

「AI 특화 파운데이션 모델」 프로젝트 참여팀으로 루닛 컨소시엄, KAIST 컨소시엄 선정

- 우리가 강점을 보유한 ①의과학, ②바이오 특화 AI 파운데이션 모델 개발·확보·확산을 통해 AI와 도메인 경쟁력 동반 성장 견인 기대
- ※ ① (루닛) 세계 최초로, 전주기 의과학 특화 AI 파운데이션 모델 개발·공개 목표
- ② (KAIST) 차세대 단백질 복합체 구조 예측 바이오 파운데이션 모델 개발·공개 목표
- 특화 분야 글로벌 최고 AI 모델을 내년까지 확보, 즉시 확산 목표

과학기술정보통신부(부총리 겸 과기정통부 장관 배경훈, 이하 과기정통부)는 「AI 특화 파운데이션 모델」 프로젝트 수행팀으로, ①루닛 컨소시엄(전주기 의과학 AI 파운데이션 모델 개발)과 ②KAIST 컨소시엄(바이오 AI 파운데이션 모델 개발)을 선정하였다.

본 프로젝트 참여팀 공모(9.5~10.13)에 총 18개의 컨소시엄이 접수하여 10:1에 육박하는 열띤 관심과 경쟁을 보여왔으며, 서면평가, 발표평가 등을 거쳐 ①루닛 컨소시엄과 ②KAIST 컨소시엄이 최종 선정됨에 따라, 오는 11월 1일부터 GPU 지원을 착수한다. 이를 통해, AI 기반 의과학·바이오 분야에서 글로벌 리더십 확보 기반을 확충하는 데 기여할 수 있을 것으로 기대된다.

< 「AI 특화 파운데이션 모델」 프로젝트 개요 >

- ◇ (지원대상) 국내 AI 기업·대학·연구기관 등(단독 또는 컨소시엄 가능)
- ◇ (지원내용) 분야별 AX 혁신을 위한 특화 AI 파운데이션 모델 개발 지원
- ◇ (지원규모) B200 512장(총 2개 과제, 과제당 256장)
- ◇ (지원기간) 2025년 11월 1일 ~ 2026년 9월 9일(5개월 + 5개월 9일)
- ◇ (추진일정) 공고(9.5~10.13.) → 선정평가협약체결(10월) → 과제수행(25.11월~26.9월초)
- ※ (단계평가) '26.3월말 → (최종평가) '26.9월초 / 우수과제는 GPU 추가 지원 검토

각 컨소시엄 별 주관·참여기관, 과제 목표와 주요 특징 등은 다음과 같다.

< 루닛 컨소시엄 >

“분자에서 인구까지 전주기 의과학 혁신을 위한 멀티스케일 의과학 특화 파운데이션 모델 개발”

주관·참여기관 루닛 컨소시엄에는, 루닛과 함께 7개 기업, 6개 대학 연구실, 9개 의료기관 등 총 22개의 국내 우수 기관이 참여하였다.

※ (기업 : 7) 트릴리온랩스, 카카오헬스케어, 아이젠사이언스, SK바이오팜, 디씨클, 리벨리온, 스탠다임
(대학 : 6) KAIST(최윤재, 김태균, 예종철, 김현우, 홍승훈 교수 연구실), 서울대(정유성 교수 연구실)
(의료기관 : 9) 국민건강보험공단 일산병원, 용인세브란스병원, 강동경희대학교병원, 경희의료원, 고려대 산학협력단(고려대병원), 건양대병원, 이화의대부속 서울병원, 계명대 동산의료원, 양산부산대병원

과제목표 이들은 분자 - 경로 - 의약품 / 안전성 - 임상시험 / 가이드라인 - 실세계 임상 등 의과학 전주기를 아우르는 증거사슬(Chain of Evidence) 기반의 의과학 특화 AI 파운데이션 모델과 멀티 에이전트 서비스를 구축·실증할 계획으로,

이를 통해, ▲임상 의사결정의 정확성·안전성, ▲연구개발 생산성, ▲국민 건강권 체감 효용을 동시에 제고하겠다는 목표를 제시하였다.

주요특징 동 컨소시엄은 ①자체 기술력을 바탕으로 글로벌 SOTA(State Of The Art, 최고수준) 모델을 뛰어넘는 최대 32B급 의과학 특화 AI 파운데이션 모델을 처음부터(From Scratch) 개발하여 순차 확보하고(목표), 3종의 도메인 재특화 모델*과 에이전틱(Agentic) 시스템으로 확장한다.

* ▲화합물·단백질 도메인 재특화 모델, ▲오믹스 도메인 재특화 모델, ▲임상 도메인 재특화 모델

②세계 최초로 의과학분야 전주기 지식을 담은 특화 파운데이션 모델 파운데이션 모델을 오픈소스(Open source, 상업용, Apache 2.0 라이선스)로 공개하는 한편, 추후 자기자본 추가 투자 등을 통해, 초거대 AI모델(1T급)로 확장하겠다는 포부를 밝혔다.

③다수 의료기관 등과 협업을 통해 현장에서의 실증과 시장성을 강화하고, 카카오헬스케어 플랫폼을 통한 대국민 서비스 실증에도 중점을 둘 계획이다.

※ 전국민 대상 다양한 건강관리 서비스 제공 및 의료 사각지대 최소화 목표

④국산 NPU(리벨리온社) 적용 등을 통해 추론 비용 절감도 추진하는 한편, ‘AI모델 + 국산반도체’ 패키지로 글로벌 진출을 도모한다. 또한, ⑤다수 대학들이 참여하고 있으며, 다양한 해커톤 개최 등도 계획 중인 만큼, 실무형 AI연구 인재 양성 등에도 다방면으로 기여할 수 있을 것으로 기대된다.

< KAIST 컨소시엄 >

“차세대 바이오 파운데이션 모델 K-Fold”

주관·참여기관 KAIST 컨소시엄에는, KAIST와 함께 유관 기업과 협회 등 총 5개의 국내 우수 기관이 참여하였다.

※ (기업 : 3) 히츠(HITS), Merck, 아토텍
(협회 : 2) 한국제약바이오협회, 한국바이오협회

과제목표 이들은 통계적 경향성에 의존이 큰 기존 글로벌 유사모델(Google Alpha Fold3) 방식(MSA* 기반 예측)에서 탈피해, 물리·화학적 상호작용의 인과구조를 학습한 바이오 특화 AI 파운데이션 모델을 개발할 계획으로,

* (MSA) Multiple Sequence Alignment, 다중 서열 정렬

단백질의 동적·다중 상태와 정량적 결합력을 예측 가능하면서도, 빠른 추론 속도의 실용적 AI 파운데이션 모델을 확보하겠다는 목표를 제시하였다.

주요특징 동 컨소시엄은 ①국내 최고 수준의 연구진 참여로, 기존 글로벌 모델의 한계를 극복하는 신기술 전략을 채택해, 동 분야 글로벌 Top1 AI모델을 처음부터(From Scratch) 개발·확보한다.

②개발된 바이오 AI 파운데이션 모델을 오픈소스(Open source, 상업용, Apache 2.0 라이선스)로 공개하는 것은 물론, 7B급 메인 AI모델에 더해, 2B급 경량 AI모델도 개발하여 국내 생태계 참여자들의 기술 접근성 극대화를 도모할 계획이다.

③SaaS(서비스형 SW, SW as a Service) 형태로도 AI서비스 배포 계획을 갖고 있어, 다양한 연구자들의 접근 장벽을 낮추는 기반 마련에도 힘쓸 예정이며,

④글로벌 신약발굴 플랫폼인 Digital Chemistry Solution(Merck社) 연계 등을 통해 글로벌 협업 강화, 글로벌 시장으로 확산을 추진하고, ⑤국내 바이오 우수 협회(한국제약바이오협회, 한국바이오협회)와의 연계로 AI 서비스 활용 확산과 AI·바이오 인력 교육·양성 강화(워크숍 정례화 등)에도 힘쓰겠다는 포부를 밝혔다.

< 종합 >

①글로벌 의과학·바이오 시장이 급성장하는 가운데, 과기정통부는 우리나라가 강점을 갖는 의과학·바이오 분야에 AI 기술을 접목함으로써, 산업 혁신을 가속화하는 전략적 프로젝트를 추진한다.

특히, AI 전문부처가 중심이 되어, AI 중심의 패러다임 전환을 이끌고, AI로 의과학·바이오 산업의 혁신과, 글로벌 경쟁력 제고를 도모할 수 있는 기반을 마련한다는 점에서 의미가 있다.

②선정된 두 컨소시엄은 자체 기술력으로 처음부터 AI모델 개발하여(From Scratch), 독자적 AI 기술력 확보는 물론, 국내 AI 생태계와 의과학·바이오 생태계의 동반 성장과 확장에 기여할 것으로 전망된다.

개발된 특화 AI 파운데이션 모델은 `26년 내 글로벌 Top 수준으로 구축되어 오픈소스(상업용 오픈소스) 형태로 즉시 활용 가능토록 공개한다는 목표이다.

③기업·대학·의료기관·협회 등 다양한 주체가 컨소시엄에 참여하며, 특화 AI 파운데이션 모델의 개발·확산을 위한 목표와 로드맵을 보유하고 있어, 개발된 AI 모델이 의료 진단, 신약 개발, 바이오 연구 등 해당 도메인 전반에 폭넓게 적용되어 AX 성과를 창출할 수 있을 것으로 기대된다.

과기정통부는 AI 고속도로 구축·확장과 이를 통한 대한민국 AI 대전환, AI 강국 도약 등을 적극 추진하고 있으며, 앞으로도 AI컴퓨팅 인프라 등을 기반으로 국내 산업 전반의 혁신을 가속화할 계획이다. 이번 프로젝트는 AI 고속도로 위에서 달리는 실질적 성과를 만드는 출발점이 될 것으로 기대된다.

본 프로젝트의 성과 확장 등을 지원하기 위해, 데이터 스페이스*(`26년 신규 사업, 정부안 기준) 등 데이터 지원 사업 연계도 검토하며, 이를 통해 의료 등 분야의 데이터 활용 촉진 도모에도 노력할 계획이다.

* AI 모델·솔루션 개발 등을 위해 분야별 고품질 데이터를 안전하게 공유하고 발생한 수익을 배분하는 데이터 유통 생태계 구축(의료 분야부터 시작해 도메인 확장 목표)

과기정통부	인공지능인프라정책관	책임자	과 장	양기성 (044-202-6560)
	인공지능기술기반정책과	담당자	사무관	이현우 (044-202-6566)


참고1

루닛 컨소시엄 과제 개요

과제명	분자에서 인구까지 전주기 의과학 혁신을 위한 멀티스케일 의과학 특화 파운데이션 모델 개발	
주관기관	주식회사 루닛	
참여기관	기업	(주)트리온랩스, (주)카카오헬스케어, (주)아이젠사이언스, SK바이오팜(주), (주)디씨클, 리벨리온(주), (주)스탠다임
	대학	한국과학기술원(최유재/김태균/예종철/김현우/홍승훈 연구실), 서울대학교(정유성 연구실)
	의료기관	국민건강보험공단 일산병원, 용인세브란스병원, 강동경희대학교병원, 경희의료원, 고려대학교 산학협력단(고려대 안암·구로·안산병원), 건양대학교 병원, 이화의대부속서울병원, 계명대학교 동산의료원, 양산 부산대학교병원
최종산출물	모델	- 의과학 파운데이션 모델(최대 32B) - 화합물·단백질 도메인 특화 모델(최대 32B) - 오믹스 도메인 재특화 모델(최대 32B) - 임상 도메인 재특화 모델(최대 32B)
	서비스	- CDSS(Clinical Decision Support System, 임상 의사결정 지원 시스템) 에이전틱 시스템 - BMCS(Bio-Medical Co-Scientist, 바이오 의학 공동 과학자) 에이전틱 시스템 - 대국민 건강관리 챗봇(카카오 플랫폼 내 케어챗 서비스)
라이선스	Apache 2.0 라이선스(상업용, 소스코드), Responsible AI EULA(가중치 등)	
학습데이터 규모	- (1단계) 범용 지식데이터 9T(L1, 일반 텍스트/코드), 특화 데이터(L5(의약/안정성), L6(문헌/임상), L7(EMR/청구/설문) 데이터 최대 500B) - (2단계) 특화 데이터(L2(화학/ADMET), L3(단백질/항체), L4(오믹스/경로), L5(의약/안정성), L6(문헌/임상), L7(EMR/청구/설문) 데이터 최대 1T), 에이전트 특화 데이터(SFT 데이터 약 10M 케이스, RL 데이터 약 1M 케이스)	
추진목표	"임상결과 향상"을 겨냥한 증거사슬로 의과학 전주기 혁신	
과제내용	모델개발	<p>① 전주기 의과학 특화 파운데이션 모델 개발·상용화(세계최초) - (1단계) CDSS/BMCS 파일럿용 모델·API 개발(최대 24B*) - (2단계) L1-L7 통합 멀티모달 모델(최대 32B**)</p> <p>② 도메인별(화학·단백질, 오믹스, 임상) 특화 모델 개발(최대 32B**) * 8B~24B 범주에서 Scaling Law 실험에 따라 결정 ** 16B~32B 범주에서 Scaling Law 실험에 따라 결정</p>
	사업화	- (국내) CDSS(참여병원 11개 기관), BMCS(참여 제약사, 의료기관 13개 기관) 에이전트 시스템 실증·상용화 적용, 국민건강관리 챗봇 상용화(카카오 케어챗 서비스 적용) - (국외) 일본·중동 각 1건이상 PoC 수행 및 해외 PoC 2건 이상, LOI 2~3건 확보 목표, 동남아 공공 의료기관에 시스템 실증
기대효과	<p>① 경제적 효과 : 의과학 전주기 파운데이션 모델이 의료지출 9.8조달러 시장에서 큰 경제적 효과를 창출할 것으로 기대 * WHO, Global Health Expenditure: 2021 update, 2023</p> <p>② 사회적효과 : 모델을 기반으로 파생된 상용화 에이전트와 서비스를 통해 전국민 대상 다양한 건강관리 서비스 제공 및 의료 사각지대 최소화</p> <p>③ 의료주권 확보 : 국산 NPU 기반, 해외 종속 없는 자립형 AI의료 생태계 구축</p> <p>④ 산업확산 : 11개 기관 실증 → 전국 상급종합병원·제약사 확산</p> <p>⑤ 인재양성 : 국내 학생·연구자 대상 분산학습 실습 및 해커톤 운영</p>	

참고2

KAIST 컨소시엄 과제 개요

과제명	차세대 바이오 파운데이션 모델 K-Fold	
주관기관	한국과학기술원	
참여기관	주식회사 히츠, Merck(머크) 주식회사, 아토랩 주식회사, 한국제약바이오협회, 한국바이오협회	
최종산출물	모델	- '복합체 구조 예측' 바이오 파운데이션 모델 개발(K-Fold) (7B, 경량화 : 2B)
	서비스	- SaaS/DaaS 형태로 K-Fold 플랫폼 상용화 - 머크의 Digital Chemistry Solution 플랫폼을 통해 글로벌 상용화 - InnoCORE-AI-CRED 프로그램과 연계하여 바이오-AI 융합 인재 양성
라이선스	Apache 2.0 라이선스(상업용)	
학습데이터 규모	- (1단계) 200K 구조 데이터, 45M 서열데이터 - (2단계) 700K 구조-실험 & 합성 데이터	
추진목표	AlphaFold3의 한계점을 극복하고 현업에서 실용적으로 사용될 수 있는 차세대 바이오 파운데이션 모델 개발	
과제내용	모델개발	<p>자유상태 구조 샘플링 모듈, 구조-서열 통합 표현형 모듈, 구조 및 결합력 예측 모듈을 유기적으로 결합해 정확하고 효율적인 바이오 파운데이션 모델 개발</p>  <p>- (1단계) 다양한 생물학적 데이터 기반, 생체 분자 상호작용 예측 바이오 파운데이션 모델 구축(2B) - (2단계) 전문가 수준 예측 바이오 파운데이션 모델로의 고도화* 및 사용자 피드백 기반 서비스 플랫폼 구축(7B, 경량화 : 2B) * MSA의 통계적 패턴에 의존하는 알파폴드3의 한계점을 보완하여 SOTA 대비 추론 속도를 획기적으로 개선 목표</p>
	사업화	<p>- (1단계) 클라우드 기반 웹 플랫폼에 K-Fold(개발 모델) 탑재 및 GUI 기반 기능 제공 - (2단계) SaaS/DaaS 형태 상용화, 머크 Digital Chemistry Solution 플랫폼으로 글로벌 서비스 추진 - (인재양성) 개발된 모델과 플랫폼을 국내 대학 및 연구소에 교육용으로 제공</p>
기대효과	<p>① 경제적효과 : K-Fold 기반 서비스형 신약개발 모델 통해 향후 5년간 5조 원 이상의 직·간접 경제적 가치*(투자 대비 23배 ROI) 창출 전망 * 신약 1개당 수천억 원 규모의 R&D 비용 절감과 블록버스터 신약 창출(연매출 1조원 이상)을 달성하고, 대한민국의 차세대 성장동력인 '테크-바이오 산업' 육성</p> <p>② 사회적효과 : 국산 파운데이션 모델을 기반으로 감염병 발생 시 유전자 정보만으로 수일 내 치료제 후보를 도출할 수 있는 플랫폼을 제공하여 '감염병 대응 100일 작전'을 실현하고 국민 보건 안보와 신약 주권을 강화</p> <p>③ 기술주권 확보 : 알파폴드3와 같은 해외 빅테크 중심의 기술 종속을 탈피하고, 개방형 Bio-AI 생태계를 확립하여 국가 기술 자립 기반을 구축</p> <p>④ 산업확산 : SaaS/DaaS 형태로 상용화 머크 Digital Chemistry Solution 플랫폼으로 글로벌 서비스 확산</p> <p>⑤ 인재양성 : 국내 바이오 협회(한국제약바이오협회, 한국바이오협회)와의 연계를 통해 AI 서비스 활용 확산과 AI-바이오 인력 교육·양성 강화</p>	