

제네시스 미션과 「과학: 새로운 황금기」에 부치는 Sci.Sci & AI4S 기반 검토의견서

A REVIEW ON THE U.S. SCIENCE STRATEGY AND THE *GENESIS MISSION*,
BASED ON SCIENCE OF SCIENCE AND AI FOR SCIENCE.

이제현. 2026. 7. 31.

과학은 개척자에게 아직 탐사되지 않은 드넓은 오지를 내어 준다,
그 일에 맞는 도구를 갖춘 이에게.

Science offers a largely unexplored hinterland for the pioneer who has the tools for his task.

— Vannevar Bush,

※ 이 글은 미국 백악관 과학기술정책실(OSTP)이 2026년 7월 펴낸 대통령 보고서 「Science: A New Golden Age」 (Michael Kratsios 著)와 그 실행체인 미 에너지부(DOE)의 제네시스 미션(Genesis Mission, 행정명령 14363호 및 공모 DE-FOA-0003612)을 읽고, 한 사람의 연구자로서 남기는 검토의견이다. 근거로는 1964년 이후 출판된 527편의 과학학(science of science) 논문과 2015년 이후 출판된 2,657편의 AI4S(AI for Science) 논문을 담은 개인 코퍼스([paper curation](<https://github.com/jehyunlee/paper-curation>))를 활용했다.

요약 Executive Summary

미 행정부는 2026년, 자국의 과학 체제를 다시 설계하겠다고 선언했다.

「Science: A New Golden Age」는 그 청사진이고, 제네시스 미션은 AI로 과학을 가속하겠다는 그 청사진의 플래그십 사업이다. 이례적인 점은 이 전략의 뼈대가 정치적인 수사보다 지난 20년간 과학학이 축적한 실증 연구결과로 짜여 있다는 것이다. 참신한 연구가 구조적으로 감소하는 현상¹, 스타 과학자 사후에 열리는 기회², 재현성의 위기³⁴, 합의 기반 심사의 보수성⁵, 사람에게 거는 장기 지원의 효과⁶가 모두 현실에 대한 명시적 진단과 처방의 근거로 쓰였다. 학문적 근거가 정책에 강하게 반영된 이례적인 모범사례라고 볼 수 있다.

그러나 한편으로 미국적 현실과 정책이 강하게 반영되어 학문적 근거와 상충되는 경우도 적지 않다. 본 보완의견은 과학학과 AI4S의 학문적 근거에 비추어 (1) 일치하는 점, (2) 일치하지 않는 점, 그리고 (3) 한국의 현실 위에서 함께 생각해 볼 거리로 구성되어 있다.

제1장 일치하는 점은 전략이 옳게 채택한 근거를 확인한다. 감속 진단, 사람 중심 장기 지원, 재현성 회복(골드 스탠더드 과학), 동료심사 다변화, 그리고 AI가 실제로 연구를 가속한다는 발견⁷⁸⁹과 자율실험의 잠재력¹⁰까지, 학문적 연구결과와 대체적으로 견고하게 정렬되어 있다.

제2장 일치하지 않는 점은 아직 반영되지 않았거나 정책 방향과 충돌하는 근거를 짚는다. 예를 들어, AI는 개인의 영향력을 키우는 동시에 학계 전체의 연구 초점을 좁힌다는 중요한 지견¹¹이 누락되어 있다. 제네시스 미션은 생산성 향상에 집중할 뿐 다양성 축소에 대한 대비가 없다. 여기에 제네시스 미션의 Phase I에 작은 팀이 혁신을 유발한다는 근거¹²가 반영된 것으로 보이지만, Phase II에서 대형 팀으로 단계적 확장때 발생할 수 있는 긴장에 대비하는 장치가 없다. 이 외에도, 원격·분산 협업의 불리함¹³, 자율실험실 성과의 재현 가능성 논쟁¹⁴¹⁵¹⁶, 이민·국제 인재의 기여¹⁷와 해외 의존 축소 기조가 주요 논문들에서 제시하는 방향과 어긋나고 있다.

제3장 생각할 거리는 이 어긋남들을 한국의 조건 위에 겹쳐 놓는다. 한국은 이미 PBS 폐지¹⁸, K-문샷과 PD 제도¹⁹²⁰, KIST의 임무중심연구소²¹, 글로벌 TOP 전략연구단²², 기초연구 생태계 육성²³, 그리고 국가과학AI연구센터(NAIS)²⁴로 같은 방향의 개혁을 진행 중이다. NAIS는 데이터·GPU·모델·표준을 모으는 허브(Hub)와 출연연이 도메인 문제를 푸는 스포크(Spoke)로 짜여, 미국이 제네시스 미션으로 하려는 일을 우리 방식으로 이미 시작한 셈이다. 다만 인구·안보·국제협력의 조건은 미국과 다르다. 이 지형 위에서 무엇을 할지는 정해진 처방이 아니라 우리가 각자, 그리고 함께 생각해 볼 물음으로 남긴다.

미국의 전략을 맹목적으로 참고하기보다 근거를 찾아 우리에게 맞는 정책을 고심할 필요가 있다.

서론: 논문 위에 세운 정책



얼마 전, 우연히 미 백악관이 발간한 「Science: A New Golden Age」 보고서를 읽으며 적지 않게 놀랐다. 지난 20여년간 많은 연구자들이 축적한 과학학(Science of Science)의 실증 연구를 다양하게 인용하고 있기 때문이다. 수년 전 Dashun Wang 과 Albert-László Barabási가 공저한 "과학의 과학(Science of Science. 한국판 제목)"²⁵을 재미있게 읽은 이후, 이 분야에 관심을 가지고 논문을 계속 따라오다 최근에는 Dashun Wang 그룹에서 나온 논문 80편을 포함해 약 130편 가량의 논문을 일주일간 출퇴근길을 이용해 공부한 터라 보고서에 나오는 문장마다 원 논문이 떠오를 정도였다.

예를 들어, 보고서에는 "과학이라는 기계가 느려지고 있다(*The Scientific Machine Is Getting Bugged Down*)"이라는 문장이 나오는데, 이는 막연한 걱정이 아니라 4,500만편의 논문과 390만건의 특허를 뒤져 연구가 갈수록 혁신보다 기존 지식을 조금 더 보태는 쪽으로 바뀌었음을 보인 연구¹에서 나온 말이다. 여기에, 분야가 성장할수록 소수의 유명 논문에만 인용이 집중되는 바람에 새 아이디어가 묻힌다는 관찰²⁶이 더해진다. "과학은 장례식마다 한 걸음씩 나아간다(*science advances one funeral at a time*)"는 오래된 격언조차, 생명과학 분야의 요절한 스타 연구자 452명을 추적한 실증 연구²에서 재조명된 말이다.

제네시스 미션은 이 진단에 대한 미 정부의 처방이다. 국가 슈퍼컴퓨터와 AI 모델, 과학 장비와 데이터를 하나의 사이언스 클라우드로 묶고, 스스로 나아지는 과학용 AI 모델을 키우는 컨소시엄을 세워, 17개 국가 난제에서 "AI는 우리가 과학을 다루는 방법을 급진적으로 바꿈으로써 과학 발견을 앞당긴다(*AI will accelerate and radically transform the way we do science*: 이관 서한, *AI will accelerate discovery*: 제5장)"는 것을 증명하고 10년 안에 미국 과학의 생산성을 두 배로 올리겠다는 구상이다. 연구비는 두 단계로 나누어 지급될 예정이다. 첫 단계에서 정확한 규모는 명시되지 않았으나 작은 팀에 9개월짜리 소액(약 5억~7억 원)을 주어 가능성을 시험하고, 6개월 시점에 지속 여부를 결정한 뒤, 살아남은 소수만 기존의 3~5배로 규모를 확대한다.

그러나 한편으로는 이민을 제한하는 정책 등 과학학의 발견과는 상반된 주장도 발견되고, 미국의 과학적 자부심을 강조하는 등 우리 실정과 다른 대목도 적지 않게 눈에 띄었다.

이에 기억에 의존하기보다 보유하고 있는 과학학 논문 527편과 AI for Science 논문 2,657편을 활용해 이 보고서의 내용을 항목별로 확인하고, 우리 실정에 맞는 좋은 제안이 있다면 찾아내고자 했다. 연구결과를 근거로 정책을 수립한다면, 마찬가지로 연구결과를 근거로 정책을 비판하거나 더 나은 정책을 제안할 수 있을 것이다. 근거자료를 AI-Ready Data 로 보유하고 이를 활용할 수 있는 기술을 확보하고 있기에 가능한 일이다.

제 1 장 일치하는 점

미국의 전략이 과학학과 AI4S의 발견과 일치하는 대목. 그대로 참고해도 좋은 항목이다.

과학의 속도가 느려지고 있다

밭에서 감자를 캔다고 하자. 처음에는 삽만 꽂아도 굵은 감자가 쏟아지지만, 밭을 한 번 다 파헤친 뒤부터는 같은 수확을 얻으려 더 깊이, 더 오래 파야 한다. 보고서는 오늘날 과학이 이같다고 진단한다. 돈과 사람을 더 넣는데도 혁신이라 부를 수 있는 발견은 오히려 줄어든다는 것이다. 크게 두 가지 근거를 기반으로 한 주장이다. 하나는 4,500만 편의 논문과 390만 건의 특허를 분석한 연구로, 연구가 갈수록 기존 지식을 뒤엎기보다 그 위에 얹히는 쪽으로 바뀌었음을 보인 연구¹다. 다른 하나는 분야가 커질수록 소수의 대표 논문에만 인용이 집중되어 새로운 아이디어가 좀처럼 주목받지 못한다는 관찰²⁶이다. 파괴적 혁신을 측정하 지표(CD 지수: Consolidation-Disruption Index)를 두고 측정이 편향됐다는 반박이 제기된적 있다. 이에 원저자들이 지표를 보완했음에 결론이 유지되어 가설을 재확인했다²⁷.

과제가 아니라 사람을 지원한다

연구비를 줄 때 우리는 보통 "무엇을 하겠느냐"를 묻는다. 계획서를 받고, 그 계획대로 했는지 채점한다. 그런데 미국의 의학연구 명문 HHMI(하워드휴즈의학연구소)는 이 방식을 바꾸었다. 계획서가 아니라 사람을 보고 장기적으로 지원하며, 초기의 실패를 눈감아 줬다. 결과적으로 HHMI의 지원을 받은 연구자는 비슷한 수준의 정부 지원 과학자보다 크게 성공한 논문도, 크게 실패한 논문도 훨씬 많이 냈다⁶. 안전하게 중간 길을 가기보다 여러 모로 모험을 했음이 증명된 것이다. 백악관 보고서에서 "과제가 아니라 사람을 지원한다"는 것은 자잘한 성공 대신 혁신을 기대하며 여기에 걸맞는 지원을 한다는 것으로 읽힐 수 있다. 전체 기간의 예산 전부를 첫해에 확정·배정하는 방식의 5년 이상 장기 지원 (cf. NIH, NSF 등 대부분은 명목상 다년 과제여도 매년 성과를 보여야 다음 해 예산이 나온다), 보고 최소화와 초기 실패에 대한 관용은 모두 이 실증 연구를 기반으로 하고 있다.

거인이 떠난 자리

"과학은 한 번의 장례식마다 한 걸음씩 나아간다"는 오래된 말이 있다. 대가가 버티고 있으면 그 그늘에 눌러 새 시도가 나오기 어렵다는 뜻이다. 이 말이 정말 맞는지 확인한 연구가 있다. 젊은 나이에 갑자기 세상을 떠난 생명과학 스타 452명을 추적해 보니, 그들이 떠난 뒤 그와 함께 일하지 않던 연구자들의 논문이 평균 8.6% 늘었고, 그 새 논문들이 유난히 많이 인용됐다². 스타가 사라지자 비로소 바깥의 새 피가 흘러든 것이다. 보고서가 말하는 "기득권 세금(the incumbency tax)", 곧 이미 자리를 잡은 사람과 아이디어가 누리는 일종의 자릿세는 비유가 아니라 숫자로 측정된 현상이다²⁸.

재현성 위기

과학의 신뢰는 "누가 다시 해봐도 같은 결과가 나온다"는 데서 온다. 그런데 유명 심리학 논문 100편을 그대로 다시 실험해 봤더니 3분의 1 남짓만 같은 결과가 나왔다³. 통계를 다루는 관행 탓에 출판된 결과의 상당수가 애초에 거짓일 수 있다는 경고⁴가 있었다. 대개 $p\text{-value} < 0.05$ 로 유의성을 판정하지만, 이는 '우연이라기엔 드문 결과'를 의미할 뿐 '이 발견이 참이다'를 증빙하지는 못하기 때문이다. 게다가 실험연구는 표본 수가 적은 경우가 많은데, 표본이 작으면 진짜 효과는 자주 놓치고 어쩌다 유의하게 나온 것이 크게 주목받는 경향이 있다. 게다가 최고 저널에 실린 사회과학 실험조차 재현에 실패한 사례²⁹, 과학 전반의 편향을 폭넓게 뒤진 분석³⁰까지 더하면, 재현성 위기는 과학학에서 이견이 가장 적은 문제에 속한다. 보고서와 '골드 스탠더드 과학 회복(Restoring Gold Standard Science)' 행정명령(Executive Order 14303)이 재현성·투명성·반증 가능성을 연방 연구 전반에 요구한 것은 이 공통된 문제 의식에 기인한 것이다.

심사위원회는 튀는 걸 싫어한다

여러 전문가가 모여 합의로 뽑으면 좋은 것을 고를 것 같지만, 실제로는 아무도 반대하지 않을 무난한 제안이 살아남기 쉽다. 튀는 아이디어는 누구 한 명만 고개를 저어도 탈락하기 때문이다. 보고서는 이 문제를 알아, 개별 심사자가 합의를 거스르고 한 건을 밀어 줄 수 있는 골든 티켓, 며칠 만에 결정하는 패스트 그랜트, 믿을 만한 연구자에게 배분 권한을 나눠 주는 재그랜팅을 처방으로 내놓는다. 이 방향은 세 갈래 연구와 맞닿는다. 심사에서 심사자의 전문성이 편향보다 크게 작동한다는 분석⁵, 동료심사가 아이디어의 확산에 미치는 영향을 다룬 최근 연구³¹, 그리고 과학계가 스스로를 더 과학적으로 들여다보자는 자기 진단³²이다.

AI는 과학을 가속화한다. 진짜다.

AI가 과학을 빠르게 한다는 말은 과장이 아니다. 실제 연구 현장에서 AI가 얼마나 도움이 되는지 정량화한 분석⁸, 사람이 잘 가지 않는 길을 일부러 골라 탐색하도록 설계한 AI가 새로운 조합을 더 잘 찾더라는 연구⁷, 과학 실무의 자동화가 무엇을 열고 무엇을 위협하는지 정리한 종합 논의⁹가 모두 같은 방향을 가리킨다. 특히 보고서가 "무언가를 만들어 내는 비용은 급격히 싸졌지만 그것이 맞는지 검증하는 비용은 그대로다"라며 검증 병목을 정면으로 지적한 대목은, AI4S 연구가 반복해서 확인해 온 통찰과 그대로 겹친다.

사람이 지켜보는 자율실험실

기계가 스스로 가설을 세우고 실험하고 결과를 본 뒤 다음 실험을 정하는 자율실험실은 발견의 시간을 크게 줄일 수 있다. 공상이 아니다. 여러 AI 에이전트가 역할을 나눠 SARS-CoV-2에 붙는 새 나노항체를 설계하고, 그것을 실제 실험으로 확인한 '가상 실험실'¹⁰이 대표 사례다. 다만 오해는 금물이다. 이 성과가 나온 것은 사람이 문제를 정의하고 에이전트들을 지휘했기 때문이지, 기계가 알아서 방향까지 정했기 때문이 아니다. 반대로 사람이 지휘의 손을 놓으면 같은 도구가 길을 잃는다. 방향을 스스로 좁히지 못해 막다른 길을 맴돌고, 없는 것을 봤다고 우기며, 자기 결과의 좋고 나쁨조차 가려내지 못하고³³, 자율 연구

시스템의 실제 성능이 광고에 못 미친다는 재평가도 잇따른다¹⁶. 요컨대 잠재력은 분명하되, 그 잠재력은 사람이 방향을 잡고 지휘하는 구조 안에서만 현실이 된다.

정리하면, 미국 전략의 진단은 과학학과 AI4S가 오래 쌓아 온 결론과 잘 맞는다. 다만 진단이 여기서 멈춘 것이 아쉬울 뿐이다.

제 2 장 일치하지 않는 점

미국의 전략이 과학학과 AI4S의 발견과 일치하않거나, 심지어 반대인 대목.

AI는 개인을 키우고 학계를 좁힌다

논문 6,790만 편을 분석한 연구¹¹에 따르면, AI를 쓰는 과학자 개인은 논문 수가 약 67% 늘고 인용은 약 3배로 뛰며 팀 리더가 되는 시점도 빨라졌다. 여기까지는 좋은 이야기지만 학계 전체를 놓고 보면 정반대의 일이 벌어진다. AI가 다루는 주제의 폭이 좁아지고, 이미 데이터가 넘치는 익숙한 영역으로 연구가 쏠린다.

이 쏠림은 추상적 우려가 아니라 두 나라 플래그십 사업의 주제 선택에서 이미 드러난다. 제네시스 미션은 17개 도전 과제로 넓게 펼쳐 있지만, AI를 실제로 앞세우는 축(자율실험실·과학 파운데이션 모델·물질 설계)은 화학·소재·바이오에 몰려 있고³⁴, NAIS의 자율형 AI 과학자 파일럿도 바이오부터 시작해 촉매·전도체 같은 화학·소재를 예로 든다²⁴. 모두 데이터가 두껍고 실험 자동화가 잘 되는, AI가 이미 몰려드는 분야다. 공급망 안보가 걸린 반도체·핵심광물(희토류)도 미국은 제네시스의 별도 과제로, 한국은 국가전략기술로 함께 다루지만, 큰 데이터셋이 없거나 실험 자동화가 어려운 분야, 표본이 얇은 이론·생태·희귀 현상 쪽은 애초에 뒤로 밀리기 쉽다. 개인은 빨라지는데 집단의 지식 지도는 오히려 좁아지는 것이다.

제네시스 미션이 성공의 잣대로 삼는 "AI 우위"는 속도와 예측력에만 맞춰져 있고, 이 다양성 축소를 재는 방법도, 그것을 상쇄할 장치도 없다. 생산성과 다양성은 AI 시대 과학정책의 두 바퀴인데, 한 바퀴만 손보면 다른 바퀴에서 대가를 치를 수 밖에 없다. 논문 안에서 LLM 사용이 급격히 늘고 있다는 측정³⁵은 이 압력이 이미 현실임을 보여 준다.

혁신은 작은 팀에서 나온다

큰 팀과 작은 팀은 역할이 다르다. 4,200만 편의 논문·특허·소프트웨어를 파괴성 지수(CD)로 분석한 연구¹²는 큰 팀일수록 기존 지식을 다지고(develop) 작은 팀일수록 판을 뒤엎는다(disrupt)는 규칙성을 드러냈고, 지난 반세기동안 지식 생산이 개인에서 점점 팀 중심으로 옮겨 온 흐름³⁶ 위에서도 이 대비는 견고하다. 여기에 영향력이 큰 논문을 분석해 보니 대부분의 통상적 조합에 파격적 조합을 살짝 엮는다는 분석 결과³⁷를 더하면, 상식을 뒤엎는 혁신은 연구팀의 규모가 아니라 이질적 결합에서 나온다는 것이 분명해진다.

공개된 제네시스 미션의 운영 방식은 Phase I (\$50만~75만 달러, 9개월)에서 육성한 소규모 팀 중 일부를 Phase II에서 3~5배 규모의 대형 팀으로 키운다³⁴. 초기 탐색은 작게, 검증된 방향은 크게라는 논리 자체는 합리적이다. 문제는 Phase I 착수 후 6개월 되는 시점에 계속 갈지, 중단할지(go/no-go) 결정을 내려 Phase II 진출 자격을 가른다는 점이다. 계획한 실험 등을 마치려면 최대 3개월 연장을 신청할 수 있어 실질 심사 시점은 Phase I이 종료되는 9개월까지 밀릴 수 있지만, 이마저도 너무 짧아 승자독식 구조가 다양한 소규모 시도의 폭을 지나치게 빨리 좁힌다는 데 문제가 있다. 보고서 역시 "오늘날 가장

중요한 문제들은 학계 연구실 하나가 다루기에는 너무 크다"면서 대형 신규 조직(FRO-ARPA)을 처방하고 있다³⁸. 여러 면에서 연구팀의 규모를 키우는 처방은 촘촘하지만 혁신을 유발할 수 있는 작은 팀을 지켜 주는 처방은 찾아보기 힘들다.

판을 뒤엎을 씨앗이 자라기도 전에 숙이면 정작 원하던 혁신의 원천이 마를 수 있다.

홀어지면 큰 발견이 어렵다

원격 근무가 자리 잡으면서 "어디서 일하든 상관없다"는 말이 흔해졌다. 그러나 논문 2천만 편과 특히 4백만 건을 분석한 연구¹³는 구성원이 물리적으로 흩어진 팀일수록 판을 바꾸는 돌파구를 만들 확률이 낮다는 것을 증명했다. 창의적인 아이디어들이 충돌하며 진화하려면 같은 공간에서 물리적으로 부대끼는 시간이 중요하다는 뜻이다. 보고서의 제4장에도 "기술 리더십은 연구와 생산이 가까이 붙어 있는 곳에서 나온다(*Technological leadership emerges from places where research and production sit close together.*)"며 밀집한 지역 혁신 클러스터를 처방한다³⁸.

그런데 정작 플래그십 사업인 제네시스 미션은 "전국 국가연구소에 걸쳐(*across national laboratories*)" 슈퍼컴퓨터·AI 모델·과학 장비·데이터를 통합하는 분산 플랫폼을 전제한다^{38,34}. 제네시스는 흔히 '제2의 맨해튼 프로젝트'로 소개되고²⁴, 보고서 역시 맨해튼 프로젝트를 수천 명의 활동을 조정한 새 제도의 모범으로 든다³⁸. 그러나 원조 맨해튼 프로젝트는 우라늄 농축(오크리지)과 플루토늄 생산(헨퍼드) 같은 인프라는 여러 곳에 나뉘어도, 정작 폭탄을 설계한 과학적 돌파는 오픈하이머가 각지의 최고 두뇌를 뉴멕시코의 외딴 로스앨러모스 한곳에 불러 모아 밤낮으로 부대끼게 한 데서 나왔다. 인프라는 분산하되 발견은 집결시킨 셈이다. 제네시스처럼 컴퓨팅·데이터·장비를 전국에 나누는 것은 옳지만, 그것만으로 대면에서 솟는 돌파구까지 저절로 생기지는 않는다. 접근성과 확산에는 분명 유리하지만, 분산을 기본값으로 두면 그만큼 대면 연구에서 나오는 혁신이 줄어들 수 있다. 새로운 도전에는 연구자가 주제를 넘나드는 과정이 따르는데, 주제를 크게 바꿀 때 성과가 한동안 주춤한다는 '전환 페널티(*pivot penalty*)' 연구³⁹도 새로운 조합에는 물리적·인지적 근접이 필요함을 같은 방향에서 지적한다.

단편적인 기준의 함정

앞서 살펴본 항목들을 다른 관점에서 다시 살펴보자. 앞서서 AI가 "무엇을 고르는가"(주제와 팀)를 물었다면, 여기서는 "어떻게 측정하는가"(평가의 잣대)를 묻는다. 보고서는 한쪽에서 "5년 이상 지속되는 장기 과제를 늘리고, 첫해에 전액을 배정해 연구자에게 시간과 자원을 주자"고 권한다³⁸. 그런데 같은 전략의 실행체인 제네시스 미션은 6개월 go/no-go와 9개월 Phase I로, 짧은 기간에 눈에 보이는 속도 우위(AI advantage)를 증명하도록 설계돼 있다³⁴. 한 전략 안에서 장기 탐색과 단기 속도가 부딪힌다. 어떤 지표가 목표가 되는 순간 그 지표는 더 이상 좋은 척도가 아니게 된다는 굿하트의 법칙(*Goodhart's law*)은, 속도를 성패의 관문으로 세우면 연구 주제가 결과를 빨리 보여 줄 수 있는 문제로 쏠릴 것임을 예고한다.

무엇보다 속도를 진보와 동일시하는 것 자체가 위험하다. 다른 관점을 가지고 독립적으로 수행된 연구들이 같은 지점을 가리킨다. 첫째, R&D 생산성을 장기 추적한 연구⁴⁰는 동일 수준의 진보성을 달성하기 위해 연구가 심화될수록 더 많은 연구자가 필요해진다는, 곧 아이디어 탐색이 점점 어려워지는

현상을 보고했다. 물살을 빠르게 가르는 물체가 받는 저항이 더 큰 것처럼, 결과를 얻는 속도가 빨라질수록 난이도가 함께 상승한다는 뜻이다. 둘째, 분야가 커져 논문이 쏟아질수록 인용이 이미 잡은 소수의 대표 논문에 몰려 새 아이디어가 통념을 뒤집기 어려워진다는 분석²⁶은 새로운 지식을 많이, 빨리 쏟아내더라도 실질적인 지식의 경계선은 거의 정지하고 있을 수 있음을 경고하고 있다. 셋째, 과학자들의 연구 전략을 분석한 연구⁴¹는 안전하고 효율적인 길이 생산성 향상에는 유리해도 혁신에는 불리하다는 것을 밝혔다. 속도와 효율을 평가 기준으로 세우면 연구자는 과감한 시도를 하지 않는 것이 합리적인 선택이 된다.

여기에 앞 절에서 본 대규모 블라인드 실험³³을 이 각도에서 다시 놓아 보자. LLM이 내놓은 연구 아이디어는 사람의 아이디어보다 참신하다는 평은 받았지만 서로 엇비슷해 다양성이 부족했고, AI는 자기 아이디어의 좋고 나쁨조차 가려내지 못했다. 생성은 폭증하는데 선별이 부실하면, 자동으로 쏟아지는 아이디어를 속도만으로 채점하는 게이트는 빠른 진보가 아니라 빠르게 쌓이는 잡음을 낳는다. 이미 연구논문 데이터베이스에는 AI 쓰레기(slop)라 부를 수 있는 것들이 쌓이고 있다.

자율실험실: 속도만큼 검증도 키워야

자율실험실(self-driving lab)의 잠재력은 과장이 아니라 이미 성과로 나타나고 있다. 화학·소재 분야 자율실험실을 종합한 대형 리뷰⁴²는 설계-제작-측정-분석-학습의 폐루프가 실험 횟수를 크게 줄이며 사람이 놓치던 최적점을 찾아낸다는 것을 보여 준다. 전자 고분자 용액 공정을 스스로 최적화한 자율 플랫폼⁴³과 앞서 본 나노항체 설계 가상 실험실¹⁰까지, 동료심사를 통과한 실물 사례가 쌓이고 있다. 게다가 약점으로 지적되던 검증도 빠르게 개선되는 중이다. 재현을 대규모로 자동화하는 워크플로, 공용 벤치마크와 데이터·코드·프로토콜의 표준화가 자율실험실 연구의 기본 절차로 들어오고 있다⁴²⁹. 요컨대 이 분야는 실현 가능성을 묻는 단계를 넘어 "어떻게 신뢰를 담보하느냐"의 단계로 넘어왔다.

그러나 제네시스 미션에서 자율실험실을 국가 미션으로 운영하는 방식은 보완이 필요하다. 이 분야의 상징적 초기 성과도 순탄치만은 않았다. 17일 만에 신물질 41개를 합성했다고 발표한 A-Lab¹⁴은, 그 신물질의 약 3분의 2가 실은 이미 알려진 무질서 상(phase)일 가능성이 높고 자동 XRD 분석이 아직 신뢰하기 이르다는 재분석 반론을 받았다⁴⁴. GPT-4 기반 Coscientist¹⁵는 이중용도 관련 안전 우려로 전체 코드와 프롬프트를 공개하지 않아 독립 재현이 제약됐고, 자율 연구 시스템의 성능 주장을 독립적으로 재평가한 연구¹⁶도 홍보와 실제의 간극을 지적한다. 보고서 자신도 "결함 있는 지식 기반 위에서 작동하는 AI는 나쁜 과학을 고착시킬 뿐(AI operating on a flawed knowledge base will only entrench bad science)"이며 "생성의 비용은 기하급수적으로 싸졌지만 검증의 비용은 그렇지 않다(While the cost of generation has decreased exponentially, the cost of verification has not.)"고 못 박는다³⁸. 그런데 정작 실행체인 제네시스 미션은 속도(AI 우위)와 규모(Phase II 확장)를 앞세우면서도 그 검증을 성패의 관문에 같은 무게로 넣지는 않았다.

다행히 위에서 살펴보았듯 검증 도구가 성숙하고 있다. 재현 패키지 제출 여부와 독립 재현성을 처음부터 연구과제 심사평가 기준에 넣는 것은 이제 비용이 아니라 선택의 문제다. 검증을 속도와 같은 속도로 키우면 자율실험실은 제조·로봇에 강한 한국에 특히 큰 기회가 되고, 반대로 검증을 건너뛰어 채 규모만 키우면 잘못된 지식을 더 빠르게 쌓는 기계를 갖게 될 뿐이다.

외국인 연구자는 연구 성과를 높인다

보고서는 "미국이 오랫동안 외국인 학생을 대규모로 길러 왔고 느슨한 보안 아래 기술이 해외로 넘어가는 것을 방치했다(*For decades, the United States has funded large cohorts of foreign students and tolerated technology transfer abroad under lax security standards.*)"며, "2024년 기준 임시 비자 소지자가 미국 컴퓨터과학·수학 박사 졸업생의 약 절반을 차지한다(*As of 2024, temporary visa holders accounted for around half of U.S. doctoral graduates in computer science and mathematics with confirmed postgraduation plans*)"는 사실을 우려했다³⁸. 미국에서 훈련받은 인재의 일부는 본국으로 돌아가 경쟁국의 부상에 기여했고, 특히 중국의 과학·기술 도약에는 미국 유학을 거친 귀국 연구자(海歸 또는 海龜, 하이구이)의 몫이 컸다. 미국 입장에서 이 위험은 부정할 수 없는 현실이다.

그러나 그렇다고 외국인에게 문을 좁히는 것으로 대처하면 역효과가 난다. 열려 있던 문을 닫으면 원래 남았을 인재까지 밀려나 경쟁국으로 돌아가고, 지정학적 긴장이 높아질수록 해외에서 온 과학자들은 협력을 접고 고국으로 떠나는 쪽을 택하기 마련이다⁴⁵. 애초에 미국이 오래 이득을 본 이유는 유학 온 인재의 상당수가 그대로 남아 그곳에서 성과를 냈기 때문이다. 당장 오늘의 미국 과학을 세운 힘의 상당 부분도 2차대전 시기 독일과 이탈리아의 배척에서 왔다. 히틀러의 나치가 유대인 학자를 대학에서 내쫓자 아인슈타인, 폰 노이만, 베테, 실라르드를 비롯한 유럽의 두뇌가 미국으로 대거 건너갔고, 이들은 미국의 물리학과 맨해튼 프로젝트, 훗날의 컴퓨터 과학을 결정적으로 끌어올렸다. 한 나라가 문을 닫자 다른 나라가 그 인재를 고스란히 얻었고, 그 유입이 실제로 미국의 발명을 늘렸음은 뒤에 계량 연구로도 확인됐다⁴⁶. 그러니 관건은 외국인 연구자를 들이느냐 막느냐가 아니라, 들여서 남게 하느냐다. 물론 개별 기술·인물의 안보 위험을 가려내는 표적 심사는 반드시 필요하다. 그러나 그런 선별과 문 자체를 좁히는 빗장은 전혀 다른 일이며, 역사는 후자가 붙잡아야 할 인재를 경쟁국에 넘겨주는 자충수임을 거듭 보여 준다. 인구가 줄고 국제 협력이 사활적인 한국에는 이 교훈이 특히 무겁다.

실제로 미국에 거주하는 외국인 연구자는 내국인 못지않은, 때로는 더 나은 성과를 보여 왔다. 첫째, 이민 인재는 제 몫 이상을 한다. 1990~2016년 미국 발명가 약 88만 명을 분석한 연구¹⁷에서 이민자는 발명가의 16%에 불과했지만 특허의 23%를 저술했고, 이민 협력자가 이른 나이에 세상을 떠나면 함께 일하던 미국 출생 동료의 생산성까지 유의하게 떨어졌다. 자기 성과만이 아니라 옆 사람의 성과까지 끌어올린다는 뜻이다. 둘째, 다양성 자체가 성과를 높인다. 수백만 편 규모의 논문을 분석한 연구⁴⁷는 민족적으로 다양한 팀일수록 인용으로 측정한 영향력이 더 크다는 것을 보였다. 요컨대 문을 여는 것은 시혜가 아니라 성과 전략이다. 셋째, 그런데 정작 그 다양성은 제값을 받지 못한다. 과학자 100만 명 이상을 추적한 '다양성-혁신 역설(*The Diversity-Innovation Paradox in Science*)' 연구⁴⁸는 과소대표 집단이 새로운 개념 조합을 더 많이 내놓지만 그 참신함이 학계에서 오히려 낮게 평가되고 경력으로 덮여버린다는 것을 밝혔다. 문을 닫는 것은 이미 저평가된 혁신의 원천을 마저 잘라 내는 일이다. 넷째, 빗장을 올릴 때의 대가도 측정된다. 국가 간 정치가 국제 협력의 평가에 끼어든다는 연구⁴⁹, 세계 과학의 협력·인용에 뿌리 깊은 지역 편향이 있다는 연구⁵⁰, 그리고 지정학적 긴장이 높아지자 미국 과학자들이 협력 상대와 주제를 바꾸며 이탈·회피로 대응하더라는 분석⁴⁵은, 장벽이 위험만 막는 것이 아니라 정상적 협력과 인재까지 함께 밀어낸다는 것을 보여 준다.

제비뽑기

보고서는 개별 심사자가 합의를 거스르고 한 건을 밀어 줄 수 있는 골든 티켓과 며칠 만에 결정하는 패스트 그랜트는 받아들였다³⁸. 심사자 합의가 무난한 제안으로 풀린다는 문제를 인정한 셈이다. 그런데 합격선 부근에서 지원 여부를 제비뽑기로 정하는 부분 추첨은 담지 않았다. 다른 처방은 받아들이면서 이것만 뺀 것은 아쉽다. 보고서의 각주는 이미 미국의 과학 추첨 실험(NSF)을 인용하고 있고³⁸, 스위스·뉴질랜드·독일 등 일부 연구재단은 우열을 가리기 어려운 구간에 실제로 추첨을 도입해 운영 중이다.

우리는 연구과제를 선정할 때 심사 점수를 공정한 잣대로 인지하고 이를 절대적인 기준으로 받아들이고 있다. 합격선에서 멀리 떨어진 확실한 합격과 불합격 지원자들에게는 사실일 수 있겠으나 합격선 부근에서는 이 믿음이 깨진다. 평가에서 상위 점수를 받은 제안서들의 점수 차이는 심사자 간 편차, 곧 잡음의 크기보다 작다. 그럼에도 정해진 예산으로 인해 어딘가에 합격과 불합격을 결정하는 선을 그어야 하는데, 이 인위적인 선이 놓이는 위치는 실력보다 편향이 결정한다는 데 문제가 있다. 우열이 사실상 동률일 때 심사자는 무의식적으로 본질과 무관한 대리 신호에 기대게 된다. 이름값과 소속의 후광(이미 유명한 사람과 명문 기관이 유리해지는 매튜 효과⁵¹), 안전한 쪽으로 기우는 보수성(참신하고 위험한 제안일수록 심사자 사이 평가가 갈려 밀려난다⁴⁸), 그리고 나와 비슷한 부류를 선호하는 동류 선호가 그것이다. 요컨대 커트라인의 줄 세우기는 실력의 미세한 차이를 채는 일이 아니라, 짚 수 없는 차이를 편향으로 메우는 일이다.

이 임의성이 사소하다면 넘어갈 수도 있지만, 대가는 크고 오래간다. 종이 한 장 차이로 합격선 위아래에 놓인, 사실상 똑같은 지원자들이 이후 8년간 받는 연구비가 두 배 넘게 벌어진다는 분석⁵¹, 그리고 초기의 아슬아슬한 탈락 하나가 이후 경력 전체의 궤적을 바꾼다는 연구⁵²가 이를 보여 준다. 잡음에 가까운 한 번의 판정이 '실력의 판결'로 포장되어 평생 복리로 불어나는 것이다. 여기에 숨은 비용도 있다. 사실상 동전 던지기인 경계선의 결정을 위해 수많은 연구자가 정교한 제안서를 쓰고 또 심사하며, 이 제안서 경쟁은 그 자체로 무난하고 심사자 취향에 맞는 연구를 부추긴다.

부분 추첨은 이 모든 것에 대한 현실적이고도 유일한 해법이다. 명백한 합격과 명백한 탈락은 지금처럼 심사로 가리되, 우열을 가릴 수 없는 경계의 무리만 대상으로 미세 서열이 허구임을 인정하는 것이다. 이렇게 하면 경계의 편향이 사라지고, 헛된 줄세우기에 드는 노력이 줄며, 우연을 실력으로 포장하는 일도 멈춘다. 실제로 이를 '추첨을 앞세운 연구비'로 제도화하자는 제안이 이미 나와 있고⁵³, 개별 심사자가 숨은 보석을 건져 올리는 골든 티켓과도 서로를 보완한다. 평가 방식이 연구자의 진로 자체를 바꾼다는 증거⁵⁴, 학제 간 연구가 오히려 분과형 연구비의 지원을 받을 때 더 깊고 넓은 영향을 남긴다는 분석⁵⁵은, 평가 설계가 곧 연구 방향의 설계임을 일깨운다. 덧붙여 보고서가 "평가는 오직 실력에 근거해야 하며 그때그때의 정치적 유행을 따라서는 안 된다(...ensure that selection rests purely on merit, not the political fashions of the day.)"³⁸고 못 박은 대목은 특정 형평 정책에 대한 미국적 가치 판단이 담겨 있다.

이 일곱 가지는 현재의 보고서에 누락되어 있거나 상반된 정책에 대한 학술적 근거를 제공하며, 한 연구진이 '과학 이후(After science)'라 부른 국면⁵⁶, 곧 AI가 발견의 방식 자체를 바꾸는 시대에는 의미있는 과학적 발견을 촉진할 근거가 될 것이다.

제 3 장 생각할 거리

미국의 처방을 그대로 들여올 수는 없다. 우리에게는 규제를 실험할 주정부도 없고, DOE 국가연구소 같은 거대한 공공 연구망도 없으며, 민간 R&D는 소수 대기업에 몰려 있어 미국식 벤처 생태계와도 다르다. 대신 우리에게는 출연연이라는 공공 연구 자산과, 미국이 오히려 부러워하는 제조업 기반이 있다. 그리고 한국은 이미 미국 전략과 비슷한 방향으로 움직이는 중이다. 외부 과제를 따라 인건비를 버는 PBS를 단계적으로 없애고 기관 출연금으로 인건비를 대기로 했고¹⁸, AI로 12대 국가 난제를 푸는 K-문샷을 띄우며 민간 전문가(PD)에게 기획과 예산 권한을 줬다^{19,20}. KIST는 임무 중심 연구소로 조직을 개편했으며²¹, 출연연들은 공공기관에서 해제되어 칸막이를 낮춘 협력단을 운영하기 시작했고²², 전임 연구원을 늘리고 기초연구 AI 센터를 세우는 계획도 나왔다²³.

무엇보다 2026년, 출연연 총괄기관인 국가과학기술연구회(NST) 산하에 국가과학AI연구센터(NAIS)가 세워져 독립을 준비하고 있다²⁴. AI for Science에 필요한 데이터·GPU·모델·표준을 한곳에 모으는 허브(Hub)를 두고, 각 출연연이 자기 분야의 문제를 푸는 스포크(Spoke)와 짝을 이루는 구조다. 미국이 제네시스 미션으로 하려는 일을, 우리는 우리 방식의 허브·스포크로 이미 시작한 셈이다. 다만 우리의 조건은 미국과 다르다. 인구는 줄고, 안보와 국제 협력이 놓인 자리도 다르며, 대학·출연연·기업의 관계도 미국과 같지 않다.

이제 우리는 주변 환경을 둘러보며 어떻게 행동할지 스스로 생각해 볼 차례다.

참고문헌

- [1] Park, M., Leahey, E., Funk, R. J. (2023). Papers and patents are becoming less disruptive over time. *Nature*, 613(7942), 138–144. <https://doi.org/10.1038/s41586-022-05543-x>
- [2] Azoulay, P., Fons-Rosen, C., Graff Zivin, J. S. (2019). Does Science Advance One Funeral at a Time? *American Economic Review*, 109(8), 2889–2920. <https://doi.org/10.1257/aer.20161574>
- [3] Open Science Collaboration. (2015). Estimating the Reproducibility of Psychological Science. *Science*, 349(6251), aac4716. <https://doi.org/10.1126/science.aac4716>
- [4] Ioannidis, J. P. A. (2005). Why Most Published Research Findings Are False. *PLoS Medicine*, 2(8), e124. <https://doi.org/10.1371/journal.pmed.0020124>
- [5] Li, D. (2017). Expertise versus Bias in Evaluation: Evidence from the NIH. *American Economic Journal: Applied Economics*, 9(2), 60–92. <https://doi.org/10.1257/app.20150421>
- [6] Azoulay, P., Graff Zivin, J. S., Manso, G. (2011). Incentives and Creativity: Evidence from the Academic Life Sciences. *RAND Journal of Economics*, 42(3), 527–554. <https://doi.org/10.1111/j.1756-2171.2011.00140.x>
- [7] Sourati, J., Evans, J. A. (2023). Accelerating science with human-aware artificial intelligence. *Nature Human Behaviour*, 7(10), 1682–1696. <https://doi.org/10.1038/s41562-023-01648-z>
- [8] Gao, J., Wang, D. (2024). Quantifying the use and potential benefits of artificial intelligence in scientific research. *Nature Human Behaviour*, 8(12), 2281–2292. <https://doi.org/10.1038/s41562-024-02020-5>
- [9] Musslick, S., Bartlett, L. K., Chandramouli, S. H., Dubova, M., Gobet, F., et al. (2025). Automating the practice of science: Opportunities, challenges, and implications. *Proceedings of the National Academy of Sciences*, 122. <https://doi.org/10.1073/pnas.2401238121>
- [10] Swanson, K., Wu, W., Bulaong, N. L., Pak, J. E., Zou, J. (2025). The Virtual Lab of AI agents designs new SARS-CoV-2 nanobodies. *Nature*. <https://doi.org/10.1038/s41586-025-09442-9>
- [11] Hao, Q., Xu, F., Li, Y., Evans, J. (2025). Artificial intelligence tools expand scientists' impact but contract science's focus. *Nature*, 649(8099), 1237–1243. <https://doi.org/10.1038/s41586-025-09922-y>
- [12] Wu, L., Wang, D., Evans, J. A. (2019). Large teams develop and small teams disrupt science and technology. *Nature*, 566(7744), 378–382. <https://doi.org/10.1038/s41586-019-0941-9>
- [13] Lin, Y., Frey, C. B., Wu, L. (2023). Remote collaboration fuses fewer breakthrough ideas. *Nature*, 623(7989), 987–991. <https://doi.org/10.1038/s41586-023-06767-1>
- [14] Szymanski, N. J., Rendy, B., Fei, Y., Kumar, R. E., He, T., et al. (2023). An autonomous laboratory for the accelerated synthesis of inorganic materials. *Nature*, 624(7990), 86–91. <https://doi.org/10.1038/s41586-023-06734-w>
- [15] Boiko, D. A., MacKnight, R., Kline, B., Gomes, G. (2023). Autonomous chemical research with large language models. *Nature*, 624(7992), 570–578. <https://doi.org/10.1038/s41586-023-06792-0>
- [16] Beel, J., Kan, M.-Y., Baumgart, M. (2025). Evaluating Sakana's AI Scientist for Autonomous Research: Wishful Thinking or an Emerging Reality Towards 'Artificial Research Intelligence' (ARI)? *arXiv preprint arXiv:2502.14297*. <https://doi.org/10.48550/arXiv.2502.14297>
- [17] Bernstein, S., Diamond, R., Jiranaphawiboon, A., McQuade, T., Pousada, B. (2022). The Contribution of High-Skilled Immigrants to Innovation in the United States. *NBER Working Paper No. 30797*. <https://doi.org/10.3386/w30797>
- [18] 과학기술정보통신부, 「과학기술분야 출연(연) 정책방향」 (제2회 과학기술관계장관회의), 2025. 12. 18. — PBS 단계적 폐지(경제·인문사회 2026 즉시, 과학기술 5년)·기관 출연금 인건비 전액 지원·임무중심 전환. <https://www.msit.go.kr/bbs/view.do?sCode=user&mPid=208&mId=309&bbsSeqNo=86&nttSeqNo=3180072>
- [19] 과학기술정보통신부, 「AI 시대 과학기술 경쟁력 대도약을 위한 K-문샷 추진전략」, 2026. 2.; K-문샷 프로젝트 개시 및 PD 위촉, 2026. 5. — 2035년 12대 국가 미션·2030년 연구 생산성 2배·PD 중심 추진체계·'AI 과학자' 미션 포함. (대한민국 정책브리핑) <https://www.korea.kr/news/policyNewsView.do?newsId=148965259>
- [20] 과학기술정보통신부, 「K-문샷 프로그램 운영·관리 규정」 (훈령) 제정·시행, 2026. 6. 30. — 프로그램 디렉터(PD) 권한·책임 명문화(미션 설정·과제 기획·예산 우선배분).

- [21] 한국과학기술연구원(KIST), 임무중심연구소 체제 재편 및 프로젝트매니저(PM) 제도 도입, 2024. 7. — 차세대반도체·AI·로봇·기후환경·정정수소융합·뇌과학 등 여섯 개 임무중심연구소(각 연구소는 다수 연구단으로 구성), PM(연구소장급)의 예산·평가 권한. (KIST 공식 조직도) <https://www.kist.re.kr/ko/intro/organization-chart.do>
- [22] 과학기술정보통신부·국가과학기술연구회, 「2025년도 글로벌 TOP 전략연구단」 선정(국가전략형·미래도전형, 출연연 칸막이를 낮춘 임무 중심 협력), 2025. 5.; 과기 출연연 22개 공공기관 지정 해제(2024). (KDI 경제정책자료) <https://eiec.kdi.re.kr/policy/materialView.do?num=252223>
- [23] 과학기술정보통신부, 「기초연구 생태계 육성 방안」(제2회 과학기술관계장관회의), 2025. 12. 18. — 대학 불록펀딩을 통한 전임연구원(Staff-Scientist)·연구지원 인력 확충, 도메인별 「기초연구 AI 센터」 40개 구축·기초연구·AI 융합 인력 2,000명 양성.
- [24] 국가과학기술연구회(NST) 국가과학AI연구센터(NAIS) 운영단(운영단장 유용균), 「AI Agent 기반 연구시스템과 국가과학AI연구센터의 역할」, 2026. — NST 산하 과학 AI 허브. Hub&Spoke 모델(NAIS=데이터·GPU·모델·표준 허브 / 출연연=도메인 실행 스포크), AI-Ready 데이터 표준·K-허깅페이스·과학AI-OS·AI 과학자 표준(재현 패키지·신뢰성/안전성 평가·AI Peer Review)·전국 자율실험실 On-site 허브, 자율성 5단계 현실 진단, 2026년 10대 목표(AI 파이오니어 1,000명·1인 10 에이전트 등). KIST는 과학AI 정책 담당.
- [25] Wang, D., Barabási, A.-L. (2021). *The Science of Science*. Cambridge University Press. <https://doi.org/10.1017/9781108610834>
- [26] Chu, J. S. G., Evans, J. A. (2021). Slowed canonical progress in large fields of science. *Proceedings of the National Academy of Sciences*, 118(41), e2021636118. <https://doi.org/10.1073/pnas.2021636118>
- [27] Park, M., Leahey, E., Funk, R. J. (2026). Robust Evidence for Declining Disruptiveness: Assessing the Role of Zero-Backward-Citation Works. *arXiv preprint arXiv:2503.00184*. <https://doi.org/10.48550/arXiv.2503.00184>
- [28] Azoulay, P., Graff Zivin, J. S., Wang, J. (2010). Superstar Extinction. *Quarterly Journal of Economics*, 125(2), 549–589. <https://doi.org/10.1162/qjec.2010.125.2.549>
- [29] Camerer, C. F., Dreber, A., Holzmeister, F., Ho, T.-H., Huber, J., Johannesson, M., et al. (2018). Evaluating the Replicability of Social Science Experiments in Nature and Science between 2010 and 2015. *Nature Human Behaviour*, 2(9), 637–644. <https://doi.org/10.1038/s41562-018-0399-z>
- [30] Fanelli, D., Costas, R., Ioannidis, J. P. A. (2017). Meta-assessment of bias in science. *Proceedings of the National Academy of Sciences*, 114(14), 3714–3719. <https://doi.org/10.1073/pnas.1618569114>
- [31] Wang, B., Ma, Z., Wang, D., Uzzi, B. (2025). Peer Review and the Diffusion of Ideas. *arXiv preprint arXiv:2507.11825*. <https://arxiv.org/abs/2507.11825>
- [32] Azoulay, P., Graff-Zivin, J., Uzzi, B., Wang, D., Williams, H., Evans, J. A., et al. (2018). Toward a more scientific science. *Science*, 361(6408), 1194–1197. <https://doi.org/10.1126/science.aav2484>
- [33] Si, C., Yang, D., Hashimoto, T. (2024). Can LLMs Generate Novel Research Ideas? A Large-Scale Human Study with 100+ NLP Researchers. *arXiv preprint arXiv:2409.04109* (ICLR 2025). <https://doi.org/10.48550/arXiv.2409.04109>
- [34] U.S. Department of Energy. (2026). *The Genesis Mission: Transforming Science and Energy with AI* — Request for Application DE-FOA-0003612 (Amendment 000003); Executive Order 14363.
- [35] Liang, W., Zhang, Y., Wu, Z., Lepp, H., Ji, W., Zhao, X., et al. (2025). Quantifying large language model usage in scientific papers. *Nature Human Behaviour*, 9(12), 2599–2609. <https://doi.org/10.1038/s41562-025-02273-8>
- [36] Wuchty, S., Jones, B. F., Uzzi, B. (2007). The Increasing Dominance of Teams in Production of Knowledge. *Science*, 316(5827), 1036–1039. <https://doi.org/10.1126/science.1136099>
- [37] Uzzi, B., Mukherjee, S., Stringer, M., Jones, B. (2013). Atypical Combinations and Scientific Impact. *Science*, 342(6157), 468–472. <https://doi.org/10.1126/science.1240474>
- [38] The White House, Office of Science and Technology Policy (Kratsios, M.). (2026). *Science: A New Golden Age — A Report to the President*, July 2026.
- [39] Hill, R., Yin, Y., Stein, C., Wang, X., Wang, D., Jones, B. F. (2025). The pivot penalty in research. *Nature*, 642(8069), 999–1006. <https://doi.org/10.1038/s41586-025-09048-1>
- [40] Bloom, N., Jones, C. I., Van Reenen, J., Webb, M. (2020). Are Ideas Getting Harder to Find? *American Economic Review*, 110(4), 1104–1144. <https://doi.org/10.1257/aer.20180338>

- [41] Foster, J. G., Rzhetsky, A., Evans, J. A. (2015). Tradition and Innovation in Scientists' Research Strategies. *American Sociological Review*, 80(5), 875–908. <https://doi.org/10.1177/0003122415601618>
- [42] Tom, G., Schmid, S. P., Baird, S. G., Cao, Y., Darvish, K., et al. (2024). Self-Driving Laboratories for Chemistry and Materials Science. *Chemical Reviews*. <https://doi.org/10.1021/acs.chemrev.4c00055>
- [43] Wang, C., Kim, Y.-J., Vriza, A., Batra, R., Baskaran, A., et al. (2025). Autonomous platform for solution processing of electronic polymers. *Nature Communications*, 16. <https://doi.org/10.1038/s41467-024-55655-3>
- [44] Leeman, J., Liu, Y., Stiles, J., Lee, S. B., Bhatt, P., Schoop, L. M., et al. (2024). Challenges in High-Throughput Inorganic Materials Prediction and Autonomous Synthesis. *PRX Energy*, 3(1), 011002. <https://doi.org/10.1103/PRXEnergy.3.011002>
- [45] Li, M., Ma, Y., Wang, Y., Wang, D. (2026). Pivoting as an Adaptive Strategy to Geopolitical Tensions in U.S. Science. *작업논문(preprint)*.
- [46] Moser, P., Voena, A., Waldinger, F. (2014). German Jewish Émigrés and US Invention. *American Economic Review*, 104(10), 3222–3255. <https://doi.org/10.1257/aer.104.10.3222>
- [47] AlShebli, B. K., Rahwan, T., Woon, W. L. (2018). The Preeminence of Ethnic Diversity in Scientific Collaboration. *Nature Communications*, 9, 5163. <https://doi.org/10.1038/s41467-018-07634-8>
- [48] Hofstra, B., Kulkarni, V. V., Munoz-Najar Galvez, S., He, B., Jurafsky, D., McFarland, D. A. (2020). The Diversity–Innovation Paradox in Science. *Proceedings of the National Academy of Sciences*, 117(17), 9284–9291. <https://doi.org/10.1073/pnas.1915378117>
- [49] Furnas, A. C., Jia, R., Roberts, M. E., Wang, D. (2026). Geopolitics in the Evaluation of International Scientific Collaboration. *NBER Working Paper* No. 34789. <https://doi.org/10.3386/w34789>
- [50] Wu, L., Huang, Y., Lu, W., Saxena, A., Traag, V. (2026). Persistent geographical biases in global scientific collaboration and citations. *arXiv preprint arXiv:2604.01602*. <https://doi.org/10.48550/arXiv.2604.01602>
- [51] Bol, T., de Vaan, M., van de Rijdt, A. (2018). The Matthew effect in science funding. *Proceedings of the National Academy of Sciences*, 115(19), 4887–4890. <https://doi.org/10.1073/pnas.1719557115>
- [52] Wang, Y., Jones, B. F., Wang, D. (2019). Early-career setback and future career impact. *Nature Communications*, 10, 4331. <https://doi.org/10.1038/s41467-019-12189-3>
- [53] Luebber, F., Krach, S., Martinez Mateo, M., Paulus, F. M., Rademacher, L. (2023). Rethink Funding by Putting the Lottery First. *Nature Human Behaviour*, 7. <https://doi.org/10.1038/s41562-023-01649-y>
- [54] Tripodi, G., Zheng, X., Qian, Y., Murray, D., Jones, B. F., Ni, C., et al. (2025). Tenure and research trajectories. *Proceedings of the National Academy of Sciences*, 122(30), e2500322122. <https://doi.org/10.1073/pnas.2500322122>
- [55] Park, M., Maity, S., Wuchty, S., Wang, D. (2026). Interdisciplinary papers supported by disciplinary grants garner deep and broad scientific impact. *PNAS Nexus*, 5(3). <https://doi.org/10.1093/pnasnexus/pgag057>
- [56] Evans, J., Duede, E. (2025). After science. *Science*, 390(6774), eaec7650. <https://doi.org/10.1126/science.aec7650>

유의사항 Caveats

- 「Science: A New Golden Age」와 제네시스 미션 문서는 미 행정부의 정책·정치 문서로, 규범적 주장을 포함한다. 본 보완의견은 그 정치적 프레이밍을 지지하거나 검증하지 않으며, 과학·AI4S 증거와의 정합·불일치에 한정해 검토한다.
- 일부 발견에는 활발한 학술 반론이 있다. 특히 연구 참신함의 감소[1]를 재는 CD 지수에는 원저자의 재확인[27]에도 불구하고 측정 편향 논쟁이 있고, 심리학 재현성 연구[3]의 재현율에도 통계적 반박이 존재한다. 본 문서는 진단 방향의 견고함과 개별 수치의 잠정성을 구분한다.
- 자율실험실 성과(A-Lab 등)[14][15]는 발표 직후 재현성 문제가 제기되었으므로 확정된 성취로 보기 어렵다. 본 문서는 이를 성취가 아니라 검증 과제의 사례로 인용한다.
- 제네시스 미션 공모(DE-FOA-0003612)의 마감·예산·단계 세부는 개정(Amendment)마다 변동되며, 본 문서는 Amendment 000003(2026. 4. 20.) 기준이다.
- 인용 문헌의 다수는 정식 게재본과 프리프린트 사이에 수치가 조정될 수 있다. 인용 시 최신본 대조를 권고한다.

이제현 · 2026년 7월 31 일 · 미국 백악관 「Science: A New Golden Age」와 DOE 제네시스 미션을 읽고 쓴 개인 보완의견.