

Trillion Parameter Consortium 2025 리뷰

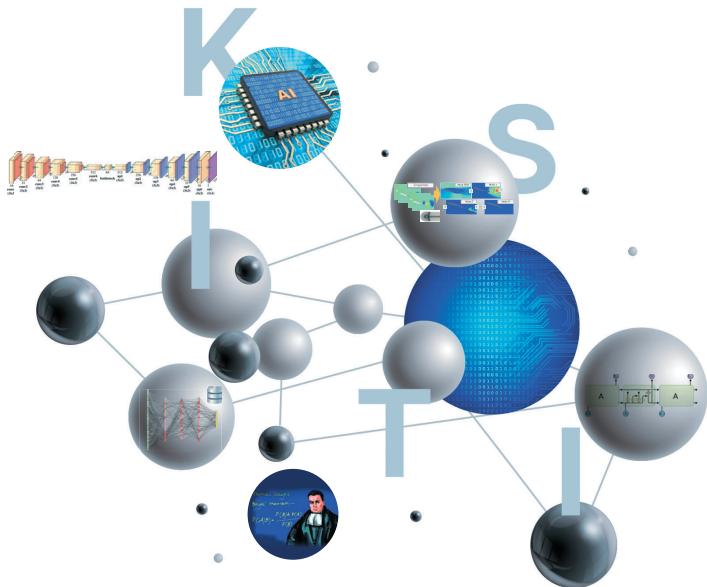
: 비전, 역사, 그리고 글로벌 Scientific AI 리더십

작성: 신정훈, 김장훈 / KISTI 오픈XR플랫폼융합연구단
초안: GPT-5 Pro

KISTI
Trend
Report

KISTI Trend Report

TPC25 Review Report



A Review on Trillion Parameter Consortium 2025

Junghun Shin¹, Janghun Kim¹

¹Korea Institute of Science and Technology Information, Daehak-ro, Yuseong-gu, Daejeon 34141

Abstract

The Trillion Parameter Consortium (TPC) is a global alliance of leading research institutions, national laboratories, and industry partners dedicated to advancing AI for scientific discovery. Established in 2023, TPC aims to build and govern shared infrastructure, models, and data frameworks that enable large-scale, trustworthy, and reproducible AI-driven science. The TPC25 Conference, held in San Jose, California, in July 2025, marked a major milestone in this mission. It brought together global experts in high-performance computing (HPC), machine learning, and domain science to explore how artificial intelligence can transform the methodology of science itself. The conference emphasized a shift from AI as a computational tool toward AI as an autonomous scientific agent capable of hypothesis generation, simulation control, and experimental validation. In the opening plenary, Rick Stevens (Argonne National Laboratory) introduced a vision of “AI-native discovery,” in which large-scale reasoning models assist scientists in formulating and testing hypotheses. Thierry Pellegrino (AWS) discussed the convergence of cloud and HPC systems as the foundation of exascale AI infrastructure for scalable science. Flora Salim (University of New South Wales) presented advances in modeling cyber-physical-social systems using multimodal, physics-informed, and prompt-based learning architectures. Subsequent plenaries and panel discussions addressed key frontiers including multimodal data integration, evaluation of non-LLM model architectures, and the design of agentic systems for science. Participants also examined the ethical and governance challenges of AI-driven research, focusing on data provenance, model interpretability, and energy sustainability. TPC25 concluded with three strategic priorities for the coming decade: (1) integration of exascale AI infrastructure across global HPC and cloud systems, (2) development of physics-informed and operator-learning models for data-physics fusion, and (3) establishment of trustworthy and autonomous AI governance frameworks. Through these initiatives, TPC is evolving beyond a research consortium into a global governance platform for AI-native science, laying the groundwork for transparent, reproducible, and sustainable scientific discovery.

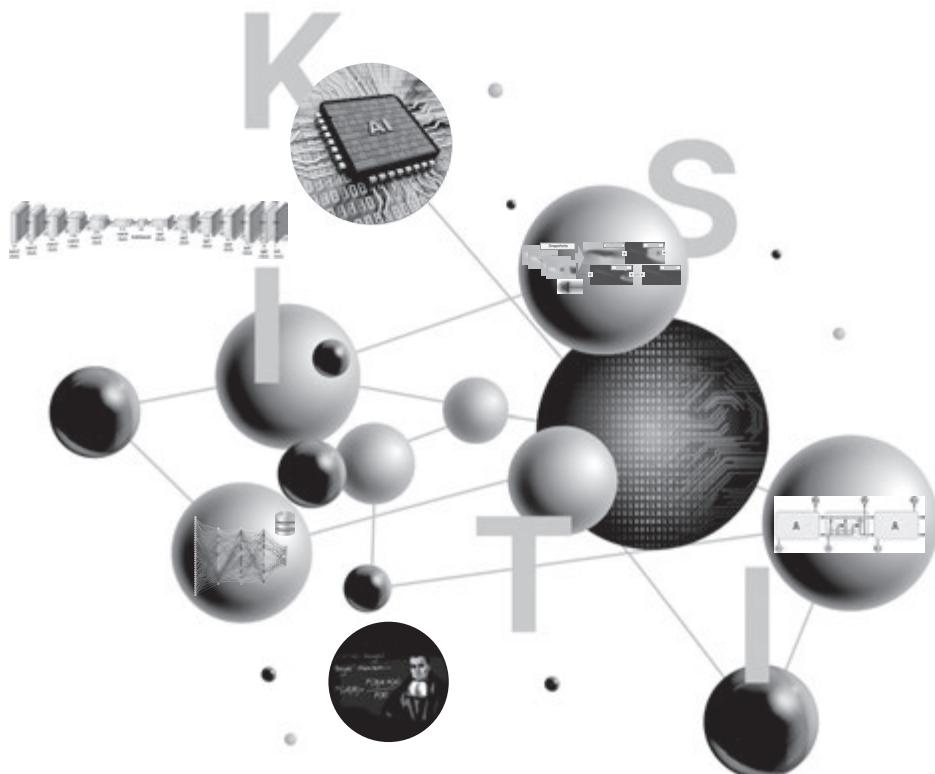
Acknowledgment

This work was supported by the National Research Council of Science & Technology (NST) grant by the Korea government (MSIT) (No. CRC21014)

Trillion Parameter Consortium 2025 리뷰

: 비전, 역사, 그리고 글로벌 Scientific AI 리더십

작성: 신정훈, 김장훈 / KISTI 오픈XR플랫폼융합연구단
초안: GPT-5 Pro



CONTENTS

01	서 문 Introduction	2
02	조직 및 운영 구조 Organization & Operation	5
03	연혁 및 연례행사 History & Annual Conferences	11
04	TPC25 행사 개요 TPC25 Overview	13
05	핵심 주제 및 기술 초점 Key Themes & Technical Focus	17
06	TPC25 프로그램 Program: Plenaries, Panels, Tutorials & Breakout Tracks	20
07	주요 발표내용 발췌 Talks	33
08	매체 하이라이트 Media Highlights	78
09	협업과 거버넌스 Collaboration & Governance	80
10	도전 과제와 윤리 Challenges & Ethics	82
11	전망 및 결론 Perspectives & Conclusions	86
	참고문헌 References	88
	부 록 Appendix	91

"TPC25는 단순한 학술행사가 아니라, 'AI와 과학의 융합'을 실습·협업·표준화로 구현한 첫 번째 글로벌 실험장이었다."

(Alwire, 2025-08-06, 「TPC25 Highlights AI's Expanding Role」)



<출처 URL - <https://tpc25.org/what-is-tpc/> >

1. 서문

Trillion Parameter Consortium(TPC)은 초거대 AI 모델의 과학적 활용 필요성이 대두되던 2023년 전후에 설립되었다. AI 기술이 단순한 언어 생성 단계를 넘어 복잡한 과학 문제 해결, 고성능 시뮬레이션, 약물 개발, 물리 모델링 등으로 확장되면서, 다양한 기관이 협업 체계를 구축할 필요가 커졌고 그 결과 TPC가 출범했다. 이러한 컨소시엄의 필요성은 대규모 AI 모델, 특히 대형 언어 모델(LLM)과 같은 생성형 모델의 급속한 발전과 과학·공학 커뮤니티가 이를 활용해 새로운 통찰을 얻고 연구를 가속화하며 복잡한 문제를 보다 효율적으로 탐구하려는 열망이 커지고 있다는 점에서 비롯된다. TPC는 설립 후 18개월 동안 구성원이 150명에서 1,200명 이상으로 늘어나는 과정에서 여러 작업 그룹을 구성해 상호 협력했고, 다음과 같은 협력 분야를 도출했다: 데이터 자원의 구축·공유, 대형 모델의 공동 학습, 과학적 역량에 대한 모델 평가 방법·도구 개발, 그리고 공통 접근(나아가 표준화)을 통한 진전 가속 영역의 식별 등이다. 특히 수천억~수조 개 파라미터를 갖는 생성형 모델(대형 언어 모델, LLM 포함)의 발전은 천문학, 입자물리학, 재료과학, 생명과학, 의료, 기후 모델링 등에서 복잡한 데이터 패턴을 효과적으로 포착하고 새로운 과학적 통찰을 제공

할 잠재력을 보여준다. 이러한 모델의 개발과 활용은 본질적으로 학제 간 접근을 요구한다. 머신러닝, 수학, 대규모 분산 컴퓨팅, 확장 가능한 소프트웨어 프레임워크, 데이터 전처리·큐레이션, 슈퍼컴퓨터 기반 성능 최적화, 그리고 모델의 신뢰성·안전성·편향·공정성·재현성·해석 가능성에 대한 정량 평가 등 다양한 전문 영역의 협력이 필수적이다. 특히 과학적 역량에 대한 모델 평가 방법론과 도구 개발은 다학제 전문가 집단의 참여를 통해 이루어져야 한다.

TPC는 개방성, 공정성, 책임성, 그리고 책임 있는 개발을 핵심 가치로 삼아 AI의 발전이 과학적 진보뿐 아니라 사회적 이익에도 기여하도록 한다. 산업 주도의 대형 모델 개발은 뛰어난 성능에도 불구하고 학습 데이터, 알고리즘, 모델 파라미터 등에서 투명성이 부족한 경우가 많다. 이에 과학 공동체는 재현성과 검증 가능성을 중시하며, 공개 데이터셋과 오픈소스 도구, 명확히 문서화된 방법론, 평가 프레임워크의 공동 개발을 요구한다. 또한 윤리 원칙, 공정성, 개인정보 보호, 지적재산권, 규제 준수와 정합적인 책임 있는 모델 개발이 중요한 과제로 인식된다.

한편, 수조 개 파라미터를 갖는 모델의 학습에는 막대한 계산 자원이 필요 하며 이를 독자적으로 수행할 수 있는 기관은 제한적이다. 따라서 첨단 컴퓨팅 자원의 공동 활용, 데이터셋 공유, 소프트웨어 도구의 공동 개발을 통해 과학 분야에서 고급 AI 기술에 대한 접근성을 민주화하는 것이 중요하다.

매년 개최되는 TPC 행사는 당해의 최신 기술 트렌드를 반영하며 점차 규모와 영향력을 키워 왔다. 오픈사이언스, 공유 인프라, 거버넌스를 학계·국립연구소·산업계 전반에 통합하고, 연례 회의를 통해 워킹그룹과 기관 간 프로젝트의 연속성을 제공하며 공통 벤치마크·평가 프레임워크로 연구 중복을 줄이기 위해 노력해 왔다. TPC25는 Trillion Parameter Consortium이 2025년에 주최한 연례 행사로, 수조(Trillion) 파라미터 규모의 대형 AI 모델을 과학 연구 및 공학적 문제 해결에 활용하기 위한 글로벌 협력의 장으로 많은 관심을 모았다. 본 행사는 미국 캘리포니아주 산호세(San Jose)에서 열렸으며, 학계·정부 연구소·산업계의 다양한 전문가가 한자리에 모여 AI의 최신 기술 발전과 적용 가능성을 논의했다. 수조 파라미터 모델은 과학적 발견의 새로운 시대를 여는 촉매가 되고 있다.

Trillion Parameter Consortium은 연구소, 대학, 산업계, 지원 기관을 정렬하여 과학을 위한 책임 있는 AI 발전을 가속화할 것으로 기대한다. TPC25는 본회의 (Plenary), 튜토리얼, 패널, 해커톤, 30개 이상의 브레이크아웃 세션으로 구성되어 수백 명이 참여한 대규모 행사였다. 본 보고서는 TPC의 목적, 역사, 기술 주제, 글로벌 임팩트를 종합적으로 정리한다. 핵심 주제로 엑사스케일 AI 인프라, 물리정보·연산자 학습, 과학용 에이전트 시스템, 평가·신뢰성, 그리고 재료·지구·생명·핵융합·소프트웨어 등 응용 도메인에서의 적용을 다룬다. 2025년의 핵심 목표 중 하나는 산업계의 참여와 통찰 확대이며, 이를 위해 집중 작업 그룹 또는 산업 협의체 구성 방안이 검토되고 있다.

TPC Has Three Overarching Objectives for AI in Science and Engineering:

The graphic consists of three overlapping circles in light blue, teal, and dark blue. Each circle contains a number and a brief description of an objective. Below the circles is a statement about the community's goals and values.

1. Building an Open Community	2. Identifying, Incubating, and Facilitating	3. Creating a Global Network of Expertise and Resources
-----------------------------------------	--------------------------------------------------------	-------------------------------------------------------------------

The TPC community pursues these objectives in alignment with scientific and government guidelines, with a commitment to open collaborations that accelerate the development of scientific AI capabilities and are characterized by:

- Transparency
- Fairness
- Ethical AI practices
- Mitigating bias
- Ensuring trustworthiness
- Protecting privacy
- Ensuring safety

< 출처 URL - <https://tpc25.org/what-is-tpc/> >

2. 조직 및 운영 구조

TPC는 초거대 AI 기술의 윤리적이고 신뢰할 수 있는 개발을 지향하면서, 글로벌 기술 주도권 확보에도 중요한 역할을 한다. 미국·유럽·일본 등 주요 국가의 연구소·대학·정부기관이 참여하며, TPC는 과학기술 분야의 AI 경쟁력을 높이고 국제 표준 수립의 기반이 되고 있다.

TPC LEADERSHIP

TPC Executive Committee

TPC is guided by an executive committee, a technical steering group, and a planning team – all comprising leaders from around the world. The executive committee was formed by TPC founders from North America (Rick Stevens, Argonne National Laboratory and The University of Chicago, USA), Asia-Pacific (Satoshi Matsuoka, RIKEN, Japan), and Europe (Mateo Valero, Barcelona Supercomputing Center, Spain).

An additional 22 leaders serve on the planning team and technical steering group, which are led by regional co-chairs. The planning team co-chairs are Charlie Catlett (USA), Fabrizio Gagliardi (Spain), and Kyoung-Sook Kim (Japan), while the technical steering group co-chairs are Ian Foster (USA), Laura Morselli (Italy), and Rio Yokota (Japan).

TPC Executive Committee	TPC Planning Team Co-Chairs	TPC Technical Steering Group Co-Chairs

TPC PLANNING TEAM

TPC Executive Committee	TPC Planning Team Co-Chairs	TPC Technical Steering Group Co-Chairs
Thierry Bidot (Inria, France)	Kyoungsook Kim (AIST, Japan)	Javier Aula-Blasco (BSC, Spain)
Jerome Bobin (CEA, France)	Taiji Makoto (RIKEN, Japan)	Franck Cappello (Argonne, USA)
Charlie Catlett (Argonne National Laboratory and The University of Chicago, USA)	Per Oster (CSC, Finland)	Miguel Vazquez (BSC, Spain)
Karthik Duraisamy (University of Michigan, USA)	Noah Smith (University of Washington and Allen Institute for AI, USA)	Arvind Ramanathan (Argonne, USA)
Nick Jones (NESI, New Zealand)	Samantika Sury (HPE, USA)	Mohamed Wahib (RIKEN, Japan)
Fabrizio Gagliardi (Barcelona Supercomputing Center, Spain)	Valerie Taylor (Argonne National Laboratory and The University of Chicago, USA)	Rio Yokota (IS Tokyo, Japan)
Nicolay Hammer (Leibniz Supercomputing Centre, Germany)		

TPC TECHNICAL STEERING GROUP

TPC Executive Committee	TPC Planning Team Co-Chairs	TPC Technical Steering Group Co-Chairs
Thierry Bidot (Inria, France)	Kyoungsook Kim (AIST, Japan)	Javier Aula-Blasco (BSC, Spain)
Jerome Bobin (CEA, France)	Taiji Makoto (RIKEN, Japan)	Franck Cappello (Argonne, USA)
Charlie Catlett (Argonne National Laboratory and The University of Chicago, USA)	Per Oster (CSC, Finland)	Miguel Vazquez (BSC, Spain)
Karthik Duraisamy (University of Michigan, USA)	Noah Smith (University of Washington and Allen Institute for AI, USA)	Arvind Ramanathan (Argonne, USA)
Nick Jones (NESI, New Zealand)	Samantika Sury (HPE, USA)	Mohamed Wahib (RIKEN, Japan)
Fabrizio Gagliardi (Barcelona Supercomputing Center, Spain)	Valerie Taylor (Argonne National Laboratory and The University of Chicago, USA)	Rio Yokota (IS Tokyo, Japan)
Nicolay Hammer (Leibniz Supercomputing Centre, Germany)		

TUTORIAL CHAIRS FOR TPC25
Neeraj Kumar (PNNL), Samantika Sury (HPE), Laura Morselli (CINECA), and Prasanna Balaprakash (ORNL)

HACKATHON CHAIRS FOR TPC25
Miguel Vazquez (BSC), Arvind Ramanathan (ANL), and Mohamed Wahib (RIKEN)

< 출처 URL - <https://tpc25.org/what-is-tpc/> >

Trillion Parameter Consortium(TPC)은 다층적 조직 구조를 갖추고 있다. 그 중심에는 집행위원회(TPC Executive Committee, TPCEC)가 있으며, 미국 아르곤 국립 연구소(Argonne National Laboratory), 스페인 바르셀로나 슈퍼컴퓨팅센터(Barcelona Supercomputing Center), 일본 이화학연구소(RIKEN) 등 세 창립 기관의 기관장으로

구성된다. Rick Stevens, Satoshi Matsuoka, Mateo Valero 등 해당 기관 대표들은 TPC의 전략적 방향을 제시하고 고위 정책을 수립하며, 조직 구성원에 대한 감독 역할을 수행한다.

TPC의 중장기 전략 수립과 운영 계획은 기획·전략팀(TPC Planning and Strategy Team, TPCPT)이 담당한다. 이 팀은 TPC가 국제 과학 공동체에 가장 효과적으로 기여할 방안을 제안하며, 연간 일정 관리, 신규 작업그룹 제안, 컨퍼런스·해커톤·튜토리얼 등 각종 행사 기획 및 운영을 지원한다. TPCPT는 다양한 기관의 전문가로 구성되어 국제적 대표성을 갖춘다.

기술적 방향성과 워킹그룹 운영을 관장하는 조직은 기술운영그룹(TPC Technical Steering Group, TPCSG)이다. 이 그룹은 중복 노력을 최소화하고 시너지 가능성을 식별하며, 해커톤·워크숍 등 기술 중심 행사를 기획한다. 또한 기술 목표를 설정하고 진행 상황을 평가하며, 컨소시엄의 기술 산출물에 대한 지속적 개선을 촉진한다.

TPC의 거버넌스 구조는 유연성과 지속적 혁신을 염두에 두고 설계되었다. 초기 리더십 임기는 통상 12~18개월로 설정되며, 이후 새 구성원의 참여를 통해 리더십 순환, 새로운 아이디어와 관점의 도입, 차세대 리더에 대한 멘토링 기회가 제공된다. 이러한 순환 구조는 역동적이고 포용적인 리더십 환경을 조성하는 데 기여한다.

TPC EXECUTIVE COMMITTEE (TPCEC)

역할: TPC를 공동 창립한 핵심 기관 대표들로 구성되며, 전반적 방향 설정과 조직 회원 프로그램을 감독한다.

이름	소속 기관	국가
Rick Stevens	Argonne National Laboratory; The University of Chicago	미국
Satoshi Matsuoka	RIKEN; Tokyo Institute of Technology	일본
Mateo Valero	Barcelona Supercomputing Center	스페인

TPC PLANNING AND STRATEGY TEAM (TPCPT)

역할: 일정·전략, 방향·행사, 기획 등 전반 운영을 담당한다. 세계 주요 연구기관·슈퍼 컴퓨팅 센터·대학의 대표들로 구성된다.

이름	소속 기관	국가
Thierry Bidot	INRIA; GENCI / CEA	프랑스
Jerome Bobin	CEA	프랑스
Karthik Duraisamy	University of Michigan	미국
Fabrizio Gagliardi	Barcelona Supercomputing Center	스페인
Nikolay Hammer	Leibniz Supercomputing Centre	독일
Nick Jones	New Zealand eScience Infrastructure	뉴질랜드
Per Oster	CSC – IT Center for Science	핀란드
Kyoung-Sook Kim	AIST	일본
Noah Smith	University of Washington / Allen Institute for AI	미국
Samantika Sury	Hewlett Packard Enterprise	미국
Makoto Taiji	RIKEN	일본
Valerie Taylor	Argonne National Laboratory, The University of Chicago	미국

TPC TECHNICAL STEERING GROUP (TPCSG)

역할: 기술 방향 조율, 프로젝트 간 중복 방지, 워킹그룹 지원을 담당한다. 초기 12~18개월 임기로 활동하며 이후 운영 방식을 정립할 예정이다.

이름	소속 기관	국가
Javier Aula-Blasco	Barcelona Supercomputing Center	스페인
Franck Cappello	Argonne National Laboratory	미국
Jens Domke	RIKEN	일본
Ian Foster	Argonne National Laboratory, The University of Chicago	미국
Neeraj Kumar	Pacific Northwest National Laboratory	미국
Laura Morselli	CINECA	이탈리아

Arvind Ramanathan	Argonne National Laboratory, The University of Chicago	미국
Miguel Vazquez	Barcelona Supercomputing Center	스페인
Mohamed Wahib	RIKEN	일본
Rio Yokota	Institute of Science Tokyo(IST)	일본

2023년 8월 미국 아르곤 국립연구소 워크숍에서 처음 결성된 여러 기술 워킹그룹(Technical working group)은, 2024년 스페인 바르셀로나 슈퍼컴퓨팅 센터에서 열린 컨퍼런스와 이후 해커톤을 거치며 연구 분야를 더욱 구체화했다.

모델 아키텍처 및 성능 평가 MODEL ARCHITECTURE AND PERFORMANCE EVALUATION

대규모 AI 모델을 위한 확장 가능한 아키텍처와 성능 벤치마크를 개발·비교하여, 최신 슈퍼컴퓨터와 고급 하드웨어 플랫폼에서 학습과 추론이 효율적으로 수행되도록 보장한다.

데이터, 교육 워크플로 및 대규모 언어 모델 전략 DATA, TRAINING WORKFLOWS, AND LARGE LANGUAGE MODEL STRATEGIES

과학적 AI 프로젝트를 위한 데이터 수집·정리·포맷팅·큐레이션을 다루고, 대규모 병렬화와 분산 최적화, 도메인별 대규모 언어 모델 적응을 처리할 수 있는 강력한 훈련 워크플로를 개발한다.

기술, 안전 및 신뢰 평가 SKILLS, SAFETY, AND TRUST EVALUATION

모델의 신뢰성·안전성을 평가하는 방법을 확립한다. 편향 감소, 공정성 검증, 해석 가능성 향상, 윤리 지침 준수, 과학적 과제를 위한 생성 모델의 신뢰성 제고 방안을 모색한다.

생물학, 생화학 및 생물정보학 BIOLOGY, BIOCHEMISTRY, AND BIOINFORMATICS

단백질 구조 예측, 약물 발견, 유전체학, 의료 영상 분석 등 생물학·생의학 응용을 대상으로 하며, 대규모 AI 접근과 잘 통합되는 새로운 데이터셋과 기법을 정식화한다.

아웃리치 및 교육 OUTREACH AND TRAINING

광범위한 커뮤니티에 초거대 매개변수 모델을 개발·활용하는 데 필요한 방법과 도구를 교육하기 위해, 교육 콘텐츠·튜토리얼·여름학교를 기획·제공한다.

과학 소프트웨어를 위한 대규모 언어 모델 LARGE LANGUAGE MODELS FOR SCIENTIFIC SOFTWARE

대규모 언어 모델이 과학 소프트웨어 생태계의 코딩·디버깅·최적화·유지관리에서 어떻게 기여할 수 있는지 조사한다.

의료용 대규모 언어 모델 LARGE LANGUAGE MODELS FOR HEALTHCARE

데이터 개인정보 보호, 규정 준수, 환자 안전을 보장하면서, 의료 영상 분석·임상 의사결정 지원·헬스 인포매틱스에 어떻게 기여할 수 있는지를 탐구한다.

환경, 생태 및 기후 ENVIRONMENT, ECOLOGY, AND CLIMATE

대규모 AI 모델을 기후 시뮬레이션, 환경 데이터 분석, 생태 모델링에 적용하여 장기 기후 패턴, 생태계 역학, 지속가능성 지표에 대한 통찰을 제공한다.

공학, 에너지 및 기계 ENGINEERING, ENERGY, AND MECHANICS

AI 를 활용해 유체역학, 재료공학, 기계설계, 에너지 시스템 최적화 등에서 현상을 모델링·예측한다.

기초 물리학 및 우주론 FUNDAMENTAL PHYSICS AND COSMOLOGY

생성형 AI 모델을 천체물리학, 입자물리학, 우주론, 양자 시뮬레이션에 적용해 이론 개선, 매개변수 공간 탐색, 대규모 관측 데이터 해석을 지원한다.

재료, 화학 및 나노 과학 MATERIALS, CHEMISTRY, AND NANOSCIENCE

AI 모델을 활용해 신소재 발견, 화학적 특성 예측, 분자 동역학 탐구, 합리적 소재·약물 설계를 가속화한다.

Trillion Parameter Consortium은 실질적 협력과 실행을 중심으로 운영된다.

이를 위해 다양한 형태의 활동을 전개하고 있으며, 주요 예시는 다음과 같다.

해커톤 및 집중형 멋업(HACKATHONS AND FOCUSED MEETUPS)

연중 다양한 해커톤을 기획해 참가자들이 온·오프라인으로 모여 특정 목표를 중심으로 집중 협업하도록 한다. 주요 목표에는 학습용 데이터셋 준비·정제, 모델 아키텍처 개선, 평가 프레임워크 고도화, 소프트웨어 파이프라인 공동 개발 등이 포함된다.

워크숍, 학술대회 및 세미나(WORKSHOPS, CONFERENCES, AND SEMINARS)

SC(International Conference for High Performance Computing, Networking, Storage and Analysis), ISC-HPC, SupercomputingAsia(SCA) 등 주요 국제 학술대회에서 정기 워크숍을 개최한다. 또한 기술 발전·도전 과제·경험 공유를 위한 독립 컨퍼런스를 주최하며, 격주 온라인 세미나를 통해 세계적 전문가의 최신 연구 방향과 기술적 돌파구에 대한 통찰을 제공한다.

교육 및 역량 강화(TRAINING OPPORTUNITIES AND SKILL BUILDING):

교육을 핵심 가치로 삼아 튜토리얼을 제공하고 관련 작업그룹이 활동 중이다. 2025년에는 여름학교 및 단기 강좌로의 확장을 계획하고 있으며, 참가자는 최신 기술을 학습하고 첨단 도구를 실험하며 멘토·숙련 실무자와 교류할 수 있다.

자원 공유 및 배분(RESOURCE SHARING AND ALLOCATION):

고성능 슈퍼컴퓨터, 차세대 가속기 기술, 대규모 정제 데이터 저장소 등 개별 기관이 독자적으로 확보하기 어려운 전문 자원에 대한 접근을 가능케 하는 협력 모델을 모색한다. 이를 통해 공동 연구팀이 첨단 인프라를 활용할 수 있는 기반을 마련한다.

TPC는 2025년부터 연간 약 4회의 대면 행사를 개최할 예정이며, 각 행사에는 해커톤을 포함하고 연 1회 대규모 커뮤니티 컨퍼런스를 연다. 이와 더불어 주요 학술대회에서 소규모 워크숍과 세션을 지속 운영한다. 개최지는 아시아·태평양, 유럽, 북미를 순환하며 전 세계적 참여를 독려한다.

3. 연혁 및 연례행사

TPC는 출범 이후 매년 컨퍼런스를 통해 기술 교류, 정책 정렬, 성과 공유를 이어왔다. 2025년 미국 산호세에서 열린 TPC25는 참여 폭을 산업계까지 확장하고, 추론 중심 발견 플랫폼과 과학 추론형 서비스 등 새로운 초점을 제시했다. 출범 후 약 16개월 동안 TPC는 다음과 같은 이벤트를 통해 거대 AI 모델 전략을 조율하고 리소스를 공유하며, 강력한 도구·방법 생태계를 구축하려는 역동적 다대륙 커뮤니티로 자리매김했다. 모임 기간 동안 워킹그룹을 대상으로 한 교육 세션, 주제별 워크숍, 전체 회의, 분과 회의를 통해 조직과 운영 구조를 지속적으로 정비해 나갔다.

2023년 8월 아르곤 국립 연구소, 미국

- 약 73 명 대면, 117 명 원격 참가
- 참가자들이 초기 워킹그룹을 자발적으로 구성하여 TPC 워킹그룹 구조의 토대 마련

2024년 2월 SupercomputingAsia, 시드니, 호주

- 약 100 명의 TPC 참가자가 컨퍼런스 참석
- 초거대 AI 튜토리얼, 관련 세션, 하이브리드 워킹그룹 토론 진행
- 협력 관계가 강화 및 초기 계획 구체화

2024년 5월 국제 슈퍼컴퓨팅 컨퍼런스, 독일 함부르크

- 약 80 명이 참가한 TPC 워크숍에서 글로벌 AI 프로젝트 강조
- 새로운 기회와 과제를 논의하는 포럼 제공으로 유럽 내 가시성 확대

2024년 6월 바르셀로나 슈퍼컴퓨팅 센터, 바르셀로나, 스페인

- 16 개국에서 190 명 이상 참가
- 전체회의, 12 개 워킹그룹 분과회의, 'AI for Science' 포괄적 튜토리얼 진행
- 커뮤니티 추진력 강화, 구체적 협력 목표 설정 및 실무 교육

Trillion Parameter Consortium 2025 (TPC25) 리뷰
: 비전, 역사, 그리고 글로벌 Scientific AI 리더십

2024년 10월 미국 아르곤 국립연구소 및 시카고 대학교

- 세 워킹그룹(모델 아키텍처·성능, 데이터·훈련, 안전·신뢰) 공동 해커톤 개최
- 50 명 이상 참여

2024년 11월 슈퍼컴퓨팅 2024, 미국 애틀랜타

- 약 200 명 참가, 다기관 협력 프로젝트와 커뮤니티 활동 현황 소개
- 확대 초록 공모 33 편 중 12 편 발표 선정
- 유럽 TPC 리더들이 유럽연합 집행위원회 신규 기금 지원 논의를 위해 회동

Trillion Parameter Consortium 2025 (TPC25) 리뷰
: 비전, 역사, 그리고 글로벌 Scientific AI 리더십

4. TPC25 행사 개요

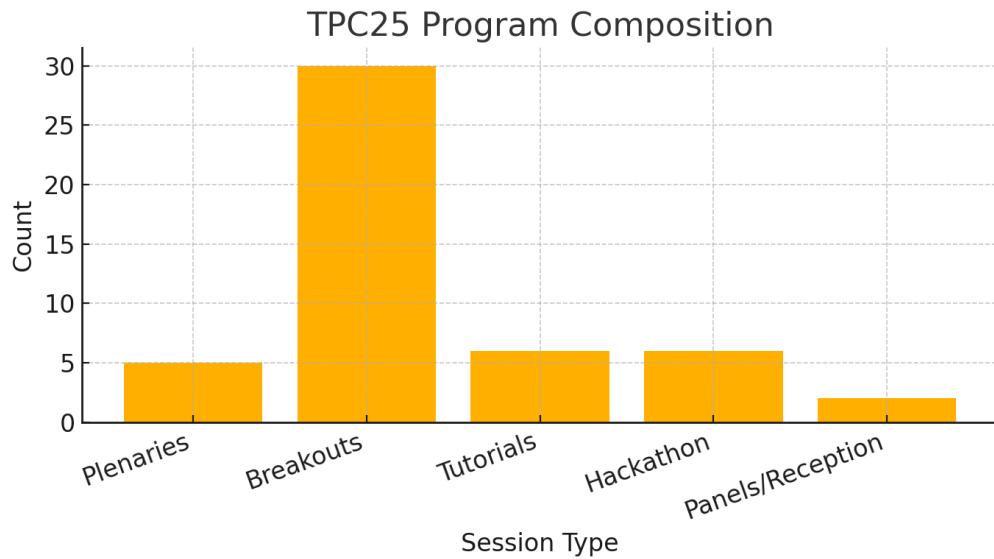


행사 정보

- 일정: 2025년 7월 28-31 일
- 장소: 더블트리 바이 힐튼 산호세 (미국 캘리포니아주)
- 참가 규모: 14 개국 100 개+ 기관 소속 도메인 과학자, AI 연구자, 정책 입안자, 산업계 리더 등 약 300 명
- 주요 프로그램: 튜토리얼·해커톤, 5 개 본회의, 2 개 런치 패널, 워킹그룹 중심 병렬 브레이크아웃

주요 프로그램

Day	Key Items
Mon, Jul 28	Hackathon/Tutorial Opening Plenary; Tutorials & Hackathon Sessions; Networking Gathering
Tue, Jul 29	Hackathon & Tutorials (AM); Opening Plenary; Plenary 2; Google Welcome Reception
Wed, Jul 30	Plenary 3; Plenary 4; Lunch Panel; Breakouts A & B
Thu, Jul 31	Breakouts C & D; Lunch Panel; Breakouts E; Closing Plenary



< TPC25 세션 유형별 모임 개수 >

참가 기관 PARTICIPATING ORGANIZATIONS

(Supercomputer Top 500 system을 붉은색으로 표시 → Top 10 중 8기)

Americas

1. Allen Institute For AI
2. Argonne National Laboratory (#3 Aurora)
3. Brookhaven National Laboratory
4. Caltech
5. Deep Forest Sciences
6. Fermilab National Accelerator Laboratory
7. Harvard University
8. Indiana University
9. Lawrence Berkeley National Laboratory
10. Lawrence Livermore National Laboratory (#1 El Capitan -- #10 Tuolumne)
11. Los Alamos National Laboratory
12. National Center for Supercomputing Applications
13. National Energy Technology Laboratory
14. National Renewable Energy Laboratory
15. Northwestern University
16. Oak Ridge National Laboratory (#2 Frontier)
17. Pacific Northwest National Laboratory
18. Princeton Plasma Physics Laboratory

Trillion Parameter Consortium 2025 (TPC25) 리뷰
: 비전, 역사, 그리고 글로벌 Scientific AI 리더십

19. Princeton University
20. Purdue University
21. Rutgers University
22. Sandia National Laboratories
23. San Diego Supercomputer Ctr
24. SLAC National Accelerator Laboratory
25. Stanford University
26. Stony Brook University
27. Texas Advanced Computing Center
28. Thomas Jefferson National Accelerator Facility
29. Université de Montréal
30. University of Arizona
31. University of Buffalo
32. University of California San Diego/San Diego Supercomputer Center
33. University of Chicago
34. University of Delaware
35. University of Illinois Chicago
36. University of Illinois Urbana-Champaign
37. University of Michigan
38. USC / Information Sciences Institute
39. University of Toronto – Acceleration Consortium
40. University of Utah
41. University of Virginia
42. University of Washington

Vendors (American)

1. AMD
2. AWS
3. Cerebras Systems
4. Groq
5. HPE
6. Intel
7. Microsoft (# 4 Eagle, Azure cloud)
8. Nvidia
9. SambaNova
10. Together AI

EMEA

Trillion Parameter Consortium 2025 (TPC25) 리뷰
: 비전, 역사, 그리고 글로벌 Scientific AI 리더십

1. Barcelona Supercomputing Center
2. ETH Zürich
3. CEA
4. CINECA (#9 Leonardo)
5. CSC – IT Center for Science (#8 LUMI)
6. CSCS, Swiss National Supercomputing Centre (#7 Alps)
7. INESC TEC
8. INRIA
9. Juelich Supercomputing Center
10. Leibniz Supercomputing Centre
11. LAION, Large-Scale AI Open Network
12. Max Planck Computing & Data Facility (MPCDF)
13. STFC Rutherford Appleton Laboratory, UKRI
14. SURF.nl

Asia Pacific

1. AI Singapore
2. A*Star
3. Australian National University
4. CSIRO, Commonwealth Scientific and Industrial Research Organization
5. Flinders University
6. Institute of Science Tokyo (formerly Tokyo Tech)
7. National Institute of Advanced Industrial Science and Technology (AIST)
8. National Supercomputing Centre, Singapore
9. NCI Australia
10. New Zealand eScience Infrastructure
11. Pawsey Institute
12. RIKEN (#6 Fugaku)
13. Seoul National University
14. University of New South Wales
15. University of Tokyo

Vendors (Asia Pacific)

1. Fujitsu Limited
2. Kotoba Technologies, Inc.
3. Sony Research

5. 핵심 주제 및 기술 초점

5-1. 엑사스케일 AI 인프라

키워드: 샤딩, 파이프라인/텐서 병렬화, IO 오케스트레이션, 하이브리드 클라우드-HPC 연계

엑사스케일급 AI 인프라는 1 조 개 이상 파라미터를 갖는 대형 모델을 효율적으로 훈련·실행하기 위한 기반 기술을 의미하며, 다음 요소로 구성된다.

- **모델 샤딩(Model Sharding):** 초기대 모델을 여러 GPU/노드에 분할 배치해 병렬로 학습하는 기술. 예) ZeRO, Megatron-LM, DeepSpeed-Ulysses
 - **파이프라인 및 텐서 병렬화:** 모델 내부 연산을 시간·공간 축에서 병렬화해 메모리 효율과 처리량을 동시에 높인다.
 - **I/O 오케스트레이션:** 시뮬레이션·관측·실험 등 대용량 과학 데이터의 입출력을 최적화해 데이터 병목(data bottleneck)을 최소화한다.
 - **하이브리드 클라우드-HPC 연계:** 온프레미스 슈퍼컴퓨터(Fugaku, Aurora 등)와 상용 클라우드(AWS, Azure) 자원을 유기적으로 통합한다.
- ◆ **목표:** 초기대 모델을 단순히 “큰 GPU 클러스터”에서 운영하는 수준을 넘어, 과학 계산과 AI 훈련을 통합 운영하는 범지구적 컴퓨팅 패브릭(fabric)을 구축한다.

5-2. 물리정보/연산자 학습

키워드: PINN, FNO, Neural ODE/FDE 등 과학 선형지식 통합

이 영역은 과학적 선형지식(물리 법칙, 제약·경계 조건 등)을 AI 학습 구조에 직접 통합하는 연구 방향이다.

- **PINN (Physics-Informed Neural Networks):** 미분방정식의 잔차(residual)를 손실 함수에 반영해 물리적으로 일관된 예측을 수행한다.
- **FNO (Fourier Neural Operator):** 연산자(operator) 자체를 신경망이 학습하도록 하여 시간·공간 해석을 효율화한다.

- **Neural ODE/FDE:** 상미분방정식(ODE) 또는 분수미분방정식(FDE)을 신경망의 동역학으로 표현해 연속 시간 예측이나 비선형 진화계 모사에 활용한다.
- ◆ **목표:** 데이터 중심 접근을 넘어 “데이터 + 방정식 + 제약”의 결합을 통해 물리적 신뢰성을 갖춘 AI 모델을 구현하는 것. 이 접근은 기류·난류, 재료 미세구조, 기후 시뮬레이션 등에서 특히 강력한 성능을 보이고 있다.

5-3. 과학용 에이전트 시스템

키워드: 검색(Search), 시뮬레이션(Simulation), 실험 제어(Experiment Control), 검증(Validation) — 계획·실행 오케스트레이션

이 분야는 TPC25에서 가장 혁신적으로 주목받은 주제 중 하나로, AI가 단순 계산을 넘어 **과학적 탐구를 주도하는 ‘자율 탐구 에이전트(Agentic System)’**로 진화하는 것을 의미한다.

- **계획(Planning):** 연구 목표를 설정하고 실험·시뮬레이션 일정을 자동 생성
- **실행(Execution):** HPC·로봇 실험장비·시뮬레이터를 제어하여 실험 수행
- **분석 및 검증(Analysis & Validation):** 결과를 요약·평가·시각화
- **자기반복(Self-Refinement):** 실패나 불확실성 정보를 반영해 가설 재구성

이러한 시스템은 DeepMind의 AlphaDiscovery, Argonne의 Scientia, Stanford의 Boimni, Microsoft의 Agentic Workflow Stack 등으로 시연되었다

◆ **목표:** AI를 “도구(tool)”가 아닌 **과학적 협력자(co-investigator)**로 발전(데이터 탐색 → 모델링 → 실험 → 검증을 통합한 ‘**자율 과학 루프(Autonomous Scientific Loop)**’)

5-4. 평가 및 신뢰성

키워드: 불확실성 정량화(UQ), 개입/반사실 테스트(Interventional & Counterfactual Testing), 출처 추적(Provenance), 재현성(Reproducibility)

AI가 과학적 의사결정에 활용되려면 **결과의 신뢰성과 검증 가능성**이 필수적이다.

- **불확실성 정량화(UQ):** 모델 예측의 신뢰 구간을 통계적으로 추정한다.
 - **개입(Intervention)-반사실(Counterfactual) 테스트:** 실제 데이터 분포를 벗어난 시나리오에서 모델의 일반화 능력을 평가한다.
 - **출처 추적(Provenance Tracking):** 데이터·코드·모델 파라미터의 계보(lineage)를 기록해 실험 재현성을 보장한다.
 - **평가 프레임워크:** TrustBench, RealityGap Tests, Causal Reasoning Benchmarks 등을 TPC 산하 워킹그룹에서 개발 중이다.
- ◆ **목표:** “정확한 AI”를 넘어 “설명 가능하고 재현 가능한 AI”, 즉 과학적 신뢰성(Scientific Trustworthiness) 확보

5-5. 응용 분야

키워드: 재료과학(Materials), 지구과학(Earth System), 생명과학(Life Sciences), 핵융합(Fusion), 소프트웨어공학(Software Engineering)

각 응용 분야에서 AI의 적용 예시

- **재료과학:** 신합금·촉매·반도체 탐색을 위한 생성형 모델 및 강화학습 기반 실험 설계
- **지구과학:** 기후·기상·지진 등 멀티스케일 예측 모델 구축(예: FourCastNet, AI4Climate)
- **생명과학:** 단백질 구조 예측, 세포 수준 시뮬레이션, 약물 발굴 자동화
- **핵융합:** 물리 시뮬레이션 가속화, 경계층 불안정성 모델링, 디지털 트윈
- **소프트웨어공학:** HPC 코드 자동 최적화, 모델 병렬화 자동 튜닝, AI 기반 코드 검증

과학 데이터의 복잡·다양성을 처리하기 위해 데이터 정제(data curation)와 멀티모달 데이터(multi-modal data)의 수용이 요구된다.

- ◆ **목표:** AI 를 단순 데이터 분석 도구가 아닌, 과학적 혁신을 가속하는 ‘인지적 인프라(Cognitive Infrastructure)’로 자리매김시키는 것.

6. TPC25 프로그램

6-1. 본회의·패널 (Plenaries & Panels)

행사 개요

- 일시: 7 월 29 일(화) ~ 31 일
- 장소: DoubleTree by Hilton San Jose 컨퍼런스룸 B

Plenaries는 인프라, 양자 AI, HPC 보안형 AI, 추론 중심 플랫폼을 다뤘으며 튜토리얼에는 AI-for-Science 파운데이션 모델의 추론 평가가 포함되었다. 환영 리셉션은 Google Inc.가 커뮤니티 빌딩 취지로 후원했다.

< 주요 본회의/패널 목록 및 핵심 주제 >

	제목	발표자(소속)	핵심 주제
Opening	Reinventing Discovery: Accelerating Science in the Age of Artificial Super-Intelligence	Rick Stevens (Argonne NL)	AI-Native Discovery Loop
Opening	HPC and Science: The Need for Hybrid	Thierry Pellegrino (AWS)	Hybrid HPC + Cloud
Opening	Modeling & Simulating Complex Behavior in Dynamic Cyber-Physical-Social Systems	Flora Salim (University of New South Wales)	Cyber-Physical-Social Modeling
Plenary 2	Scaling Science with Large AI Models and HPC Convergence	Satoshi Matsuoka (RIKEN R-CCS)	AI + HPC Convergence
Plenary 2	Enabling Scientific Discovery With Generative Quantum AI	Steve Clark (Quantinuum)	Quantum Generative AI
Plenary 2	Overview of AI for Technical Computing Workloads	Earl Joseph (Hyperion Research)	HPC Market Trends
Plenary 2	Secure AI Infrastructure for	Jens Domke	AI Security & Trust

	제목	발표자(소속)	핵심 주제
	AI for Scientific Computing	(RIKEN R-CCS)	
Plenary 3	Scaling Reasoning, Scaling Science: AI-Native Scientific Discovery Platform	Moderator: Ian Foster (Argonne NL)	Scaling Reasoning
Plenary 3	Agents, Autonomy, and Agency: A Brave New World	Preeth Chengappa (Microsoft)	Agentic AI Systems
Plenary 3	Automation of Biological Discovery with LLM Agents	Siddharth Narayanan (FutureHouse)	LLM Lab Automation
Plenary 3	Active Inference AI Systems for Scientific Discovery	Karthik Duraisamy (University of Michigan)	Active Inference AI
Panel 1	Industry, Academia, and Government Collaboration: Trustworthy AI for Science	Hal Finkel (DOE); Raj Hazra (Quantinuum); Pradeep Dubey (Intel); Molly Presley (Hammerspace); Karthik Duraisamy (University of Michigan, Mod.)	Trustworthy AI Governance
Plenary 4	Irresponsible AI and the Need for Transparent Evaluation	Ricardo Baeza-Yates (Barcelona Supercomputing Center)	프레임워크로 AI 모델의 편향·허위정보 문제 비판
Plenary 4	ORNL's AI Initiative: Advancing Secure, Assured, and Efficient AI for Scientific Discovery	Prasanna Balaprakash (Oak Ridge NL)	Oak Ridge AI Initiative (AAI) 프레임워크 설명
Plenary 4	OLMoTrace: Tracing Language-Model Outputs Back to Their Training Data in Real Time	Jiacheng Liu (University of Washington / Allen Institute for AI (AI2))	OLMoTrace 시스템 소개 - LLM 출력 토론을 실시간으로 원천 학습 데이터로 역추적
Plenary 4	Fairness of Geospatial Fou	Kyoung-Sook Kim	위치 및 지리공간 데

	제목	발표자(소속)	핵심 주제
	ndation Models	(AIST Japan)	이터를 학습한 모델의 편향 문제 분석
Panel 2	From Agents to Agency: Toward Collaborative Scientific Reasoning	Elahe Vedadi (Google DeepMind), Preeth Chembappa (Microsoft Discovery), Kexin Huang (Stanford University), Arvind Ramanathan (Argonne National Laboratory), Siddharth Narayanan, (FutureHouse), Addison Snell (Intersect360 Research, mod.)	Building Agentic Systems for Science: Reports From the Field
Plenary 5	AI for Science 2030 Vision	Rick Stevens (Argonne NL)	TPC 2025 성과 및 차기 로드맵 발표 "AI for Science 2030 Vision" 초안 공개
Plenary 5	Scaling Science with AI + HPC Convergence	RIKEN R-CCS (Satoshi Matsuoka)	AI + HPC 융합 프로젝트 (기후·재료 시뮬레이션) 중간 성과 보고
Plenary 5	European AI for Science and the Role of TPC	Mateo Valero (Barcelona Supercomputing Center)	유럽형 AI 인프라 (EuroHPC) 및 TPC 협력 계획 발표
Plenary 5	Open Science Data and TPC Data Hub Initiative	Thierry Bidot (INRIA)	오픈사이언스 데이터 공유 정책 및 TPC 데이터 허브 진행 상황 보고
Plenary 5	Scaling AI Collaboration Across Continents	Ian Foster (Argonne NL)	Plenary 총괄 토론: "Scaling AI Collaboration Across Continents"

1) Rick Stevens (Argonne National Laboratory)

Title: *Reinventing Discovery: Accelerating Science in the Age of Artificial Super-Intelligence*

Key Points:

- 초기대 AI 를 "과학적 질문을 함께 사유하는 존재"로 정의
- AI-Native Discovery Loop(실험 → 시뮬레이션 → 분석 → 지식화)의 필요성 제시
- 엑사스케일 컴퓨팅 + AI 융합이 핵심 전략이며, 투명성·신뢰성 확보가 중요
- ◆ 요약: AI 는 이제 '연구 도우미'가 아닌 '공동 탐구자'다.

2) Thierry Pellegrino (AWS)

Title: *HPC and Science: The Need for Hybrid*

Key Points:

- 초대규모 AI 를 위한 **Hybrid HPC + Cloud** 아키텍처의 필수성 강조
- 온프레미스 HPC 와 클라우드 간 워크로드 분산 전략 공유
- AWS 가 TPC 와 협력해 "AI for Science Reference Stack" 개발 중
- ◆ 요약: "엑사스케일 이후 시대에는 하이브리드가 표준이다."

3) Flora Salim (University of New South Wales)

Title: *Modeling Complex Behavior in Dynamic Cyber-Physical-Social Systems*

Key Points:

- 도시 에너지·교통·인간 행동 예측을 통합하는 복잡계 모델링 제시
- **Hybrid Graph Neural Networks + Temporal Attention** 응용
- 사회적 변수까지 통합하는 설명가능 AI 필요성 강조
- ◆ 요약: "AI 는 데이터 너머 '사회적 맥락'을 이해해야 한다."

4) Satoshi Matsuoka (RIKEN R-CCS)

Title: *Scaling Science with Large AI Models and HPC Convergence*

Key Points:

- RIKEN R-CCS 의 Fugaku 슈퍼컴퓨터 활용 AI 훈련 사례 소개
- 대형 LLM 을 기후·재료·유체 등 과학 시뮬레이션에 적용
- 멀티피직스 · 멀티스케일 문제 해결을 위한 **AI + HPC 융합 로드맵** 제시

Trillion Parameter Consortium 2025 (TPC25) 리뷰
: 비전, 역사, 그리고 글로벌 Scientific AI 리더십

- 2024년 시범 결과: 기후 시뮬레이션 성능 최대 $6.9\times$ 향상
- "학습 데이터 출처 투명성"과 인간 감독 체계의 중요성 강조
- ◆ 요약: "AI와 HPC의 결합은 과학 탐구의 '규모 한계'를 무너뜨린다."

5) Steve Clark (Quantinuum)

Title: *Enabling Scientific Discovery with Generative Quantum AI*

Key Points:

- 양자 컴퓨팅과 생성형 AI 결합으로 분자·재료 탐색 혁신 모색
- **Quantum Transformer Architecture**, Hybrid Variational Circuit 실험 공개
- ◆ 요약: "양자 AI는 탐색 공간의 차원을 확장한다."

6) Earl Joseph (Hyperion Research)

Title: *Overview of AI for Technical Computing Workloads*

Key Points:

- HPC 워크로드의 AI 통합률 급상승(42% 진입 예상)
- 2028년까지 AI for HPC 지출 연 18% 성장 전망
- TPC를 "산업·학계 연결 거버넌스 모델"로 평가
- ◆ 요약: "AI는 HPC 산업 패러다임의 핵심으로 이동 중이다."

7) Jens Domke (RIKEN R-CCS)

Title: *Secure AI Infrastructure for Scientific Computing*

Key Points:

- **AI-Fabric Security Framework** 소개 (암호화 + 서명 체계)
- "과학적 신뢰성(Scientific Trustworthiness)" 강조
- ◆ 요약: "AI 인프라는 성능보다 신뢰 위에 세워져야 한다."

8) Ian Foster (Argonne National Laboratory)

Session: *Scaling Reasoning, Scaling Science*

Key Points:

- 패널 모더레이터로서 "AI-Native Scientific Discovery Platform" 비전 논의
- 자율 탐구 루프와 협업 인지(collaborative cognition) 강조

9) Preeth Chengappa (Microsoft)

Title: *Agents, Autonomy, and Agency: A Brave New World*

Key Points:

- Azure 기반 **AI Agent for Science Framework** 소개
- Autonomy(자동)와 Agency(의도적 행위) 구분
- ◆ 요약: "과학 AI 에이전트는 도구가 아닌 '동료'다."

10) Siddharth Narayanan (FutureHouse)

Title: *Automation of Biological Discovery with LLM Agents*

Key Points:

- LLM 기반 생물학 실험 자동화 시스템 시연
- 실험 → 분석 → 논문화 자동 파이프라인
- ◆ 요약: "LLM 에이전트는 지능형 연구 실험실을 만든다."

11) Karthik Duraisamy (University of Michigan)

Title: *Active Inference AI Systems for Scientific Discovery*

Key Points:

- **Abstraction / Reasoning / Reality Gap** 3 대 한계 분석
- Self-Modeling + Counterfactual Reasoning + Reality Coupling 필요
- ◆ 요약: "AI 가 과학을 이해하려면 현실과 논리를 동시에 모델링해야 한다."

12) Panel Discussion 1

Participants: Hal Finkel (DOE), Raj Hazra (Quentinum), Pradeep Dubey (Intel), Molly Presley (Hammerspace), Karthik Duraisamy (University of Michigan, Moderator)

Topic: *Industry–Academia–Government Collaboration for Trustworthy AI in Science*

Key Points:

- 데이터 거버넌스, 모델 투명성, 공유 표준의 필요성 논의
- "Computational Trust is the next frontier in AI for Science." – Pradeep Dubey
- "데이터 모빌리티와 출처 투명성은 정확도 만큼 중요하다." – Molly Presley

13) Ricardo Baeza-Yates (BSC AI Institute)

Title: *Irresponsible AI and the Need for Transparent Evaluation*

Key Points:

- AI 시스템의 비윤리적 결정 (편향, 허위 정보) 사례 분석
- 책임성·설명가능성·데이터 출처 투명성 확보 프레임워크 제안
- 과학 AI에 필요한 "Responsible Evaluation Pipeline" 정의
- ◆ 요약: "AI는 정확도 만큼 책임성과 투명성을 평가해야 한다."

14) Prasanna Balaprakash (Oak Ridge National Laboratory)

Title: ORNL's AI Initiative: Advancing Secure, Assured, and Efficient AI for Scientific Discovery

Key Points:

- Oak Ridge National Laboratory의 AI Initiative (AAI) 개요 소개 — "Secure, Assured, Efficient"를 핵심 원칙으로 하는 과학 AI 프로그램
- 초대형 과학 AI 시스템이 신뢰성 있게 작동하기 위한 보안형 인프라(Security-Aware Infrastructure) 설계 개념 제시
- Assured AI for Science 프로젝트에서 사용되는 "AI workflow validation pipeline" 공개 — 모델 훈련, 실험, 배포 단계의 자동 검증 체계
- AI 실험 자동화와 HPC 시스템 간의 인터페이스 문제, 데이터 신뢰성 확보를 위한 Assurance Protocol 사례 발표
- 모델 성능보다 안전성(Safety)과 재현가능성(Reproducibility)을 우선하는 DOE 접근 전략 제시.
- ◆ 요약: "AI for Science의 다음 단계는 성능이 아니라 '신뢰할 수 있는 시스템 설계'에 달려 있다."

15) Jiacheng Liu (University of Washington / Allen Institute for AI)

Title: OLMoTrace: Tracing LM Output Back to its Multi-trillion-token Training Data in Real Time

Key Points:

- OLMoTrace 시스템 시연 — 대규모 언어모델(LLM)의 출력을 훈련 데이터로 역추적하는 실시간 트레이서 도구
- 데이터 프로비너스(provenance) 시각화 기술을 통해 LLM의 할루시네이션(hallucination)과 출처 불명 데이터 의존성 진단

- 모델 출력 신뢰도 평가("trust rating") 메트릭 제안 — 과학적 텍스트 생성의 검증 가능성 확보
- 연구자와 기관이 AI 감사(auditing)를 수행할 수 있는 개방형 평가 프레임워크 강조
 - ◆ 요약: "OLMoTrace 는 LLM 이 생성한 과학 결과의 근원을 추적해 AI 의 투명성과 검증성을 높인다."

16) Kyoung-Sook Kim (National Institute of Advanced Industrial Science and Technology, Japan)

Title: *Fairness of Geospatial Foundation Models*

Key Points:

- Geospatial Foundation Model (GFM) 들의 학습 데이터가 지역·인구·기후 편향을 내포하고 있음을 지적
- 위성 영상, 도시 인프라, 사회경제 지표의 지리적 불균형이 모델 예측 공정성에 미치는 영향 분석
- Fairness Evaluation Pipeline 제안: 데이터 샘플링, 공간적 커버리지, 예측 편향을 통합적으로 정량화
- AIST 'AI for Societal Resilience' 프로젝트 사례 공유: 재해 예측 및 기후 리스크 평가 모델의 지역 편향 완화 기법 설명
- 공정성(fairness) 을 단순한 기술적 보정이 아니라 사회적 신뢰(social trust) 확보를 위한 필수 요소로 규정
 - ◆ 요약: "지리공간 AI 의 공정성은 기술적 편향 제거를 넘어 사회적 신뢰를 구축하는 문제다."

17) Panel Discussion 2

Participants: Elahe Vedadi (Google DeepMind), Preeth Chengappa (Microsoft Discovery), Kexin Huang (Stanford University), Arvind Ramanathan (Argonne National Laboratory), Siddharth Narayanan (FutureHouse), Addison Snell (Intersect360 Research, Moderator)

Title: Building Agentic Systems for Science: Reports from the Field

Key Points:

- a. Agentic AI 의 정의와 필요성

- Snell (Moderator): “우리는 이제 자동화(automation)에서 에이전시(agency)로 이동하고 있다.”

- 패널 공통 의견: Agentic System 은 ‘사용자 지시 이행 AI’ 가 아니라 ‘과학적 목표를 스스로 세우고 검증하는 탐구 주체’이다.

b. 프로젝트 사례 (Reports from the Field)

- Elahe Vedadi (DeepMind): AlphaDiscovery Platform — 화학·단백질 합성 실험 자동화. LLM 기반 가설 생성 + 로봇 실험 피드백 루프 시연.

- Preeth Chengappa (Microsoft Discovery): Agentic Workflow Stack — Azure 기반 과학 탐구 에이전트 플랫폼. 대형 LLM · 시각 · 시뮬레이션 모듈 통합.

- Kexin Huang (Stanford, Boimni Project): 생명과학 실험을 통합하는 멀티에이전트 협력 시스템. “LLM as Planner + Simulation Agent” 구조.

- Arvind Ramanathan (Argonne, Scientia Project): HPC 환경에서 에이전트가 실험 계획과 데이터 분석을 자동 조정하는 ‘Scientia’ 공개.

- Siddharth Narayanan (FutureHouse): 생명과학용 Agentic Pipeline — 세포 이미징 분석 및 논문화 자동화.

c. 핵심 기술 이슈

- 실험 제어 루프와 시뮬레이션 루프의 통합 (Experiment-in-the-loop).

- 에이전트 간 의사소통 및 결정 투명성.

- “Active Inference Loop” – 가설 → 실험 → 평가 → 가설 갱신 사이클 구현.

d. 거버넌스 및 책임성

- Vedadi: “과학 AI 에이전트는 ‘왜 이 가설을 세웠는가’를 설명할 수 있어야 한다.”

- Ramanathan: “Scientia 는 모든 결정 경로를 로그화해 재현가능성을 보장한다.”

- Snell: “투명한 Agentic System 로그 가 과학적 신뢰의 핵심 지표가 될 것이다.”

e. 향후 계획 및 공동 제안

- “TPC Agentic Systems Working Group” 정식 발족 예고.

- Boimni, Scientia, FutureHouse 프로젝트 간 데이터·코드 연동 테스트 추진.

- AgentBench-Science (공통 성능 평가 벤치마크) 초안 논의.

◆ 요약: “Agentic AI 는 이제 실험 실행을 넘어, 스스로 가설을 세우고 검증하는 ‘자율 탐구 주체’로 진화하고 있다.”

18) Wrap-Up (Plenary 5): “Science Updates from Key TPC Leaders”

컨소시엄 핵심 리더들이 각 기관·워킹그룹의 성과와 로드맵을 공유했다.

3대 축: AI 모델 신뢰성 평가 표준화, 글로벌 데이터 허브, AI-for-Science 교육 이니셔티브

연사	발표 제목	Key Points	한 줄 요약
Rick Stevens (Argonne National Laboratory)	AI for Science 2030 Vision	<ul style="list-style-type: none"> TPC25 전체 결과 요약 및 차기 전략 제시 "AI for Science 2030" 로드맵 공개 —모델·데이터·평가 표준화를 위한 글로벌 구상 DOE-EU-Asia 간 상호협력 인프라 계획 	TPC는 '글로벌 과학 AI 거버넌스 연합'으로 진화 중.
Satoshi Matsuoka (RIKEN R-CCS, Japan)	Scaling Science with AI + HPC Convergence	<ul style="list-style-type: none"> AI-HPC 융합이 멀티피직스 과학문제 해결을 가속함을 실증 Fugaku 기반 AI-보강 시뮬레이션 성과(기후·재료·유체) 엑사스케일 AI 인프라를 위한 RIKEN-TPC 공동 프레임워크 제안 	AI-HPC 융합은 '규모의 한계'를 허문다.
Mateo Valero (Barcelona Supercomputing Center, Spain)	European AI for Science and the Role of TPC	<ul style="list-style-type: none"> EuroHPC 및 유럽형 파운데이션 모델 프로젝트 현황 BSC-RIKEN-Argonne 협력 네트워크 ("Tri-Hub") 발표 데이터 윤리·컴퓨팅 거버넌스의 유럽 모델 제안 	유럽은 TPC와 함께 개방형 AI 과학 인프라를 구축 중.
Thierry Bidot (INRIA / Genci, France)	Open Science Data and TPC Data Hub Initiative	<ul style="list-style-type: none"> TPC Data Commons 운영 계획 발표 FAIR 원칙 기반 데이터 관리 정책 초안 미·유럽 간 연구데이터 상호 접근성 확 	TPC Data Hub가 글로벌 오픈사이언스의 허브로 부상.
Ian Foster (Moderator, Argonne National Laboratory)	Scaling AI Collaboration Across Continents	<ul style="list-style-type: none"> Plenary 5 좌담 진행 및 논의 요약 협업 거버넌스·윤리·표준화 프레임워크 정리 차기 연례행사(TPC26) 의제와 워킹 그룹 방향 제시 	AI 협력의 미래는 '신뢰와 조율' 위에서 확장된다.

6-2. 튜토리얼 (Tutorials)

행사 개요

- 기간: 2025년 7월 28일(월) – 7월 29일 오전 (총 1.5 일)
- 형태: 90 분~2 시간 단위의 실습 중심 강의 (Hands-on Workshop)
- 대상: 초보자~전문가까지, 도메인 과학자와 AI 연구자 모두 참여 가능
- 운영방식: 사전 등록자 대상 실습용 환경(클라우드 + HPC 접속권한) 제공

목표

- AI for Science 핵심 기술 학습: 초기대 모델의 과학 응용 개념과 워크플로 구축법 교육
- TPC 표준 프레임워크 및 오픈소스 툴 실습: Reference Stack 구조, 데이터 파이프라인, 파운데이션 모델 API 이해
- 협업형 실습 환경 제공: 참가자 간 코드 공유 및 공동 모델 구축 유도

주요 트랙별 세부 내용

트랙	주요 강사(기관)	핵심 내용
Foundations of AI for Science	Hal Finkel (DOE), Ian Foster (Argonne)	<ul style="list-style-type: none">- 과학문제 중심 AI 접근법- 모델 학습, 검증, 불확실성 정량화 실습- Exascale 학습환경 튜토리얼
Building Agentic Workflows for Scientific Computing	Preeth Chengappa (Microsoft), Raj Hazra (Quantinuum)	<ul style="list-style-type: none">- Agentic AI 설계와 오케스트레이션 방법- Scientific Reasoning Agent 실습- "AI 연구 조교 만들기" 실습
Evaluation and Trustworthy AI	Rick Stevens (Argonne), Karthik Duraisamy (U. Michigan)	<ul style="list-style-type: none">- 모델 평가 벤치마크 설계- Active Inference 기반 테스트 구조- 데이터 투명성과 모델 해석 도구
Data-Physics Fusion Methods (PINN, FNO, DeepONet)	S. Matsuoka (RIKEN), K. Duraisamy (Michigan)	<ul style="list-style-type: none">- PINN, FNO, Operator Learning의 이론과 구현 실습- 물리 제약 신경망 실험
AI Infrastructure and Hybrid HPC-Cloud Deployment	AWS, NERSC 기술팀	<ul style="list-style-type: none">- 하이브리드 HPC- 클라우드 설정 튜토리얼- AWS ParallelCluster, SageMaker 연동

Trillion Parameter Consortium 2025 (TPC25) 리뷰
: 비전, 역사, 그리고 글로벌 Scientific AI 리더십

튜토리얼 운영 결과

총 250 명 이상 참가(등록자의 80%+가 과학자/엔지니어)

각 세션은 Colab 유사 환경 및 HPC 접속 계정으로 실습 제공

다수 참가자가 "Agentic Workflow 튜토리얼"을 가장 혁신적으로 평가

튜토리얼의 핵심 의의

"TPC25 는 발표 중심 학술대회가 아니라, '공유된 도구와 지식'을 직접 다뤄보는 실습형 협력 학회였다." (Alwire, 2025-07-31 보도 인용)

6-3. 해커톤 (Hackathon)

행사 개요

- 일시: 7 월 29 일(화) 오전 – 오후 전일
- 형태: 팀별(3~5 인) 미션 수행형 경진 프로그램
- 장소: DoubleTree by Hilton San Jose 컨퍼런스룸 B

목표

- AI for Science 의 실제 문제를 **'작은 실험형 프로젝트'**로 구현
- 협업 중심의 에이전트, 모델 평가, 데이터 파이프라인 설계 경험 제공
- 우수 팀은 TPC GitHub 에 코드와 문서 공개

주요 트랙별 세부 내용

트랙	예시 과제	기술 포커스
AI Agents for Science	"과학 논문 요약 및 데이터 추출 에이전트"	LLM + Vector DB + API Agent 설계
Model Evaluation	"LLM 모델 신뢰도 자동 벤치마킹 도구 제작"	Python 평가 모듈 + Metrics Visualization
Hybrid HPC-Cloud Pipeline	"Fugaku + AWS S3 기반 하이브리드 워크플로 구축"	HPC Job Scheduler + CloudBridge
Data Fusion & PINN	"기상 관측 데이터로 PINN 학습 및 예측"	TensorFlow PINN 구현
Active Inference Simulation	"반사실적 시나리오 모델링 시뮬레이터"	Bayesian Neural Net + Causal Graph

결과 요약

- 최우수상: 최우수상: "AgentSmith-Sci" 팀(Microsoft + UC Berkeley), 논문 1 만 편 메타데이터를 추출·질의 가능한 에이전트 구축
- 기술혁신상: "HybridFlow"(AWS + RIKEN), HPC-Cloud 간 잡 마이그레이션 도구 개발
- 커뮤니티상: "ExplainAI"(Argonne 대학원생 팀), 시각화 기반 LLM 결과 해석 툴 제작
- 해커톤 우수 작품은 TPC 공식 GitHub 리포지토리에 공개됨.

6-4. Breakout 트랙과 워킹그룹

행사 개요

- 기간: 7 월 30–31 일 (이틀간), 5 개 타임블록 × 6 개 트랙 × 약 30 세션
- 참가 인원: 평균 세션당 35~50 명
- 형태: 번개 발표(5 분) + 토론(60 분) + 공동정리(20 분)
 - 목적: 아이디어를 공유하고 실제 연구 협력 주제를 도출
- 산출물: 12 개 세션에서 "후속 공동연구 그룹(Working Group)" 결성
- 성과 예: "TrustBench" (AI 모델 평가 WG), "BioAgent" (LLM 기반 생물학 워크플로 WG), "Hybrid-HPC Taskforce" (인프라 최적화 WG)

주요 트랙별 세부 내용

트랙명	대표 세션	주요 논의 포인트
Life Sciences	Foundation Models for Biology, Agentic Bio Workflows	단백질 점 힘, 생물학적 시뮬레이션 자동화, Bio-LLM 구축 사례
Applications (Earth, Fusion, Materials)	AI4Climate, AI4Fusion, AI4Materials	기후모델·핵융합·재료 탐색에 AI 적용 사례, 멀티피델리티 학습
Evaluation & Reasoning	Benchmarking Reasoning Models, Trust and Uncertainty Quantification	추론평가, 신뢰성, 검증 가능한 AI 실험설계
Workflows & Agents (DWARF)	Data, Workflows, Agents, Reasoning Framework (DWARF)	워크플로 자동화, 오케스트레이션 툴, HPC 연동
Scale & Services (MAPE)	Model Architecture, Performance, Efficiency	효율적 모델 병렬화, 학습 최적화, 하드웨어-SW 공진화

트랙명	대표 세션	주요 논의 포인트
Community & Outreach	AI Education for Science, Inclusive Access to Compute	인력 양성, 개발자 교육, 오픈사이언스 협력 플랫폼

#대표 브레이크아웃 요약 사례

- ◆ Life Sciences – “AI for Molecular Discovery” (Flora Salim 주재)
 - LLM 으로 분자구조·단백질 결합 특성 예측 사례 발표.
 - 생물학자와 AI 엔지니어 협업 모델 논의.
 - “Biological foundation model” 구축 로드맵 초안 마련.
- ◆ Evaluation & Reasoning – “Trustworthy AI Evaluation” (Karthik Duraisamy 주재)
 - 과학적 추론력 검증을 위한 Causal Reasoning Benchmark 초안 제시.
 - “Reality Gap Test Set” 구성 제안.
 - 향후 TPC 평가위원회 설립 논의로 연결됨.
- ◆ Workflows & Agents – “DWARF Architecture Discussion” (Ian Foster 주재)
 - 데이터·워크플로·에이전트·추론 통합 프레임워크 설계.
 - 공통 API 제안서 초안 작성.
 - “DWARF v0.1 White Paper” 발간 예정.
- ◆ Applications – “AI for Fusion Simulation” (S. Matsuoka, RIKEN)
 - 핵융합 시뮬레이션에서 Neural Operator 를 통한 경계층 모델링 성과 발표.
 - 엑사스케일 대응 학습 아키텍처 제안.

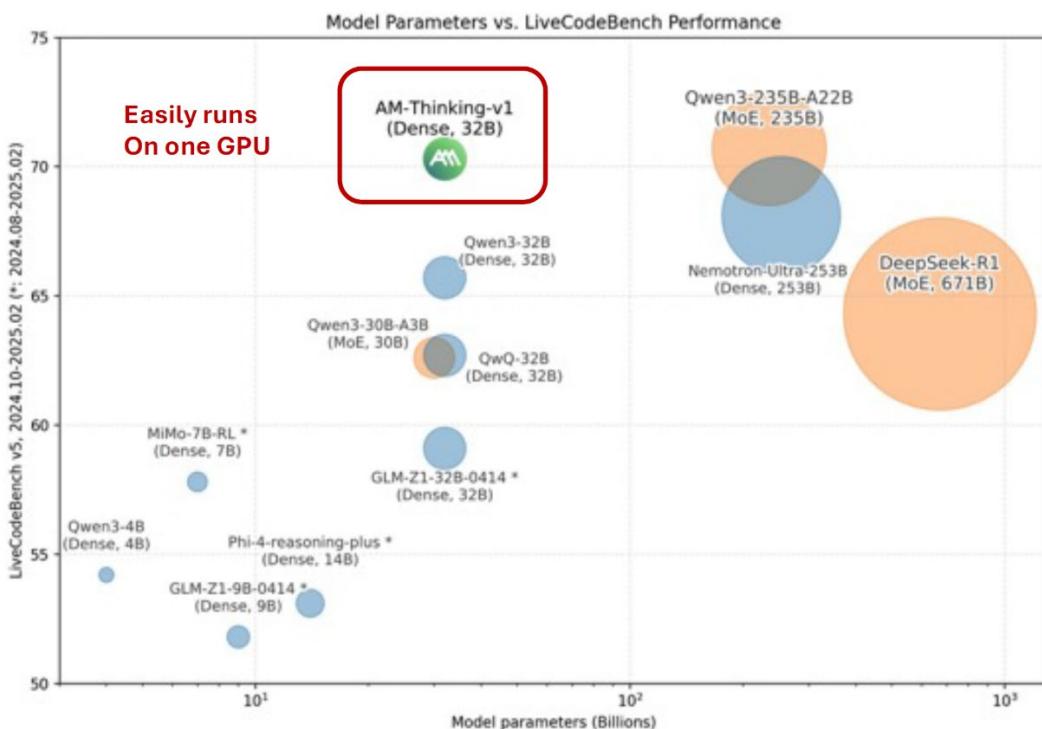
7. 주요 발표내용 발췌

Opening Session (Plenary Session 1): Transforming Science: Frontier Models, Hybrid Systems, and Agentic Systems (Tuesday, July 29, 14:00)

TPC25의 시작을 알린 오프닝 세션에는 Rick Stevens(Argonne National Laboratory), Thierry Pellegrino(AWS), Flora Salim(UNSW) 등 세 연사가 참여했다. 주제와 관점은 달랐지만 공통 메시지는 명확했다: “**AI는 이제 계산을 넘어 과학의 파트너가 된다.**”

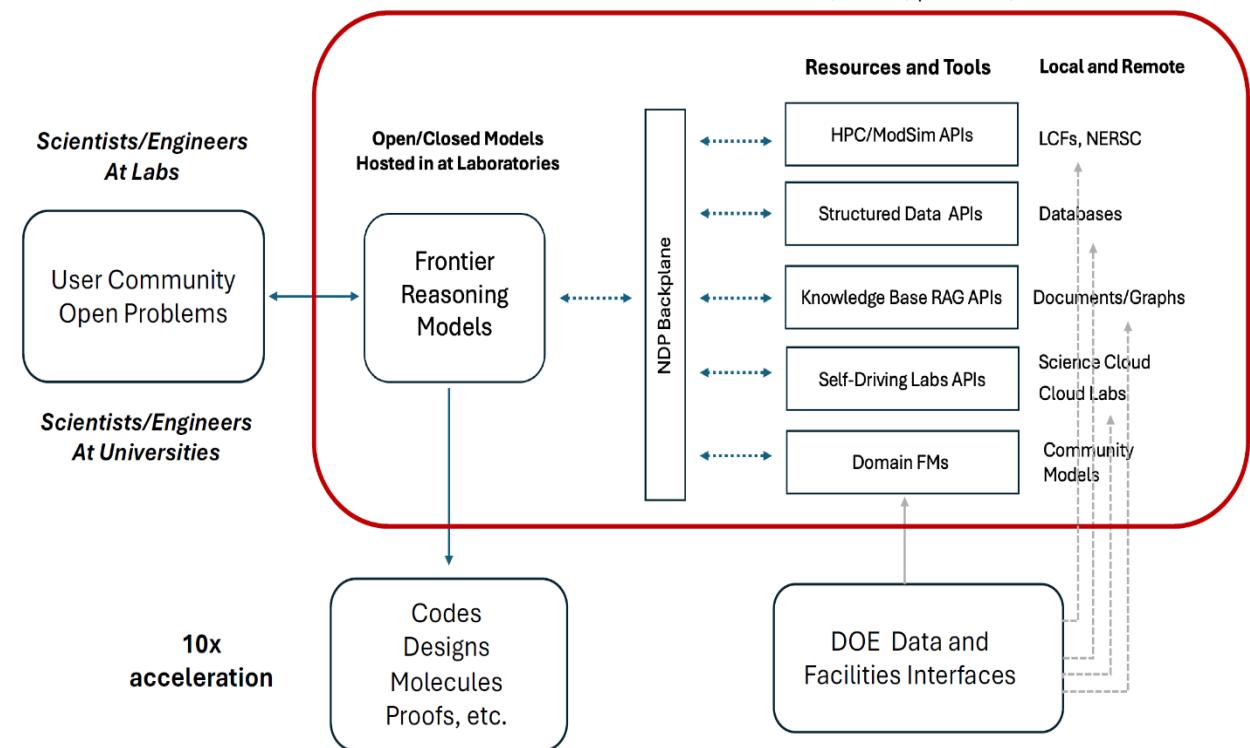
Reinventing Discovery: Accelerating Science in the Age of Artificial Super-Intelligence

Rick Stevens는 AI가 단순 계산 도구를 넘어 **스스로 과학적 질문을 생성하는 방향**으로 진화하고 있다고 밝혔다. 또한 GPT-4o, Claude 3 등 최신 모델이 이미 인간 박사 수준의 문제 해결 능력을 보이고 있다고 언급했다.



< 거대 AI 모델 성능 지표 예시 - Rick Stevens 제공 슬라이드 >

The “Scientific AI Platform” hosted at labs, cloud, partners, etc.



Rick Stevens 제공 슬라이드

< AI-Native Discovery Loop (모델-데이터-시뮬레이션 사이클 통합 개념) >

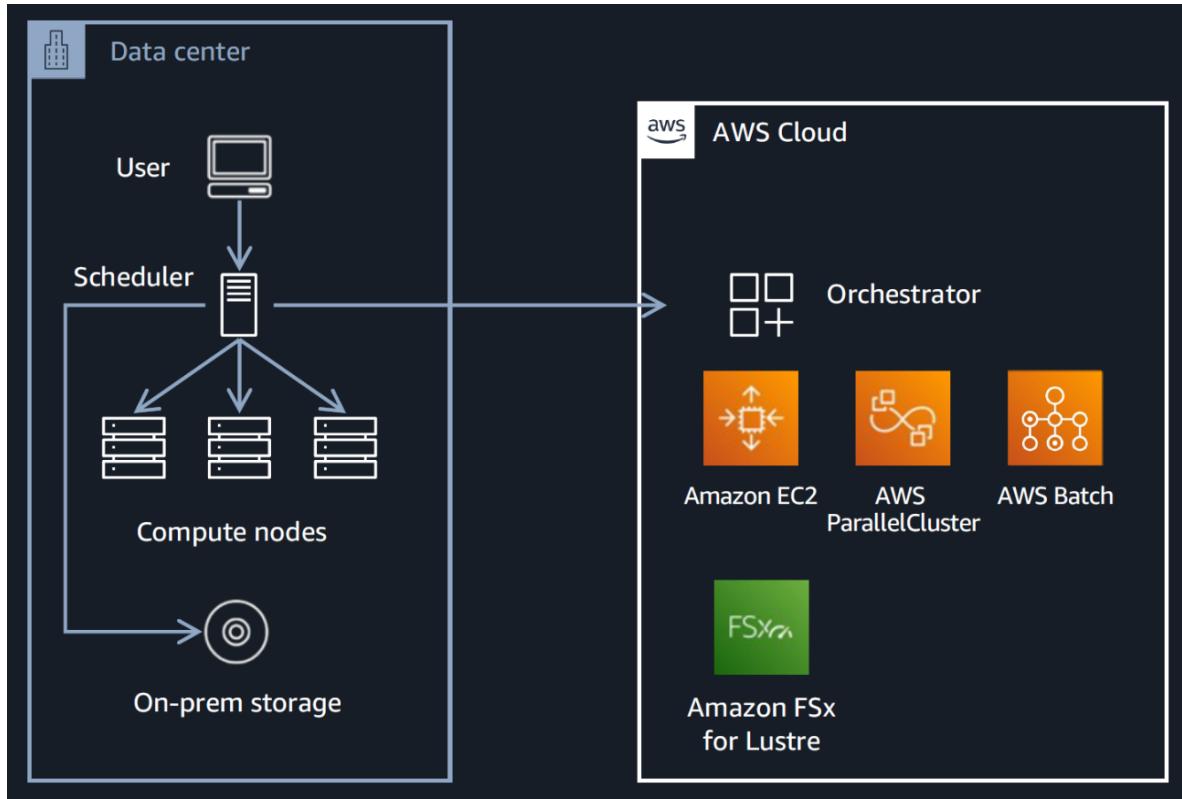
핵심 요점

- 초거대 모델 훈련 + 실험 자동화 + 데이터-시뮬레이션 통합 루프
- “과학의 다음 도약은 더 빠른 계산이 아니라 더 빠른 사고다.”

HPC and Science: The Need for Hybrid

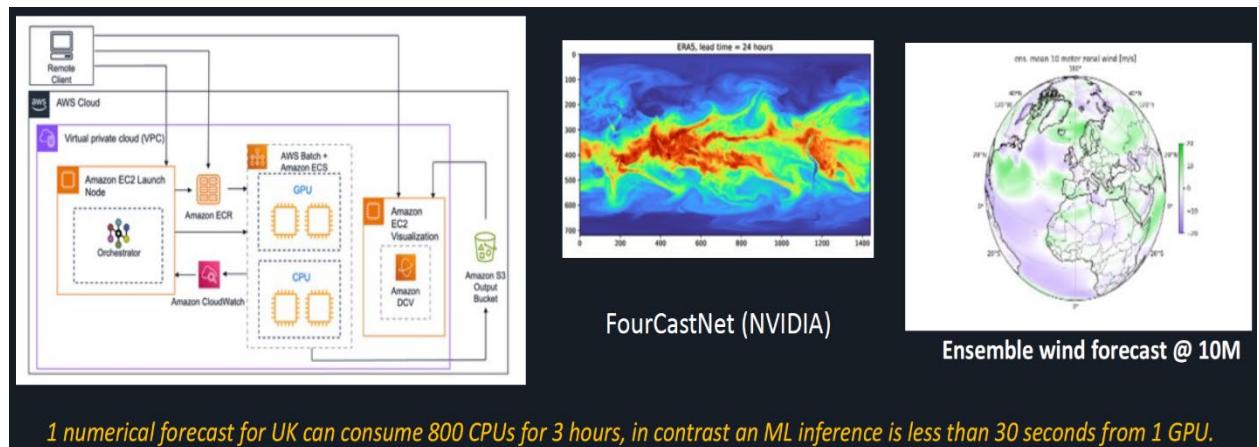
Thierry Pellegrino는 하이브리드 HPC-클라우드 인프라의 필요성을 강조했다. 초대형 모델을 단일 환경에서만 운영하는 시대는 끝났으며, 이제 HPC·클라우드·양자컴퓨팅 자원의 통합이 요구된다는 것이다.

Trillion Parameter Consortium 2025 (TPC25) 리뷰
: 비전, 역사, 그리고 글로벌 Scientific AI 리더십



Pellegrino 제공 슬라이드

< HPC-Cloud 오케스트레이션 개념도 >



Pellegrino 제공 슬라이드

< FourCastNet 예측 속도 향상 결과 (NVIDIA-AWS 통합 사례) >

핵심 요점

- 클라우드-온프레미스 연계 워크로드 배포.
- 데이터 지연 최소화와 자원 유연성 확보.
- “하이브리드 없이는 확장도 없다.”

Modeling and Simulating Complex Behavior in Dynamic Cyber-Physical-Social Systems

Flora Salim 교수는 복잡계 지능 모델링을 다뤘다. 도시·인간·기계가 얹힌 CPSS를 이해하려면 데이터만으로는 부족하며, 물리·사회·행동 데이터를 함께 학습하는 AI 접근이 필요하다고 설명했다.

PromptCast: A Prompt-Based Learning Paradigm for Time Series Forecasting



Dataset & Benchmark

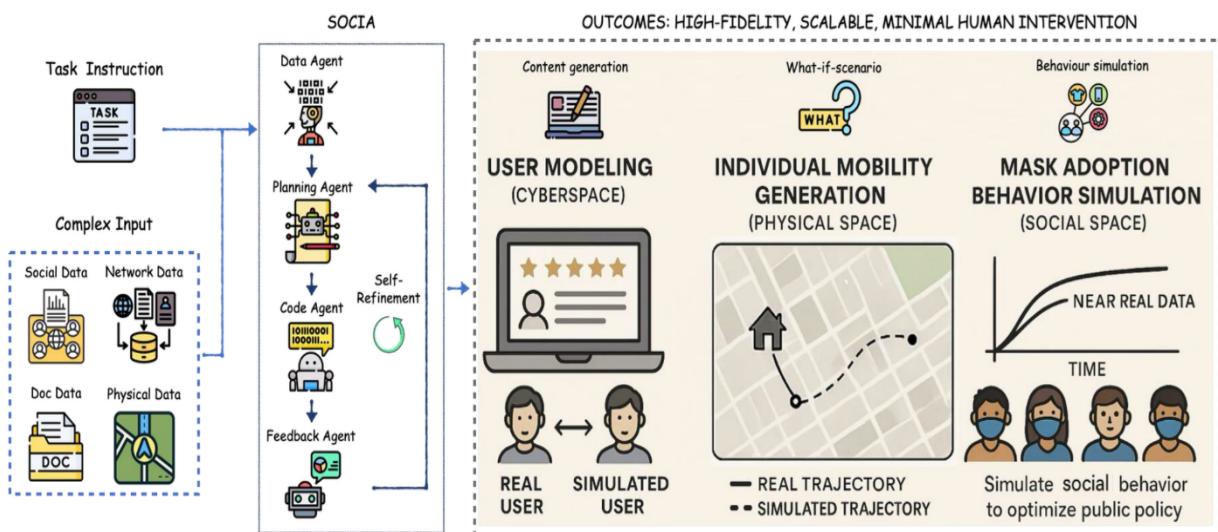
- 3 forecasting domains: mobility, weather, electricity
- 311,932 data instances in total
- 20 methods & language models included



Code & Dataset

< PromptCast 프롬프트 기반 시계열 예측 프레임워크 >

SOCIA: an Agentic Simulation Generation Engine for Cyber, Physical, and Social Behaviours



< SOCIA Agentic Simulation Framework (도시 행동 시뮬레이션 예시) >

핵심 요점

- 멀티모달 시공간 학습 및 Neural ODE/FDE 활용.
- 시뮬레이션 + 실험 데이터 융합으로 복잡계 모델 정확도 향상.
- "AI는 데이터를 넘어 세상 그 자체를 배워야 한다."

테이크어웨이

Rick Stevens 가 "AI 가 질문하는 과학"을 얘기했다면,
Pellegrino 는 "AI 가 계산을 확장하는 방법"을 보여줬고,
Salim 은 "AI 가 세상을 이해하는 방법"을 보여줬어.

→ "AI 는 과학의 속도를 높이는 엔진이 아니라, 과학의 사고 구조를 바꾸는 촉매다."

Plenary Session 2: TPC and AI for Science Two Years Later: New Directions in Convergence of AI and HPC (화요일, 7/29, 16:00)

TPC25는 최첨단 AI를 실용적 과학 도구로 전환한다는 공통 목표를 가진 연구자들이 모인 자리였다. 1일차(7/29, 화) 오후 본회의에서는 진행 중인 다양한 연구가 소개됐다. RIKEN의 Satoshi Matsuoka는 엑사스케일 시스템 발전이 과학 문제에 맞춘 모델 개발을 어떻게 견인하는지 보여줬고, Quantinuum의 Steve Clark은 양자 회로와 트랜스포머의 상호 학습 가능성을 제시했다. Hyperion Research의 CEO Earl Joseph은 고성능 컴퓨팅(HPC) 사이트에서 AI 활용이 빠르게 확산되고 있음을 최신 설문 데이터로 공유했다. RIKEN의 Jens Domke는 데이터 유출과 취약한 소프트웨어의 위험을 지적하며, 연구 그룹이 보다 안전한 온프레미스 솔루션을 구축하지 않으면 발전이 지연될 수 있다고 경고했다. 이번 세션은 가능성과 함께 직면한 실질적 장애물을 분명히 드러냈다.

Exascale 만으로는 과학에 적합한 AI를 구현할 수 없는 이유

RIKEN 계산과학연구센터(R-CCS) 소장 마츠오카 사토시는 TPC25 강연에서 오늘날 상용 파운데이션 모델은 과학 연구의 '출발점'에 불과하다고 강조했다. RIKEN AI-for-Science 팀을 대표해 그는 대규모 생성 모델이 신뢰할 수 있는 과학 도구가 되려면 데이터 처리, 모델 설계, 워크플로 오패스트레이션 측면의 난제를 해결해야 한다고 지적했다. 마츠오카는 먼저 RIKEN 인프라를 업데이트했다. 후가쿠(Fugaku)는 6만 개 CPU 노드를 갖춘 세계적 슈퍼컴퓨터이지만, RIKEN은 여기에 약 1,500개의 NVIDIA Blackwell 가속기와 수백 개의 GPU·TPU를 추가 탑재할 GPU 클러스터를 증설 중이다. 또한 156큐비트 IBM Q System Two, 20큐비트에서 156큐비트로 확장 예정인 Quantinuum 시스템, 자체 개발 64큐비트 초전도 장치 등 세 가지 양자 시스템을 운용하고 있다. 그는 향후 제타(zeta) 스케일, 즉 10^{21} op/s급 시스템 전망도 제시했다. "우리는 다양한 공급업체 및 파트너와 함께 2029년 가동을 목표로 차세대 기계를 설계하고 있으며, DOE 등 기관과 협력하고 있습니다. 이를 통해 2029년에는

AI와 전통적 HPC 모두에서 '제타스케일 컴퓨팅'이라 부를 수 있는 성과를 달성할 것입니다."

하드웨어 로드맵을 제시한 뒤, 마츠오카는 '최고 성능'에서 '실용성'으로 초점을 옮겼다. 그는 순수한 속도는 그 위에서 실행되는 모델이 과학 데이터와 워크플로를 이해할 때에만 가치가 있다고 강조했다. 예컨대 상업용 생성 모델은 일반 사용자를 대상으로 설계되어 물리·화학·생물 분야에 필수적인 기능이 부족한 경우가 많다. 과학 데이터는 텍스트·방정식·이미지·센서 스트림이 혼합된 데다 종종 테라바이트 규모다. 현재의 토크나이저(tokenizer)는 웹 텍스트에는 강하지만 도메인 기호와 단위, 매우 긴 시퀀스에는 취약하다. 고해상도 현미경 스캔과 같은 과학 이미지는 표준 비전 트랜스포머를 쉽게 과부하시킨다. 이를 해결하기 위해 과학 AI 시스템에는 맞춤형 토큰 어휘, 희소 어텐션(sparse attention), 백만 토큰을 훨씬 넘는 컨텍스트 윈도우를 처리할 수 있는 물리 기반 디코더 등의 기능이 필요하다고 그는 말했다. 이러한 요구를 충족하기 위해 RIKEN은 거대 모델의 메모리 사용량을 줄이면서 과학 데이터를 더 잘 이해하도록 만드는 실용적 기법을 실험 중이다. 예를 들어 쿼드트리 타일링, 공간충전 곡선(space-filling curve), 초고해상도 이미지의 시퀀스 길이를 줄이면서 공간적 일관성을 유지하는 적응형 패칭(adaptive patching) 등을 테스트하고 있다. 초기 결과는 정확도를 훼손하지 않으면서 계산량을 기하급수적으로 절감할 가능성을 보여주지만, 여전히 새로운 컴파일러 지원과 메모리 레이아웃이 요구된다. 다중 모달 융합을 위해 연구팀은 신경 서로게이트(surrogate)와 편미분방정식(PDE) 솔버를 결합한 연산자도 연구하고 있다. 이러한 효율화 실험은 모델 전략의 변화—초거대 단일 모델에서 과업 맞춤형 다중 모델의 조합으로—를 반영한다. 마츠오카는 전문가 혼합(MoE) 아키텍처, 도메인 미세조정 모델, AI 추론 에이전트의 부상을 강조하며, 무한한 매개변수 확장보다는 추론 단계에서의 리즈닝(reasoning)을 통해 계산 비용을 줄이면서도 견고성을 높일 수 있다고 주장했다.

마츠오카는 TPC 이니셔티브를 통해 연구자들의 참여를 장려하며, 오는 1월 오사카에서 열리는 HPC Asia 2026에서 TPC 전담 트랙을 공개한다고 밝혔다. 이 트랙의 목표는 실험·시뮬레이션 파이프라인 전반에 AI 에이전트를 내장한 실행 가능한

시스템을 선보여, 신뢰할 수 있는 과학용 AI를 통해 가설 생성, 실험 계획, 시뮬레이션, 분석, 논문 작성까지 전 과정을 가속화하는 환경에 한 걸음 더 다가가는 것이다.

생성적 양자 AI를 통한 과학적 발견 실현

양자 컴퓨팅과 AI라는 두 혁신 기술이 결합하면 무엇이 가능한가? Quantinuum의 AI 책임자 Steve Clark는 강연에서 양자 컴퓨팅과 AI가 상호 보완을 넘어, 단독으로는 달기 어려웠던 가능성을 여는 시스템을 어떻게 개발 중인지 소개했다. 그는 이를 “생성적 양자 AI(generative quantum AI)” 전략으로 규정하고, 세 가지 축을 중심으로 접근한다고 설명했다: (1) AI로 양자 컴퓨팅을 최적화하기, (2) 양자 시스템으로 새로운 형태의 AI 구동하기, (3) 양자 컴퓨터가 생성한 데이터로 머신러닝 모델 학습하기. 첫 번째 축에서 Clark는 AI를 활용해 양자 컴퓨팅을 더 실용적으로 만드는 방법을 제시했다. 최신 양자 하드웨어조차 노이즈, 제어 정밀도, 오류율 등 현실적 한계를 안고 있으며, 여기서 머신러닝이 큰 역할을 한다. 그의 팀은 심층 강화학습 등을 사용해 양자 알고리즘을 하드웨어 명령으로 변환하는 회로 컴파일을 지원하고, 값비싼 양자 게이트 수를 줄여 회로 실행 효율을 높이는 방법을 학습시켰다. 또한 최적 제어와 오류 정정 문제에도 AI를 적용하고 있으며, “하드웨어에 적용해 매우 유망한 결과를 얻고 있다”는 설명처럼 시뮬레이션을 넘어 실제 양자 컴퓨터에서도 도구를 검증하고 있다.

클라크의 전략에서 두 번째 방향은 ‘방정식을 뒤집어’ 양자 컴퓨팅이 어떻게 새로운 유형의 AI를 열 수 있는지 탐구하는 것이다. 그의 팀은 신경망을 양자 시스템을 단순 시뮬레이션하는 데 그치지 않고, 애초에 양자 시스템 위에서 동작하도록 재설계할 때 무엇이 달라지는지 연구하고 있다. 이러한 모델은 전통적 트랜스포머처럼 표준 어텐션을 쓰는 대신 중첩 등 양자적 속성을 활용해 정보를 전혀 다른 방식으로 처리한다. 클라크의 설명대로 이는 “고전적 아날로그가 없는 완전히 새로운 유형의 모델”이며, 실제 양자 하드웨어에서 초기 버전을 이미 시험했다. 세 번째 방향은 양자 컴퓨터와 AI의 완전 통합이다. 그의 팀은 기존 데이터만으로 학습하는 대신 양자 장비가 생성한 데이터를 AI 시스템에 공급해, 기존 방법으로는 얻을 수 없던

패턴을 학습하도록 한다.

한 사례로 분자의 바닥상태를 찾는 생성 양자 고유해석기(Generative Quantum Eigensolver)가 있다. 트랜스포머 모델이 양자 회로를 제안하고, 에너지를 측정한 뒤 결과를 개선하도록 조정한다. 초기 실행에서는 수소의 바닥상태를 확인했다. 클라크는 이 접근이 화학뿐 아니라 재료과학, 최적화, 나아가 사기 탐지 등에도 유용할 수 있다고 지적했다. 이 지점에서 고전 시스템과 양자 시스템이 서로로부터 진정으로 학습할 수 있다.

AI 가 HPC 에서 주류로 자리 잡았지만 데이터 품질과 예산 압박은 여전히 존재

TPC25에서 Hyperion Research의 CEO Earl C. Joseph은 기술 컴퓨팅 분야 AI 도입에 관한 최신 설문 결과를 발표하며, 과학·공학 조직이 직면할 과제와 변화 속도를 강조했다. 최근 몇 년 사이 HPC 전반의 AI 도입이 급격히 가속되어 Hyperion은 2~3개월 주기로 신규 조사를 진행할 정도다. 2020년에는 HPC 사이트의 약 3분의 1이 AI를 일부 활용한다고 답했으나, 2024년에는 그 비율이 90%를 넘어섰다. 오늘날 과학 워크로드에서의 AI는 더 이상 실험 단계가 아니다. 정부 연구소, 학술 센터, 산업계 사용자들은 시뮬레이션 보강부터 대규모 데이터 분석까지 다양한 작업에 AI를 적용하고 있다.



<하이페리온 리서치의 CEO 얼 조셉의 연설 모습>

클라우드 도입은 이러한 급증과 밀접히 연결된다. Joseph은 최첨단 하드웨어, 특히 GPU의 조달·유지 부담을 해소하기 위해 기업들이 클라우드를 점점 더 선호한다고 밝혔다. 차세대 NVIDIA 가속기가 나올 때마다 전력·냉각 비용까지 고려한 총비용이 상승하면서, 4~5년간 유지해야 할 온프레미스 시스템에 투자하는 것이 비현실적이라고 판단하는 경우가 많다. 탐색 단계에서 현세대 하드웨어와 유연성을 활용하려는 수요에 힘입어 클라우드 채택은 계속 빠르게

증가하고 있다.

그러나 그는 여전히 만연한 몇 가지 장벽을 강조했다. 가장 빈번한 과제는 학습 데이터 품질이었다. Hyperion은 데이터셋의 부족·불일치로 수많은 AI 프로젝트가 중단된 사례를 확인했으며, 한 예로 메이요 클리닉은 자체 검증 데이터만을 사용해 소규모이지만 고품질 언어 모델을 학습함으로써 이러한 위험을 완화했다. 이 밖에도 사내 AI 전문 인력 부족, 학습 데이터 규모의 한계, AI를 HPC 환경에 통합하는 과정 전반의 복잡성이 주요 장애로 지적됐다. Joseph은 이러한 복잡성이 도메인 특화 AI 소프트웨어 및 컨설팅 서비스 시장의 신규 성장을 촉진할 것으로 전망했다.

Hyperion 연구에 따르면, 설문 참여 기업의 97%는 비용 증가에도 불구하고 AI 활용을 확대하고 운영 규모·데이터 양·인력을 늘릴 계획이었다. 그는 온프레미스와 클라우드 구축 모두에서 AI 인프라 비용이 상승함에 따라 예산을 대폭 증액해야 한다고 강조했다. Hyperion이 2025년을 앞두고 주시하는 세 가지 핵심 과제는 복잡성, 비용, 하드웨어 가용성이다. 발표는 곧 공개될 Hyperion의 투자수익률(ROI)과 “연구 수익률(Research ROI)” 연구를 미리 소개하며 마무리됐다. 이 연구는 단순 수익성과 손익분기점을 넘어, 정부·학계·산업계 전반에서의 AI 도입이 과학에 미치는 영향을 측정해 재정적 성과뿐 아니라 발견과 혁신까지 평가할 예정이다.

RIKEN은 Gen AI 위험 완화를 위한 온프레미스 접근 방식을 구축하고 있습니다.

RIKEN 계산과학센터 슈퍼컴퓨팅 성능연구팀의 Jens Domke 팀장은 과학 분야에서 AI 도입을 서두르는 과정에서 위험 완화가 종종 후순위로 밀린다고 경고했다. “모두가 황금을 향해 달려가지만, 결국 보안을 처리해야 합니다.” 그는 다음의 5가지 위험을 제시했다: (1) 인간, (2) AI 소프트웨어, (3) 소프트웨어 공급망, (4) 모델 자체, (5) “법무 리스크와 악의적 행위자”. Domke는 클라우드 기반 생성형 AI 사용 시 기밀 유출 사례와 인증 정보 유출 이슈가 보고된 바 있음을 언급하며, 부주의한 데이터 업로드 관행의 위험을 지적했다. 또한 특정 서비스·모델과 관련해 제기된 국가 간 데이터 유출 가능성 경고 사례들을 소개하며, 새 AI 기능에는 새로운 위험이 동반된다는 점을 청중에게 상기시켰다.

AI 소프트웨어와 관련해 그는 이렇게 말했다. “모두가 최신·최고 소프트웨어를 내놓기 위해 서두르지만, 보안을 제대로 들여다보는 이는 거의 없습니다. 예를 들어 최근 도입된 모델 컨텍스트 프로토콜(MCP)을 보안 연구자들이 분석한 결과, MCP 서버가 모든 인터페이스와 포트에서 수신 대기만 하며 연결을 기다리는 것으로 나타났습니다. 따라서 노트북에서 MCP를 실행한 채 네트워크에 연결되어 있다면 예상치 못한 상황이 발생할 수 있음을 명심해야 합니다. MCP는 안전하지 않으며, 다른 소프트웨어 인프라도 마찬가지입니다.”

그는 모델 실행에는 보통 20~30개의 소프트웨어 패키지가 필요해 위험이 복잡해지고 확대된다는 점을 상기시켰다. 강연은 AI 위험에 대한 좋은 입문서였다. 이어 Domke는 RIKEN이 자체 AI 관리 역량을 사내에 구축하기로 한 결정을 소개했다. 이는 다수 연구기관에도 실행 가능한 해법이라고 그는 주장한다. “문제가 많기 때문에 우리는 온프레미스 솔루션을 택했습니다. 기본적으로 안전하고 프라이빗한 OpenAI 대안으로, 외부 서비스에서 하던 대부분의 작업을 내부에서 재현하면서도 데이터 유출·해킹 위험 없이 수행할 수 있습니다.”

아래 두 슬라이드는 개요를 제공한다.

Secure and privatizes “OpenAI”-replica for Internal Research

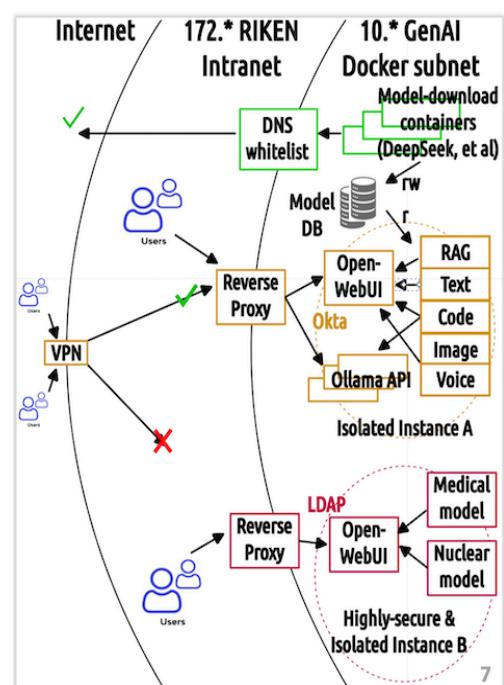
Issue with landscape of serving generative AI models

- Commercial offerings: inflexible w.r.t (new) models
- Cloud offerings: problematic (danger of leaking data)
 - Only few in this room can afford and strong-arm OpenAI into giving them an on-prem ChatGPT
- Open-source/free: many unaudited packaged (risky)

→ RIKEN needs one or multiple private genAI instances for different security/confidentiality levels behind appropriate authentication and reverse-proxies

Solution is open-source with adequate security concept

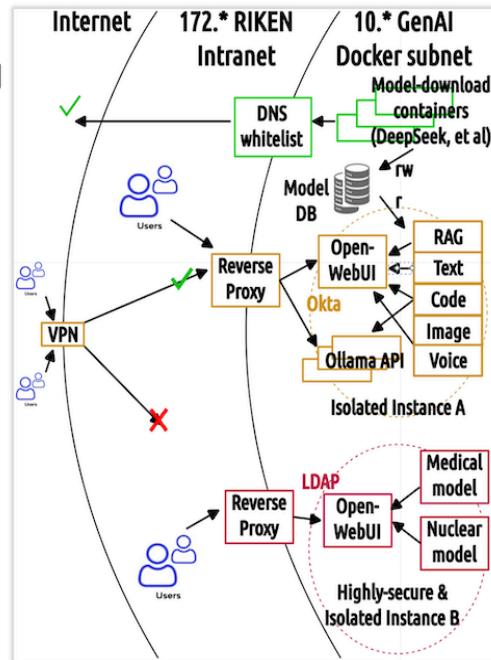
- Create 1 instance for broad usage with many features and common models; accessible via intranet and VPN
- Create additional instances for higher security levels





Our Security Concept for RIKEN – More Details

- Assumption: **don't trust anything/anyone**
(some containers from dockerhub aren't audited and/or include hundreds of 3rd-party packages)
- **Don't use same container to download models** and to server models → prevents data leak
- **Containers serve one single purpose**, read-only access to data they don't need to modify
- **No genAI container** should have **internet access**
- **Reverse proxy** can be config'd to limit access depending on user's location and IP range
- Download containers further secured with **white-list DNS server** → e.g. only huggingface
- **https/ssl** for openwebui et al (instead of http)



엔스 둠케 제공 슬라이드

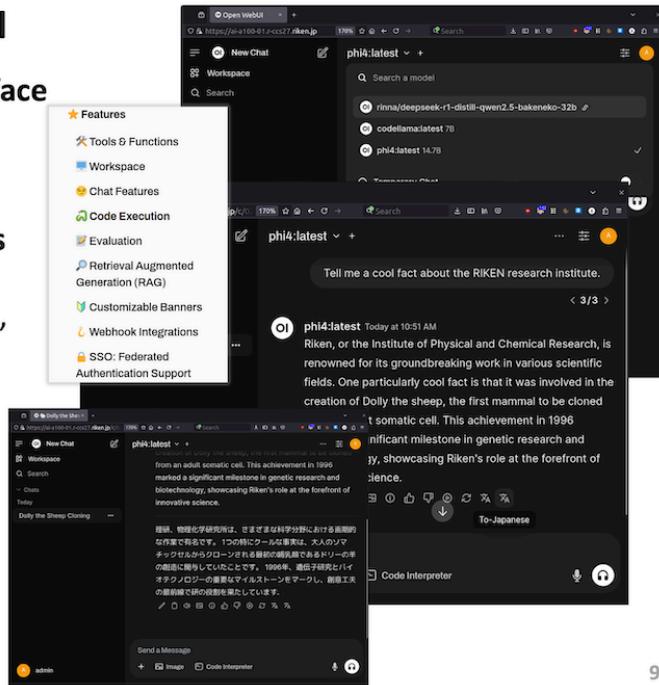
"우리는 오픈소스를 기반으로 인프라를 구축할 계획이며, 현재 다층 보안 엔클레이브 개념을 구현 중입니다. 하나는 준개방형으로, OpenAI 제품과 유사하지만 방화벽 뒤에서 내부 운영되어 광범위한 사용이 가능하고 상당한 수준의 안전성을 제공합니다. 다른 하나는 고도 보안 계층으로, 기밀 업무·의료 연구·RIKEN 내 각종 연구를 위한 고보안 인스턴스를 갖춥니다. 이렇게 계층을 나눠 필요와 편의에 따라 접근 패턴을 설정할 수 있습니다. 핵심 원칙은 '아무것도 신뢰하지 않는다'입니다. 모델 다운로드에는 컨테이너를 사용하며(제공용 컨테이너와는 분리), 모든 구성요소는 기본적으로 컨테이너화된 상태로 프라이빗 네트워크에서 동작합니다. 모든 컨테이너는 프라이빗 네트워크에 연결되고, 웹 UI·API는 역방향 프록시로만 접근하게 합니다. 이를 통해 VS Code 등 익숙한 에디터로도 접속할 수 있습니다. 여기에 유용한 기능을 추가할 예정이지만, 전반적 개념은 이렇습니다." Domke는 이 접근이 OpenAI 사용 경험과 매우 유사하다고 덧붙였다. "다양한 채팅 기능과 번역을 제공하고, 관리자가 설치한 모델 범위 내에서 모델을 전환할 수 있습니다. Hugging Face의 모델을 손쉽게 가져오고, 자체 미세조정 모델도 배포할 수 있죠. OpenAI처럼 특정 모델에 종

속되는 제약이 없으며, 오픈소스와의 호환이 뛰어납니다. 원한다면 오픈 웹 UI를 사용할 수 있고, 모델 비활성화, 가입·그룹 관리, 권한 부여 기능도 모두 내장돼 있습니다.”



User-facing Open Web UI

- Very similar design to [OpenAI interface](#)
- Supports **various features** and **multiple chats/contexts**
- Easy to select from available **models** via dropdown
- Customizable via scripts/"functions" (e.g. our eng<->jap translations)
- Access rights for models: everyone, groups, or user
- Active community (main repo and custom functions) and **open-source** (support purchasable)
<https://docs.openwebui.com/>
<https://openwebui.com/functions>



9

엔스 둠캐 제공 슬라이드

Domke는 이것이 RIKEN의 장기적 비전이라고 밝혔다. 그는 데이터를 클라우드로 전송할 필요가 없는 온프레미스 보안 환경을 언급하며 “현재 SPring-8을 활용해 프로토타입을 구축하고 있습니다”라고 말했다.

그의 강연은 하드웨어·알고리즘·사용자 도입·보안 전반에 걸쳐 일관된 메시지를 전달했다. 규모만으로는 충분하지 않으며, 과학용 AI는 도메인 튜닝 모델, 하이브리드 고전-양자 워크플로, 고품질 데이터, 그리고 신중한 위험 관리에 의존하게 될 것이다. 향후 12개월은 연구자들이 이러한 아이디어를 공유 도구와 커뮤니티 표준으로 얼마나 빠르게 구현할 수 있는지 시험하는 시간이 될 것이다. TPC25에서 제시된 진전이 이어진다면, 이 분야는 신뢰를 훼손하지 않으면서 발견 속도를 높이는 AI 시스템에 한층 가까워질 것이다.

Plenary Session 3: AI and The Future of Scientific Discovery (수요일, 7/30, 9:00)

수요일 첫 번째 본회의는 국가 연구소, 하이퍼스케일 클라우드, 학계 등 다양한 관점에서 같은 질문을 던졌다. AI가 '과학을 설명하는 도구'에서 '과학을 실제로 수행·설명하는 주체'로 나아가려면 무엇이 필요한가?

Argonne의 Ian Foster는 가설에서 실험실로 전환하기 위한 LLM의 "사고-행동(Thought–Action) 구조"를 개략적으로 제시했다. Microsoft의 Preeth Chengappa는 에이전트 플랫폼이 냉각수 화학부터 백신 설계까지 데이터·모델·도구를 어떻게 연결하는지 시연했다. FutureHouse의 Siddharth Narayanan은 언어 에이전트가 초급 연구 파트너처럼 추론할 수 있는지를 시험해, 새로운 벤치마크와 툴 콜링 스팡폴드(tool-calling scaffolds)가 그 격차를 메우기 시작했음을 보였다. University of Michigan의 Karthik Duraisamy는 AI가 토큰 예측자와 인과적 사고자를 구분하는 '추상화–추론–현실' 간의 간극을 메우지 못하면 모든 진전이 무의미할 수 있다고 지적했다.

이 본회의는 프린티어 규모 모델을 '완성품'이 아니라, 추론 엔진을 실험에 연결하고, 출처와 정책을 인코딩하며, 인간을 루프에 포함시키는 더 크고 협력적인 시스템의 구성 요소로 재정의했다.

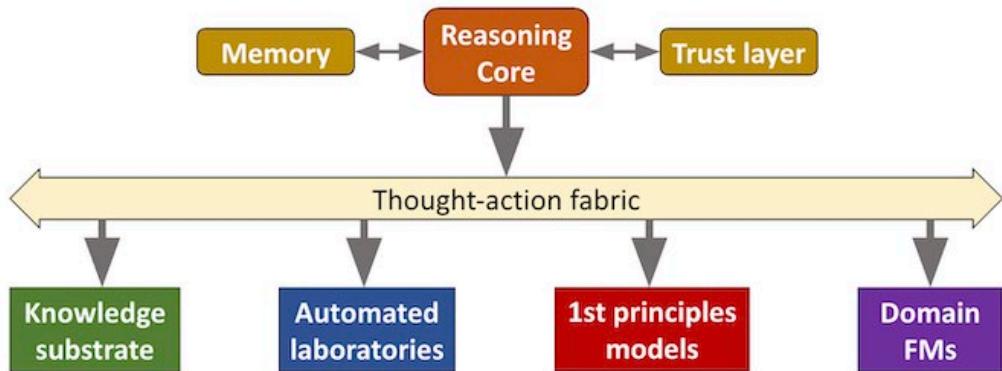
AI 기반 발견을 위한 청사진: 사고, 행동 및 인프라

Argonne의 Ian Foster는 대규모 AI 모델이 패턴 인식에서 실제 과학적 추론·발견으로 전환하는 길을 논의하며 본회의를 시작했다. 언어 모델이 GPQA(A Graduate-Level Google-Proof Q&A Benchmark) 등에서 인간 전문가를 능가할 수 있지만, 그는 이러한 추론 수준이 아직 과학이 요구하는 기준에 미치지 못한다고 지적했다. 진정한 과제는 단순히 '생각'하는 것을 넘어, 이 추론 시스템을 세계와 연결해 '행동'하게 만드는 일이다.

이를 위해 Foster가 'AI 기반 과학적 발견 플랫폼'이라고 부르는 체계가 필요하다. 핵심에는 추론 엔진이 있으나, 이는 데이터 저장소, 시뮬레이션, 실험실, 정

책 제작, 수행·학습 내역을 추적하는 메모리 시스템 등에 접근하는 더 큰 프레임워크 안에 내장돼야 한다. 그는 이를 “사고–행동(Thought–Action) 구조”라고 명명했으며, AI가 가설을 생성하고 실험을 실행·평가하며 그 결과로 학습하도록 하는 틀이라고 설명했다.

AI-native Scientific Discovery Platform



0/안 포스터 제공 슬라이드

“추론 핵심(reasoning core)을 어떻게 ‘사고–행동 구조(thought–action fabric)’에 연결할까요? 이는 모델의 생각을 세계에 대한 행동으로 전환하도록 하는 것입니다.”라고 Foster는 말했다. “이 구조를 어떻게 마련해 모델이 접근해야 하는 다양한 요소들과 연결할 수 있을까요?”

이 시스템의 작동 방식을 설명하기 위해 Foster는 Argonne에서 수행한 과제—이산화탄소를 에탄올로 전환하는 더 효율적 촉매 탐색—to를 예로 들었다. 이 가상 시나리오에서 연구자가 질문을 제기하면, 시스템은 구조화 데이터와 과학 문헌을 분석해 가설을 제시하고, 시뮬레이션·실험실 분석·반복 학습을 포함한 계획을 수립한다. 약 3시간 동안 시스템은 수천 개 후보 물질을 선별하고, 슈퍼컴퓨팅에서 고충실도 시뮬레이션을 수행하며, 자동화 실험실에서 실험을 진행하고, 새 결과에 따라 서

러게이트 모델을 재학습한다. 최종적으로 유망한 측면을 식별하고, 모든 작업을 기록하며, 정책을 점검한 뒤 결과를 연구자에게 투명하게 전달한다.

Foster는 이것이 아직은 가설적 사례이며, 어느 연구실에서도 처음부터 끝까지 완벽히 수행할 수 있는 단계는 아니라고 분명히 했다. 다만 필요한 구성 요소 상당수는 이미 독립적으로 존재한다. Argonne을 비롯한 여러 연구실은 원격 제어 가능한 자동화 연구실, 서러게이트(surrogate)-도메인(domain) 모델과 통합되는 시뮬레이션 파이프라인, 워크플로 조율 및 정책 집행을 위한 초기 인프라를 구축했다. 이제 과제는 이 구성 요소들을 대규모로 추론·행동·학습할 수 있는 일관된 시스템으로 연결하는 것이다.

그는 또한 현행 모델의 한계를 인정했다. 일부 모델은 타당한 가설을 도출하지만, 그 아이디어를 어떻게 평가하고 어떤 경로가 추구할 가치가 있는지 판단하는 방법에 대한 명확한 근거가 아직 부족하다. 많은 과학 데이터의 구조화와 접근성 부족도 주요 병목으로 남아 있다. 더 근본적으로는, 수십 년에 걸쳐 인간 사용자를 위해 구축된 과학 인프라를, 인간이 따라가기 어려운 속도와 규모로 작동하는 지능형 에이전트를 위해 재구성해야 할 가능성도 제기된다.

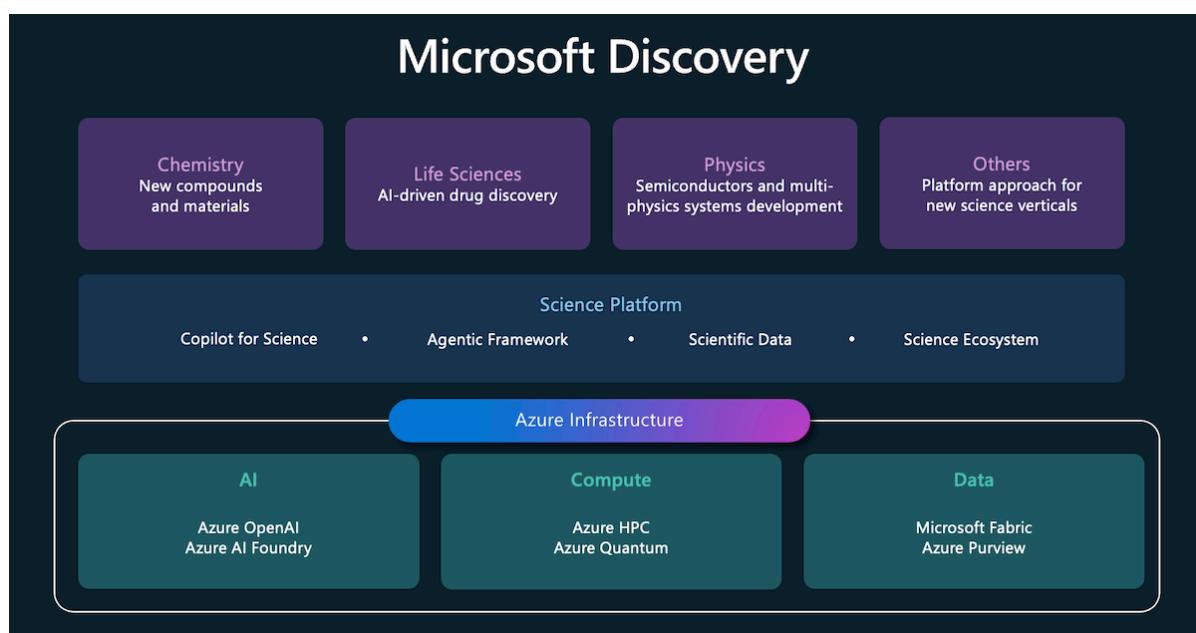
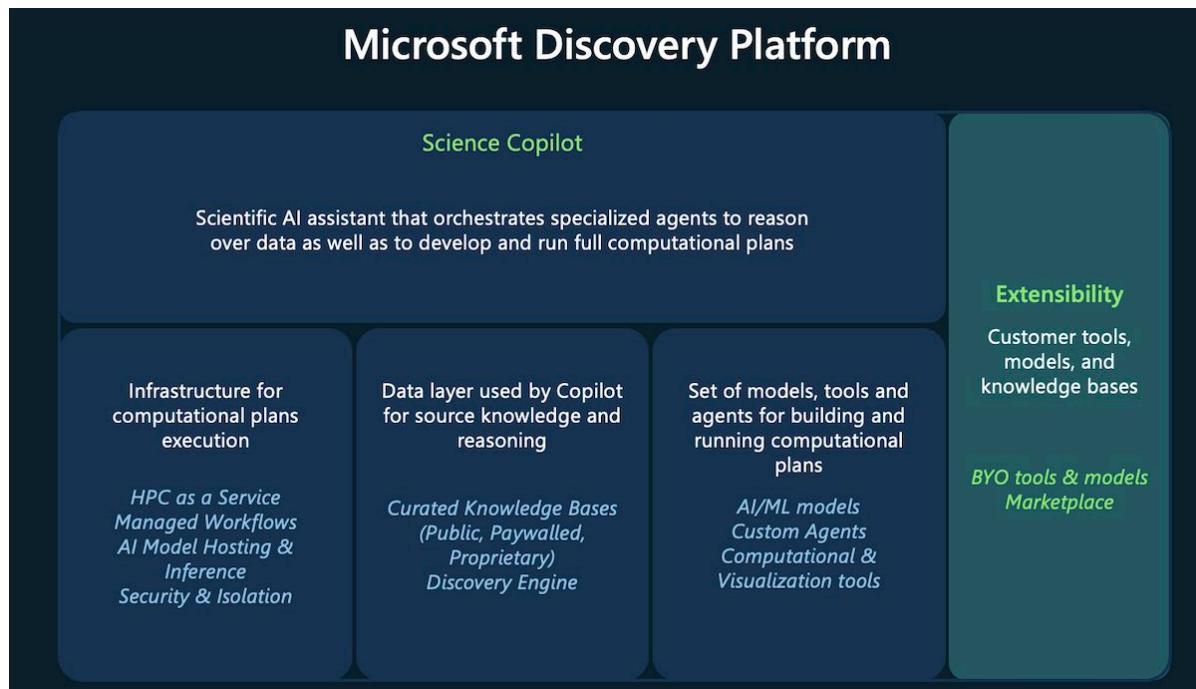
Foster는 “TPC 내부에서 과학적 추론 플랫폼을 구축하고, 우리가 필요한 사고-행동 구조를 마련하며, 접근 가능한 다양한 유형의 리소스와의 인터페이스를 함께 만들려는 분들이 참여하길 바랍니다”라고 말하며 TPC 커뮤니티의 동참을 촉구했다. 일부 기초 구성요소는 이미 여러 국가 연구소와 기관에 구축되어 있지만, 진정한 도약을 위해서는 AI, 시뮬레이션 과학, 자동화, 데이터 관리, 정책을 잇는 긴밀한 학제 협력이 필요하다는 점을 그는 강조했다. 그는 AI 기반 환경에서 과학이 어떤 형태를 갖춰야 하는지 정의하는 것이 하나의 기회라고 보았다. 성공한다면, 세상을 단순히 분석하는 데 그치지 않고 변화에 기여하는 시스템—더 빠르고 투명하며, 매 새로운 실험에서 지속적으로 학습하는 시스템—을 탄생시킬 수 있을 것이다.

클라우드에서 랩으로 자율성 확대

"마이크로소프트 전체가 이제 AI를 우선시하는 회사"라고 밝힌 산업·반도체·물리학 부문 책임자 Preeth Chengappa는 TPC 강연 'Agents, Autonomy, and Agency: A Brave New World'에서 2025년 5월 공개된 Microsoft Discovery Platform과 AI 구성 요소를 활용한 초기 사례를 소개했다. 마이크로소프트는 OpenAI 지분 참여를 통해 AI 혁명의 초기 단계부터 관여해 왔으며, 생성형 AI 연구 조직을 세 갈래로 운영한다: (1) LLM(학습·평가) 중심, (2) AI 인프라, (3) 인프라와 LLM을 기반으로 한 서비스 구축. Chengappa는 "저는 세 번째 그룹에 속해 있습니다"라고 덧붙였으며, 현재 이 작업의 상당수가 초기 고객과 협력하는 Discovery Platform에서 진행되고 있다고 설명했다.

"아래 슬라이드에서 보시다시피, 세 가지 주요 관심 분야는 물리학, 화학, 생물학입니다. 네 번째 분야는 자율 로보틱스에 집중될 전망으로, Ian이 언급한 습식 실험실(wet lab)과 실험실 자동화 등 매우 흥미로운 영역을 포함합니다. 이 모든 것은 Azure 핵심 인프라 위에 구축됩니다. 다만 오늘 제가 중점적으로 다룰 내용은 그 사이에 위치한 계층입니다."

Chengappa는 "HPC 계층은 다양한 형태의 컴퓨팅에 대한 오케스트레이션과 대규모 접근을 제공합니다. 중간 데이터 계층은 고객 데이터와 서드파티 데이터를 포함해 데이터의 전 영역을 다루며, 모델·도구 저장소도 포함합니다. 이 플랫폼은 데이터·에이전트·모델을 불러와 나열·등록할 수 있고, 이 모든 것이 최상단의 'Science Copilot'과 연결됩니다. 바로 이 부분에서 다른 구성요소들과 실제로 상호작용하죠. 몇 가지 사례로 보여드리겠지만, 여기서 가장 매력적인 점은 에이전트 아키텍처와 에이전시 프레임워크 덕분에 일정 수준의 자율성—즉, 계획과 실행—이 가능하다는 것입니다."라고 설명했다.



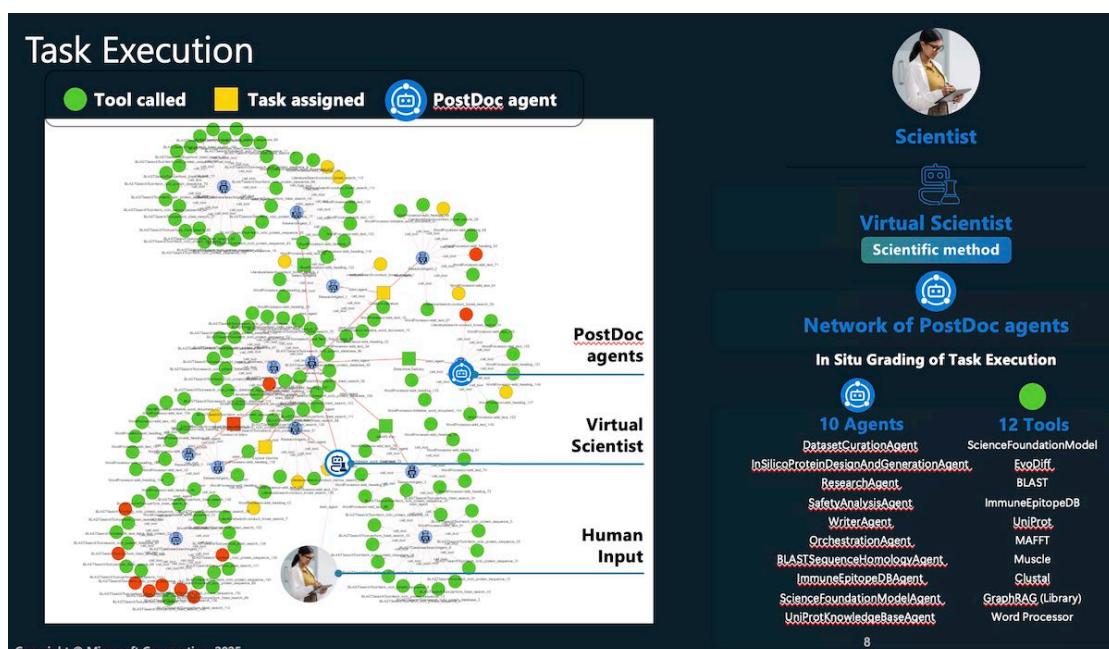
Preeth Chengappa 제공 슬라이드

첫 번째 사례 연구는 데이터센터 냉각이었다. "Microsoft는 초대형 데이터센터를 운영하는데, 냉각에 문제가 있습니다. 기존 솔루션에는 유해 화학물질인 PFA가 많이 포함돼 있다는 점도 문제였죠. 그래서 PFA 함량이 낮은 대안을 찾고자 했습니다. 먼저 '조사해 보고 내가 알아야 할 내용을 알려줘'라는 간단한 프롬프트를 던졌습니다. 시스템은 관련 자료를 검토해 유용한 요약을 만들었고, 다음으로 '계획을 세

우고 실행 방법을 알려줘'라고 요청했습니다. 그 뒤 사람이 개입해 '이건 비등점이 적절하지 않다', '이건 내가 원하는 분산이 아니다'처럼 편집·피드백을 합니다. 이 지점에서 인간 집단의 역할이 매우 중요합니다. 그렇게 계획이 수정되면 이제 '실행' 단계로 넘어갑니다. 애니메이션은 많이 넣지 않았지만, 여러 에이전트가 호출되고 후보 물질들이 도출됩니다. 연구자가 이를 검토해 '유망한 후보군 같다'고 판단하면, 논리적 다음 단계로 실제 합성을 진행해 현재 대규모 응용을 위한 테스트에 들어갑니다."라고 Chengappa는 설명했다.

다음 사례는 제약 분야의 약물 전달 메커니즘과 관련된다. WHO는 현행 항바이러스제·백신의 투여 방식이 최적 효율을 내지 못하고 부작용을 유발할 수 있다고 지적했다. 제기된 질문은 "투여 메커니즘의 상태를 개선할 수 있는가?"였다.

"다시 같은 과정을 거쳐, 반복을 통해 더 정교해진 프롬프트를 구성합니다. 그 결과, 결합 가능한 특정 단백질을 신속히 식별해 더 나은 결과로 이어집니다. 아래 슬라이드에 보이듯 다양한 에이전트와 우리가 '도구'라고 부르는 구성요소가 동원됩니다. 이는 발견 과정에서 무엇이 어떻게 일어났는지를 설명 가능하게 시각화할 때 반복적으로 사용하는 형태입니다. 이 모든 도구와 파이프라인을 활용해 전달에 사용할 수 있는 8가지 스파이크 단백질 후보를 찾아냈고, 현재 이 문제의 실제 해법을 모색하기 위한 후속 연구가 진행 중입니다."



Preeth Chengappa 제공 슬라이드

Chengappa는 몇 가지 최신 트렌드도 언급했다.

- **개인 맞춤형 상담원** – “최근 기업에서는 이메일 ID 와 Teams ID 를 가진 고도의 개인화 에이전트를 많이 도입하고 있습니다. 동료와 대화하듯 에이전트와 소통하고 업무를 위임할 수 있습니다.”
- **신뢰** – “우리는 인간–에이전트, 에이전트–도구, 에이전트–LLM 지표 등 여러 관점에서 신뢰를 봅니다. 이제 사람들은 LLM 이 무언가를 ‘할 수 있다’는 사실 자체에는 의심이 적지만, 각 구성요소의 동기와 에이전트가 실제로 무엇을 하고 있는지 이해하는 것이 중요합니다.”

그는 다음과 같은 경고를 남겼다. “마무리하며 말씀드리자면, 에이전트는 어디에나 존재하게 될 것입니다. 아직은 아니지만, 결국 우리의 많은 일을 대체하게 될 겁니다. 이는 시간의 문제입니다. 과거의 수많은 기술적 혼란 속에서도 우리는 인간으로서 다음에 무엇을 할지 예측할 수 있었습니다. TPC가 앞으로 하게 될 일과, LLM 및 생성형 AI 분야의 에이전트들이 보여줄 놀라운 성과에 모두가 기대를 걸고 있다는 것을 알고 있습니다. 그러나 에이전트가 지배하는 세상에서 우리가 어디로 향하게 될지에 대해서는 아직 해답이 없습니다. TPC는 이 문제뿐 아니라 사회적 영향과 정치적 구조에도 주의를 기울여야 합니다.”

과학자처럼 생각하도록 언어 에이전트를 가르치기

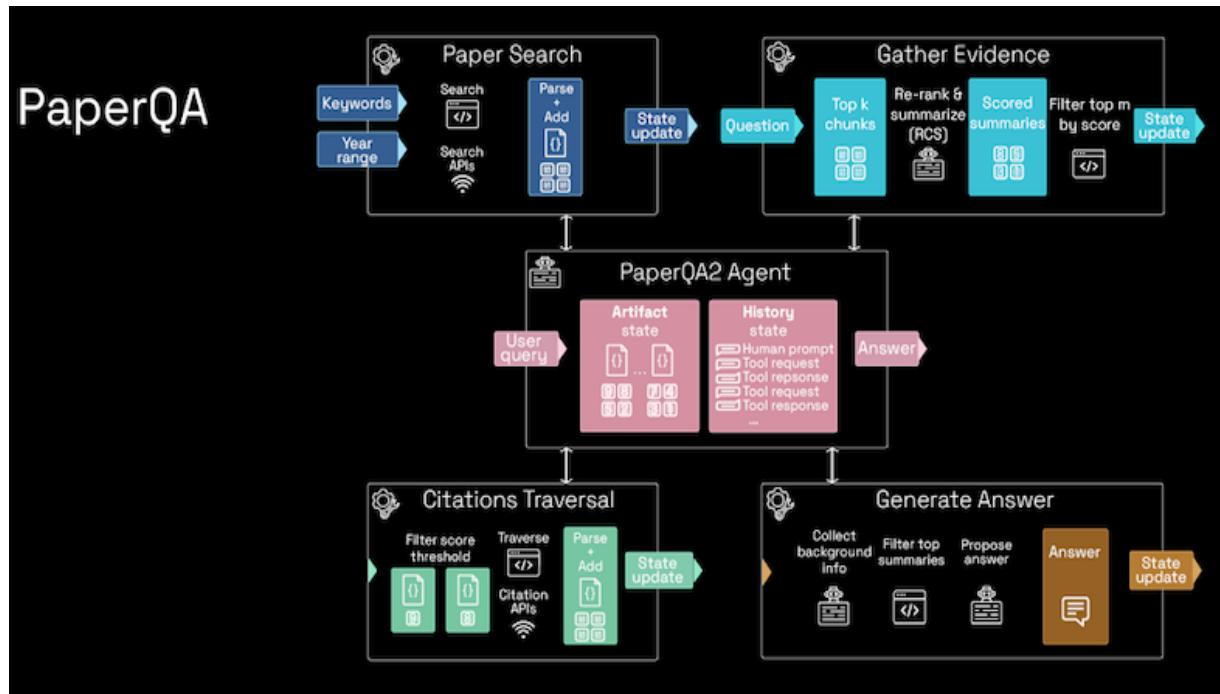
AI에 질문을 던지는 것과, 과학자처럼 사고하고 실험을 설계하며 다음에 무엇을 검증할지 제안하게 만드는 일은 전혀 다른 문제다. Siddharth Narayanan 박사는 바로 이런 시스템을 개발 중이다. TPC25에서 그는 언어 모델 에이전트가 데이터 분석, 논문 검토, 신약 후보 설계 등 복잡한 과학 과업을 어떻게 수행하기 시작했는지 공유했다. Narayanan 박사는 과학적 발견 자동화에 초점을 둔 비영리 연구 기관 FutureHouse의 물리학자이자 연구원으로, 입자물리와 머신러닝을 배경으로 암흑물질

·단백질 설계 연구를 수행해 왔다. 현재는 과학적 추론을 확장하고 연구자의 효율을 높이는 AI 시스템 구축에 집중하고 있다.

"오늘날 과학을 따라잡는 일은 그 어느 때보다 어렵습니다. 대부분의 연구자들이 감당하기 어려울 정도로 더 많은 데이터, 더 많은 도구, 더 많은 논문이 쏟아져 나오고 있습니다." Narayanan 박사는 이러한 압박을 지적하며 "정보의 양이 엄청나게 방대해졌고, 과학이 수행되는 방식 또한 더욱 복잡해졌습니다"라고 말했다.

그는 AI를 연구를 단순화하는 수단이 아니라, 연구자가 발전 속도를 따라잡도록 돋는 도구로 본다. 적절히 설계된 언어 모델 에이전트는 논문을 읽고 아이디어를 생성하며, 과학자의 사고를 확장하는 협력자로 기능할 수 있다. 다만 이러한 모델이 과학에서 실제로 무엇을 할 수 있는지 파악하는 일 자체가 어렵다. Narayanan 박사 팀은 다수 벤치마크가 모델이 정답을 맞히기만 하면—단순 암기일지라도—보상한다는 점을 발견했다. 그러나 과학은 암기가 아니라 단계적 추론이다. 이에 팀은 과학자가 실제로 일하는 방식을 반영한 테스트를 설계했다. 모델은 아직 인간을 능가하지는 못하지만 점차 근접하고 있다. 무엇을 할 수 있는지 파악하는 것은 시작일 뿐이다. 연구에 진정 유용하도록, 팀은 언어 에이전트를 구축하고 있다. 이 시스템은 모델이 도구를 실행하고 과업을 수행하며, 구조화된 환경에서 데이터와 상호작용하도록 설계됐다. 즉 AI가 단순 텍스트 예측을 넘어, 과학자의 작업 방식을 모사하며 추론·실험·문제 해결을 수행하도록 하는 것이다.

이러한 에이전트들은 이미 연구에서 가장 시간이 많이 소요되는 부분을 돋고 있다. PaperQA는 과학 논문을 검색하고 핵심 내용을 추출한 뒤, 특정 주제에 대한 상세 질문에 답하는 스마트 비서처럼 동작한다. ProteinCrow라는 또 다른 에이전트는 신규 분자 설계에 초점을 맞추며, 선행 연구를 검토하고 구조 예측을 수행해 컴퓨터에서 시험 가능한 결합제를 제안한다. 이 도구들은 단순한 실험용 데모가 아니라 실제 사용을 염두에 두고 설계되었다.



Preeth Chengappa 제공 슬라이드

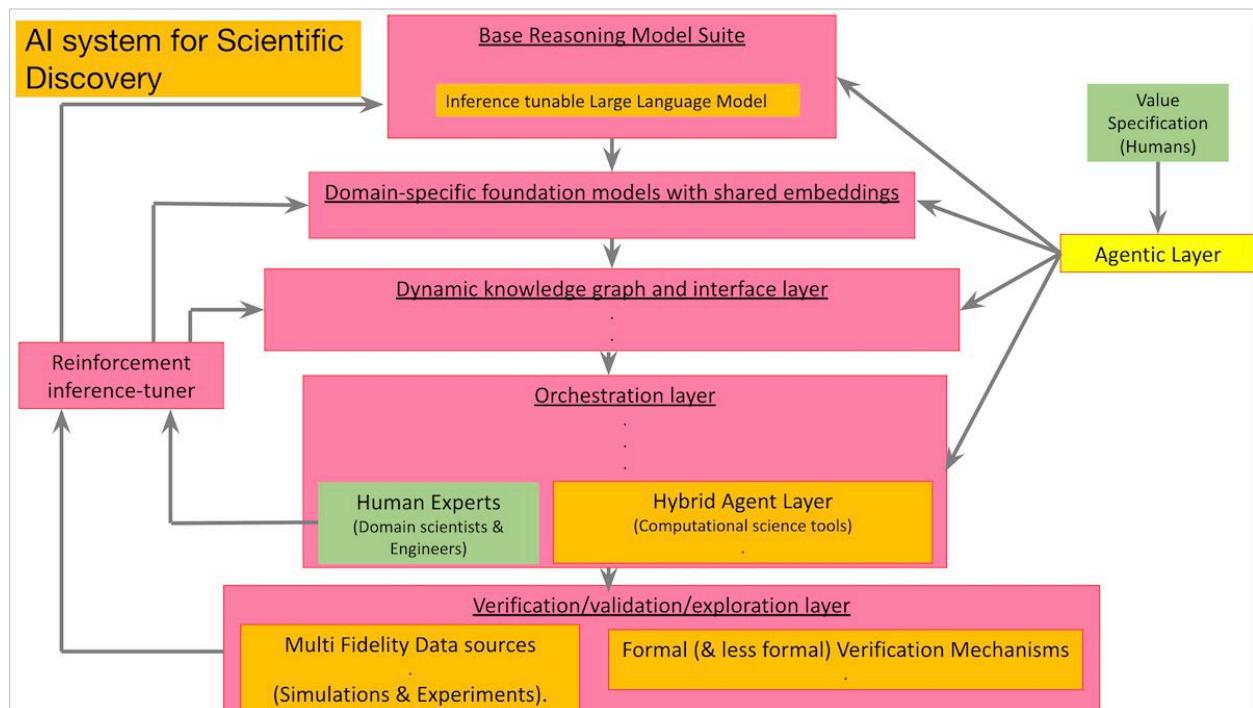
이러한 에이전트의 접근성을 높이기 위해 연구팀은 과학자가 에이전트를 실행하고 결과를 검토하며, 각 단계의 추론 과정을 추적할 수 있는 인터랙티브 플랫폼을 구축했다. 현재까지 가장 진보된 시스템 중 하나인 'Robin'은 AI를 활용해 약물 가설을 생성·개선하도록 돋는다. Narayanan 박사는 세션에서 실제 사례, 초기 결과, 향후 연구 방향을 상세히 소개했다.

발견 루프로서의 능동적 추론 Active Inference as a Discovery Loop

University of Michigan의 Karthik Duraisamy 교수는 TPC25 발표에서 '과학적 발견을 위한 능동적 추론 AI 시스템'을 통해, 더 큰 모델만으로는 과학이 ChatGPT가 불러온 충격과 같은 전환점을 맞기 어렵다고 설명했다. 그는 오늘날 AI가 '발견 엔진'으로 기능하는 데 방해가 되는 세 가지 지속적 간극—**추상화 간극, 추론 간극, 현실 간극**—에 직면해 있다고 지적했다. "추상화 간극은 과학자가 의존하는 고수준 개념 대신 토큰과 픽셀로 추론하는 모델에서 비롯됩니다. 현재의 파운데이션 모델은 통계적 패턴을 잡아내지만, 과학자가 필요로 하는 방식으로 핵심 도메인 개념을 파악하는 경우는 드뭅니다. 추론 간극은 작업이 수개월의 맥락이나 명확한 인과 관계를 요구할 때, 취약한 논리 사슬이 붕괴되는 현상을 뜻합니다. 추론 사슬이 몇 단계를 넘

어가면 많은 모델이 여전히 비틀거리며, 한 결과와 다음 결과를 잇는 인과가 사라집니다. 매일 이 모델들의 능력에 감탄하지만, 동시에 그 취약성도 분명합니다. 우리에게 필요한 것은 확립된 인과 관계 위에서 추론하고, 지금보다 훨씬 더 추상적 수준에서 유추를 전개할 수 있는 모델입니다. 또한 과학은 몇 주, 몇 달, 몇 년—어쩌면 수십 년—에 걸쳐 진행됩니다. 매우 긴 추론 과정이 필요합니다.” 또한 그는 “현실 간극은 모델이 훈련 데이터 범위를 벗어나 미지의 실험 결과를 예측할 때 드러납니다. 최신 실험 데이터가 없다면, 분포 외 예측이나 익숙지 않은 데이터로부터의 예측은 다음 실험을 위한 신뢰할 만한 지침이 아니라 추측일 뿐입니다.”라고 설명했다. Duraisamy는 이 세 간극이 서로를 강화하므로, 매개변수 수나 추가 연산을 늘리는 것보다 간극 해소가 중요하다고 강조한다. 이를 위해서는 모델링·실험·인간 판단을 하나의 피드백 루프로 통합하는 아키텍처가 필요하다고 주장한다.

Duraisamy 교수의 해법은 모델·실험·인간이 끊임없이 대화하는 능동적 추론 스택이다. 최상단에는 상용 또는 오픈소스 범용 언어 모델이 위치하고, 이들 위에 재료·생물학 등 도메인 기반 모델이 배치된다. 공유 임베딩 공간을 통해 도메인 모델들은 아이디어를 비교하고, 예상치 못한 연관성을 발견할 수 있다.



Karthik Duraisamy 제공 슬라이드

다음 계층은 문헌·시뮬레이션 결과·새 관측을 담는 동적 지식 그래프다. 그 아래에는 불확실성 정량화(UQ), 최적 실험 설계, 제어 이론에 기반한 오케스트레이션 계층이 있어 새로운 시뮬레이션이나 벤치 테스트를 계획한다. 최하단에서는 실험·시뮬레이션·정형화된 증명이 증거를 생성해 상위로 전달된다. Duraisamy는 정형 검증된 코드·정리가 신뢰 가능한 구성 요소로서 오류 확산을 줄일 수 있다고 본다. 각 계층이 서로를 조정함으로써, 시스템은 행동을 통해 학습하는 '월드 모델(world model)'로 발전할 수 있다.

그러나 이 추론 스택은 자동조종으로 운용될 수 없다. 발견을 주도하고 가드레일을 설정하기 위해서는 여전히 인간의 감독이 필요하다. 연구자는 언제 이론을 수정할지, 언제 시뮬레이션을 신뢰할지, 언제 잘못된 경로를 버릴지 결정해야 한다. 이러한 판단은 구조적 차원에서 항구적 요소다. Duraisamy는 과학이 '계산 가능한 현실'의 주머니 속에 존재한다는 점을 들어, 모델이 다양체의 경계—알려진 영역의 끝과 미지의 시작—to를 인지해야 한다고 경고한다. 실제로 AI 시스템은 자신의 맹점을 표시하고, 검증된 범위를 벗어날 때마다 새로운 데이터 수집이나 인간 검토를 따르도록 설계되어야 한다.

Duraisamy는 AGI를 당장의 획기적 돌파구가 아닌 장기적 목표로 보며 강연을 마무리했다. "연산을 통해 지능을 재구성하는 일은 실용적이라고 생각합니다. 하지만 그것이 '순수하게 연산적'으로만 이뤄지려면, 우리가 아직 갖추지 못한 올바른 연산적 추상화가 필요합니다. 여기서 핵심은 반사실적 추론입니다. 즉, 우리의 근사적 월드 모델이 현실 세계와—인간의 감독 하에—상호작용하는 과정 사이의 상호작용이 도움이 됩니다."

테이크어웨이

세션 전반에 걸쳐 반복된 메시지는 하나였다. '더 크게'만으로는 충분하지 않다. 과학용 AI의 다음 도약은 지식 그래프, 오케스트레이션 계층, 자동화된 랩, 그리고 계획–실행–학습을 폐쇄 루프로 수행하는 도메인 인식 에이전트 같은 아키텍처

요소에서 나온다. 동시에 모든 연사는 취약한 추론 사슬, 불충분한 구조화 데이터, 신뢰와 책임에 관한 의문 등 현재의 과제를 강조했다. TPC 커뮤니티 앞에 놓인 기회는 분명하다. 1조 매개변수 모델을 신뢰할 수 있는 과학적 파트너로 만드는 '연결 조직'을 구축하는 동시에, 빠르고도 책임 있는 발견을 가능하게 하는 가드레일을 설계하는 일이다.

Plenary Session 4: Multimodal Data, Evaluation, and Non-LLM Model Architectures (수요일, 7/30, 11:00)

통제력을 희생하지 않고 AI 기반 과학적 발견의 속도를 높이려면 어떻게 해야 할까? 언어 모델의 답을 학습 데이터로 추적할 수 있을까? AI가 텍스트가 아닌 지도를 해석할 때 공정성은 어떻게 보장할까? 그리고 이런 시스템이 인간과 전혀 다른 방식으로 실패한다면 어떻게 대비해야 할까? 이는 TPC25 본회의 세션 4에서 네 연사가 던진 질문들이다.

ORNL의 Prasanna Balaprakash는 AI가 과학 연구를 빠르고 정확하게 견인하는 방법을 논의했다. AI2의 Jiacheng Liu는 모델 출력을 학습 데이터에 연결하는 추적 도구를 소개했다. BSC AI 연구소의 Ricardo Baeza-Yates는 피해·허위정보·낭비에 대한 분야의 책임을 강조했다. AIST의 김경숙(Kyoung-Sook Kim)은 지공간 시스템의 공정성—데이터 격차와 지리적 편향이 왜곡·불평등을 어떻게 낳는지—에 관한 근본적 질문을 제기했다.

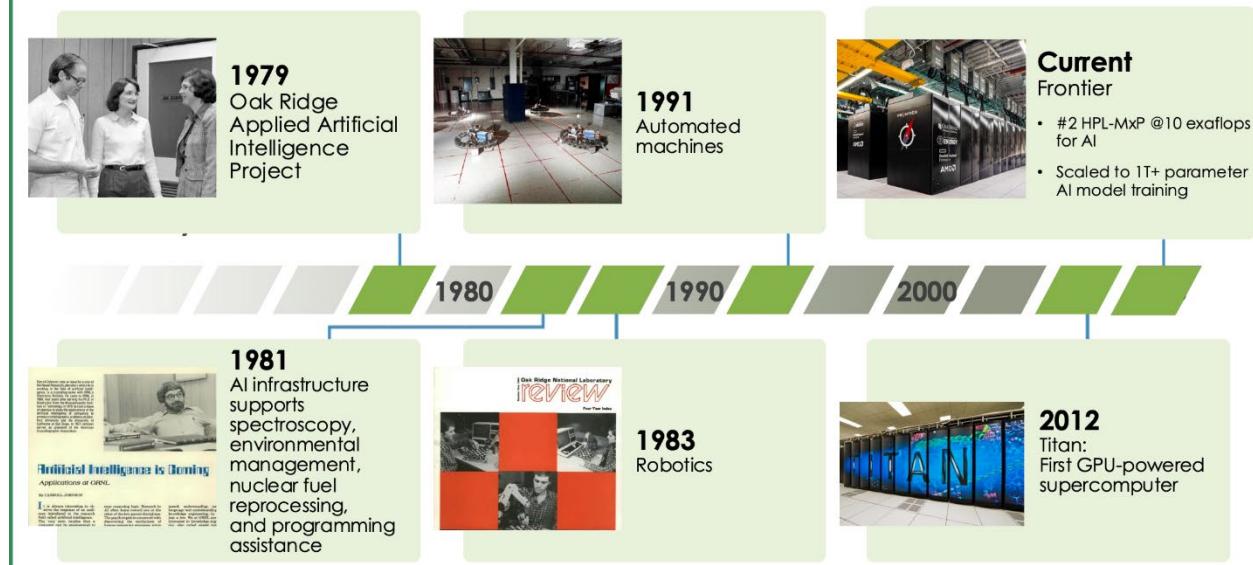
ORNL의 다초점 AI 이니셔티브 ORNL's Multi-Focused AI Initiative

ORNL의 AI 연구는 1979년 최초의 응용 인공지능 프로젝트(Applied Artificial Intelligence Project)로 거슬러 올라가며, 그 과정에서 GPU를 탑재한 최초의 DoE 슈퍼컴퓨터 Titan(2012)도 구축했다. Titan은 8,688개의 CPU와 동일한 수의 GPU를 탑재했다. 현재 ORNL의 최상위 머신은 Frontier로, 약 37,632개의 AMD GPU를 장착하고 있다. 2025년 6월 기준 Top500 목록에서 2위를 차지했다. ORNL의 AI 프로

그램 책임자이자 데이터·AI 시스템 부문 책임자이며 현재 ORNL AI 이니셔티브 리더인 Prasanna Balaprakash는 발표에서 ORNL의 AI 유산을 간략히 소개하고 현재 우선 순위를 논의했다. 그는 “오크 리지는 과학 분야에 AI를 활용해 온 풍부한 역사를 갖고 있습니다. 최초의 AI 프로젝트는 1979년에 시작됐고, 1981년에는 AI가 과학적 발견, 분광학, 환경 관리 등을 어떻게 혁신할지에 대한 리뷰 논문이 발표됐습니다”라고 말했다. 당시 시스템은 오늘날의 데이터 기반 모델이 아니라 규칙 기반 전문가 시스템으로 구상됐다.

LLM—특히 2022년 11월 공개된 ChatGPT—의 등장은 IT 전반에 충격을 주며 재검토를 촉구했다. Balaprakash는 “LLM 분야에서 정말 많은 변화가 일어나고 있다고 느꼈습니다. 2년 전 일입니다. 우리는 우선순위를 재정립하고, 업계 발전 속에서도 중요성을 잊지 않으면서 장기적으로 지속 가능한 근본적 진전이 무엇인지 논의하고자 했습니다”라고 밝혔다

ORNL has a rich history leveraging AI for science

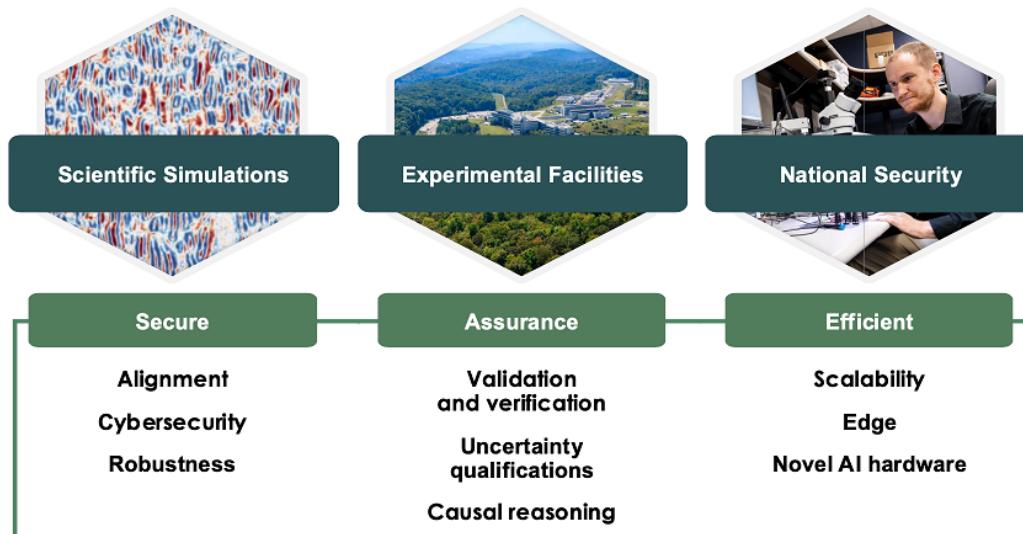


Prasanna Balaprakash 제공 슬라이드

“우리는 확신을 가지고 AI 모델을 구축해야 하며, 그 과정에서 검증·확인, 불확실성 정량화(UQ), 인과 추론, 그리고 궁극적으로 효율성까지 고려해야 합니다. 거대한 머신을 보유하고 있지만, 동시에 이러한 머신에서 대규모 모델을 효율적으로

확장하는 방법도 고민해야 합니다. 또한 엣지에서 작동할 수 있는 소형 모델—배포가 용이하고, 과학 실험을 가속화하며, 새로운 AI 하드웨어를 탐색해 과학 워크플로에 통합할 수 있는 모델—도 중요합니다. ORNL AI 이니셔티브는 과학 시뮬레이션, 실험 시설, 국가 안보의 세 영역에서 '안전하고 확실하며 효율적인 AI'를 발전시키는데 초점을 맞추고 있습니다. 현재 약 50명의 연구자가 참여해 15개 첨단 AI 프로젝트를 수행 중이며, 매우 학술 지향적인 팀입니다. 우리는 연구실과 커뮤니티 전반에서 활용될 역량을 구축하고자 합니다." Balaprakash는 다중 모달 데이터 통합, 불확실성 측정, 견고성 관련 여러 사례를 제시했다.

ORNL's AI initiative Secure, assured, and efficient



The initiative's portfolio comprises 15 advanced AI projects and involves over 50 researchers from 5 different directorates across the lab.

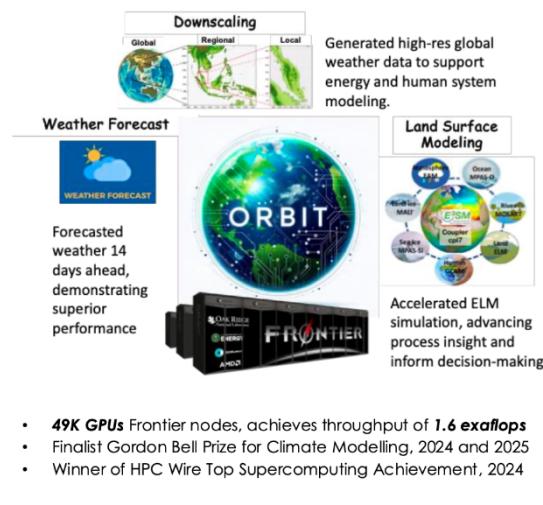
Prasanna Balaprakash 제공 슬라이드

"흥미로운 분야 중 하나는 과학용 파운데이션 모델, 특히 과학 시뮬레이션을 가속할 수 있는 모델입니다. 2년 전 우리는 산업계가 주요 역할을 하지 않는 다른 경로를 모색하는 데 과감히 투자했습니다. 예컨대 핵융합 시뮬레이션처럼 대규모 시공간 데이터를 다루는 방향입니다. 모두가 시공간적 거대 데이터이며, 이 영역에서 어떻게 도약할지를 결정하는 일입니다."

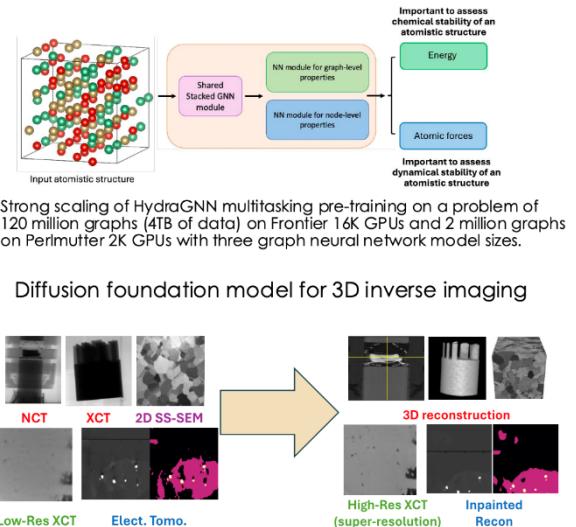
그 성과 중 하나가 ‘지구 시스템 예측성을 위한 오크리지 기반 파운데이션 모델(Oak Ridge Base Foundational Model for Earth System Predictability)’ 개발입니다. “작년 우리는 49,000개의 GPU로 대규모 학습을 수행해 엑사스케일급 처리량을 달성(또는 초과)했습니다. 고속 시공간 모델로는 전례가 없는 일입니다. 최대 100억 개 매개변수의 모델을 구축했으며, 이 유형으로는 처음입니다. 올해는 이들 시뮬레이션을 축소(surrogatization)하는 작업을 확대했고, 계산 시간을 수 자릿수(orders of magnitude) 단축하는 데 주력하고 있습니다.”

AI foundation models are accelerating scientific applications

Spatiotemporal foundation models for earth system predictability



Graph foundation models for atomistic materials modeling



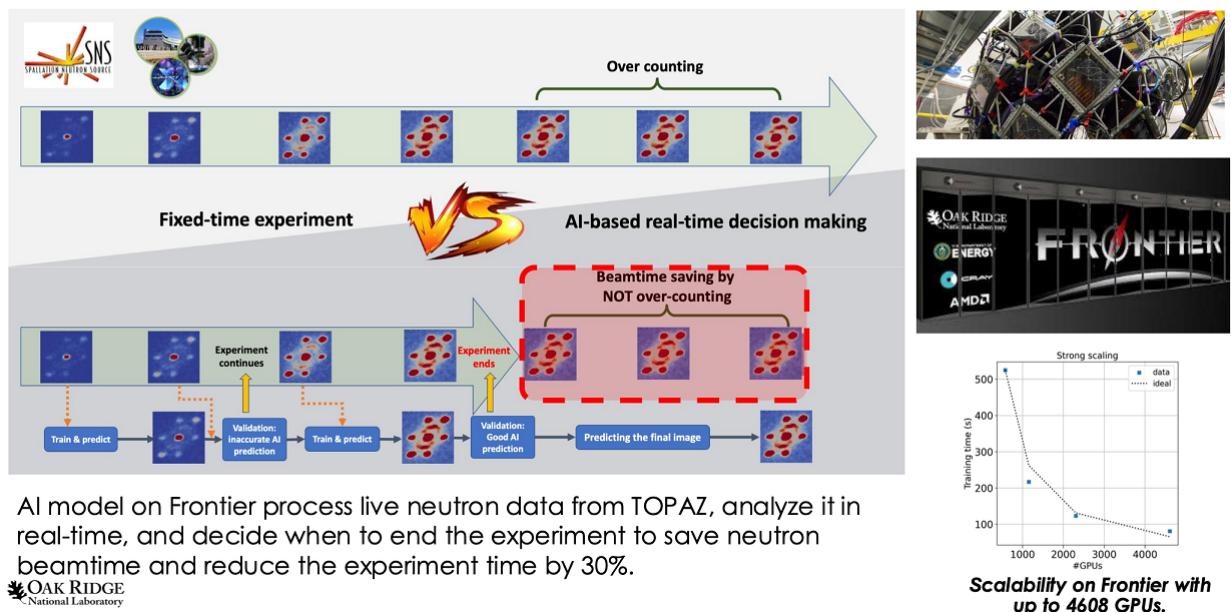
Prasanna Balaprakash 제공 슬라이드

Balaprakash는 그래프 모델, 불확실성 정량화, 그리고 실험 장비를 Frontier에 연결해 실시간으로 데이터를 처리·관리하는 노력 등을 특별히 언급했다.

- 그래프 모델: “우리는 테라바이트 규모의 재료과학 오픈데이터를 기반으로 하는 대규모 그래프 기반 모델에 크게 투자하고 있습니다. **Hugging Face**에는 수백만 개의 LLM 파생 모델이 있지만, 재료과학용 그래프 기반 모델은 극히 드뭅니다. 우리는 그중에서도 가장 큰 축에 속하는 그래프 모델을 구축 중이며, 모든 스크립트와 함께 누구나 쉽게 사용할 수 있도록 공개하고 있습니다.”

- 불확실성(Uncertainty): “불확실성 정량화(UQ)는 과학용 AI 모델의 필수 요소입니다. 누구도 ‘UQ를 쓰고 싶지 않다’고 하지는 않죠. 복잡한 AI 모델에 UQ 를 적용할 수 있도록, 가져다 바로 쓸 수 있는 UQ 라이브러리·기술을 개발하고 있습니다.”
- Instrument-to-Frontier 링크: “**ORNL 파쇄 중성자원 시설의 TOPAZ(검출기)** 실험은 방대한 데이터를 생성합니다. 우리는 TOPAZ를 Frontier 슈퍼컴퓨터에 연결해 4,000 개의 GPU 로 수 초 내 처리하고, 실험에 대한 실시간 인사이트를 제공합니다. 이 분석으로 실험 시간을 절약할 수 있습니다. 전통적 방식대로면 끝까지 진행하느라 불필요한 계산이 발생할 수 있지만, AI 기반 조종(control)을 적용하면 모델이 향후 결과를 예측하고, 모델이 유효한 한 이를 지속적으로 활용할 수 있습니다. 필요한 시점에 실험을 종료해 다른 사용자를 위해 자원을 회수할 수 있습니다.”

Frontier-scale real-time AI enables smarter, faster experiments



Prasanna Balaprakash 제공 슬라이드

거대 언어 모델의 블랙박스 열기

Allen Institute for AI(AI2)의 연구원 지아청 류(Jiacheng Liu)는 LLM 응답을 그 기반이 된 수조 토큰 규모의 코퍼스와 연결하는 추적 시스템 '**OLMoTrace**'를 공개했

Trillion Parameter Consortium 2025 (TPC25) 리뷰
: 비전, 역사, 그리고 글로벌 Scientific AI 리더십

다. 이 도구는 완전 공개 OLMo 모델군을 호스팅하는 AI2 Playground에 탑재되어 있다. 사용자가 답변을 생성한 뒤 한 번만 클릭하면, 학습 세트에 '그대로' 존재하는 긴 텍스트 구간이 하이라이트되고 주변 소스 문서가 나열된다. 일반적인 450토큰 응답의 경우 약 4.5초 내 결과가 산출되어 대화형 탐색에 충분히 빠르다.

OLMoTrace는 이전 검색 프로젝트에서 개발된 최적화 접미사 배열 인덱스를 사용한다. 이 인덱스에는 사전학습·중간학습·사후학습 단계에서 추출된 30억 편 이상의 문서와 4조 토큰이 포함된다. 모델 출력의 각 위치에서 '가장 긴 정확 일치'를 탐색함으로써 비용이 큰 부분문자열 검색을 피하고, 다중 코어 병렬 처리가 가능하다. 순위화 단계에서는 길고 고유한 스팬을 우선 표시한 다음, 전체 문서를 불러 맥락을 파악한다.

The screenshot shows the AI2 Playground interface. On the left, there's a sidebar with a 'Playground' logo, a 'New chat' button, a 'Thread history' section, and a 'FAQ' button. The main area has a search bar at the top with 'Model: OLMo 2.32B Instruct'. Below it, a question is asked: 'Who is Celine Dion?'. The response is provided by the OLMo model, which describes Celine Dion as a Canadian singer known for her powerful voice and wide vocal range. It mentions her birth on March 30, 1968, in Charlemagne, Quebec, Canada, and her career highlights, including winning the 1988 Eurovision Song Contest. The interface also shows 'OLMoTrace' results on the right, listing documents from the training data that have exact text matches with the model's response, such as pre-training documents from 'olmo-mix-1124' and post-training (SFT) documents from 'tuly-3-sft-olmo-2-mixture'.

Jiacheng Liu 제공 슬라이드

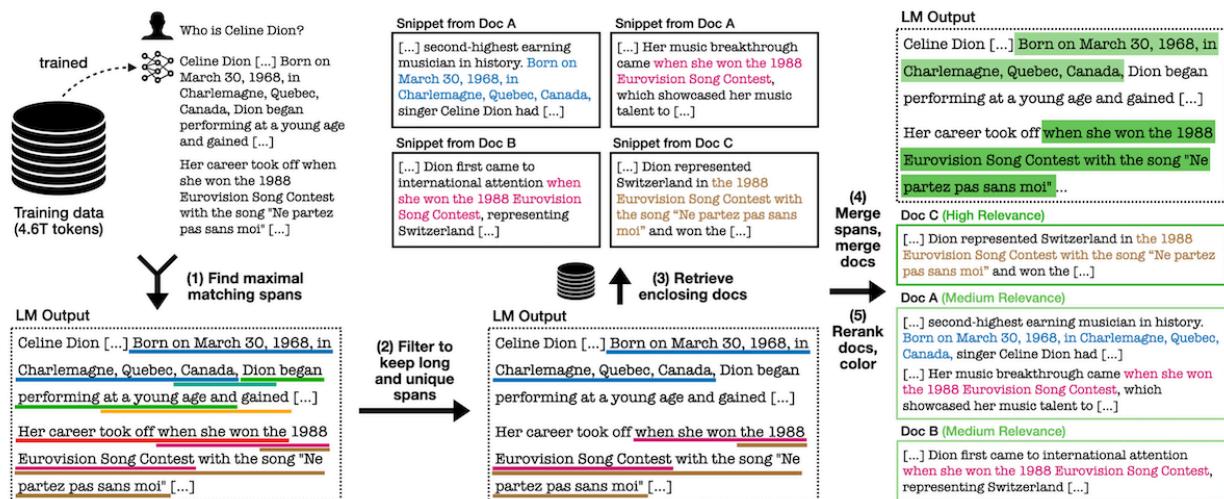
Liu는 이런 데이터 중심 관점이 모델 동작에 관한 실질적 요구를 어떻게 충족하는지 보여줬다. "사실 확인이 가장 직접적입니다. 예를 들어 모델이 시애틀의 스페이스 니들이 1962년 세계박람회를 위해 건설됐다고 답하면, OLMoTrace는 신뢰할

수 있는 출처에서 일치 문장을 찾아 신속히 검증할 수 있게 합니다. 또 이 도구는 환각(hallucination)의 근원을 드러냅니다. 한 사례에서 모델은 조작된 코드 실행 결과를 생성했는데, 추적 결과 학생이 코드를 실제로 실행하지 않고 출력만 제시한 학습 대화가 확인됐습니다. 이는 모델이 데이터로부터 나쁜 습관을 학습했음을 시사합니다.”

“과학자에게 중요한 가치는 투명성입니다. 연구팀은 답변을 감사하고 출처를 검토하며, 인용 문서가 해당 분야의 기준을 충족하는지 판단할 수 있습니다. 데이터 추적 가능한 출력은 의료, 기후 모델링 등 민감 영역에서 결과의 정당성을 요구하는 새로운 AI 거버넌스 규정 준수를 간소화합니다.” 그는 이 연구가 ‘행동을 신경 회로에 매핑’하는 기계론적 해석가능성 연구를 보완한다고 덧붙였다. 가중치 수준 분석과 코퍼스 수준 추적을 결합하면, 연구자는 모델의 추론 방식과 오류 발생 지점을 더 잘 이해할 수 있다.

“이 프로젝트는 AI2의 오픈사이언스에 대한 헌신 위에 서 있습니다. 가중치 부터 학습 스크립트까지 OLMo 2의 모든 구성요소가 공개돼 있어, 누구나 추적 파이프라인을 복제하거나 다른 모델에 적용할 수 있습니다. 이러한 개방성 덕분에 OLMoTrace는 올해 오스트리아 빈(비엔나)에서 열린 ACL에서 최우수 데모상을 수상했습니다.” Liu는 앞으로 관련성 순위를 개선하고, 의역된 출처에 대한 근사 매칭을 통합하며, 역추적이 데이터셋 큐레이션에 기여하는 방식을 탐구할 계획이라고 밝혔다.

The full pipeline



Jiacheng Liu 제공 슬라이드

차별에서 허위 정보까지: '무책임한 AI'에 대한 베이자-예이츠의 견해

BSC AI 연구소 소장 Ricardo Baeza-Yates는 오늘날 AI 시스템의 신뢰성·안전

성·공정성을 지속적으로 훼손하는 실패와 간과의 양상을 '무책임한(Irresponsible) AI'라는 틀로 비판했다. 그는 자동화된 차별, 사이비 과학, 불공정 전자상거래, 자원 낭비, 인간의 무능 등 AI 오용을 분류하며, 부실한 모델 평가에서부터 생성형 AI가 초래하는 전 지구적 결과에 이르기까지 폭넓은 비판의 토대를 제시했다.

"가장 먼저 기억해야 할 것은 데이터와 모델은 세상의 대리물이라는 점입니다. 그것들은 근사치일 뿐입니다."라고 Baeza-Yates는 서두에서 강조하며, '이해했다'고 해서 곧바로 '정확하다'고 여기는 혼동을 경계했다. 그는 AI가 종종 인간 추론의 거울로 오해되지만, 실제로는 틀릴 수 있고 때로는 위험한 예측을 반복적으로 축적하는 엔진일 뿐이라고 지적했다. 또한 이러한 기술을 의인화하는 데 대해 경고하며, '윤리적(ethical) AI'나 '신뢰할 수 있는(trustworthy) AI' 같은 표현은 근본적으로 결함이 있다고 주장했다. "기계는 윤리적일 수 없습니다."라고 그는 말하며, 신뢰와 윤리는 본질적으로 인간의 속성임을 강조했다. 이런 식의 정의는 시스템 설계자의 책임을 사용자에게 전가해 오해를 낳고 위험하다고 그는 덧붙였다.

Baeza-Yates는 생성형 AI가 야기하는 고유한 피해를 열거했다. “이러한 실패의 대가는 점점 커지고 있습니다. 허위 정보의 급속한 확산부터 해결되지 않은 저작권 분쟁, 심화되는 정신 건강 문제까지 말입니다.” 강연에서 가장 소름 돋는 순간은, 그가 챗봇이 자살 사건에 연루된 실제 사례들을 소개하며 책임과 개입의 시급성을 제기했을 때였다. “비인간적 오류는 AI가 저지를 수 있지만, 인간은 결코 저지르지 않을 오류입니다. 문제는 우리가 그런 오류에 대비하지 못했다는 것입니다.” 그는 자율주행차가 보행자를 치고도 주행을 계속하는 등 인간이 내릴 수 없는 결정을 내리는 사례를 언급하며, 이러한 오류가 통계적 이상치가 아니라 AI의 설계·평가·배포 전 과정에 내재한 시스템적 위험이라고 강조했다.

그는 성공 지표에 대한 현재 접근이 문제의 일부라고 지적했다. “정확도를 측정하는 데서 그치지 말고, 실수가 발생했을 때 무엇이 일어나는지를 측정해야 합니다.”

또한 그는 ‘AI 민주화’라는 통념에 반박했다. 전 세계에는 7,000개가 넘는 언어가 존재하지만, 대부분의 주요 모델이 다루는 언어는 200개 미만이어서 지구 인구의 상당수가 사실상 AI 접근에서 배제된다. 여기에 디지털 격차, 연령 제한, 문화적 차이까지 겹치면 ‘공평한 AI’라는 개념은 빠르게 흔들린다.

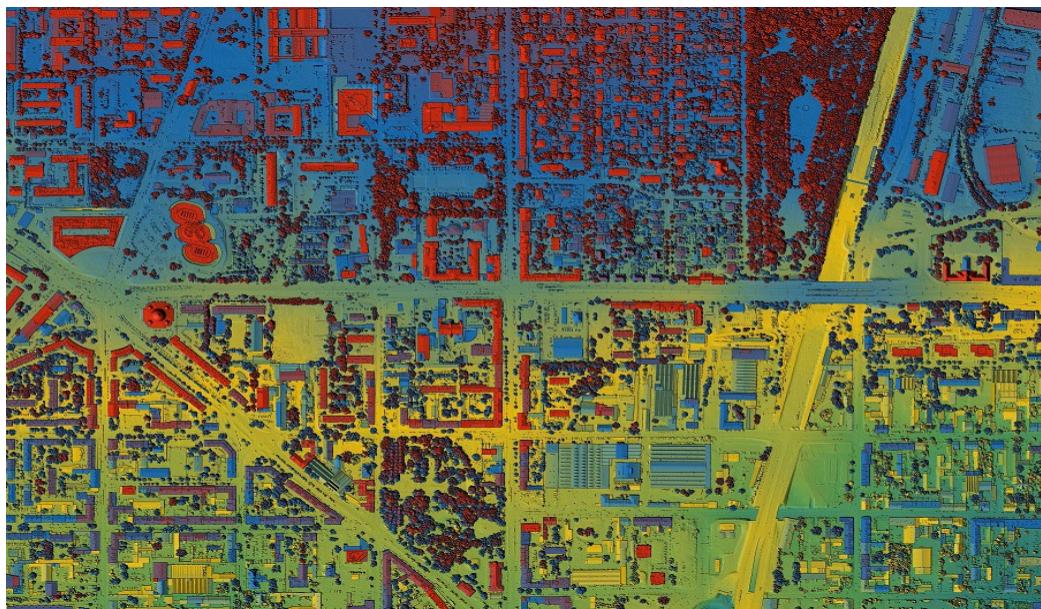
Baeza-Yates는 AI 시스템의 한계와 우리가 투사하는 착각을 분명히 하며 강연을 마무리했다. “명심해야 합니다. 그들은 보고·읽고·쓰지 못합니다. 그러면 ‘이해’가 전제되어야 합니다.” 그는 이어 말했다. “그들은 모든 것을 ‘예측’합니다. 본질적으로 늘 환각합니다. 다만 대부분의 경우, 그 환각이 우연히 정확할 뿐입니다.”

지리공간 기초 모델의 공정성

지공간 AI(GeoAI)는 위성 이미지와 도시 인프라에서부터 환경 모니터링, 재난 대응에 이르기까지 물리 세계를 이해하는 데 점점 더 폭넓게 활용되고 있다. TPC25에서 일본 산업기술종합연구소(AIST)의 김경숙 부소장은 지리공간 기초 모델의 공정성에 관한 설득력 있는 강연을 전했다.

김 박사는 지리공간 데이터베이스, 위치 기반 서비스, 사이버-물리 클라우드

컴퓨팅 분야에서 경력을 쌓았으며, GIS를 빅데이터·AI와 통합하는 방법을 10년 이상 연구해 왔다. 이번 발표에서 그는 지공간 AI 시스템 개발의 시급 과제로, 모델의 '전체 수명 주기'에 걸친 공정성 확보를 강조했다. 여기서 공정성이란, 데이터가 풍부한 지역뿐 아니라 다양한 지역과 조건에서도 지공간 기반 모델이 일관되고 신뢰롭게 작동하도록 보장하는 것을 뜻한다.



김경숙 박사 제공 슬라이드

그는 학습 과정에서의 불균등한 데이터 수집, 공간적 범위 격차, 편향된 가정이 현실에 어떤 영향을 미칠 수 있는지—특히 AI가 계획·인프라·자원 배분에 개입 할 때의 파급효과—를 탐구했다. 핵심 메시지는 분명했다. GeoAI의 형평성은 단순히 '결과'의 문제가 아니라, 시스템을 어떻게 설계하느냐의 문제다.

언어·이미지 AI의 공정성에는 많은 관심이 쏠렸지만, Kim은 지리공간 시스템이 고유한 문제에 직면해 있음을 지적했다. "정확한 정의가 없습니다. 측정 지표도 충분하지 않습니다. 텍스트·이미지 데이터와 달리 위치 데이터는 수집 지점과 현장에서 벌어지는 상황에 강하게 결부됩니다. 한 지역에서 학습된 모델이 다른 지역에서는 핵심 패턴을 오인하거나 놓칠 수 있습니다. 따라서 GeoAI의 공정성은 모든 곳에 동일하게 적용될 수 없으며, 지역·인구 특성, 그리고 이를 표현하는 데 사용 가능한

데이터의 차이를 반영해야 합니다.”

GeoAI가 사용하는 데이터 특성상 우려는 더 커진다. 위성 이미지, 휴대폰 신호, 센서 판독, GPS 로그, 위치 태그가 붙은 동영상 등 다양한 입력에 기반하기 때문이다. Kim은 이러한 데이터가 태양광 인프라 파악, 재난 대응 조정, 자율 배송, 도시 계획 지원 등 폭넓은 분야에서 활용됨을 설명했다.

반면 그는 이러한 시스템이 초래할 수 있는 사각지대도 우려했다. “데이터 선택, 라벨링, 처리 방식 같은 초기 결정은 모델을 만들기 훨씬 전부터 편향을 야기할 수 있습니다. 공정성이 의미 있으려면, 데이터에 ‘누가 반영되고 누가 배제되는지’를 면밀히 살펴봐야 합니다.”



Relationship between Fairness and Data Quality



- Removing biased data in the training dataset → Is it good data quality?
- Cleaning inaccurate, incomplete, and inconsistent data → Is it a fair dataset?
- Trade-off
 - Bias mitigation can degrade data quality; data quality improvement can reduce fairness.
 - Need for guidance
 - Data-driven decision-making requires domain-specific strategies balancing the two objectives.
 - **An international standard provides a common understanding and the same interpretation.**

김경숙 박사 제공 슬라이드

김박사는 AI 시스템이 공간 데이터를 해석·활용하는 책임이 커질수록, 공정성 평가를 돋는 공유 프레임워크 구축이 필요하다고 강조하며 발표를 마무리했다. 그는 공정성과 데이터 품질에 대한 보다 일관된 정의를 향한 중요한 단계로, 새로운 ISO 추진을 포함한 국제 표준을 제시했다. “다만 이런 도구들은 해법의 일부일 뿐입

니다. 공정성은 지리·역사·사회적 복잡성 등 다양한 요인에 의해 형성되는 맥락의존적 요소입니다. AI가 인프라·계획·공공 의사결정에 더 깊이 스며들수록, 이러한 시스템의 설계와 적용 과정에서 형평성을 보장하는 일이 그 어느 때보다 중요해질 것입니다.”

테이크어웨이

이번 논의는 AI 연구 전반의 변화를 시사한다. 모델이 커질수록 신중한 설계의 필요도 커진다. 벤치마크 성능을 넘어, 데이터의 출처, 결과 평가 방식, 실제 환경에 미치는 영향을 이해하는 일이 중요하다. 앞으로의 진전은 더 ‘똑똑한’ 알고리즘뿐 아니라, 시스템을 얼마나 책임감 있고 포괄적으로 설계·구축·적용하느냐에 달려 있다.

Plenary Session 5: Science Updates from Key TPC Leaders (목요일, 7/30, 14:00)

TPC25 세션 '핵심 TPC 리더들의 과학 업데이트'에서 두 명의 저명한 연사가 과학 분야에서 대규모 언어 모델의 미래에 대해 서로 다르지만 상호보완적인 관점을 제시했다. Argonne National Laboratory의 Franck Cappello는 AI 연구 보조원을 평가하는 새로운 프레임워크 EAIRA를 소개하며, 지속적 감독 없이 복잡한 과학 과업을 처리할 수 있도록 신뢰를 구축하기 위해 추론·적응성·분야별 기술을 어떻게 측정 할지에 초점을 맞췄다. 한편 일본에서는 Tokyo Institute of Technology의 Rio Yokota 교수가 야심 찬 투트랙 LLM 개발 계획을 설명했다. LLM-jp 컨소시엄은 일본 내 최상급 슈퍼컴퓨터에서 대규모 모델을 학습하고, 소규모 Swallow 프로젝트는 더 가벼운 아키텍처와 빠른 반복을 실험한다. 두 프로젝트는 과학에서 LLM의 미래가 '더 큰 모델' 그 이상에 달려 있음을 보여준다—핵심은 신뢰 가능한 모델을 만들고, 이를 활용할 인프라와 협업 체계를 구축하는 것이다.

LLM 연구 어시스턴트를 신뢰하는 데 필요한 것은 무엇입니까?

LLM이 과학 분야 연구 어시스턴트로서 무엇을 해주길 바라는가? 그리고 이러한 새로운 AI 연구 어시스턴트를 어떻게 효과적으로 평가할 것인가?

Characteristics of an “AI scientific assistant” that we need to/must evaluate

An AI-based system with:

- **Scientific skills**
 - Reasoning, math, literature understanding, integrity
- **Effective assistance (no hallucination!, consistency in responses)**
 - **Correct** for all different tasks related to scientific activities
- **Relevance to human and environment interaction modalities (communication skills)**
 - Understanding command (semantic of it), interface with tools and devices
- **Degree of autonomy**
 - From repeating learned workflows to developing the workflow.
 - **Capable of hypothesis generation**
- **Safety for the community**
 - Cannot be used to harm others: e.g. design harmful substances

Frank Capello 제공 슬라이드

Argonne National Laboratory의 AuroraGPT 평가팀을 이끄는 Franck Cappello는 TPC 본회의 발표 'EAIRA: 연구 어시스턴트로서 LLM을 평가하는 방법론'에서 이 두 가지 핵심 질문을 파고들었다.

전반적으로 우리는 AI 동료에 대해 갈수록 더 많은 것을 기대한다. 초기에는 과학 문헌을 빠르게 걸러 유용한 정보를 제공하는 용도로 AI를 상상했다면, 이제는 문헌 선별, 가설 생성, 코드 작성, 실험 워크플로 제안(나아가 실행)까지 수행하는 거의 완벽한 파트너를 원한다.

"하지만 그들의 추론 능력과 지식 능력을 어떻게 검증할 수 있을까요? 정말로 문제를 '이해'하고 있는지 어떻게 시험할 수 있을까요?"라고 Cappello는 물었다.

"망원경·현미경·광원 같은 도구를 만들 때 우리는 그 작동 원리를 잘 압니다. 그러나 여기서는 그렇지 않습니다. 아직 블랙박스이기 때문입니다."

"모델이 제공하는 정보를 일일이 검증하며 시간을 쓰고 싶지 않습니다. 우리는 그 결과를 신뢰하고 싶습니다."라고 Cappello는 말했다. "모델은 인간의 지시를 이해할 뿐 아니라 실험실의 도구·장치와 연동되어야 하며, 일정 수준의 자율성도 갖춰야 합니다. 워크플로를 반복·학습하는 접근도 가능하지만, 우리가 진정으로 원하는 것은 '양질의 가설'을 생성하는 능력입니다." 그 지점에 이르려면 새로운 평가 도구가 필요하다. 최근 LLM 진전을 간략히 짚은 뒤, Cappello는 효과적인 평가 방법론을 심층 검토했다. 현재 주된 평가는 객관식(MCQ)과 주관식 서술형 두 측이지만, 두 방식 모두 한계를 안고 있다고 지적했다.

Evaluation Methodology: What to measure

What to measure:

- **Knowledge Extraction, Retrieval, Distillation, Synthesis :** LLM is provided with a question and a truthful answer is expected
- **Text grounded:** Answer is expected fully grounded on peer reviewed references to support response.
- **Reasoning:** LLM is expected to solve deductive (Prove a theory or hypothesis from formal logic and observations), inductive (validate/explain observations from theories) problems
- **Creativity:** A creative answer is expected from a question or instruction: e.g. find a solution to open scientific questions.
- **Thoughtful dialogue, Coding, etc. → Agentic aspects**
- **safety, usability, robustness (adversarial attacks), cost/energy (training, inference)**

Criteria for all of the above:

- **Correctness** (of facts, results),
- **Accuracy** of solutions AND reasoning,
- **Reliability** (consistently good in quality or performance),
- **Speed** (how fast to produce a response), O1-pro can take minutes to respond to a single prompt
- **Consistency** (slight prompt variations should not generate significantly different responses),
- **#shots** (how many examples are needed for good quality),
- **Extent of Prompt Engineering.**

Challenges:

- **AI model capabilities (knowledge, reasoning, correctness) are increasing at extreme pace**
- **Evaluation should be done in enclaves** (AI models should not be trained on the tests)
- **Generation and validation of large corpus** of difficult enough tests



Frank Capello 제공 슬라이드

"연구자에게 객관식 문제를 대량 생성하도록 요구하면 시간이 많이 듭니다. 그럼에도 이 과제들은 매우 중요하므로 계속 고려해야 합니다."라고 Cappello는 말했다. "현재 이용 가능한 벤치마크는 지나치게 일반적이고, 특정 도메인에 국한되지 않

으며, 정적이라 시간 경과에 따른 ‘오염’ 문제가 발생합니다. 즉, 모델 학습에 벤치마크가 포함되면 그 자체로 의미가 퇴색할 수 있습니다. 이 점을 반드시 감안해야 합니다.” 개방형 응답 역시 채점이 어렵지만 여전히 중요하다. 그는 평가 방법론이 고려해야 할 다양한 기준(아래 슬라이드)을 상세히 설명했다.

End-to-End Eval: Field Style Experiment

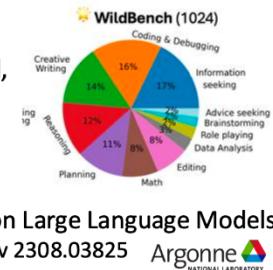


Lab style experiments: Human evaluation, tries to solve 1 specific problem, compare different models, guide LLMs (requires efforts: some prompt engineering),

Field style experiments: Automatic evaluation, capture what researchers actually ask, much broader diversity of Q&As, large diversity of prompt engineering, statistical evaluation

Several papers on this topic (but not for Science activity)

- **WildBench:** Benchmarking LLMs with Challenging Tasks from Real Users in the Wild, B. Y. Lin and Y. Deng and K. Chandu and F. Brahman and A. Ravichander and V. Pyatkin and N. Dziri and R. Le Bras and Y. Choi, 2024, arXiv 2406.04770
- **HaluEval-Wild:** Evaluating Hallucinations of Language Models in the Wild, Zhiying Zhu and Yiming Yang and Zhiqing Sun, 2024, arXiv, 2403.04307
- **“Do Anything Now”:** Characterizing and Evaluating In-The-Wild Jailbreak Prompts on Large Language Models Xinyue Shen and Zeyuan Chen and Michael Backes and Yun Shen and Yang Zhang, 2024, arXiv 2308.03825



Frank Capello 제공 슬라이드

Cappello는 이어 Argonne에서 개발 중인 EAIRA를 소개했다. EAIRA는 엄격하고 재현 가능한 평가 절차를 확립하기 위한 프레임워크다. 2월, Cappello와 여러 기관의 동료들은 arXiv에 사전 공개 논문(EAIRA: Establishing a Methodology for Evaluating AI Models as Scientific Research Assistants)을 게시했다. 그는 “여기(아래 슬라이드)에 제안 방법론이 나와 있습니다. 여러 기법을 결합합니다. 즉, 객관식 문항을 활용합니다—네, 객관식 문제를 생성합니다. 또한 개방형 응답(오픈-엔디드) 벤치마크를 포함합니다. 그리고 새롭게 두 가지, ‘실험실 스타일’ 평가와 ‘현장 스타일’ 평가가 있습니다.”라고 설명했다. 이 방법론은 네 가지 주요 평가 클래스를 통합한다. 사실적 기억력을 평가하기 위한 객관식 문항:

1. 고급 추론 및 문제 해결 기술을 평가하기 위한 열린 응답

2. 통제된 환경에서 연구 보조원으로서의 역량을 정밀 분석하는 '실험실 스타일' 평가
3. 다양한 과학 분야·응용에서 연구자-LLM 상호작용을 대규모로 포착하는 '현장 스타일' 평가

EAIRA: Multi-faceted Eval Methodology

Proposed Methodology				
Techniques	MCQ Benchmarks	Open Response Benchmarks	Lab Style Experiments	In the Wild Field Style Experiments
Main Goal	Testing knowledge breadth, basic reasoning	Testing knowledge depth, planning, reasoning	Realistic testing	Realistic trend analysis and weakness diagnosis
Problem Type	Predetermined , Fixed Q&As with known solutions	Predetermined , Fixed Free-Response Problems with known solutions	Individual Human Defined Problems with unknown solutions	Many Human Defined Problems with (un)known solutions
Verification	Automatic response verification	Automatic or Human response verification	Humans detailed response analysis	Scalable automatic summary of human response
Examples	Astro, Climate, AI4S (multi-domain), Existing Benchmarks	SciCode, ALDbench	see "lab style experiments"	see "field style experiments"
Cross Cutting Aspects	← Trust and Safety (ChemRisk), Uncertainty Quantification, Scalable Software Infrastructure (STAR) →			

Methodology consisting of 4 complementary evaluation techniques to comprehensively assess the capabilities of LLMs as scientific assistants:

- purple text shows prior contributions by the researchers participating in AuroraGPT
- blue text shows AuroraGPT contributions.
- Black text aspects adapted from existing work are included for a complete approach.

<https://arxiv.org/abs/2502.20309>

Argonne
NATIONAL LABORATORY

Frank Capello 제공 슬라이드

이 논문의 초록은 EAIRA를 잘 요약하고 있다.

"과학 연구를 위한 혁신적 도구로서 대규모 언어 모델(LLM)은 추론, 문제 해결, 의사결정이 필요한 복잡한 과제를 처리할 수 있다. LLM의 뛰어난 역량은 과학 연구 보조원으로서의 잠재력을 시사하지만, 실제 과학 응용에서의 효과를 평가하려면 전체론적·엄격·도메인 맞춤형 평가가 필요하다. 본 논문은 Argonne National Laboratory에서 개발한 'EAIRA(과학 연구 보조원으로서의 AI 모델 평가)'를 위한 다면적 방법론을 설명한다. 이 방법론은 네 가지 평가 유형을 포함한다: 1) 사실 기억력 평가를 위한 객관식 문항, 2) 고급 추론 및 문제 해결 능력 평가를 위한 주관식 문항, 3) 통제된 환경에서 연구 보조원으로서의 역량을 세부 분석하는 실험실(studio)

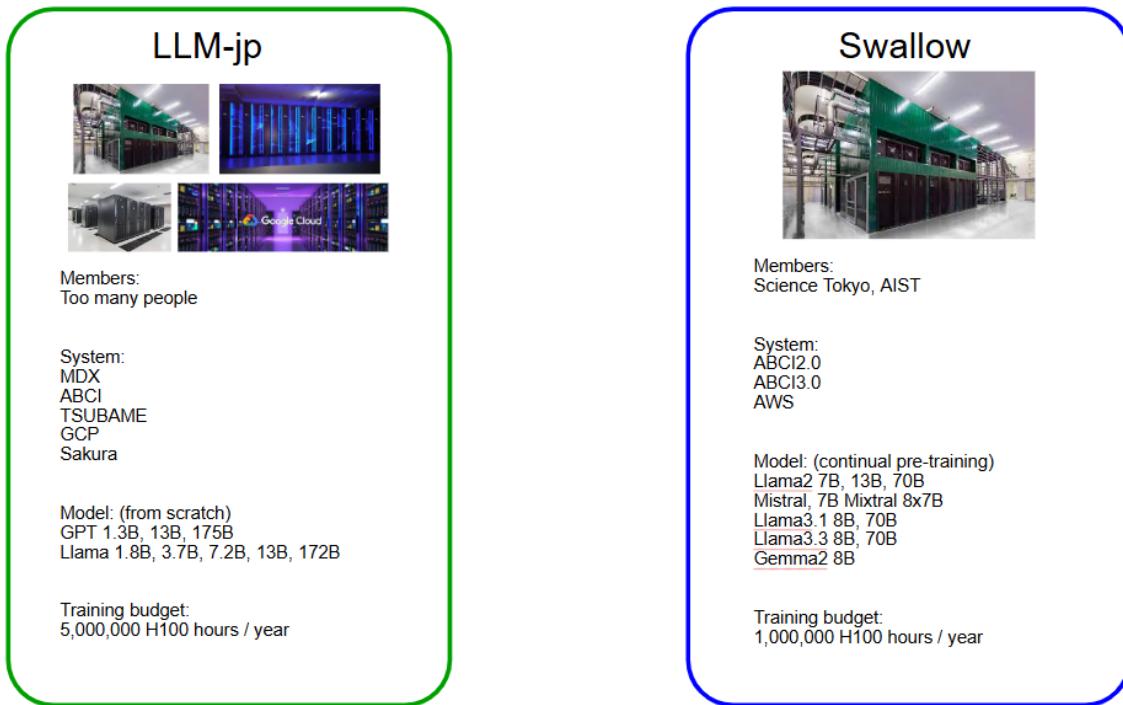
평가, 4) 광범위한 과학 분야와 응용에서 연구자-LLM 상호작용을 대규모로 포착하는 현장(field) 평가.”

이 보완적 방법론을 통해 LLM의 과학 지식, 추론 능력, 적응성 측면의 강·약점을 종합적으로 분석할 수 있다. LLM 발전 속도가 매우 빠르다는 점을 고려해, 본 방법론은 지속적 관련성과 적용성을 보장하도록 계속 발전·적용할 수 있게 설계되었다. 본 논문은 2025년 2월 말 기준으로 방법론의 현황을 설명한다. 특정 과학 분야에서 개발되었지만, 광범위한 과학 영역에 일반화될 수 있도록 고안되었다.

Cappello는 또한 ASTRO MCQ(천문학)와 SciCode Open Response 벤치마크 등 다른 벤치마크를 검토했으며, ANL-HPE 협업(DoReMi: ‘과학 문제의 난이도를 지향한 추론 노력 모델링’—Reasoning LLM을 위한 프레이밍)에 대해서도 간략히 언급했다.

일본 LLM의 최근 진행 상황

일본 AI 커뮤니티는 글로벌 LLM 분야에서의 역할 확대를 위해 과감한 발걸음을 내딛고 있다. TPC25 기조연설에서 도쿄공업대학의 리오 요코타 교수는 일본의 야심찬 양대 계획—대규모 LLM-jp 컨소시엄과, 집중적 연구 의제를 추진하는 소규모 Swallow 프로젝트—을 소개했다. 두 프로젝트는 방대한 다국어 데이터세트를 구축하고, 1,720억 매개변수급 고밀도 모델에서 민첩한 MoE(Mixture-of-Experts) 설계까지 폭넓게 탐구하며, 수백만 GPU-시간의 H100 자원을 투입해 글로벌 선두와 보조를 맞춘다. 주요 학습은 ABCI 슈퍼컴퓨터와 Fugaku 등을 포함한 일본 최고 수준의 컴퓨팅 인프라에서 수행되어, LLM 연구에 필요한 규모와 유연성을 동시에 확보한다.



Rio Yokota 제공 슬라이드

Yokota는 이런 규모를 달성하려면 하드웨어와 데이터만으로는 부족하며, 정교한 조정, 체계적 실험, 그리고 위험·트레이드오프에 대한 지속적 인식이 필요하다고 강조했다. 그는 곧바로 현실 문제로 전환해 “이 수준의 훈련에는 수백만 달러가 듭니다. 매개변수 하나만 잘못 설정해도 수백만 달러가 허공으로 사라질 수 있습니다.”라고 지적했다. 또한 더 크고 더 영리한 모델을 위해서는 데이터 정제와 중복 제거가 핵심 병목이며, 가장 중요한 작업 중 하나라고 덧붙였다.

본 프로젝트는 추상적 비전을 넘어, 일본의 AI 야망을 어떻게 ‘조율된 국가 프로그램’으로 구체화할지에 초점을 맞췄다. LLM-jp는 대학·정부·연구센터·기업 파트너를 자금 지원과 개발 우선순위를 정렬하는 공동 프레임워크 아래 통합했다. 이 구조는 단일 기관이 감당하기 어려운 대규모 실험을 가능하게 하고, 한 분야의 진전을 지역사회 전체와 신속히 공유할 수 있게 한다. Yokota의 말처럼, 목표는 “가능한 한 빨리 모든 것을 공유해, 다른 이들이 즉시 그 위에서 발전할 수 있도록 하는 것”이다.

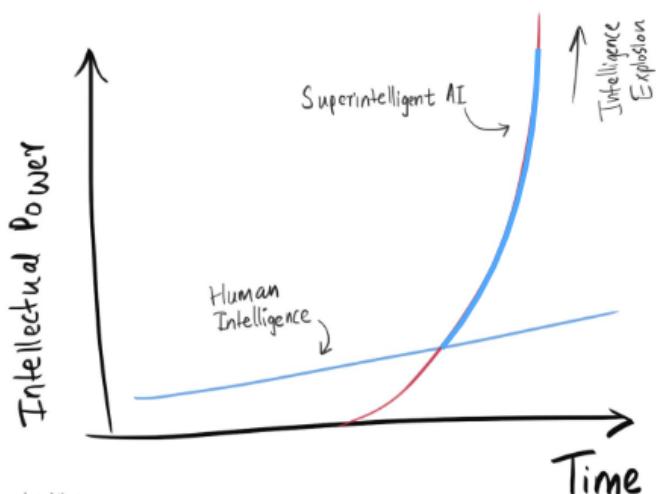
Yokota는 컨소시엄의 거버넌스를 ‘신속성’에 맞춰 설계했음을 설명했다. 팀

들은 중간 결과를 수시로 교환해 기술적 문제를 조기에 발견하고, 장기 승인 절차에 묶이지 않은 채 접근을 신속히 조정할 수 있다. 그는 이런 즉각적 적응력이, 급변하는 글로벌 이니셔티브와 경쟁할 때 컴퓨팅 용량 못지않게 결정적이라고 강조했다.

"LLM-jp가 규모와 조정을 중시한다면, Swallow는 다른 길을 택합니다. 소규모 이니셔티브로서 효율적 학습법과 간결한 모델 아키텍처에 집중하며, 특정 목적에 맞춘 실험을 지향합니다."

Yokota는 Swallow가 LLM-jp의 최대 모델보다 훨씬 적은 매개변수로 작동하지만, 데이터 필터링부터 최적화된 하이퍼파라미터 탐색까지 전 프로젝트에 적용 가능한 혁신을 추구한다고 설명했다. 그는 "여기는 1,720억 매개변수 규모에서는 시도하기 어려운, 위험하지만 잠재력 있는 아이디어를 실험하는 공간"이라고 말했다.

Intelligence Explosion



Rio Yokota 제공 슬라이드

"Swallow의 MoE 실험은 희소 활성화를 사용합니다. 즉, 주어진 입력에 대해 소수의 전문가만 활성화되어, 정확도를 유지하면서 FLOPs를 크게 줄입니다. 또한 이 프로젝트는 MoE 설계의 시험장이기도 하며, 특수 하위 모델은 필요할 때에만 켜집니다. 이 접근은 복잡한 작업 성능을 유지하면서 계산 비용을 절감하므로, 제한된

GPU 예산을 가진 팀들에 특히 매력적입니다. 그리고 이렇게 얻은 교훈은 거의 즉시 대규모 모델로 이전됩니다.” Yokota에 따르면, Swallow의 민첩성은 빠른 반복이 요구되는 과제에 적합하다.

Yokota는 LLM-jp와 Swallow를 하나의 전략을 이루는 ‘두 축’으로 제시하며 발표를 맺었다. 하나는 규모의 한계를 돌파하는 것이고, 다른 하나는 그 규모를 실현하는 데 필요한 기술을 예리하게 다듬는 일이다. 두 축 모두 결과를 신속히 공개해 더 넓은 커뮤니티가 혜택을 누리게 한다는 강한 원칙으로 서로 연결된다. 그는 일본의 LLM 분야 도전이 쉽지 않음을 인정하며, 특히 컴퓨팅 비용 상승과 급변하는 기준 변화 속에서 난도가 높다고 말했다. 그럼에도 일본은 국가 차원의 협력, 목표 지향적 혁신, 그리고 개방적 교류를 통해 글로벌 AI 시장에서 경쟁력을 확보·유지할 수 있다고 강조했다.

핵심 요점

두 논의는 하나의 지점으로 수렴한다. LLM이 과학 분야에서 잠재력을 온전히 발휘하려면 ‘신뢰’와 ‘규모’가 함께 진화해야 한다. 세계에서 가장 큰 모델이라도 결과를 검증할 수 없다면 영향력은 제한적이며, 가장 엄격한 평가 절차도 복잡한 현실 문제를 풀 만큼 강력한 시스템에 적용되지 않으면 가치가 희석된다.

Cappello의 EAIRA 프레임워크는 여러 평가 방식을 결합해 AI가 실제로 무엇을 할 수 있는지에 대한 명확한 시야를 제공함으로써 신뢰의 문제를 다룬다. Yokota의 LLM-jp-Swallow 이니셔티브는 국가 차원의 협력, 효율적 아키텍처, 신속한 지식 공유 문화를 통해 ‘규모’의 문제를 정면 돌파한다.

공유된 메시지는 명확하다: “과학에서 가장 중요한 LLM은 야심 찬 역량을 갖추되, 엄격하고 투명한 테스트 위에 세워진 LLM이다.”

8. 매체 하이라이트 (HPCwire/AIwire)

매체	제목	날짜
AIwire	TPC25 Preview: Inside the Conference Shaping Frontier AI for Science https://www.aiwire.net/2025/07/22/tpc25-preview-inside-the-conference-shaping-frontier-ai-for-science/?utm_source=chatgpt.com	2025년 7월 22일
HPCwire	TPC25 Preview: Inside the Conference Shaping Frontier AI for Science (上同)	2025년 7월 23일
HPCwire / Off-the-Wire	Trillion Parameter Consortium Partners with Tabor Communications to Launch Global AI for Science Conference, TPC25 https://www.hpcwire.com/off-the-wire/trillion-parameter-consortium-partners-with-tabor-communications-to-launch-global-ai-for-science-conference-tpc25/?utm_source=chatgpt.com	2025년 3월 28일
AIwire	Everything You Always Wanted to Know About the Trillion Parameter Consortium and TPC25 But Were Afraid to Ask https://www.aiwire.net/2025/06/25/everything-you-always-wanted-to-know-about-the-trillion-parameter-consortium-and-tpc25-but-were-afraid-to-ask/?utm_source=chatgpt.com	2025년 6월 25일

매체	제목	날짜
Alwire	Beyond Speed: TPC25 Maps the Road to Science-Ready AI https://www.aiwire.net/2025/07/31/beyond-speed-tpc25-maps-the-road-to-science-ready-ai/?utm_source=chatgpt.com	2025년 7월 31일
Alwire	TPC25 Highlights AI's Expanding Role: Multimodal Data, Model Evaluation, and Non-LLM Architectures https://www.aiwire.net/2025/08/06/tpc25-highlights-ais-expanding-role-multimodal-data-model-evaluation-and-non-llm-architectures/?utm_source=chatgpt.com	2025년 8월 06일
Alwire	From Hypothesis to Hardware: Four Voices on the Future of Scientific AI at TPC25 https://www.aiwire.net/2025/08/01/from-hypothesis-to-hardware-four-voices-on-the-future-of-scientific-ai-at-tpc25/?utm_source=chatgpt.com	2025년 8월 01일
Alwire	From Trust to Scale in LLMs: Insights from Key Leaders at TPC25 https://www.aiwire.net/2025/08/12/from-trust-to-scale-in-llms-insights-from-key-leaders-at-tpc25/?utm_source=chatgpt.com	2025년 8월 12일

9. 협업과 거버넌스

TPC는 80개 이상의 회원 기관을 결집해 오픈사이언스, 공유 인프라, 윤리·정책 전반의 우선순위를 정렬하고 있다. 산하 워킹그룹은 로드맵, 데이터 거버넌스, 평가 프레임워크를 통해 HPC와 AI 커뮤니티를 연결한다. TPC는 단순한 학회나 프로젝트 네트워크를 넘어, 초기대 AI 모델의 과학 연구 활용을 위한 글로벌 규범과 인프라를 공동 설계하는 ‘**국제 과학-AI 생태계의 거버넌스 플랫폼**’이라 할 수 있다. 이 거버넌스 체계 안에서 협업할 때의 장점은 크게 네 가지로 요약된다.

1) 공유 인프라 접근 (Shared Infrastructure Access)

- TPC는 DOE(미국 에너지부), RIKEN(일본), BSC(스페인), Argonne, AWS, Intel 등 국가 연구소와 산업체의 컴퓨팅 자원 및 AI 훈련 환경을 연결하는 허브 역할
- 컨소시엄 참여자는
 - **공유 데이터 허브(Data Commons)**
 - **모델 허브(Model Zoo for Science)**
 - **공통 학습 프레임워크 (Exascale AI Infrastructure)**에 접근 가능
- 예를 들어, 미국 Argonne이나 일본 RIKEN의 슈퍼컴퓨팅 클러스터에서 동일 모델을 재현하거나 후속 학습을 이어받을 수 있는 ‘연합 학습 환경’이 제공된다.

2) 표준화된 연구 생태계 (Standardized Scientific AI Ecosystem)

- TPC는 AI for Science 분야에서 **모델, 데이터, 평가의 표준화 규범**을 제정
- 예:
 - **TrustBench** - 모델의 신뢰성·추론력 평가 프레임워크
 - **RealityGap Test** - 시뮬레이션과 실제 데이터 간 간극을 검증하는 지표
 - **Causal Reasoning Benchmark** - 과학적 추론 평가 세트
- 이러한 표준 프레임워크를 공동으로 마련해 기관 간 연구 결과의 **비교·재현·통합**을 용이하게 한다.

- 논문·데이터·모델을 공통 구조로 정리하여 국가 간 협업 시 기술 언어를 통일한다.

3) 지식 공유 및 기술 확산 (Collaborative Knowledge Transfer)

- TPC 내 워킹그룹(Working Group)들은 분야별로 운영돼요.
- 예: AI Infrastructure WG, Data Governance WG, Evaluation WG, Life Sciences WG 등.
- 참여자는 이 그룹 내에서
 - 최신 AI·HPC 기술 교류
 - 공동 논문 및 백서 작성
 - 연구소 간 인턴십/교류 프로그램 참여 등의 기회를 얻어요.
- 특히 'TPC25 → TPC26'로 이어지는 연례회의를 통해 각 기관의 성과를 공동 백서 형태로 보존하고 후속 프로젝트로 발전시킨다.

4) 책임 있는 AI 구축 (Trustworthy & Ethical AI Governance)

- TPC의 핵심 미션 중 하나는 신뢰할 수 있는 과학 AI(**Trustworthy AI for Science**) 구축
- 각국의 정책 기관, 윤리 위원회, 산업체가 함께 데이터 거버넌스와 AI 윤리 기준 정립
- 예를 들어,
 - 모델이 사용한 데이터 출처를 투명하게 공개(Provenance Tracking),
 - AI 의사결정 로그를 보존(Reproducibility Logging),
 - AI 실험 자동화에 인간 감독 절차를 포함(Human-in-the-loop).
- AI가 연구 결정에 관여해도 신뢰를 확보할 수 있는 제도적 안전망 확보

즉, 기술적으로는 초대형 AI 인프라를 공동으로 활용하고, 과학적으로는 표준화된 검증 체계 안에서 협업하며, 사회적으로는 신뢰 가능한 AI 연구 거버넌스를 공동 구축하는 것이다.

10. 도전 과제와 윤리

TPC는 막대한 잠재력을 지니고 있으나, 모델 해석 가능성, 계산 자원 제약, 윤리 이슈, AI-전문지식 통합 등 복합 과제에 직면해 있다. 개방성을 유지하면서 보안을 확보하고, 국가 간 협력을 조율하는 일 역시 지속적인 거버넌스 과제로 남아 있다.

1) 데이터 거버넌스와 접근 불균형 (Data Governance & Access Inequality)

문제

- TPC의 목표는 '글로벌 과학 AI 모델을 위한 공통 데이터 인프라' 구축이지만, 국가별 데이터 보안법·연구윤리 규정·산업 기밀 보호 조항의 차이로 **데이터 공유가 일관되게 작동하지 않는다.**
- 특히 생명과학·의료·국방 분야는 국외 반출 제한 등으로 **모델 학습·검증에 불균형이 발생한다.**
- 일부 일부 대형 기관(DOE, RIKEN, AWS 등)은 초대형 HPC를 보유하지만, 개발도상국·중소 연구기관은 접근성이 낮아 '**AI for Science divide(과학 AI 격차)**'가 커질 수 있다.

윤리적 쟁점

- '공동 과학 모델'의 데이터 소유자는 누구인가?
 - 데이터 기여량이 다를 때 결과물(모델·논문)의 소유권·인용권은 어떻게 분배할 것인가?
- ◆ **핵심:** "글로벌 과학 AI의 공정한 접근과 소유 구조를 어떻게 보장할 것인가?"

2) 모델 투명성·추론 가능성 (Transparency & Interpretability)

문제

- TPC 모델은 수십억~수조 파라미터의 '**블랙박스 AI**'에 가깝다.
- 물리 법칙 내장형 모델이라도, 결과의 근거나 데이터 편향을 설명하기 어렵다.

- 과학 연구 본질이 '이유(reasoning)의 검증'인 만큼, 설명 불가능한 결과는 증거로 채택되기 어렵다.

윤리적 쟁점

- AI의 예측이 과학적 결론에 영향을 줄 때, 책임은 AI 인가 인간 연구자인가?
 - 모델 오류가 재현되지 않을 경우, 연구 진실성(scientific integrity)이 흔들릴 위험.
- ◆ **핵심:** "AI가 과학의 언어로 설명할 수 있을 때만, 진정한 과학 파트너가 된다."

3) 신뢰성과 재현성 (Trustworthiness & Reproducibility)

문제

- 초기대 모델을 공동 학습하더라도, 각기 다른 하드웨어·환경·데이터셋으로 인해 동일 모델의 결과가 달라질 수 있다.
- 엑스케일 시스템 간 통합 로그·버전 관리 표준이 미흡해 '반복 가능한 과학(Repeatable Science)' 확보가 어렵다.

윤리적 쟁점

- 과학적 결과가 'AI 실험 로그'로만 남으면, 인간이 검증 가능한 증거 체계가 약화된다.
 - 모델이 자율 실험을 수행할 경우, 결과를 '누가 언제 승인했는가'의 추적이 어렵다.
- ◆ **핵심:** "AI가 만든 과학은 인간이 검증 가능한가?"

4) 에너지 소비와 환경 윤리 (Environmental & Energy Ethics)

문제

- 수조 파라미터급 모델은 학습 1회에 수천 MWh의 전력을 소모할 수 있다.

- 다수 HPC 자원이 화석연료 기반 전력망에 연결되어 있어 환경 부담이 커질 수 있다. → "AI-driven science"가 역설적으로 환경 부담을 높일 수 있음

윤리적 쟁점

- 초대형 학습의 막대한 전력 사용은 '지속 가능한 과학'과 합치하는가?
 - 환경 비용을 정당화하려면, AI 결과가 실제 사회적 가치로 환원되는 구조가 필요함.
- ◆ **핵심:** "AI 가 지구를 이해하는 도구라면, 지구를 해치지 않는 방식으로 작동해야 한다."

5) 인간-AI 역할 경계와 책임 (Human–AI Boundary & Scientific Agency)

문제

- 'Agentic Systems for Science'는 AI 가 가설 수립과 실험 수행까지 맡는 자율 과학 시스템을 지향한다.
- 그러나 이 시스템에서 과학적 '책임 주체'가 누구인지는 아직 불명확하다.
- 예를 들어, AI 가 제안한 실험이 실패했거나 오염된 데이터를 사용했다면 연구 책임은 사람인가, AI 시스템인가?

윤리적 쟁점

- AI 는 공동 연구자(co-author)가 될 수 있는가?
 - AI 가 실험을 제안하거나 자동으로 논문을 작성할 때, 저작권·인용·공헌도는 어떻게 명시할 것인가?
- ◆ **핵심:** "AI 가 과학을 수행해도, 인간은 여전히 '책임의 중심'이어야 한다."

6) 글로벌 정책·규제의 비대칭성 (Policy and Regulation Asymmetry)

문제

- 미국·EU·일본은 EU AI Act, NIST AI RMF, RIKEN AI Charter 등 규제를 신속히 수립 중이나, 기준 차이로 TPC 내 규제 충돌(regulatory inconsistency)이 발생한다.
- 어떤 국가는 'AI-generated data'를 공식 연구 데이터로 인정하지만, 다른 국가는 비신뢰 자료로 간주한다.

윤리적 쟁점

- AI 과학 결과의 국제 상호 인증(mutual recognition)이 어려워질 위험.
- '글로벌 과학 AI의 기준을 누가 정할 것인가'라는 **거버넌스 주도권 경쟁**.
- ◆ **핵심:** "국제 과학 협력의 신뢰를 훼손하지 않으면서, 규제의 균형을 잡아야 한다."

➔ TPC 가 직면한 윤리적 전환점에 대한 고찰

- TPC 의 목표는 'AI 를 통한 과학 혁신'이지만, 그 혁신은 '책임 있는 과학(Responsible Science)'의 틀 안에서 논의되어야 한다.
- 현재 TPC 는 세 가지 원칙에 따라 대응 중이다: **Transparency (투명성)**
- 모델·데이터·코드 공개, **Accountability (책임성)**
- 인간 중심 승인 체계 유지, **Sustainability (지속가능성)**
- 환경·사회적 영향 평가 포함

11. 전망 및 결론

Trillion Parameter Consortium(TPC)은 초거대 인공지능 모델 개발을 넘어, AI가 과학의 방법론을 어떻게 바꿀 수 있는지 전 지구적 차원에서 실험하는 거버넌스 플랫폼이다. TPC는 앞으로도 과학 AI 분야를 선도하며, **AI를 통해 발견 프로세스를 재설계하고 이를 신뢰 가능한 체계 안에서 작동하도록 하는 데 기여할 것으로 기대된다.** 논의는 물리 기반 AI, 멀티모달 모델, 자기지도학습, 인과 추론, 인간-AI 협력 구조 등으로 더욱 확장될 것이다.

2025년 산호세에서 열린 TPC25는 이러한 방향성을 선명하게 보여준 행사였다. 데이터·시뮬레이션·실험이 하나의 루프 안에서 통합되고, AI는 단순한 연산 도구를 넘어 **가설을 수립하고 검증하는 '탐구 주체'**로 등장했다. 이 변화는 개별 연구자의 성과를 넘어 과학 전반의 구조적 패러다임이 바뀌고 있음을 시사한다.

앞으로 TPC의 협력 구조는 국가·산업·학문 경계를 넘어서는 지속 가능하고 상호운용 가능한 글로벌 연구 네트워크로 발전할 것이다. 미국 DOE, 일본 RIKEN, 유럽 BSC-INRIA, 호주 NCI, 한국 KISTI 등 주요 HPC 기관이 참여하는 다자 협력체를 통해 데이터 공유, 모델 표준화, 평가 지표 통합이 현실화될 것으로 기대된다. 이러한 움직임은 과학을 더욱 개방적이고 재현 가능한 체계로 이끌 것이다.

기술적 측면에서 TPC의 미래는 세 가지 축으로 정리된다. 첫째, 하이브리드 클라우드-HPC-양자 인프라를 통합한 **엑사스케일 AI 인프라 구축**. 둘째, 물리정보 및 연산자 학습 기반의 **과학용 파운데이션 모델 개발**. 셋째, AI가 스스로 탐구 과정을 수행하는 **에이전트형 과학 시스템 구현**. 이 세 방향은 AI를 단순한 예측 도구에서 과학적 사고의 협력 주체로 확장시키는 토대가 될 것이다. 최근 부상하고 있는 **Physical AI와 AI4PDE에 대한 발표가 부재하다는 한계**가 TPC25에 있었다. 이는 이 분야를 주도하는 중국이 TPC에서 빠져 있는 것과 관련이 있을 것으로 사료된다.

그러나 발전 속도만큼 중요한 것은 **신뢰와 책임**이다. AI가 과학의 의사결정에 깊이 관여할수록, 결과를 검증·통제할 수 있는 인간 중심 거버넌스가 필요하다. TPC는 투명성(Transparency)·책임성(Accountability)·지속가능성(Sustainability)을 핵심

원칙으로 삼아 데이터 공개, 결과 검증, 환경 영향 평가를 병행하고 있다.

향후 TPC는 이러한 원칙 아래에서 글로벌 데이터·모델 허브의 본격 가동, 자율 과학 시스템의 성능 평가 체계 구축, 그리고 '2030 비전' 문서 발표를 목표로 한다. 이를 통해 TPC는 초거대 모델 개발 단체를 넘어, **'AI 기반 과학의 국제 표준'**을 주도하는 거버넌스 기관으로 자리매김할 가능성이 크다.

결국 TPC가 지향하는 것은 단순히 빠른 과학이 아니다. 더 신뢰할 수 있고 더 지속 가능한 과학이다. 초거대 AI와 자율 시스템이 아무리 빠르게 발전하더라도, 그 발전이 책임과 윤리의 틀 안에서 이루어질 때에만 진정한 의미의 'AI-driven Science'가 완성된다.

참고문헌

- 1) HPCwire (2025-03-28): TPC25–Tabor Communications 파트너십 공지.
- 2) Alwire (2025-07): TPC25 프리뷰 및 현장 보도.
- 3) Argonne NL (2024): Trillion Parameter Consortium 개요.
- 4) NeSI/Cineca (2025): TPC25 참여 공지.
- 5) "The Trillion Parameter Consortium (TPC) Launch Announcement," ANL News, Argonne National Laboratory, 2024.
- 6) "TPC25—Trillion Parameter Consortium 2025 Conference," HPCwire, 2025.
- 7) "Global Scientists Join Forces for AI Breakthroughs through TPC," The Decoder, 2024.
- 8) Goodfellow, Ian, et al. Deep Learning. MIT Press, 2016.
- 9) Raissi, Maziar, Paris Perdikaris, and George Em Karniadakis. "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations." Journal of Computational Physics, 2019.
- 10) Lu, Lu, et al. "Learning nonlinear operators via deep neural networks: Proposal of the DeepONet." Nature Machine Intelligence, 2021.
- 11) Peng, Lei, et al. "Neural ODE for System Identification in Physics." Proceedings of XYZ Conference, 2023.
- 12) Smith, J., et al. "Challenges of Large-Scale Scientific AI Models: Transparency, Efficiency, and Governance." AIWire, 2025.
- 13) Zhang, Wei, et al. "Operator Learning and its Applications in Fluid Dynamics." HPCwire, 2024.
- 14) Brown, Tom, et al. "Language Models are Few-Shot Learners." NeurIPS, 2020.
- 15) TPC25 Organizing Committee. TPC25 Program and Agenda. TPC25.org, 2025.
- 16) "Trillion Parameter Consortium Joins Forces with Research Institutions Worldwide," AIWire, 2024.
- 17) Argonne National Laboratory. (2025, August). *TPC25 inaugural conference highlights scientific discovery in the age of AI*. U.S. Department of Energy. <https://www.anl.gov>

- 18) Argonne National Laboratory. (2023). *AI for science 2030 vision*. Office of Science AI Roadmap Document.
- 19) Alwire. (2025, August 6). *TPC25 highlights AI's expanding role: Multimodal data, model evaluation, and non-LLM architectures*. <https://www.aiwire.net>
- 20) Baeza-Yates, R., Balaprakash, P., Liu, J., & Kim, K.-S. (2025, July). *Plenary 4: Multimodal data, evaluation, and non-LLM model architectures*. In *TPC25 Conference Proceedings* (San Jose, CA).
- 21) Barcelona Supercomputing Center. (2025). *European AI for science and the role of TPC*. Press release.
- 22) Bidot, T., & Valero, M. (2025). *Open science data and model hub: Toward FAIR collaboration within TPC*. INRIA Technical Note.
- 23) Cineca HPC. (2025). *TPC25 consortium meeting announcement: July 28–31, San Jose*. <https://www.hpc.cineca.it>
- 24) Department of Energy (DOE). (2024). *AI for science strategic plan 2024–2030*. U.S. Department of Energy, Washington, DC.
- 25) DOE & RIKEN Joint Taskforce. (2025). *AI + HPC convergence for scientific discovery: Interim whitepaper*.
- 26) Duraisamy, K. (2024). *Scientific machine learning and physics-informed AI models*. University of Michigan Lecture Series.
- 27) Elahe Vedadi, Chengappa, P., Huang, K., Ramanathan, A., & Narayanan, S. (2025, July). *Panel discussion: Building agentic systems for science – Reports from the field*. In *TPC25 Conference Proceedings* (San Jose, CA).
- 28) Foster, I. (2023). *Designing AI workflows for scalable science*. Argonne National Laboratory Technical Report.
- 29) FutureHouse Labs. (2025). *Agentic life sciences framework*. Technical report presented at TPC25, San Jose.
- 30) HPCwire. (2025, July 23). *Inside the conference shaping frontier AI for science: TPC25 preview*. <https://www.hpcwire.com>

- 31) Hyperion Research. (2025). *Global AI for science and HPC market outlook 2025–2030*. Washington, DC.
- 32) Ian Foster, Stevens, R., Matsuoka, S., Valero, M., & Bidot, T. (2025, July). *Plenary 5: Science updates from key TPC leaders*. In *TPC25 Closing Plenary* (San Jose, CA).
- 33) NCI Australia. (2025). *NCI Australia joins the Trillion Parameter Consortium to create reliable AI for science*. <https://nci.org.au>
- 34) Quantinuum. (2025). *Quantum-AI collaboration in scientific discovery: Quantinuum and TPC partnership announcement*.
- 35) Rick Stevens. (2025, July). *Reinventing discovery: Accelerating science in the age of artificial super-intelligence*. Opening plenary talk, TPC25, San Jose.
- 36) RIKEN Center for Computational Science. (2025, July). *RIKEN joins the Trillion Parameter Consortium*. Kobe, Japan.
- 37) Salim, F. (2024). *PromptCast: Prompt-based forecasting for spatio-temporal data*. *NeurIPS 2024 Workshop Proceedings*.
- 38) Salim, F. (2025, July). *Modeling and simulating complex behavior in dynamic cyber-physical-social systems*. Opening plenary talk, TPC25, San Jose.
- 39) The Decoder. (2025, July). *Global scientists join forces for AI breakthroughs: The Trillion Parameter Consortium*. <https://the-decoder.com>
- 40) Times of India. (2025, August). *Global consortium aims for trillion-parameter AI models in scientific research*.
- 41) Trillion Parameter Consortium (TPC). (2025). *TPC25 conference program and lightning talks insert*. <https://tpc25.org>
- 42) Trillion Parameter Consortium (TPC). (2025, August). *Science at exascale: TPC data and model hub launch briefing*.
- 43) Vedadi, E., Chengappa, P., Huang, K., Ramanathan, A., & Narayanan, S. (2025). *Building agentic systems for science: Reports from the field*. *TPC25 Panel Discussion Summary*.
- 44) Wolfram, S. (2024, March). *Can AI solve science?* <https://writings.stephenwolfram.com>

부록 1 – 리플릿 Leaflet



GENERAL CONFERENCE •

Plenary Sessions will take place July 29-31 and are open to all conference attendees. Topics and speakers are TBD as per TPC Steering and Program committees, and will be announced over the coming months. Open public breakouts will take place between plenaries, July 30-31. Breakouts are organized like mini-workshops with 1/3 of the time for lightning talks and 2/3 of the time for discussion. Please see the breakout abstracts and lightning talk line-up from [TPC's Barcelona Conference](#) for examples of breakout topics for TPC25.

Other breakout topics might include:

- **GraphRAG for Drug Discovery:** How drug companies are using GraphRAG to accelerate drug discovery by seeding foundation models with ground truth from knowledge graphs.
- **The State of AI Reasoning Models in Science:** How reasoning models such as Deepseek-R1, OpenAI 03-mini, and Anthropic Claude 3.5 Sonnet have the potential to expand the search for scientific discovery.
- **GenAI Needs New Hardware:** As AI workloads evolve, so do their hardware requirements, and bigger might not be better. How might the future be more distributed?

TPC25 HACKATHONS •

The primary mode of engagement within TPC is participation in working groups: self-organized, community-driven teams that focus on specific technical or domain areas. Many working groups have identified collaborative projects, and pre-conference hackathons provide opportunity to synchronize, plan, and make progress on those projects in person. These sessions are open to currently active working group participants, as well as others by invitation; registration for TPC25 allows registrants to apply for an invite. Please see the [Fall 2024 Hackathon Report](#) for examples of hackathon topics for TPC25.

TPC25 TUTORIALS •

TPC places special emphasis on education by offering tutorials. These training opportunities help participants learn new techniques, experiment with state-of-the-art tools, and engage with mentors and experienced practitioners. TPC will be providing a full-day tutorial on Monday, July 28, and sponsors are welcome to provide appropriate half-day tutorials one Tuesday, July 29. Tutorials are open to all conference attendees, for an additional fee.

TPC25 HACKATHON AND BREAKOUT TOPICS

(current TPC working groups)

- | | |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <ul style="list-style-type: none">• Model Architecture and Performance Evaluation• Data, Training Workflows, and Large Language Model Strategies• AI Hardware Strategies at Scale• Skills, Safety, and Trust Evaluation• Outreach and Training | <ul style="list-style-type: none">• Large Language Models for Scientific Software• Large Language Models for Healthcare• Biology, Biochemistry, and Bioinformatics• Climate, Environment, and Ecology• Engineering, Energy, and Mechanics• Fundamental Physics and Cosmology• Materials, Chemistry, and Nanoscience |
|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

SPONSORSHIP CATEGORIES

ACADEMIC: \$5,000 •

- One table in either the exhibits area or at the Job Fair
- Two conference tickets

CORPORATE: \$15,000 (\$20,000 after May 31) •

- One table in the exhibits area
- Meal sponsorship for one meal (signage and acknowledgement from stage)
- Two conference tickets

Note: Meal sponsorship is only available to the first 13 sponsors

PLATINUM: \$50,000 (limited to four) •

- Participation on plenary panel on Science + AI Industry
- One table in the exhibits area
- One table at the Job Fair
- Four conference tickets

FOUNDATION: \$75,000 (limited to two) •

- 30-minute keynote (featuring conference-approved content)
- One table in the exhibits area
- One whisper suite for private meetings
- One table at the Job Fair
- Eight conference tickets
- One pre- and one post-event promotional email to TPC25 registrants
- First right of refusal to maintain Foundation Sponsorship in coming years

ADDITIONAL SPONSORSHIP OPPORTUNITIES

Only available to existing Academic, Corporate, Platinum, or Foundation sponsors.

JOB FAIR TABLES •

- Academic: \$2,000
- Corporate: \$5,000

EVENING EVENT SPONSORSHIPS •

- Hackathon gathering on Monday: \$15,000 (limited to one)
- Welcome cocktail party on Tuesday: \$25,000 (limited to one)
- Wednesday evening events: \$5,000 (limited to three)

Note: Wednesday evening events will be organized and paid for by their sponsors, but will be promoted as official TPC25 parties in the TPC25 agenda.

MISCELLANEOUS OPPORTUNITIES •

- Half-day Tutorial: \$10,000 (limited to two)
- Lanyards: \$2,500 (limited to two)

TERMS: 50% due at signing, 50% due 30 days prior to the event.

For more information or to sponsor TPC25, please email
tpc25@taborcommunications.com. •

부록 2 - 전체 아젠다

Day		Key Items	
Mon, Jul 28		Hackathon/Tutorial Opening Plenary; Tutorials & Hackathon Sessions; Networking Gathering	
Tue, Jul 29		Hackathon & Tutorials (AM); Opening Plenary; Plenary 2; Google Welcome Reception	
Wed, Jul 30		Plenary 3; Plenary 4; Lunch Panel; Breakouts A & B	
Thu, Jul 31		Breakouts C & D; Lunch Panel; Breakouts E; Closing Plenary	
Plenary	Theme / Talk	Speaker	Affiliation
Opening	Reinventing Discovery: Accelerating Science in the Age of Artificial Super-Intelligence	Rick Stevens	Argonne National Laboratory
Opening	HPC and Science: The Need for Hybrid	Thierry Pellegrino	AWS
Opening	Modeling & Simulating Complex Behavior in Dynamic Cyber-Physical-Social Systems	Flora Salim	University of New South Wales
Plenary 2	Enabling Scientific Discovery With Generative Quantum AI	Steve Clark	Quantinuum
Plenary 2	Overview of AI for Technical Computing W	Earl Joseph	Hyperion Research

	orkloads		
Plenary 2	Secure AI Infrastructure for Scientific Computing	Jens Domke	RIKEN R-CCS
Plenary 3	Scaling Reasoning, Scaling Science: AI-Native Scientific Discovery Platform	Moderator: Ian Foster	Argonne National Laboratory
Plenary 3	Agents, Autonomy, and Agency: A Brave New World	Preeth Chengappa	Microsoft
Plenary 3	Automation of Biological Discovery with LLM Agents	Siddharth Narayanan	FutureHouse
Plenary 3	Active Inference AI Systems for Scientific Discovery	Karthik Duraisamy	University of Michigan
Panel	Industry, Academia, and Government Collaboration: Trustworthy AI for Science	Hal Finkel; Raj Hazra; Pradeep Dubey; Molly Presley (Mod.)	DOE; Quantinuum; Intel; Hammertime

부록 3 – 주요 Scientific AI 모델 목록

< 현재 발표된 주요 Scientific AI 모델들 >

모델 이름	설명
SchNet	<ul style="list-style-type: none">- 원자 간 상호작용을 예측하기 위해 고안된 화학·재료 시뮬레이션용 신경망 모델- 분자 구조의 좌표를 직접 입력으로 사용하며, 양자역학 기반 물성 예측(예: 에너지, 힘)
AlphaFold	<ul style="list-style-type: none">- DeepMind 가 개발한 단백질 3 차원 구조 예측 모델- 주어진 아미노산 서열로부터 단백질 접힘 구조를 거의 원자 수준 정확도로 예측- 생명과학, 의약 등에서 혁명적인 도구로 평가됨
ChemBERTa	<ul style="list-style-type: none">- 화학 분자 구조를 자연어처럼 처리하는 BERT 기반 트랜스포머 모델- SMILES 형식(분자 문자열)을 학습하여 화합물의 특성 예측, QSAR, 독성 분석 등에 활용
FourCastNet	<ul style="list-style-type: none">- NVIDIA 가 개발한 날씨 예측용 AI 모델로, 물리기반 수치예보(NWP)를 대체 또는 보완할 수 있는 트랜스포머 기반 딥러닝 모델- 전지구적인 고해상도 예측을 수 초 내 수행 가능
AlphaGeometry	<ul style="list-style-type: none">- DeepMind 가 개발한 기하학 문제 해결 AI로, 대수 연산 + LLM 기반 증명 생성을 결합하여 수학 기하 증명을 스스로 생성- IMO(국제수학올림피아드) 문제에서 인간 수준 성능을 달성
Reasoning LLMs	<ul style="list-style-type: none">- 추론 능력을 강화한 대형 언어 모델 계열을 뜻함- 단순한 텍스트 생성이 아니라 다단계 추론, 논리적 연역, 수리 계산 등에 특화된 모델- 대표적으로 GPT, Claude , Minerva 등이 있음
Commoditized Agents	<ul style="list-style-type: none">- 범용화된 AI 에이전트 도구 또는 프레임워크를 의미함- 특정 목적 없이도 다목적으로 활용 가능한 agent 들을 지칭- 예: AutoGPT, CrewAI, LangGraph 등- 워크플로우 자동화 및 다중 모델 협력 등에 사용됨

KISTI Trend Report

비매품/무료



ISBN 978-89-294-0226-6

(PDF)