

관리번호	2026-차세대바이오-01-지정공모-12		RFP 유형코드	목적·내용	성과물 특성	지원유형
				R	0	2
국가전략연구 기획평가전문분야	PM분야	차세대바이오	RB분야	원천연구	이론·기술의 정립 및 검증(TRL 2-4)	컨소시엄형
사업명	디지털 AI세포 구축사업					
RFP명	디지털 AI세포 구축 및 실증(컨소시엄)					
	(TRL : [시작] 2단계 ~ [종료] 4단계)					
지원 정보	지원기간	2026.04. ~ 2029.12.	정부지원금	18,750백만원		
	1단계 (1차년도)	2026.04. ~ 2027.12. (2026.04. ~ 2026.12.)	1단계 (1차년도)	8,750백만원 (3,750백만원)		
	2단계	2028.01. ~ 2029.12.	2단계	10,000백만원		
	주관기관유형	<input type="checkbox"/> 제한없음 ■ 대학/출연(연)/국공립연/특정연 ■ 기업 ■ 기타 비영리법인(병원 등) <input type="checkbox"/> 외국법인				
주관기관 외 필수참여기관	<input type="checkbox"/> 제한없음 ■ 기업 <input type="checkbox"/> 기타 비영리법인(병원 등) <input type="checkbox"/> 외국법인					
키워드	한글	AI세포, 멀티모달 오믹스 데이터, 신약개발, 파운데이션 모델, 유효물질				
	영문	AI-powered Cell, Multimodal Omics Data, Drug Discovery, Foundation Model, Hit				

<b>1. 추진배경</b>	
<p>○ 추진근거</p> <ul style="list-style-type: none"> <li>- 과학기술기본법 제11조(국가연구개발사업의 추진), 생명공학육성법 제11조(연구개발사업의 추진)</li> <li>- 「제4차 생명공학육성 기본계획」 (생명공학종합정책심의회, '23~'32)</li> <li>- 「첨단바이오 이니셔티브(안)」 (국가과학기술자문회의, '24.4.)</li> <li>- 「2026년도 국가연구개발 투자방향 및 기준」 (국가과학기술자문회의, '25.3.)</li> <li>- 「AI 바이오 확산전략」 (생명공학종합정책심의회, '25.4.)</li> <li>- 「국정과제」 목표2. 세계를 이끄는 혁신경제 (대한민국정부, '25.9.)</li> <li>- 「AI 바이오 국가전략」 (과학기술정보통신부, '25.12.)</li> </ul> <p>○ 세부 추진배경</p> <ul style="list-style-type: none"> <li>- AI는 더 이상 보조 도구가 아닌 생명과학 혁신의 핵심으로 AI-바이오 융합은 필연적인 흐름임. 특히 분자-세포 수준에서 발생하는 복잡한 생명현상과 질병 기전을 규명하려면, 인간의 인지능력을 초월하는 AI의 패턴인식과 인과관계 분석 능력이 필수적임. 이미 글로벌 바이오 기업들은 AI를 핵심 혁신 엔진으로 채택하여 신약개발 경쟁력을 확보하고 있음.</li> <li>- 디지털 AI 세포*는 분자 상호작용부터 세포 신호전달까지 통합적으로 모델링하여, 질병 기전의 전체적인 그림을 제공함. 분자 수준에서 발생하는 오류가 세포 기능 장애로 이어져 질병이 발생하는데, 이러한 복잡한 과정을 정밀하게 시뮬레이션할 수 있는 플랫폼이 필요함.</li> </ul> <p>* Virtual Cell Challenge(Arc Institute), VirtualCellModles(CZI), Deepmind의 Virtual Cell 개발 선언, AlphaCell(SciLifeLab) 등</p> <ul style="list-style-type: none"> <li>- 현재 <i>in silico</i>에서 예측한 약물의 경우 <i>in vitro</i>, <i>in vivo</i>로 이어지며 비임상, 임상실험에서 대부분 실패하고</li> </ul>	

있는 상황임. 디지털 AI 세포는 분자-세포 수준의 정밀한 시뮬레이션을 통해 실패 가능성을 최소화하여, 초기 단계부터 성공 가능성이 높은 신약 유효, 선도, 후보물질 선별을 가능케 함. 분자 타겟 발굴부터 세포 수준 검증까지 통합된 디지털 AI 세포 플랫폼은 신약개발 기간이나 비용을 혁신적으로 절감할 것으로 기대됨.

- 디지털 AI 세포 플랫폼은 암, 대사질환, 신경퇴행성 질환 등 다양한 질병 연구에 적용될 것임. 나아가 재생 의학, 맞춤형 의료, 합성생물학까지 확장하여 국가 바이오산업 전반의 경쟁력을 획기적으로 강화하는 핵심 인프라 역할을 할 것으로 기대되며, 국제 흐름에 발맞추어 국내에서도 자체적 모델을 선도적으로 개발할 필요성이 있음.

○ **기획의 주안점**

① 세계 최고 수준의 범용 디지털 AI 세포 모델 구축

- 특정 질환에 국한되는 기존 AI 모델들의 한계를 극복하는, 여러 질환에 범용적으로 적용 가능한 통합 AI 모델 구현
- 범용 디지털 AI 세포 모델은 세포 내 계층적 생물학적 구조(멀티스케일: 분자-세포소기관-세포막-세포 전체)와 다양한 분자 상호작용(멀티모달: DNA, RNA, 단백질, 이미징 등 다양한 모달리티 데이터)를 AI 모델 아키텍처에 직접 반영하고 수십 억개 이상 파라미터로 구성된 대규모 멀티모달-멀티스케일 파운데이션 모델을 의미함.
- 모델 경량화 기술(증류 등), 모델 병렬 학습을 적용하여 효율적으로 학습 및 추론이 가능한 최적화된 아키텍처를 설계할 필요가 있음.
- AI 모델의 재현성 및 신뢰성을 담보하기 위해서 체계적인 학습-검증-테스트 프로토콜을 수립

② 고품질 표준 데이터셋 구축

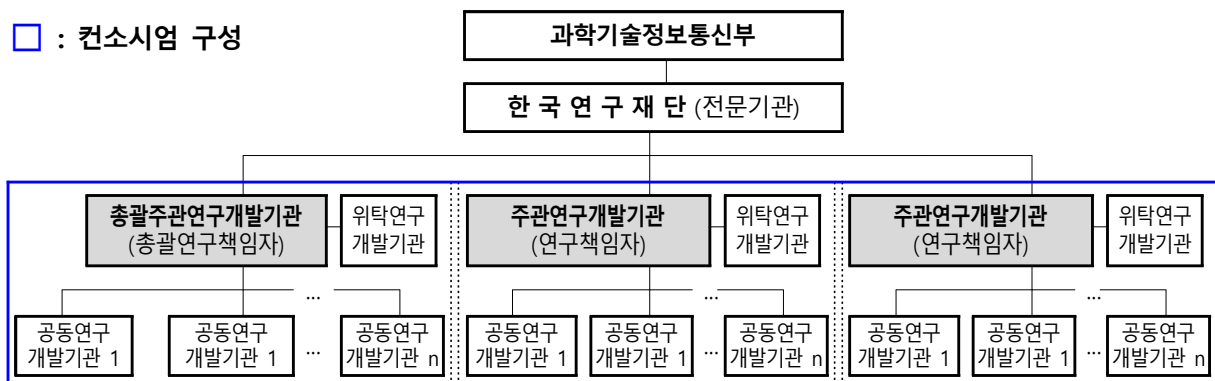
- 다양한 모달리티 데이터(예시: 유전체, 전사체, 단백질체, 대사체, 이미징, 단일세포/공간오믹스 등)를 확보하여 분자로부터 세포로 이어지는 다차원 시뮬레이션이 가능한 디지털 AI 세포 모델을 구축함
- 이를 통합한 고품질 학습 및 검증용 벤치마크 데이터셋을 구축하여 AI 플랫폼의 신뢰성과 실용성을 입증하고 국내 연구자들에게 공개

③ 신약 개발 연계

- 구축된 AI 모델과 데이터셋을 활용하여 실제 신약 개발 파이프라인(신약 타겟 발굴, 유효/선도물질 도출, 독성 예측 등)에 적용하여 시간 및 비용 절감 효과를 정량적으로 입증하고 성공사례를 도출

④ 최적의 컨소시엄 구성 및 운영

□ : 컨소시엄 구성



- 사업의 복잡성과 다학제적 성격\*을 고려할 때, 단독 연구보다는 각 분야 최고의 전문성을 가진 기관들이 주관연구개발기관으로 참여하는 컨소시엄 방식의 추진이 필수적임.

\* 예시: ① 데이터 생산 및 실험 책임 기관, ② AI 모델 및 계산 인프라 책임 기관, ③ 약리, 독성, 동물실험 검증 기관, ④ 플랫폼, 표준, 데이터 연계 기관

※ 연구개발기관 구성(컨소시엄 방식)과 관련된 내용은 특기사항 내용 참고 필수.

## 2. 과제목표

- **최종 목표** : 분자-세포 수준의 생명현상을 정밀하게 시뮬레이션하는 세계 최고 수준의 확장 가능한 범용 디지털 AI 세포 플랫폼(멀티모달-멀티스케일 파운데이션모델)을 구축하여, 이를 통해 신약 개발 성공률을 획기적으로 향상시킬 수 있는 핵심 원천기술 확보 및 질환 특이적 유효물질 발굴 및 실증함.

### ○ 단계별 목표

<p><b>1단계('26~'27)</b></p>	<ul style="list-style-type: none"> <li>○ 디지털 AI 정상세포 구축을 위한 학습용 공개데이터 수집, 생산된 정상세포의 세계 최대 규모 고밀도, 고품질 멀티모달, 멀티스케일 데이터셋 구축</li> <li>○ 정상세포(Untreated, treated) 내 생물학적, 생화학적 현상 및 유전자 조절 회로를 모사하는 세계 최고 수준(State of The Art, SOTA)의 범용적 디지털 AI 세포(멀티모달-멀티스케일 파운데이션 모델) 구축</li> <li>○ 구축된 디지털 AI 정상세포 모델의 기초 성능 및 재현성을 세포 모델 실험을 통해 <i>in silico/in vitro</i> 검증</li> <li>○ 차단계에서 활용할 타겟 질환(3개 이상) 선정</li> <li>○ 범용 모델의 질환 특이적 파인튜닝을 위한 질환 특이적 데이터셋 수집, 생산</li> </ul>
<p><b>2단계('28~'29)</b></p>	<ul style="list-style-type: none"> <li>○ 목표 질환에 특화된 디지털 AI 세포 모델 구축(3종 이상) 및 전문성 향상 검증</li> <li>○ 디지털 AI 질환세포 모델을 활용한 목표 질환의 신규 표적 도출 및 <i>in vitro/in vivo</i> 수준의 표적 제어 효능 평가를 통한 성능 입증</li> <li>○ 웹서비스 기반의 디지털 AI 세포 활용 플랫폼 구축 및 배포</li> </ul>
<p><b>1, 2단계 공통</b></p>	<ul style="list-style-type: none"> <li>○ AI바이오 관련 학회 성과발표 및 전문가 의견수렴 반영</li> <li>○ 디지털 AI세포 프로젝트 소개 논문(Perspective 등) 출판 및 주요 메이저학회(해외) 발표</li> </ul>

## 3. 연구개발내용 및 성과지표

### ○ 연구개발내용

구분	연구개발내용	비고
<p><b>1단계 ('26~'27)</b></p>	<ul style="list-style-type: none"> <li>○ <b>디지털 AI 정상세포 구축을 위한 학습용 공개데이터 수집, 생산된 정상세포의 세계 최대 규모 고밀도*, 고품질 멀티모달, 멀티스케일 데이터셋**</b> 구축                             <ul style="list-style-type: none"> <li>* 동일시료 멀티모달 데이터</li> <li>** 1) 세포 수준 멀티오믹스 데이터: 단일세포/공간오믹스, 후성유전체, 단백질체, 대사체, 이미징 등</li> <li>2) 유전자 조작 데이터: CRISPR/Perturb-seq 기반</li> </ul> </li> <li>- Public 데이터베이스 및 신규 생산된 동일 시료 다중 데이터를 통합하여 세계 최대 규모의 정상세포 상태 데이터(오믹스, 이미지 등) 확보                             <ul style="list-style-type: none"> <li>※ 국내·외 공개 데이터를 우선 수집하여 기본 모델 학습에 최대한 활용</li> </ul> </li> <li>- 멀티모달, 멀티스케일 데이터의 종류와 샘플 수 제시                             <ul style="list-style-type: none"> <li>※ 단일세포 데이터 : 정상세포 ≥ 1,000만 개, 단일세포(Perturb-seq) ≥ 100만 개</li> <li>※ 멀티모달 데이터 : 4종 이상(데이터 간 연결성 강화를 위한 paired sample을 확보)</li> </ul> </li> <li>- 다양한 조직에서 유래된 10종 이상의 세포주에 대한 유전체 수준 perturbation 실험기반의 정상세포 상태 데이터 생산(High contents screening, Perturb-seq 등)</li> </ul>	

구분	연구개발내용	비고
	<ul style="list-style-type: none"> <li>- 데이터 일관성 확보 및 배치효과 최소화를 위한 표준운영절차(SOP 수립) <ul style="list-style-type: none"> <li>※ 데이터 생산 시 실험 디자인, 실험 프로토콜, 품질관리, 표준화 등 표준 프로토콜 제시</li> </ul> </li> <li>○ <b>정상세포(Untreated, treated) 내 생물학적, 생화학적 현상 및 유전자 조절 회로를 모사하는 세계 최고 수준의 범용적 디지털 AI 세포(멀티모달-멀티스케일 파운데이션 모델) 구축</b> <ul style="list-style-type: none"> <li>- 디지털 AI 정상세포를 구현하기 위한 인공지능 기반의 방법론 개발 및 구현</li> <li>- 멀티모달, 멀티스케일 데이터의 통합적 전략 및 인공지능 모델의 아키텍처 제시</li> <li>- 유전자 서열 뿐 아니라 세포 내 신호전달 경로와 위치 정보를 학습하여, 특정 유전자 조절 시 세포의 변화를 예측가능하도록 구현</li> <li>- 최고 성능 AI 모델은 최소 10B 이상의 파라미터로 구성</li> <li>- 실서비스 가능성을 고려해 경량화, 지연시간, 메모리 사용량을 지표로 관리하고 필요시 모델 압축을 적용하여 범용적으로 활용 가능하도록 구현</li> </ul> </li> <li>○ <b>구축된 디지털 AI 정상세포 모델의 기초 성능 및 재현성을 세포 모델 실험을 통해 <i>in silico/in vitro</i> 검증</b> <ul style="list-style-type: none"> <li>- 모델의 올바른 작동유무를 판단하기 위한 방법론 제시 및 구현 <ul style="list-style-type: none"> <li>※ 예시: 정상세포의 생물학적·생화학적 기능에 대한 올바른 perturbation 확인, 정상적인 세포를 모델에 입력 시 정상적인 세포로 올바르게 인식함을 확인 등</li> </ul> </li> <li>- 기 확보된 데이터셋 또는 추가 데이터셋 확보 및 <i>in silico/in vitro</i> 실험을 통해 검증</li> <li>- 기존의 디지털 AI 세포 모델들과의 비교를 위한 벤치마킹 전략 제시 및 성능 비교 수행(성능지표 및 성능목표 제시)</li> </ul> </li> <li>○ <b>차단계에서 활용할 타겟 질환(3개 이상) 선정 및 범용 모델의 질환 특이적 파인 튜닝을 위한 질환 특이적 데이터셋 수집, 생산</b> <ul style="list-style-type: none"> <li>- 디지털 AI 정상세포 구축 후 선정한 타겟 질환에 대한 perturbation 실험을 수행하여 범용적 활용 가능성을 제시</li> <li>- 선정된 3개 이상 질환에 연관된 세포 유래 오믹스 데이터의 생산</li> <li>- 질환에 연관된 기도출 환자 유래 오믹스 아틀라스 정보와 질환 모델 데이터를 통합하여 학습데이터 구축 <ul style="list-style-type: none"> <li>※ 각 질환별 약물 반응, 독성, 바이오마커 데이터(<i>in vitro</i> 수준)</li> </ul> </li> <li>- H&amp;E 이미지를 donor 별 6개 이상의 검증된 슬라이드 확보</li> </ul> </li> </ul>	
<p style="text-align: center;"><b>2단계</b> (‘28~’29)</p>	<ul style="list-style-type: none"> <li>○ <b>목표 질환에 특화된 디지털 AI 세포 모델 구축(3종 이상) 및 전문성 향상 검증</b> <ul style="list-style-type: none"> <li>- 구축된 범용 디지털 AI 정상세포 모델을 생산된 질환 특이적 데이터셋을 통해 파인튜닝하고, 분자 수준에서 세포 표현형 예측까지 연계되는 디지털 AI 질환세포 모델 제시</li> <li>- 도출된 오믹스 아틀라스 데이터와 공간 정보 데이터, 세포간 상호작용 데이터를 통합하여 질환 발생의 공간적 시뮬레이션 모듈 제시</li> <li>- 모델 전문성 평가: 정상세포 모델과 질환세포 모델 비교를 통한 성능 향상 평가</li> </ul> </li> <li>○ <b>디지털 AI 질환세포 모델을 활용한 목표 질환의 신규 표적 도출 및 <i>in vitro / in vivo</i> 수준의 표적 제어 효능 평가를 통한 성능 입증</b></li> </ul>	

구분	연구개발내용	비고
	<ul style="list-style-type: none"> <li>- 디지털 AI 질환세포 모델을 활용한 수만 종의 가상 약물 화합물 또는 CRISPR 등을 활용한 <i>in silico</i> 스크리닝을 통한 신규 표적인자 발굴</li> <li>- 유전체 수준 perturbation 실험(예시: 전장 CRISPR 스크리닝)에서 도출한 결과 (학습에 사용되지 않은 결과)와 모델 데이터를 교차 검증</li> <li>- 질환세포 모델로부터 도출한 표적에 대해 동물 모델 혹은 환자 유래 오가노이드를 활용하여 효능이 입증된 유효물질 질환별 5개 이상 제시</li> </ul> <p>○ 웹서비스 기반의 디지털 AI 세포 활용 플랫폼 구축 및 배포</p> <ul style="list-style-type: none"> <li>- 개발된 디지털 AI 세포 모델을 클라우드 환경에서 직접 가상 실험 수행이 가능한 웹기반 인터랙티브 플랫폼으로 구현하고 K-BDS/K-BDI 등 국가 데이터 인프라에 제공</li> <li>- 사용자가 인터페이스를 통해 세포 유형을 선택하고, 약물 처리나 유전자 조작 등의 조건을 입력하면, 디지털 AI 세포가 가상으로 반응하는 결과를 시각화하는 시스템 구현</li> </ul>	

○ 성과지표

항목	1단계	2단계 (최종목표)	성과수준			비고	
			국내 최고수준	세계 최고수준	기타		
필수	디지털 AI세포 구축과 검증을 위한 데이터셋 구축(정상세포)	○	-	V	세계 최대 수준		
	디지털 AI세포 구축과 검증을 위한 데이터셋 구축(질환세포)	○	-	V	세계 최대 수준		
	범용 디지털 AI세포 모델 개발 원천기술 확보	≥1건	-	-	멀티모달-멀티스케일 파운데이션 모델		
	디지털 AI세포 모델의 성능 검증	○	-	V	현시점(SOTA)보다 우수한 성능을 입증		
	질환 특화 AI 세포 모델	-	≥3개	-	-		
	질환별 유효물질	-	≥5개	-	-		
	디지털 AI 세포 활용 플랫폼	-	○	-	유관 학회 의견 수렴 (예시: 만족도 조사)		
	디지털 AI세포 프로젝트 소개 논문(Perspective등) 출판	○	○	-	-		
	주요 결과 국내외 유관학회 발표	○	○	-	-		
자율	SCIE 논문	단계별 자율제시		-	-	JCR 상위 10% 이내 (본 프로젝트의 결과를 소개하는 플래그십페이퍼 포함)	
	기술이전	단계별 자율제시		-	-	선급금 기준	

4. 특기사항			
기본 특성분류	주요 항목별 담당여부	국가전략기술	<input checked="" type="checkbox"/> Y (첨단바이오/디지털 헬스데이터분석·활용) <input type="checkbox"/> N
		혁신도전형 R&D	<input type="checkbox"/> Y <input checked="" type="checkbox"/> N
		특허로 R&D(舊 IP-R&D)	<input type="checkbox"/> Y <input checked="" type="checkbox"/> N
		경쟁형 R&D	<input type="checkbox"/> Y <input checked="" type="checkbox"/> N
		보안과제	<input type="checkbox"/> Y <input checked="" type="checkbox"/> N
		기술료 징수	<input checked="" type="checkbox"/> Y <input type="checkbox"/> N
		3책5공 적용	<input checked="" type="checkbox"/> Y <input type="checkbox"/> N
		국제공동연구 의무	<input type="checkbox"/> Y <input checked="" type="checkbox"/> N
		지자체 예산매칭 의무	<input type="checkbox"/> Y <input checked="" type="checkbox"/> N
	ESG	<input type="checkbox"/> E(환경) <input type="checkbox"/> S(사회) <input type="checkbox"/> G(지배구조) <input checked="" type="checkbox"/> 해당없음	

- 선정 과제 수 : 1개 과제 내외
  - 1차년도 37.5억원 내외, 2차년도 이후 50억원 내외
- 총 연구개발 기간은 4년(2+2)이며, 1차년도 연구개발 기간은 9개월임.
- 총괄주관연구개발기관을 포함하여 **3개 이상**의 주관연구개발기관을 **필수로 구성하는** 연구개발과제(컨소시엄) 형식으로 제안하여야 하며, **산업체를 필수로** 하여 민관 협력 구조 방안을 제시하여야 함.
  - ※ 명시된 연구개발과제 구성 형식을 반드시 준수하여야 하며, 미준수 시 평가대상에서 제외.
  - ※ 총괄주관, 주관연구개발기관의 역할을 명확히 정의하고 정기적인 소통과 성과 공유를 통해 효율적인 협력체계를 운영해야 함. 특히 주관연구개발기관은 AI모델 개발, AI모델 실증 및 응용연구(신약 유효물질 발굴), 데이터 생산, AI모델 응용연구, 총괄 등의 연구 기능을 구분하여 구성해야 하며, 총괄주관연구개발기관은 총괄기능을 포함한 추가적인 연구 기능을 수행할 수 있음. (1개의 연구과제에서 다수의 기능 수행 가능)
- 주관 및 공동연구개발기관 책임자는 디지털 AI세포 구축 및 실증(컨소시엄) 내 1개의 과제만 신청할 수 있음. (중복 신청 시 해당과제 모두 평가 제외)
  - ※ 참여연구원, 위탁연구개발기관은 중복 신청 시 평가 제외하지 않으나 지양 권고
- 유사과제 수행 또는 참여하고 있는 경우는 중복지원을 지양함.
- 단계평가 후 계속지원 여부를 결정하고, 연구기간 및 연구예산은 변경될 수 있음.
- 11가지 유형의 연구성과에 해당하는 연구개발성과물\*은 전담기관에 등록 또는 기탁하여야 함.
  - \* 제품, 시설·장비, 논문, 특허 등 지식재산권, 연구보고서의 원문, 기술의 요약정보, 생명자원, 소프트웨어, 화합물(化合物), 신품종, 표준 등
- (비)임상 또는 인간 유래 시료를 이용하는 경우 동물실험윤리위원회(IACUC) 또는 기관생명윤리위원회(IRB)의 승인을 받아야 함.
- 인체유래 데이터 생산 시 데이터 활용이 가능하도록 생명윤리법, 개인정보보호법 등 관련 법적 검토를 통하여 제3자 자료 제공 및 활용 동의를 획득하여야 함.
- 본 과제 신청 시, 수집된 데이터에 대한 K-BDS 등록 및 개발된 AI 모델에 대한 K-BDI 연계·공개 계획을 구체적으로 제시하여야 하며, 본 연구에서 확보된 모든 데이터(원 데이터, AI 학습용/평가용 데이터)와 AI 모델은 K-BDS 또는 K-BDI에 등록하여야 함.
- 본 연구에서 개발하는 AI 모델 학습을 위한 데이터의 범위, 모델 개발 수준(범용성, 특화성 등), 모델 평가방법, 필요한 예상 연산량을 구체적으로 제시할 것.

- 본 과제 연구책임자는 AI바이오 혁신 연구거점 조성 (시범)사업과의 협조를 통해 성과교류 및 공동연구 협력을 추진하여야 함.
- 본 과제 연구책임자는 개발된 정상세포 AI 파운데이션 모델을 부처 및 전문기관의 요청 시, 타 연구과제에서 활용할 수 있도록 적극적으로 협조하여야 함.
- 본 과제의 연구개발비로 GPU 등 AI 컴퓨팅 인프라 구축을 지양함(부처에서 추진하는 공공클라우드 활용을 권장함).
- 본 과제 신청 시 실험 데이터 생산, AI 모델 구축을 위한 연구 인력, 장비 등 세부 산정 근거와 데이터 품질을 보장할 수 있는 기준을 마련하여 계획서에 제시하여야 함.