연구기관 자체 인프라에 둔 채 필요한 맥락만 전송하고, 코드·메시지 이력·설명 정보를 함께 남겨 데이터 보안과 결과의 감사 가능성을 높이는 데 초점
 - OpenAI가 무료 개방을 통해 연구자 접근을 넓히는 전략이라면, Anthropic은 데이터 보호와 감사 가능성을 앞세워 규제 및 보안 요건이 엄격한 연구기관 수요 확보에 초점

(2) (조직 운영) 기업 R&D 절차와 판단 기록의 플랫폼화

- (Microsoft) 연구 과정과 판단 근거를 조직 자산으로 축적
 - 기업 R&D에서는 데이터 접근과 연구 승인, 분석 절차와 검토 기록이 여러 시스템에 흩어져 있어, 이를 조직이 검토·재사용할 수 있는 공통 절차로 통합하려는 플랫폼 경쟁이 확대
 - Microsoft는 '26년 6월 Azure 기반의 기업용 과학 워크플로 플랫폼 Microsoft Discovery를 정식 출시하며 조직 단위의 R&D 운영 절차를 통합하는 시장에 진입(Microsoft, '26.6)
 - 문헌과 데이터 탐색부터 가설 수립·분석·검증까지 연구 과정을 연결하고 근거와 추론 경로를 함께 기록해, 특정 연구자에게 머물던 판단 방식을 조직이 검토·재사용하는 공통 절차로 전환
 - 이는 조직의 연구 승인 및 검토 절차를 플랫폼 안에 그대로 이식하고 워크플로로 표준화해, 기관의 R&D 운영 방식 자체를 선점하려는 전략
- (AWS) 분석 절차를 조직 자산으로 전환하고 외부 실험까지 연결
 - AWS의 Amazon Bio Discovery는 코드 작성 없이 여러 생명과학 모델과 조직 내부 데이터를 결합해 신약 후보 탐색 절차를 구성·실행할 수 있도록 설계된 연구 플랫폼 (AWS, '26.4)
 - 개별 연구자가 구성한 분석 절차와 모델 활용 방식을 조직 내에서 재사용 가능한 형태로 축적해, 개인의 분석 노하우를 공동연구 자산으로 전환
 - 이에 더해 발굴한 후보를 외부 인프라로 보내 검증하고 그 결과를 다시 분석에 반영하는 과정을 동일 체계에 연결해, 조직이 직접 보유하지 않은 실험 자원까지 통합 운영 가능

- 이로써 AWS는 조직이 직접 보유한 내부 자원과 외부 기관의 실험·검증 자원을 자사 클라우드 위에서 단일 체계로 통합함으로써, 신약 개발 과정 전반을 자사 인프라 안으로 편입

(3) (전문 실행 기반) 전문 모델과 연산 인프라가 공통 실행 기반으로 부상

● (NVIDIA) 상위 플랫폼에 구애받지 않는 공통 실행 기반

- 연구자 접점과 조직 운영을 둘러싼 경쟁이 확대되는 가운데, 그 아래에서는 분야별 전문 모델과 연산 기능을 여러 작업환경에 공통으로 공급해 연구실행의 기반을 선점하려는 경쟁도 전개
- NVIDIA는 '26년 6월 BioNeMo Agent Toolkit을 공개하고, 생명과학 분야의 전문 모델과 계산 도구를 여러 에이전트가 호출할 수 있는 스킴 형태로 제공 (NVIDIA, '26.6)
- 특정 연구 플랫폼에 종속되지 않는 개방형 구조를 채택해 제약·바이오·연구 데이터 분야 50여 개 기업이 이를 활용하거나 자체 시스템에 통합할 수 있도록 구성
- 이를 통해 최종 사용자 인터페이스를 직접 제공하지 않더라도, 여러 연구 플랫폼이 자사 모델과 GPU 연산을 활용하도록 하며 전문 실행 기반 영역에서 영향력을 확대

(4) (생태계 통합) 모델·데이터·도구·클라우드를 묶은 연구 생태계 통합

● (Google) 연구 전반을 자체 스택으로 연결하는 통합 전략

- 연구자 접점과 조직 운영, 전문 계산 기반에서 각각 경쟁이 전개되는 가운데, 여러 계층을 자체 모델, 데이터, 클라우드로 연결해 연구 전반을 하나의 생태계에 묶으려는 전략도 등장
- Google은 '26년 5월 문헌 분석과 가설 생성, 계산 탐색을 연결하는 Gemini for Science를 공개하며 여러 연구 단계를 자체 생태계 안에서 통합하는 경쟁에 진입(Google, '26.5)
- 자체 범용 모델에 단백질·유전체 분야의 특화 모델과 데이터베이스, 에이전트 플랫폼과 클라우드 인프라를 결합해 앞선 단계에서 생성된 결과를 후속 연구에 연속적으로 활용할 수 있도록 구성
- 이는 모델, 데이터, 클라우드를 한 스택에 묶어 계층 간 연동 부담을 없앴 편의성을 강점으로, 기업 및 기관의 연구 전 과정을 자사 생태계 안에 묶어두려는 전략
- 이처럼 R&D 플랫폼 경쟁은 개별 계층의 우위를 넘어, 과학데이터 접근권과 지식재산 창출 경로 전체를 누가 통제하느냐의 문제로 확대



➔ AI Scientist, 연구 인프라 재설계 요구

● 가설·계산·실험의 반복이 지속적 연산 수요 창출

- AI Scientist는 한 번의 응답 생성으로 끝나는 일반 생성형 AI와 달리, 가설 수립·계산 검증·실험 반영을 반복하면서 연구 진행 기간 전반에 걸쳐 지속적 연산 수요 발생
- 하나의 연구과제 안에서 언어모델 추론, 코드 실행, 분자·소재 시뮬레이션, 실험 데이터 분석이 순차적으로 수행되며 계산 흐름이 장시간 이어지는 구조 형성
- 자율 수행 범위가 넓은 영역일수록 산출물의 기계적 재실행과 검증이 빈번해, 동일 과제 내에서도 연산 수요가 일회성 학습이 아닌 연구 진행 기간 전체로 분산
- 이에 AI Scientist용 인프라는 대규모 GPU 확보에 그치지 않고, 성격이 다른 연산을 끊김 없이 연결하고 장기간 안정적으로 운영하는 구조를 요구

● 가설·계산·실험의 반복이 지속적 연산 수요 창출

- AI Scientist의 워크플로에서는 언어 추론, 수치 해석, 분자·소재 시뮬레이션이 과업 단계마다 교차하므로 단일 연산장치만으로는 효율적 대응에 한계
- 이에 CPU·GPU·AI 가속기·HPC가 과업별로 역할을 분담하고, AI 추론과 과학 계산을 하나의 연구 흐름 안에서 연계하는 이종 컴퓨팅 구조가 중요
- 최근에는 AI가 기존 시뮬레이션 계산을 가속하거나, 복잡한 계산을 대리 모델로 근사해 실험 후보를 빠르게 좁히는 방식도 연구 현장에서 함께 확산
- 전통적 HPC 환경과 클라우드 기반 AI 환경이 분리될 경우, 연구 흐름이 끊길 수 있어, 두 환경의 연산 자원을 통합적으로 배분·운영하는 체계 필요
- 장기적으로 양자컴퓨팅이 일부 분자·소재·화학 계산을 보완할 가능성이 있으나, 단기적으로는 AI·HPC 통합 기반 위에서 특정 계산에 선택적으로 결합하는 수준에 머물 전망

● 가설·계산·실험의 반복이 지속적 연산 수요 창출

- AI Scientist가 반복적으로 생성하는 실험·시뮬레이션 데이터를 후속 연구에 재활용하려면, 연산장치 외에도 고성능 스토리지와 데이터 관리 체계가 함께 필요
- 연구 데이터는 규모·보안·비용 제약으로 이동이 어려운 경우가 많아, 데이터를 중앙으로 옮기기보다 데이터가 있는 곳에 모델과 연산을 배치하는 설계가 중요
- 실험 현장에서 생성된 측정값을 즉시 분석해 다음 실험 조건에 반영하려면, 장비 인근의 엣지 컴퓨팅과 연구망 기반의 고속 데이터 연결이 함께 필요
- 결국 AI Scientist의 실행 성능은 개별 연산장치의 능력뿐 아니라, 계산·저장·실험 시설 사이에서 데이터가 얼마나 빠르고 안정적으로 순환하느냐에 의해 좌우

➔ 자율 연구 확산, 검증·기록·안전 체계가 핵심 과제로 부상

● 형식적 완성도와 과학적 타당성의 분리

- AI Scientist는 논리적으로 정돈된 논문, 가설, 분석 결과를 빠르게 생성할 수 있으나, 형식적 완성도가 곧 과학적 타당성과 재현성을 보장하지는 않음
- 특히 AI가 생성한 산출물은 문장과 구조가 자연스러워 보이더라도, 존재하지 않는 문헌 인용, 부정확한 수치, 검증되지 않은 결론을 포함할 가능성 존재
- 이러한 오류가 후속 연구의 전제로 재활용될 경우, 초기 단계의 작은 오류가 가설 설정·실험 설계·논문 작성 전반으로 연쇄 확산될 우려가 존재
- 따라서 AI Scientist의 산출물은 사람이 작성한 연구 결과보다 낮은 기준으로 볼 것이 아니라, 출처·근거·재현 가능성을 별도로 확인하는 검증 절차 필요

● 재현성 확보를 위한 연구 경로 기록 확대

- 전통적인 연구에서는 방법론과 데이터 기록이 재현성 확보의 핵심이었으나, AI Scientist 환경에서는 사용된 모델 버전·프롬프트·도구 호출 순서·중간 판단까지 보존해야 결과를 재구성 가능
- 특히 AI의 내부 추론 과정이 불투명한 경우가 많아, 특정 가설이 왜 제안되었고 어떤 실험이 왜 우선 배정되었는지를 사후에 추적할 수 있는 기록 체계가 연구의 신뢰성을 좌우
- 이에 따라 산출물만이 아닌 산출물에 이르는 전체 경로를 기록하고, 주요 판단 지점에서 근거와 대안을 함께 보존하는 감사 기반의 연구 기록 방식이 요구
- 다만, 기록 범위가 확대될수록 관리 비용도 증가하므로, 재현성과 검증 가능성 확보에 필요한 최소 기록 기준을 분야별로 정립하는 것도 과제

● 자율 실험의 물리적 안전과 이중용도 위험 관리

- AI Scientist가 로봇 실험실과 결합해 화학물질을 취급하고 장비를 제어할 경우, 디지털 오류가 물리적 피해로 이어질 수 있어 실험 수행 능력과 안전성을 별도로 검증할 필요
- 자율 실험 환경에서는 디지털 트윈을 활용해 실제 장비 가동 전에 배치·도달 범위·충돌 여부·안전 인터록을 사전 검증하고, 이상 발생 시 즉시 중단하는 안전 절차가 운영의 전제 조건
- 나아가 AI가 독자적으로 실험을 설계하고 새로운 화합물을 합성하는 단계에 이르면, 독성 물질 생성이나 이중용도 연구 등 의도하지 않은 위험에 대한 통제 체계 마련도 과제



- 따라서 자율 실험 확대는 장비 제어 능력의 고도화뿐 아니라, 실험 승인·물질 반입·허용 조건·중단 기준을 포함한 운영 설계를 병행할 필요

● 자율성 수준별 책임·승인 절차 차등 적용

- AI Scientist가 수행하는 과업의 범위가 넓어질수록, 연구 결과의 진위와 실험 안전에 대한 최종 책임 주체를 연구자·기관·플랫폼 사업자별로 명확히 구분 필요
- 문헌 검색과 데이터 정제처럼 위험이 낮은 과업과, 가설 확정·실험 설계·장비 제어처럼 후속 영향이 큰 과업에 동일한 승인 기준을 적용하기는 곤란
- 이에 자율성 수준과 위험도에 따라 출처 검증, 사용 로그 보존, 사람 승인, 사전 심사, 즉시 중단 요건을 차등 적용하는 단계적 연구 거버넌스 체계가 필요
- 결국 AI Scientist의 신뢰성은 AI가 얼마나 많은 연구를 수행하는가보다, 수행 과정과 결과를 얼마나 검증 가능하게 남기고 책임 있게 통제하는가에 좌우

출처 : Stanford HA 외(2026.7.)

<https://arxiv.org/html/2605.23204v1>

<https://arxiv.org/html/2605.18661v1>

<https://hai.stanford.edu/news/how-ai-is-accelerating-scientific-discovery>

<https://www.preprints.org/manuscript/202601.0405>

<https://www.preprints.org/manuscript/202607.0102>

<https://link.springer.com/article/10.1007/s10994-025-06955-2>

<https://pubs.acs.org/doi/10.1021/acsmaterialslett.6c00224>

<https://www.nature.com/articles/s43246-026-01219-5>

<https://www.nature.com/articles/s42004-026-01932-9>

<https://news.stanford.edu/stories/2026/07/biomni-ai-powered-biomedical-co-scientist>

<https://blog.google/intl/ko-kr/company-news/technology/coscientist-io-2026-kr/>

<https://phys.org/news/2026-03-ai-paper-peer.html>

<https://www.drugdiscoverynews.com/the-2026-ai-power-shift-17020>

<https://codephusion.com/blog/ai-models-for-life-sciences>

<https://link.springer.com/article/10.1007/s43681-025-00908-0>

<https://www.techradar.com/pro/anthropic-launches-ai-workbench-for-scientists>

[-using-claude](#)

<https://www.theverge.com/ai-artificial-intelligence/914463/openai-sora-bill-pees-kevin-weil-leaving-departing>

[bles-kevin-weil-leaving-departing](#)

<https://blog.google/innovation-and-ai/technology/research/gemini-for-science-io-2026/>

<https://blogs.microsoft.com/blog/2026/06/02/microsoft-build-2026-be-yourself-at-work/>

<https://www.reuters.com/business/healthcare-pharmaceuticals/amazon-launches-ai-research-tool-speed-early-stage-drug-discovery-2026-04-14/>