
2024년도 초거대AI 확산 생태계 조성 사업(2차) 공고문

2024. 05월



과학기술정보통신부

NIA

지능정보원
한국지능정보사회진흥원

과학기술정보통신부 공고 제2024-0645호

2024년도 초거대AI 확산 생태계 조성사업(2차) 공고

과학기술정보통신부와 한국지능정보사회진흥원(NIA)은 AI 성능 향상 및 서비스 개발을 위한 초거대AI 확산 생태계 조성사업(2차)을 공고하오니 참여를 희망하는 기업·기관은 신청하여 주시기 바랍니다.

2024년 5월 31일

과학기술정보통신부 장관

한국지능정보사회진흥원 원장

< 목 차 >

1. 지원목적 및 배경	1
2. 지원사업 주요내용	2
3. 수행기관 선정방안 및 평가기준 등	8
4. 사업 요구 사항	13
5. 사업 추진체계 및 절차 등	25
6. 제안서 접수 및 방법, 서류 등	27
7. 주요 추진일정(안)	31
(붙임)초거대AI 확산 생태계 조성 사업 세부 과제별 요구사항	32

1 지원목적 및 배경

□ 지원 배경

- 챗GPT의 등장으로 인공지능(AI) 기술의 활용이 일상화·보편화되고, 기존의 '분류형' 중심의 기술에서 '생성형 AI'로 대전환의 계기 마련
- 대통령의 뉴욕 구상 및 국정과제 구체화를 위한 「대한민국 디지털 전략(22.9)」을 비롯하여 「초거대AI 경쟁력 강화방안(23.4)」, 「서비스 산업의 디지털화 전략(23.7)」 등 초거대AI 모델 구축·활용 계획의 구체화
- 민간 기업들은 글로벌 AI 산업 선점을 위해 AI인프라와 연구개발에 대한 꾸준한 투자와 함께 플랫폼(AI 허브 등)을 통한 데이터 공유 요구

□ 지원 목적

- 디지털 전환의 핵심 자원이자 AI서비스 경쟁력 제고의 관건이 되는 초거대AI 데이터 구축·개방을 통해 AI 생태계 조성 및 일상화 실현
 - 모든 분야의 AI 도입 확산과 기술 발전을 선도할 수 있는 대규모 데이터 확보 및 민간 데이터 구축 사업 촉진
 - 체계적인 데이터 자원을 확보하여 AI용 데이터 부족 문제를 해소하고 국내 기업·기관 등의 AI 도입·개발에 대한 진입장벽 완화
 - 고품질·대규모 AI 데이터를 구축·개방하여 초거대AI 발전 기반 조성을 통해 세계 최고의 AI강국 실현 도모

⇒ 거대언어모델(LLM) 등 초거대AI 확산과 AI 현장 수요에 대응하여 AI 산업 고도화 및 디지털 플랫폼 정부 기반 제공

2 지원사업 주요내용

□ 공고 및 접수 기간 : '24. 5. 31(금) ~ 7. 3(수)

※ 분야별 접수마감 기한(7.2~7.3)은 상이하므로 접수일정은 27페이지 필히 참조

□ 지원 기간 : 협약일로부터 ~ '24. 12. 31.

※ 과제의 협약 시작일은 협약체결 완료 월의 1일로 소급 적용하여 협약체결 예정

□ 지원 대상 : 기업, 대학, 공공기관, 정부·지자체, 협회, 개인사업자 등

□ 지원 규모 : '24년 총 82억원 (1차 재공고 8종 및 2차 신규 7종 포함)

○ (지원규모) 정부지원금 5~6억원 수준 (데이터 종별)

※ 정부 지원금은 과제조정위원회 심의·조정에 따라 일부 조정될 수 있음

○ (지정 및 자유공모) 10개 분야 12개 수행기관 선정·지원

- (지정공모/재공고) 6개 분야(8종 데이터) / 6개 수행기관 선정·지원

구분	연번	분야(과제)	데이터	수행기관	예산* (억원)
제조·로보틱스	1	사람 행동 인식 로봇 자율 행동 데이터	1종	1개	5
	2	대규모 물리 환경 로봇 조작 데이터	1종	1개	5
교통·물류	3	사고위험 환경에서의 운전습관 데이터	1종	1개	5
	4	물류공간 예측 데이터	1종	1개	5
교육	5	교과 데이터	3종	1개	15
미디어·콘텐츠	6	다화자 자연 발화 음성 데이터	1종	1개	5

※ 상기 6개 분야(과제)는 초거대AI확산 생태계 조성사업 1차 공고에 대한 재공고임(내용 일부 변경)

※ 상기 6개 분야(과제)가 금번 재공고에서도 유찰 또는 평가 부적격이 발생하는 경우, 자유공모로 전환하여 재차 공고될 수 있음

- (자유공모) 4개 분야(7종 데이터) / 6개 수행기관 선정·지원

구분	연번	분야(과제)	데이터	수행기관	예산* (억원)
글로벌 데이터	7	글로벌 규범·문화 데이터	2종	1개	12
	8	글로벌 규범·문화 평가 데이터	1종	1개	6

구분	연번	분야(과제)	데이터	수행 기관	예산* (억원)
온디바이스 AI 서비스 지원	9	온디바이스 AI 서비스를 위한 멀티모달 데이터	2종	2개	각 6
자유주제	10	AI 기술 및 산업 혁신에 필요한 데이터	2종	2개	각 6

- ※ 벤치마크 데이터셋은 하위 측정 항목 10개 이상 포함하여 각 항목별 1,000개 이상의 셋 구성
- ※ 벤치마크 데이터셋은 품질 검증시 다양한 LLM, LMM으로 벤치마크 성능 측정 실험 구성
- ※ 데이터셋 중 10%는 향후 AI허브 LLM, LMM 리더보드에 포함할 수 있도록 구성
- ※ 현재 제시된 예산은 최대 범위이며, 과제조정위원회 등을 통해 조정될 수 있음
- ※ 2024년도 예산 상황 변화에 따라 선정 규모 및 평가 일정 등은 변경될 수 있음
- ※ 기관생명윤리위원회(IRB)와 데이터심의위원회(DRB) 운영이 필요한 보건의료 데이터는 제안 불가
- ※ 글로벌 데이터는 데이터 구축개방 이후, 해외에서 협력 요청이 있을 경우 개방한 규범문화 데이터를 해외에 제공할 수 있도록 협조해야 함

□ 지원 조건

- 초거대AI 데이터 구축 역량을 갖춘 2개 이상(수요기관 제외)의 기업·기관이 컨소시엄 형태로 수행기관 구성 필수

- ※ 초거대AI 학습에 필요한 다량의 비라벨링 데이터 중심으로 구축하되, 미세조정에 필요한 라벨링 데이터를 추가하는 형태로 추진

- ▶ (수행기관) 주관기관 + 참여기관
- ▶ (주관기관) 과제를 주관하여 수행하는 대표 기관(기업)
- ▶ (참여기관) 주관기관과 공동으로 해당 과제에 참여하여 사업을 수행하는 기관(기업)
- ▶ (수요기관) 과제에 참여하여 과제수행의 결과 발생하는 유·무형의 결과물 또는 서비스 수요자로서 이용하거나 활용하는 기관(기업)으로 정부출연금 지원 없이 과제에 참여

- 주관기관 또는 참여기관 자격으로 최대 2개 분야까지 지원 가능하며 동일 분야 내 주관기관 또는 참여기관으로 중복지원은 불가

- ※ 1개 기업·기관은 총 2개 분야에 지원 가능하나 주관기관 자격으로는 최대 1개까지만 지원 가능(예시: 주관기관 자격 1개, 참여기관 자격 1개 지원 가능 / 참여기관 자격 2개 지원 가능, 주관기관 자격 2개 지원 불가)

* 지정공모 및 자유공모 구분 없이 최대 2개 분야까지 지원 가능(대학 산학협력단 주의 필요)

- ※ 동일 기관이 동일 분야(예: 사람 행동 인식 로봇 자율 행동 데이터)에 주관기관으로 지원 후 다른 컨소시엄을 구성하여 주관기관 또는 참여기관으로 재차 지원 불가