

국책과제 아이템 제안 (Physical AI & 4D Scene Generation)

[과제명 : 인간지향적 차세대 도전형 AI 기술 개발] (창의도전 R&D, 25년 하반기 과제시작)

- 연구개발비 : 1.15억(총 40.15억)
- 연구개발기간 : 총 4년(1+3)
- (개념) 새로운 환경에 스스로 적응/성장하는 "인간지향적 차세대 도전형 범용 AI기술개발"로 現 AI의 지적수준 고도화 및 범용적 활용성 극대화 추진
- (필요성) 범용인공지능(AGI)는 현재 정의가 확립되지 않은 미지의 영역으로 연구자의 자발적 기획을 중심으로 혁신성/도전성이 높은 연구영역을 발굴/지원 필요

주요내용

- ❶ 연구자가 해당 조건(핵심전략 R&D)에 맞는 과제 자율 제안 → ❷ 과제 기획 지원 → ❸ 우수한 연구방법론 후속 R&D 지원
- 디지털세계, 물리세계 등 연결 가능한 AGI 실세계 활용기술 분야, Decision Intelligence, AI-Ready Data, Causal AI, Embodied AI 등 5년 내 안정기에 이를 것으로 예상(Gartner, Hype Cycle for Artificial Intelligence 2024)되는 기술 분야 제안
- (1단계 : R&D 방법론 기획) 다수의 연구자가 AGI R&D 방법론 제안, 지원 후보과제 선별 후 기획 및 방법론 도출
- (2단계/3단계 : R&D 수행) 1단계/2단계 기획과제 중 우수과제 선별, 연구비 지원을 통한 본격 R&D 수행

[아이템] Physical AI / Embodied AI

물리적 세계를 인식하고 이해하며 상호작용하는 AI 기술

시뮬레이터라는 3D 가상환경에 Agent를 생성하여 여러가지 task를 수행시켜 학습시킨 후, 현실의 로봇과 같은 기계에 전이(Sim2Real)하여 현실에서도 task를 잘 수행할 수 있도록 하는 분야

[Physical AI 플랫폼]

1. Cosmos (NVIDIA) : <https://blogs.nvidia.com/blog/cosmos-world-foundation-models/>
 - 텍스트, 이미지, 비디오 등의 입력을 받아 현실 세계의 물리적 상호작용을 이해하고 예측할 수 있는 고품질의 시뮬레이션 영상과 구체적인 행동 결과(의사결정, 힘의 크기, 경로 데이터 등)를 함께 제공하여 AI 모델이 현실 세계에서 어떻게 동작할지를 명확하게 보여줌
 - 이를 통해 실제 환경이 아닌 가상 환경에서 AI 모델을 학습 및 테스트 가능
 - 예) 자율주행 시뮬레이션
 - 입력: 도로의 이미지 데이터 + 교통 상황(차선, 보행자, 신호등 상태) + 차량 센서 정보
 - 출력

- 비주얼 시뮬레이션 : 가상의 도시에서 차량이 실제로 주행하는 장면이 비디오로 생성됨. 차량이 브레이크를 밟거나, 속도를 올리거나, 보행자를 피하는 등의 행동을 보여줌
- 의사결정 데이터 : 특정 조건에서 충돌이 일어날 수 있는 상황을 예측하고 이를 시뮬레이션 비디오로 제공

2. Genesis (NVIDIA) : <https://genesis-embodied-ai.github.io/>

- Physical AI 개발을 위한 물리 현상 시뮬레이션 플랫폼
- 입력
 - 환경 정의 데이터
 - 시뮬레이션 공간의 구조 (평면, 장애물, 경사로 등)
 - 재료 속성 (마찰력, 질량, 밀도, 경도 등)
 - 기타 물리 데이터 (중력, 공기 저항 등)
 - 외부 입력 데이터
 - 센서 데이터 (카메라, LiDAR 등)
 - 명령 데이터 (예: 로봇이 물체를 들어올린다)
- 출력
 - 비주얼 시뮬레이션 (3D 렌더링) 비디오
 - 충돌, 접촉력 등 시뮬레이션 된 물리 데이터

3. VR-GS (World Labs) : <https://yingjiang96.github.io/VR-GS/>

[제안 아이템] : 4D Fully Interactive Scene Generation by Physical/Embodied AI

물리법칙이 적용되어 객체와 상호작용이 가능한 실사 혹은 가상 공간을 생성

[KAIST Research Interest]

김태균 교수

- 실사영상 기반 초실감 3D 공간 재구성
- 3D 공간 물리 시뮬레이터를 활용한 synthetic data 생성 및 학습

오태현 교수

- VLM 기반 물리 현상 이해
- Semantic 4D Gaussian Splatting
- Spatial 3D audio generation

PhyScene: Physically Interactable 3D Scene Synthesis for Embodied AI (CVPR 2024)

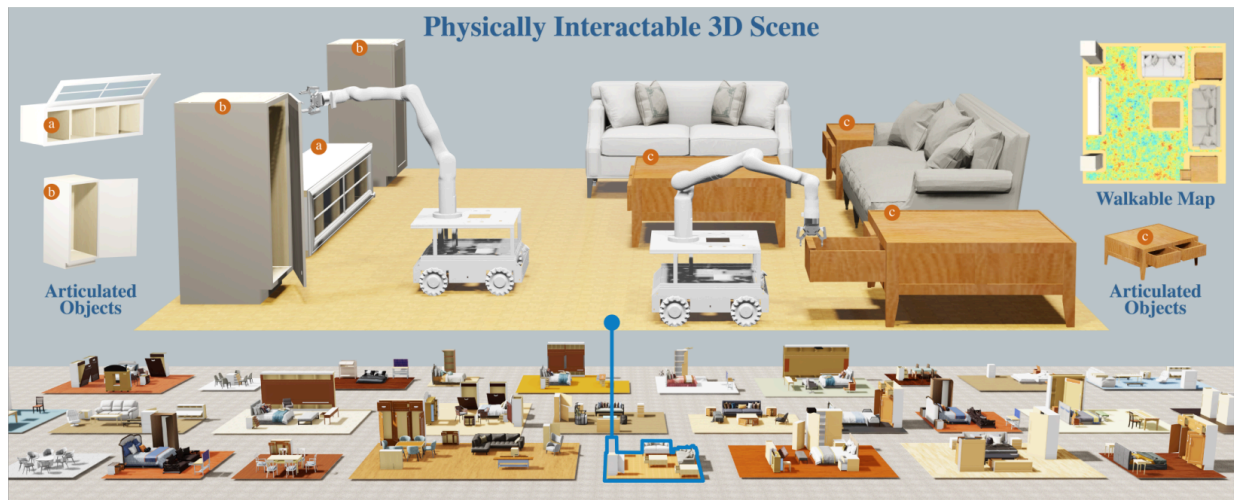


Figure 1. **Illustration of the PHYSCENE**, physically interactable scene synthesis method to generate interactive 3D scenes characterized by **realistic layouts, articulated objects, and rich physical interactivity** tailored for embodied agents.

KAIST 역할 : 핵심원천기술 연구 (Physical AI / VLM / 3D 멀티모달 모델)

KT의 역할 : 실증가능한 플랫폼 개발

KT 사업 정렬성 : 로봇, AI Agent, LLM