사업 전략 참고 자료

공간과 세상을 인지하여 완전 자율을 지향하는 AI

산업 분야의 AI 시스템 구현 방안

- 공간 지능과 월드 모델을 중심으로 -

2026. 3. 31.

작성 : AI융합원(RMH)

심스리얼리티
SIMSREALITY

목 차

1. 배경 및 현황 ---- 1
가. 문제 제기 및 대안 모색 ---- 1
나. 배경 및 필요성 ---- 3
다. 목적 및 주요 내용 ---- 3
라. 국내외 추진 현황 ---- 4

2. 핵심 AI 기술의 도출 ---- 7
가. 핵심 AI 기술의 도출 및 범위 ---- 7
나. 생태계 측면의 핵심 AI 기술 ---- 9
다. 활용성 측면의 핵심 AI 기술 ---- 12

3. 공간 지능 및 월드 모델의 개념과 피지컬 AI와의 관계 ---- 16
가. 공간 지능의 개념과 특징 ---- 16
나. 월드 모델의 개념과 특징 ---- 18
다. 공간 지능 및 월드 모델과 피지컬 AI와의 관계 ---- 20

4. 산업 분야의 최적 AI 시스템 구현 방향 ---- 22
가. 산업 분야 AI 시스템의 구성 ---- 22
나. 디지털 트윈의 고도화 - 일반적 DT vs 생성적 DT ---- 24
다. 최적 AI 시스템의 대안 분석 ---- 31
라. 자율 시스템의 핵심 AI 기술 적용 ---- 35
마. 기능별 핵심 AI 기술의 매핑 ---- 40
바. 단계별 최적 AI 시스템의 구현 방향 ---- 41

5. AI 기반 자율 시스템의 산업 적용 ---- 44
가. 전기차 모터 부품의 AI 자율 제조(예시) ---- 44
나. 항만 컨테이너터미널의 자율 항만(예시) ---- 47

[참고 문헌] ---- 50

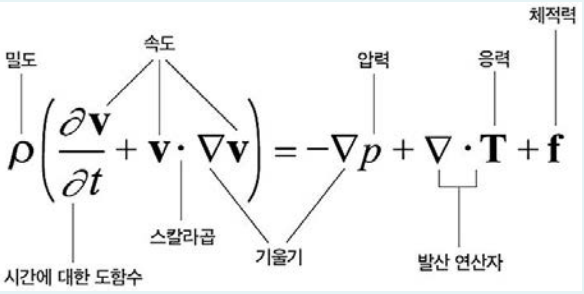
[부록 AI 용어] ---- 54

1 배경 및 현황

가. 문제 제기 및 대안 모색

○ 산업 분야에 AI 적용이 확산되면서 특히 **제조・로봇, 자율공장・다크팩토리 관련 피지컬 AI**가 세계적인 화두이며, 이를 구현하기 위한 각종 AI 모델에는 고등 수학과 물리 이론이 필요한데, **나비에-스토크스 방정식(Navier-Stokes equations)처럼 아직 해결되지 않은 난제를 현장의 피지컬 AI에 어떻게 구현해야 하나?**

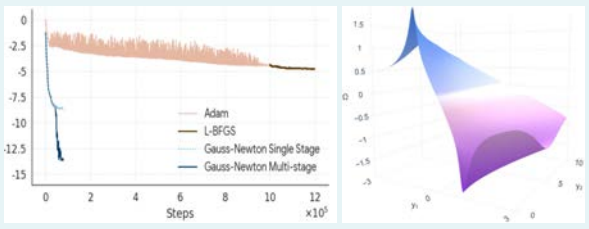
나비에-스토크스 방정식(Navier-Stokes equations)은 점탄성을 고려한 유체(뉴턴 유체, Newtonian fluid)에 대한 운동량 수지식(balance)의 비선형 편미분 방정식임. 1850년에 완성된 이 방정식은 뉴턴의 운동 제2법칙(F=ma)을 유체역학에서 사용하기 쉽게 운동량을 기준으로 하였으며, 기상학, 기계공학, 항공우주, 물리학 분야에서 유체 거동을 기술하는 데 핵심적인 도구로 사용함. 그러나 **이 방정식에 대한 3차원 해(Strong Solution)의 존재(Existence)와 매끄러움(Smoothness)이 증명되지 않은 수학의 7대 난제(Millennium Prize Problems)** 중 하나이며, 이로 인해 현장 적용에 한계



좌변(관성 항)은 시간에 따른 속도 변화와 유체 이동에 따른 대류 가속도를 나타내며, **우변(힘의 합)**은 압력 경도(압력 차이에 의한 힘), 점성항(유체의 끈끈함에 의한 마찰력), 체적력(중력 등 외부에서 작용하는 힘)을 의미함
좌변 2번째 항과 우변 마지막 항이 해를 구하기 어려운 **비선형 편미분 방정식으로 아직까지 일반해가 있는지 없는지조차 모름**

이에 따라 수학적으로 풀 수 없는 물리적 규칙의 나비에-스토크스 방정식을 컴퓨터가 계산할 수 있는 형태의 **대수 방정식(Algebraic equations)으로 변환하여 근사해(Approximate solution)를 구하는 전산유체역학(CFD, Computational Fluid Dynamics)의 계산 방법론을 통해 현실 세계에 적용**

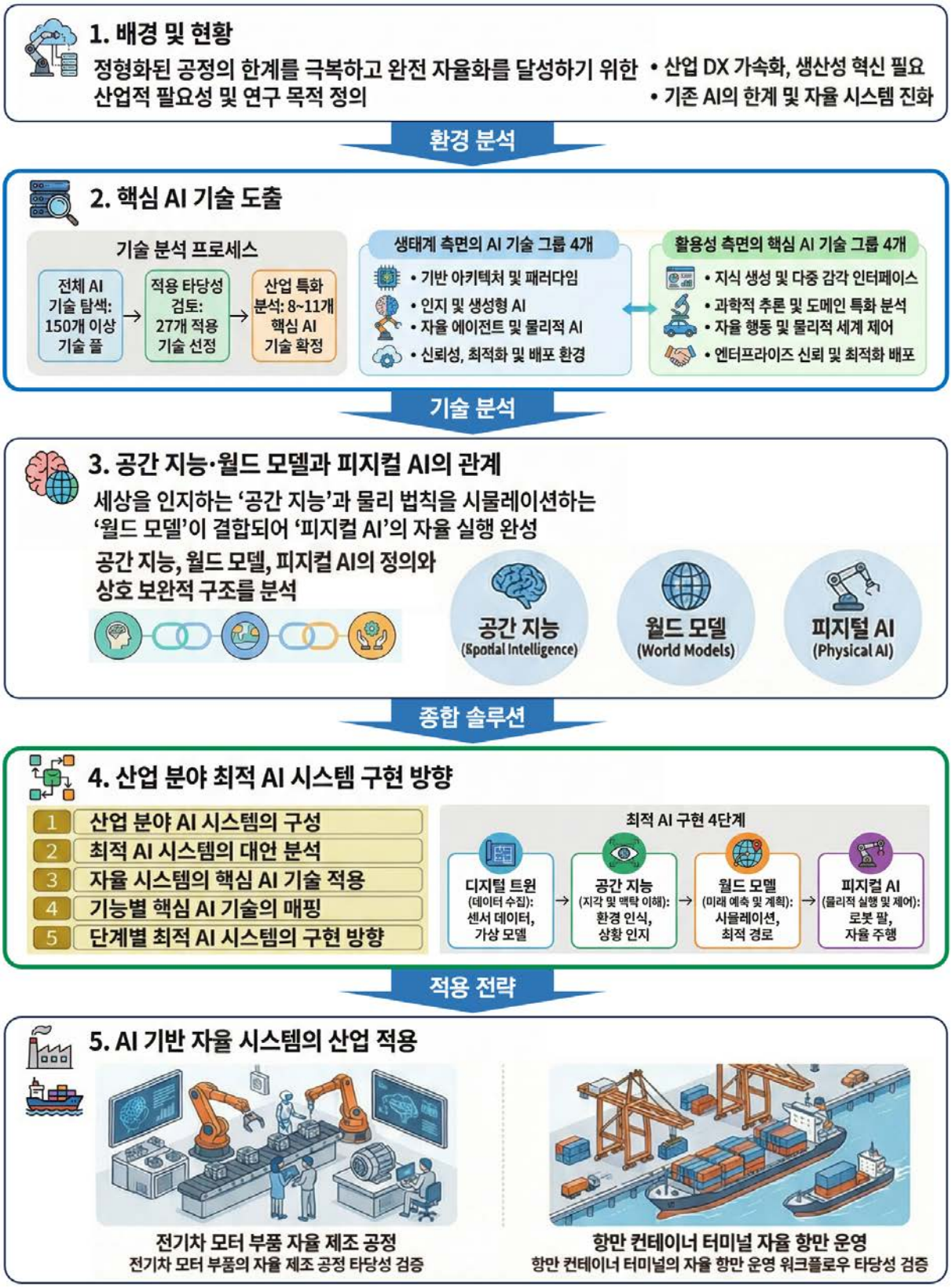
○ 한편, 코넬 대학 arXiv의 “Discovery of Unstable Singularities” 논문(2025.9.17)과 구글 딥마인드의 "Discovering new solutions to century-old problems in fluid dynamics" 발표(2025.9.18)는 물리학의 난제인 유체역학 문제를 AI 기법으로 해결하려는 혁신적 시도로 평가되며, **피지컬 AI의 Spatial Intelligence와** World Models **활용에 중대한 역할**



기존 CFD는 복잡한 비선형 유체 시스템을 계산하기 위해 막대한 컴퓨팅 자원과 시간이 필요했으나, **AI 기반 엔진**은 시뮬레이션 결과를 학습하고 수식의 본질을 내재화하여, **물리적 정확도를 유지하면서도 연산 속도를 기존 대비 수만 배 가속화하는 '대체 모델'의 실효성을 입증함**

이에 따라 **피지컬 AI 적용의 공간 지능(Spatial Intelligence)은 정적(Static) 인식에서 '동적(Dynamic) 유체 공간'까지 3차원상에서 실시간으로 매핑**할 수 있도록 확장되며, 피지컬 AI의 행동 예측 시뮬레이터인 **월드 모델(World Models)은 시각적 패턴의 예측을 넘어, 정교한 물리적 상호작용을 수행**

○ 본 연구는 산업 분야에 **공간 지능과 월드 모델의 기술혁신을 피지컬 AI에 동시에 적용**하는 방안을 검토하는 차원에서 **핵심 AI 기술과 최적 AI 시스템의 매칭**을 분석



(그림 1) 연구 개요 및 분석 체계

나. 배경 및 필요성

1) 배경

❍ 기존 산업 현장의 시스템은 정형화된 공정과 단순 관제에 머물렀으나, 최근 AI 기술은 텍스트나 이미지 생성을 넘어, **스스로 상황을 판단하고 물리적 세계에서 행동하는 '자율형 피지컬 AI(Physical AI)'로 급속하게 진화**

❍ 구글 딥마인드, 엔비디아 등 글로벌 빅테크 기업들은 **3차원 공간을 이해하는 '공간 지능(Spatial Intelligence)'과 물리 법칙을 내재화하여 미래 상태를 예측하는 '월드 모델(World Models)'**을 앞다투어 상용화하며 차세대 산업 혁신을 주도

❍ **현실 데이터를 모으는 '디지털 트윈'이나 제한된 환경에서 작동하는 '로봇 비전' 등 개별 기술의 도입만으로 비정형적이고 복잡한 실제 산업현장의 병목현상 해소 및 완전 자율화에 한계**

2) 필요성

❍ 산업 현장의 생산성, 경제성, 안전성을 극대화하기 위해 현실 세계를 복제하는 **디지털 트윈,** 환경의 의미를 해석하는 **공간 지능,** 물리적 인과관계를 시뮬레이션하는 **월드 모델,** 그리고 이를 현실에서 실행하는 **피지컬 AI가 유기적으로 결합된 최적의 통합 아키텍처 설계를 요구**

❍ 현장 특화의 핵심 AI 기술인 VLM(시각-언어 모델), PINN(물리 정보 신경망), LAM(대형 행동 모델) 등 쏟아져서 **수많은 AI 기술들을 산업 시스템의 각 기능적 단계(인지-예측-실행-통제)에 맞게 정확히 매칭하고 파이프라인으로 연결**하는 전략적 접근이 필요

다. 목적 및 주요 내용

1) 목적

❍ 본 보고서는 **핵심 AI 기술(공간 지능, 월드 모델, 피지컬 AI 등)의 발전 동향을 분석**하고, 이를 **산업 현장에 적용하기 위한 이상적이고 최적화된 AI 시스템 프레임워크를 설계**하는 데 목적

❍ 또한, **도출된 최적 AI 시스템 아키텍처에 기능별 핵심 AI 기술을 세밀하게 매핑**하여, 제조 및 물류 등 **실제 산업 분야에서 생산성과 안정성을 극대화할 수 있는 실행 방안과 모델을 제시**

2) 주요 내용

❍ **(산업 분야 핵심 AI 기술의 도출)** 방대한 AI 생태계를 분석하여 기반 아키텍처, 인지 및 생성형 AI, 자율 에이전트 및 피지컬 AI, 신뢰성 및 최적화 배포 환경 등의 영역으로 그룹핑하고 **산업 자율 시스템의 적용에 필수적인 핵심 AI 기술을 도출**

❍ **(공간 지능 및 월드 모델과 피지컬 AI의 관계 정립) 피지컬 AI 구현의 목표이자 능력인 '공간 지능'과 시뮬레이션 엔진인 '월드 모델'의 개념**을 명확히 하고, 이들이 어떻게 상호 보완되어 사이버-물리 시스템을 완성하는지 구조적 관계를 분석

○ **(산업 분야의 최적 AI 시스템 구현 방향 제시) 디지털 트윈, 공간 지능, 월드 모델, 피지컬 AI의 다양한 결합 대안을 분석**하여 4대 영역이 모두 통합된 완전 자율 제조 시스템으로 아키텍처의 타당성을 입증

○ **(기능별 핵심 AI 기술의 매핑 및 단계별 구현 방향)** VLM, PINN, RL, LAM, XAI 등 구체적 AI 기술을 **데이터 수집(디지털 트윈), 지각 및 맥락 이해(공간 지능), 예측 및 시뮬레이션(월드 모델), 실행 및 제어(피지컬 AI)의 4단계 루프에 매핑**하여 상세 구현 로직을 설계

○ **(AI 기반 자율 시스템의 산업 적용 모델 예시)** 제안된 최적 AI 시스템이 실제 전기차 모터 부품 제조 공정(설계~출하)과 항만 컨테이너터미널 물류 공정(게이트~선석)에 어떻게 적용되어 자율화를 달성하는지 구체적인 워크플로우를 제시

라. 국내외 추진 현황

1) 국내 AI 핵심 기술과 자율 시스템의 개발 및 적용 현황

가) 국내 AI 핵심 기술 및 자율 시스템 트렌드

○ **(글로벌 AI 3강 도약을 위한 '피지컬 AI' 집중 육성)** 한국은 대규모 언어 모델(LLM) 중심의 소프트웨어적 경쟁을 넘어, **세계 최고 수준의 ICT 인프라와 고도화된 제조·산업, 선박·자동차, 물류 등을 기반** 삼아 '피지컬 AI(Physical AI)'를 국가적 전략 자산으로 집중 육성

○ **(기술 주권(Sovereign AI) 확보를 위한 '월드 모델' 독자 개발 가속화)** 엔비디아 등 글로벌 빅테크의 플랫폼 종속을 탈피하기 위해 한국형 월드 모델 개발에 국가적 역량을 투입

- 특히, 가상 시뮬레이션에서 완벽히 학습한 로봇이 현실의 미세한 마찰이나 변수 앞에서 오작동을 일으키는 'Sim2Real(시뮬레이션-현실 격차)' 문제를 해결하기 위해, 디지털 트윈 기반의 고정밀 합성 데이터(Synthetic Data) 생성 기술이 핵심 트렌드로 부상

○ **(산업 도메인 특화 하향식(Top-down) 자율 시스템 실증)** 범용적인 AGI(범용 인공지능) 구현보다는 **제조, 선박, 자동차, 물류 등 특정 산업 현장의 병목을 즉각적으로 해소할 수 있는 맞춤형 시스템 최적화와 빠른 상용화 전략을 추진**

나) 국내 AI 관련 저명 인사 및 학계의 연구 동향

○ **이연수(NC AI 대표)의 WFM(World Foundation Model) 연구:** 피지컬 AI의 두뇌 역할을 하는 독자적 월드 파운데이션 모델을 선도적으로 연구하며, 무거운 영상 생성 단계를 과감히 생략하고 잠재 공간(Latent Space)에서 로봇의 물리적 행동을 직접 생성하여 연산 효율과 제어 정확도를 동시에 끌어올리는 혁신적 해법을 제시

○ **장영재(KAIST 산업및시스템공학과 교수)의 디지털 트윈 기반 자율제조 및 다개체 로봇 제어 연구:** 제조 공장 내 수백 대의 자율이동로봇(AMR)과 무인이송장비(AGV)가 디지털 트윈

공간에서 강화학습(RL)을 통해 스스로 최적의 동선과 작업 순서를 학습하는 '메타 팩토리(Meta-Factory)' 연구를 주도

- 이는 정해진 규칙을 따르는 단순 자동화를 넘어, 설비 고장이나 병목 현상 등 물리적 제약과 동적인 변화를 AI가 실시간으로 인지하고, 공장 전체의 물류 흐름을 자율적으로 최적화하는 '피지컬 AI' 및 '자율제조' 시스템의 핵심 기반 기술로 평가

❍ **손지연(ETRI 책임연구원)의 디지털 트윈 기반 물리 검증 연구:** 피지컬 AI가 현장에서 안전하게 행동하기 위한 필수 학습 인프라로서 디지털 트윈의 역할을 강조하며, 제조 현장의 2D 도면이나 비정형 데이터를 입체적 3D 환경으로 변환할 때 발생하는 시뮬레이션 오차를 수학적/물리적으로 최소화하는 연구를 주도

다) 국내 AI 선도 기업의 시스템 개발 및 적용 현황

❍ **NC AI (고효율 피지컬 AI 및 대규모 합성 데이터 상용화):** 국내 유일의 3D 생성 모델을 결합한 자체 월드 모델(WFM)을 구축하여, 글로벌 최고 성능(SOTA) 모델이 요구하는 컴퓨팅 자원의 단 25%만으로도 로봇 팔 제어 등 고난도 물리 태스크에서 70~80%에 달하는 높은 성공률을 입증하고 효율성을 극대화했으며, 또한, 극단적 시나리오를 대비한 로봇 학습용 합성 데이터를 단기간에 대량으로 생성해 내는 파이프라인을 가동 중

❍ **HD현대 및 물류 선도 기업 (자율운항 및 지능형 물류 제어):** 해상 디지털 트윈과 AI를 결합한 자율운항 선박 상용화에 박차를 가하고 있으며, 선박에 탑재된 통합 AI 시스템과 육상의 원격운항센터(ROC)가 항로 감시, 충돌 회피, 기상 예측 데이터를 실시간으로 교환하며 해양 물류의 완전 무인화를 단계적으로 실증 중

❍ **마키나락스(MakinaRocks) 등 제조특화 기업 (산업용 AI 기반 자율제어 및 예지보전 상용화):** 자동차, 반도체, 배터리 등 고도로 복잡한 제조 공정에 특화된 산업용 AI(Industrial AI) 솔루션을 선도적으로 적용하고 있으며, 이들은 딥러닝과 강화학습을 결합하여 복잡한 다관절 로봇 팔의 제어 궤적을 자율적으로 최적화

2) 해외 AI 핵심 기술과 자율 시스템의 개발 및 적용 현황

가) 해외 AI 핵심 기술 및 자율 시스템 트렌드

❍ **(피지컬 AI(Physical AI)로의 패러다임 전환)** 대규모 언어 모델(LLM) 중심의 텍스트 기반 추론을 넘어, AI가 3차원 공간을 지각하고 실제 물리적 세계에서 로봇 등을 통해 **직접 상호작용하는 피지컬 AI가 글로벌 트렌드의 핵심으로 부상**

❍ **(공간 지능과 월드 모델의 결합)** 정적인 3D 맵핑(SLAM 등)에 머물렀던 과거 기술에서 진화하여, 사물의 기하학적 형태와 의미를 파악하는 **'공간 지능(Spatial Intelligence)'**과 시간 흐름에 따른 물리 법칙 및 상호작용 결과를 시뮬레이션하는 **'월드 모델(World Models)'**이 통합되어 **'공간적 지능을 갖춘 월드 모델(Spatially Intelligent World Model)'로 고도화**

❍ **(물리 난제 해결을 통한 물리 법칙 내재화)** 기존 전산유체역학(CFD) 방식으로는 막대한 컴퓨팅 자원이 필요했던 나비에-스토크스 방정식(Navier-Stokes equations) 같은 물리/수학적 난제를 AI 기반 모델로 해결하려는 시도가 성공을 거두어, 이는 **연산 속도를 기존 대비 수만 배 가속화하면서도 물리적 정확도를 유지하여 피지컬 AI에 정교한 물리 법칙을 내재화**

❍ **(완전 자율 폐쇄 루프(Autonomous Closed-Loop) 아키텍처)** 산업 현장에서는 **'디지털 트윈(가상화 및 검증) + 공간 지능(인지) + 월드 모델(예측) + 피지컬 AI(실행)'**의 4대 요소가 결합된 아키텍처가 완전 자율 제조를 위한 궁극적인 최적 구조로 평가

나) 해외 AI 관련 저명 인사 및 학계의 연구 동향

❍ **페이페이 리(Fei-Fei Li, 스탠퍼드 대학교 교수):** 컴퓨터 비전의 세계적 석학인 페이페이 리 교수는 에세이 **"From Words to Worlds: Spatial Intelligence is AI's Next Frontier"**를 통해 **'공간 지능'**이 스토리텔링, 로봇 공학, 과학적 발견 등에서 인간이 가상 및 현실 세계와 상호작용하는 방식을 혁신할 **AI의 다음 개척지라고 발표**

❍ **얀 르쿤(Yann LeCun, 메타 수석 AI 과학자):** 얀 르쿤 교수는 **월드 모델**을 AI가 주변 환경의 물리적 법칙과 행동에 따른 결과를 내부적으로 시뮬레이션할 수 있는 가상세계의 두뇌로 정의하며, 이를 **인간 수준의 범용 인공지능(AGI) 구현을 위한 핵심 요소로 강력히 주장**

❍ **미국 대학 & 구글 딥마인드**: NYU, 스탠포드, 브라운 대학 등과 구글 딥마인드 연구진의 "Discovery of Unstable Singularities(2025.9.17)", "Discovering new solutions to century-old problems in fluid dynamics(2025.9.18)" 논문은 유체역학의 난제를 AI 기법으로 해결하려는 혁신적 성과를 입증했으며, 이는 **피지컬 AI가 동적 유체 공간을 3차원 상에서 실시간으로 매핑하고 정교한 물리적 상호작용을 수행**하는 데 중대한 역할

다) 해외 AI 선도 기업의 시스템 개발 및 적용 현황

❍ **구글 딥마인드 (Google DeepMind): 월드 모델의 새로운 지평을 여는 'Genie 3'를 개발**하여, 텍스트 프롬프트를 실시간으로 탐험 가능한 동적 가상 세계로 변환하는 시스템을 선보였으며, **Genie 3는 사전 정의된 라벨 없이도 행동 공간의 구조와 상호작용(인과) 구조를 스스로 학습하여 개념적 표상을 형성하는 것이 특징**

❍ **엔비디아 (NVIDIA):** 물리 기반 가상 3D 영상과 시뮬레이션 학습(WFM)으로 **Omniverse 및 COSMOS 플랫폼을 통해 피지컬 AI 구현을 주도**하고, 특히 **'Cosmos-Predict2.5' 모델은 행동이 주어졌을 때 미래 상태를 고신뢰도로 예측하고 정책 기반의 의사결정을 지원**하며, 또한 대규모 공간 지능 프레임워크인 'fVDB'를 출시하여 고품질 그래픽 데이터 구조와 AI 연산을 결합해 도시 규모의 대형 물리 시뮬레이션이 가능

❍ **월드 랩스 (World Labs):** 페이페이 리 교수가 공동 창업한 스타트업으로 **공간 지능을 구현할 월드 모델 구축을 목표**로 설립되었으며, 설립 1년 만에 기업가치가 급등하여 막대한 규모의 투자를 유치하는 등, **공간 지능 기반의 피지컬 AI 시스템에 대한 혁신을 주도**

2 핵심 AI 기술의 도출

가. 핵심 AI 기술의 도출 및 범위

❍ 오픈AI의 ChatGPT 출시(2022.11.30) 이후 AI 기술은 급속하게 진화하고 있으며, 과학, 경제, 산업, 국방, 의료, 교육 등 사회 전반에 활용되어 효율성 및 생산성 향상에 기여

※ **AI와 직접 관련 있는 기술 및 관련 시스템/솔루션은** AI Vision & Paradigm(GenAI 등), Model Architectures(RNN 등), Model Types(LLM 등), Learning Paradigms(RL 등), GenAI Techniques(Prompt Engineering 등), Reasoning & System Design(RAG 등), Agents & Autonomy(AI Agents 등), Trust & Safety(XAI 등), Causality & Probabilistic (Causal AI 등), Infrastructure & Deployment(On-device AI 등), Governance & Policy (Sovereign AI 등), Applications(Image Segmentation 등), Engineering(MLOps 등), Data & Knowledge Systems(Vector Database 등), Evaluation & Metrics(BLEU 등) **150 종류 이상 등장하고 있음**

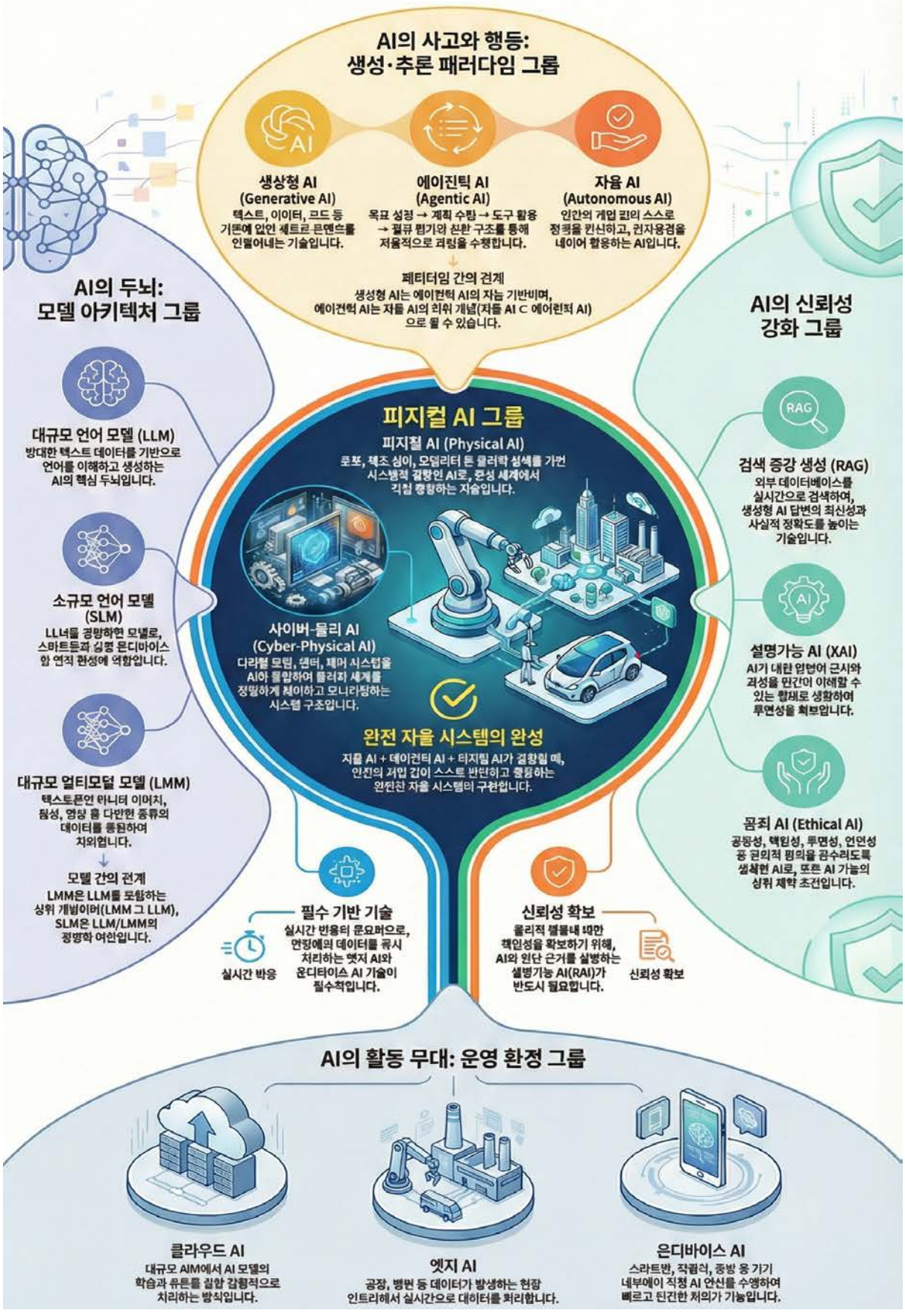
❍ 본 보고서에서는 AI 기술 중 **"디지털트윈▸공간지능▸월드모델▸피지컬 AI"**에 초점을 맞추어 산업 현장에 적용되는 핵심 기술인 LM, SLM, LMM, VLM, VLA, LAM, GAN, GNN, RNN, PINN, RL, Foundation Models, World Models, Diffusion Models, Transformer Models, VAE, 공간 지능, GenAI, 에이전틱 AI, 자율 AI, 피지컬 AI, RAG, XAI, Causal AI, 온디바이스 AI, 엣지 AI, 소버린 AI 등을 연관 매트릭스 기법으로 도출

- 수학적 기반인 **Transformer** 모델은 방대한 데이터 학습을 통해 **Foundation Model**로 발전했으며, **LLM**에서 시작해 시각과 청각을 결합한 **LMM/VLM**으로 인지 능력을 넓히고 있음
- 단순 **생성형 AI**는 **RAG**를 통해 지식의 정확성을 얻고, 목표를 설정하는 추론 능력이 더해지면서 **Agentic AI**와 컴퓨터를 조작하는 **LAM**으로 진화하고 있으며, 여기에 **World Models**를 통한 물리 법칙 이해와 **RL(강화학습)**이 결합하면, **VLA** 모델을 탑재한 로봇 즉 **Physical AI**가 되어 현실 세계에서 작동
- 이 모든 과정은 사용자 보안과 반응 속도를 위해 클라우드에서 **Edge AI**와 **On-device AI**로 구동 환경이 분산되며(**SLM**의 적용), 궁극적으로 각 국가의 특성을 반영하는 **Sovereign AI** 형태로 구축되는 추세

❍ **피지컬 AI 중심**으로 AI 기술들을 그룹핑하면, **'AI의 사고와 행동: 생성·추론 패러다임', 'AI의 두뇌: 모델 아키텍처', 'AI의 활동 무대: 운영 환경', 'AI의 신뢰성 강화'**로 구분할 수 있으며, 피지컬 AI 그룹에는 사이버-물리 AI를 포함

❍ **AI 생태계 측면**에서는 이러한 기술들이 데이터의 패턴을 학습하는 **'기반 아키텍처'**에서 출발하여, 텍스트와 이미지를 이해하고 생성하는 **'인지 생성'**을 거쳐, 실제 환경에서 스스로 판단하고 행동하는 **'물리적/자율적 지능'**으로 확장되고, 실제 산업 및 현장에 효율적으로 적용되는 **'신뢰 배포'** 영역으로 진화

❍ **활용성 측면**에서는 AI 기술들이 독립적으로 쓰이지 않고, 실무 현장에서 강력한 파이프라인으로 연결되어 융합형 시스템으로 구축되는데, 크게 **'콘텐츠/지식 생성', '도메인 특화 분석', '자율적 행동/제어', '현장 신뢰성 확보 및 최적화'**라는 4가지 가치사슬로 그룹핑할 수 있음



(그림 2) 피지컬 AI 중심의 AI 그룹

나. 생태계 측면의 핵심 AI 기술

1) 기반 아키텍처 및 패러다임 (Foundational Architecture)

기반 아키텍처 및 패러다임 분야는 AI 모델이 데이터를 처리하고 학습하는 근본적인 수학적 및 구조적 뼈대로서 피지컬 AI의 핵심

❍ **신경망 아키텍처 분야**

- **RNN (Recurrent Neural Networks):** 시계열 데이터나 순차적 데이터(텍스트, 음성)를 처리하는 전통적인 모델
- **Transformer Models:** 어텐션(Attention) 메커니즘을 사용하여 데이터의 문맥을 병렬로 처리하는 현대 AI의 핵심 구조로서 LLM의 기반
- **GNN (Graph Neural Networks):** 소셜 네트워크, 분자 구조 등 노드와 엣지로 이루어진 그래프 데이터를 분석
- **PINN (Physics-Informed Neural Networks):** AI의 학습 과정에 열역학, 유체역학 등의 물리 법칙(미분 방정식)을 제약 조건으로 추가하여, 과학/공학적 시뮬레이션의 정확도를 높이는 모델

❍ **패러다임 및 기반 모델 분야**

- **Foundation Models (파운데이션 모델):** 방대한 양의 레이블 없는 데이터를 자율 학습(Self-supervised learning)하여 광범위한 다운스트림에 활용할 수 있는 거대 범용 모델
- **World Models (월드 모델):** AI가 주변 환경의 물리적 법칙, 인과관계, 공간적 동역학을 내부적으로 시뮬레이션하고 예측할 수 있도록 하는 모델로 에이전트가 행동을 계획하는 데 필수적인 기술

2) 인지 및 생성형 AI (Perception & GenAI)

인지 및 생성형 AI 분야는 텍스트, 이미지, 소리 등의 데이터를 이해(인지)하고 새로운 콘텐츠를 만들어내는(생성) 기술

❍ **언어 및 모달리티 모델 분야**

- **GenAI (Generative AI):** 데이터의 분포를 학습하여 새로운 텍스트, 이미지, 비디오 등을 만들어내는 모든 AI 기술의 총칭
- **LM (Language Model) / LLM (Large Language Model):** 인간의 언어를 이해하고 생성하는 모델
- **SLM (Small Language Model):** 파라미터 크기를 줄여 특정 도메인에 특화시키거나, 모바일 기기 등 제한된 환경에서 구동 가능하도록 효율화한 언어 모델

- VLM (Vision-Language Model): 이미지와 텍스트를 동시에 이해하여 이미지 캡셔닝, 시각적 질의응답 등을 수행하는 모델
- LMM (Large Multimodal Model): 텍스트, 비디오, 오디오 등 두 개 이상의 다양한 데이터 형태(Modality)를 동시에 입력받고 처리하는 거대 모델

❍ **생성형 아키텍처 분야**

- GAN (Generative Adversarial Networks): 생성자와 판별자가 서로 경쟁하며 진짜 같은 가짜 데이터를 만들어내는 구조
- VAE (Variational Autoencoders): 입력 데이터를 잠재 공간(Latent space)으로 압축한 뒤 복원하는 과정에서 데이터의 확률적 분포를 학습하여 새로운 데이터를 생성
- Diffusion Models (디퓨전 모델): 데이터에 노이즈를 점진적으로 추가한 뒤, 이 노이즈를 다시 제거하는 과정을 학습하여 고품질의 이미지나 비디오를 생성(예: Midjourney)

3) 자율 에이전트 및 물리적 AI (Autonomy & Physical AI)

자율 에이전트 및 물리적 AI 분야는 단순한 답변 생성을 넘어, 스스로 계획을 세우고 디지털 또는 물리적 세계에서 행동(Action)을 진행

❍ **디지털 자율성 분야**

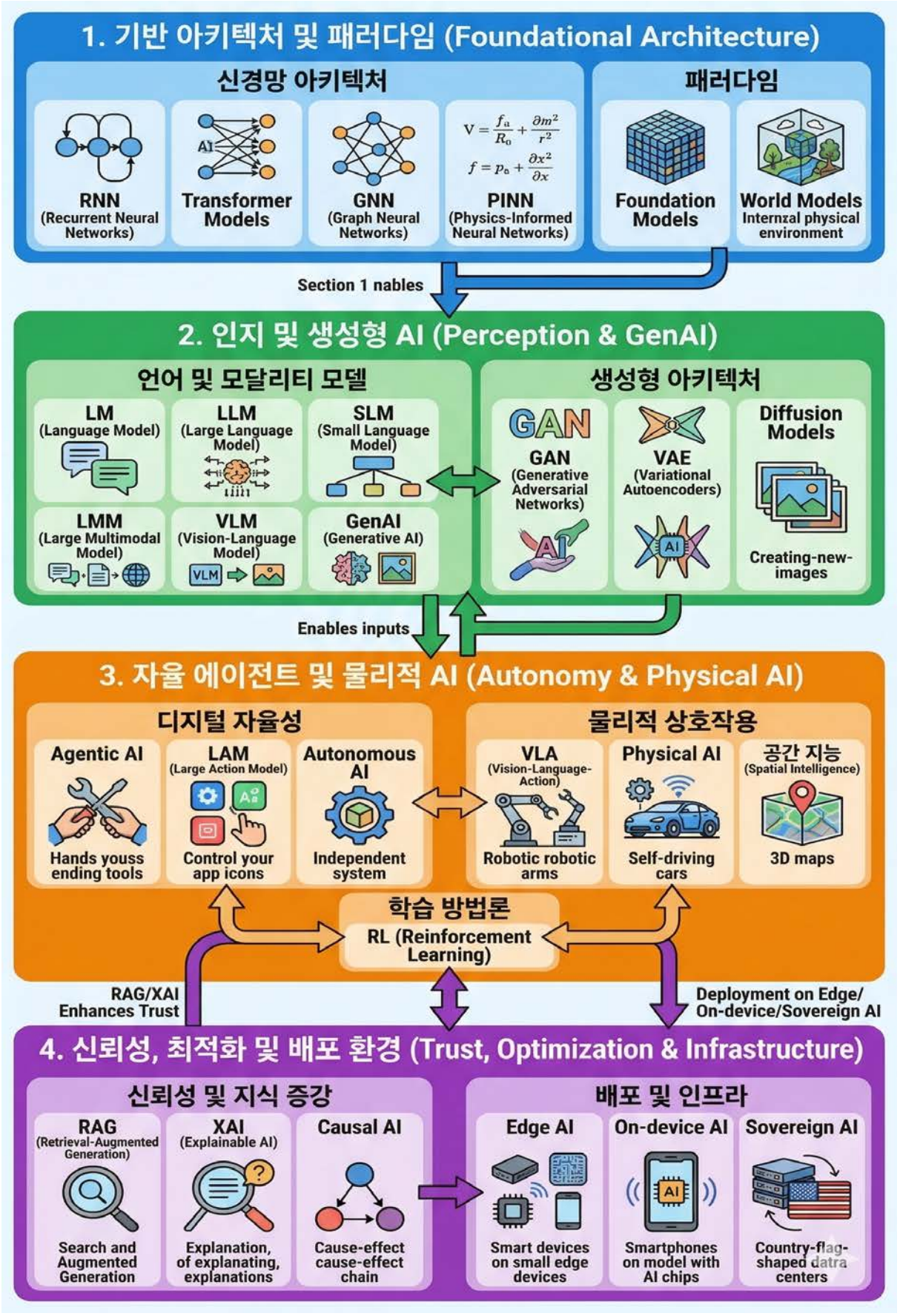
- Agentic AI (에이전틱 AI): 목표가 주어지면 스스로 계획을 분할하고, 검색이나 계산기 등의 도구를 활용하며, 환경과 상호작용하여 임무를 완수하는 주도적 AI
- LAM (Large Action Model): 사용자의 의도를 언어로 이해한 뒤, GUI 환경이나 API를 통해 앱을 조작하는 등 구체적인 컴퓨터상의 '행동'으로 변환하는 모델
- Autonomous AI (자율 AI): 에이전틱 AI의 확장 개념으로, 인간의 개입 없이 장기간 복잡한 시스템을 운영하고 최적화하는 수준의 AI를 의미

❍ **물리적 상호작용 분야**

- Spatial Intelligence (공간 지능): 3D 공간을 인식하고, 사물의 위치와 물리적 속성을 이해하여 현실 세계를 디지털로 모델링하고 상호작용하는 지능형 AI
- Physical AI (피지컬 AI): 공간 지능을 바탕으로 실제 물리적 세계에서 작동하는 로봇, 자율주행차 등의 형태 AI
- VLA (Vision-Language-Action): 로보틱스의 핵심으로, 시각 정보와 언어 명령을 입력받아 로봇 팔의 관절 제어 등 직접적인 '물리적 행동(Action)' 데이터를 출력하는 모델

❍ **학습 방법론 분야**

- RL (Reinforcement Learning – 강화학습): 에이전트가 환경 내에서 행동을 취하고 보상을 받으며 최적의 행동 정책을 학습하는 알고리즘으로 자율 AI와 피지컬 AI의 핵심



(그림 3) 생태계 분야별 AI 기술

4) 신뢰성, 최적화 및 배포 환경 (Trust, Optimization & Infrastructure)

신뢰성, 최적화 및 배포 환경 분야는 AI를 실제 산업 및 일상에 안전하고 효율적으로 적용하기 위한 기술 생태계

❍ **신뢰성 및 지식 증강 분야**

- RAG (Retrieval-Augmented Generation): 언어 모델이 답변을 생성하기 전, 외부의 신뢰할 수 있는 데이터베이스에서 최신/전문 정보를 검색하여 답변의 정확성을 높이고 환각(Hallucination)을 줄이는 기술
- XAI (Explainable AI - 설명 가능한 AI): AI가 특정 결과를 도출한 과정과 이유를 인간이 이해할 수 있도록 시각화하거나 설명하는 기술
- Causal AI (인과 AI): 단순한 데이터 간의 상관관계를 넘어, "A가 B의 원인인가?"라는 인과관계를 추론하여 의사결정의 신뢰성을 높이는 기술

❍ **배포 및 인프라 분야**

- Edge AI (엣지 AI): 중앙 클라우드가 아닌, 데이터가 발생하는 네트워크 가장자리(Edge, 예: 라우터, IoT 허브)에서 AI 연산을 처리하여 지연 시간을 줄이는 기술
- On-device AI (온디바이스 AI): 스마트폰, PC 등 사용자 기기 내부에서 AI가 직접 구동되는 기술로, 개인정보 보호와 오프라인 작동이 강점(SLM과 밀접하게 결합됨)
- Sovereign AI (소버린 AI): 국가나 기업이 외산 클라우드나 글로벌 빅테크 모델에 의존하지 않고, 자체 데이터와 인프라를 바탕으로 독립적인 AI 주권을 확보하는 전략 및 기술

다. 활용성 측면의 핵심 AI 기술

1) 지식 생성 및 다중 감각 인터페이스(Knowledge Generation & Multimodal Interfaces)

지식 생성 및 다중 감각 인터페이스 분야는 사용자와 상호작용하며 콘텐츠를 기획, 창작, 요약, 번역하는 등 창의적이고 지식 집약적인 업무를 지원하는 활용 그룹

❍ **범용 지식 및 언어 처리 분야**

- Transformer Models / Foundation Models: 범용적인 엔진 역할을 수행하며, 대규모 텍스트와 지식을 사전 학습하여 범용적인 작업의 기반을 제공
- GenAI / LM: 텍스트 기반의 카피라이팅, 코드 작성, 문서 요약 등 일상 사무 자동화에 활용

❍ **시청각 미디어 생성 및 편집 분야**

- GAN / VAE / Diffusion Models: 마케팅용 이미지 생성, 게임 에셋 디자인, 영상 합성 등 시각 예술 및 미디어 산업에서 새로운 데이터를 합성하는 데 활용

❍ **다중 감각 경험 제공 분야**

- **VLM / LMM:** 이미지 해설, 영상 콘텐츠의 자동 자막 및 내용 분석, 고객 센터의 멀티모달 챗봇 등 사용자의 다양한 입력(음성, 텍스트, 이미지)을 동시에 처리하는 서비스에 적용

2) 과학적 추론 및 도메인 특화 분석 (Scientific Reasoning & Domain-Specific Analysis)

과학적 추론 및 도메인 특화 분석 분야는 단순한 텍스트나 이미지 생성을 넘어, 복잡한 데이터의 인과관계를 밝히거나 과학/공학적 난제를 해결하는 산업 특화의 분석 도구

❍ **복잡계 및 시계열 분석 분야**

- **GNN:** 데이터가 정형화되지 않은 요소들간 연결 관계가 핵심일 때, 신약 개발(분자 구조 결합 예측), 소셜 네트워크 사기 탐지, 물류망 최적화 등에 활용
- **RNN:** 데이터의 순서가 중요할 때, 금융 시장의 주가 예측, 공장 설비의 센서 데이터를 활용한 예지 보전 등 시간 흐름에 따른 데이터 분석에 활용

❍ **물리 법칙 및 인과성 결합 분야**

- **PINN:** 항공우주, 기계 공학 등에서 막대한 컴퓨팅 자원이 필요한 유체/열역학 시뮬레이션을 AI로 수행하여 빠르고 정확하게 대체하는 데 활용
- **Causal AI:** 마케팅 ROI 분석, 의료 역학 조사 등에서 단순 상관관계를 넘어 "특정 약물이 질병 치료의 진짜 원인인가?"를 규명하여 비즈니스 의사결정의 근거를 제공

3) 자율 행동 및 물리적 세계 제어 (Autonomous Action & Physical World Control)

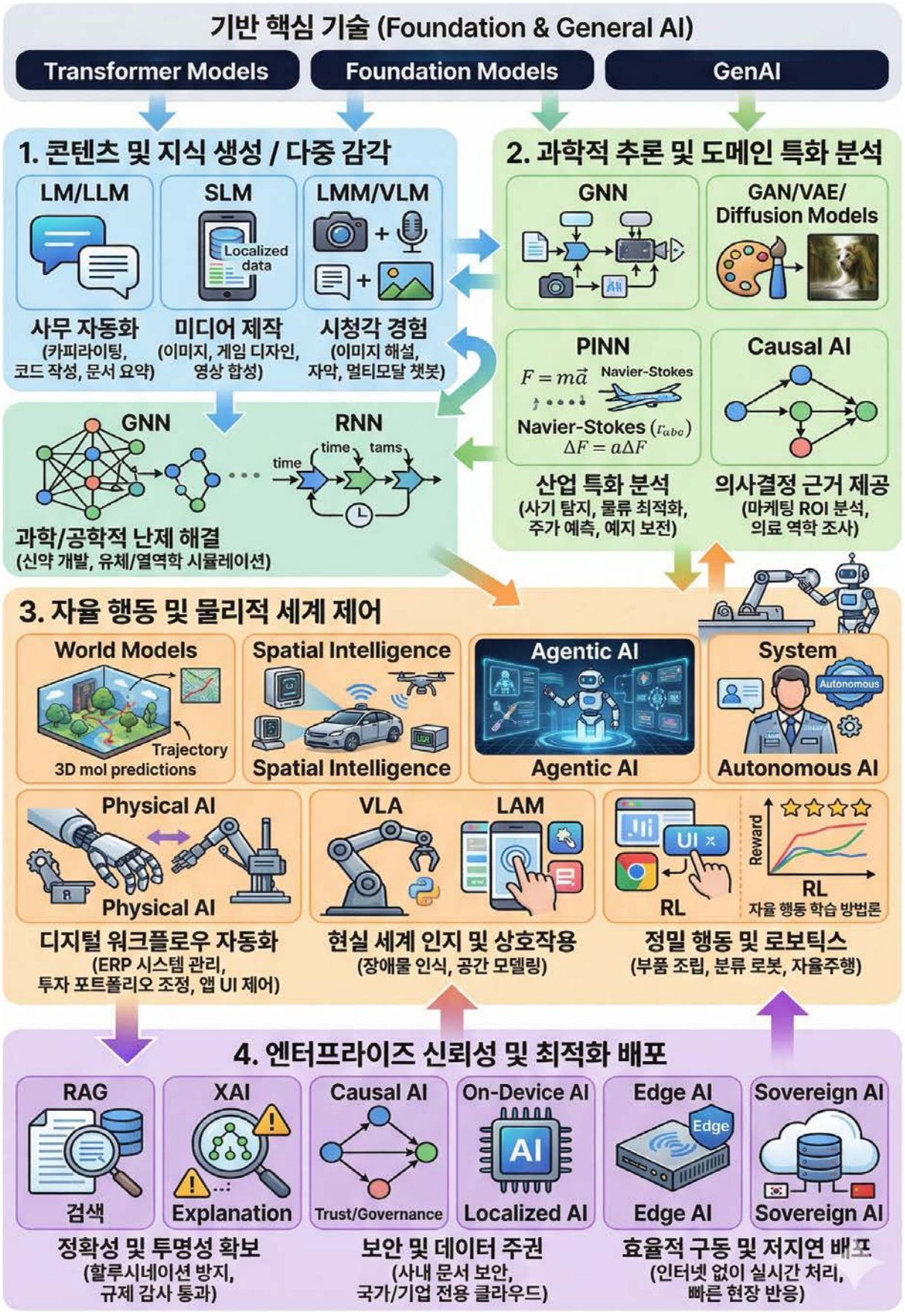
자율 행동 및 물리적 세계 제어 분야는 인간의 지시를 기다리지 않고, 현실 세계나 디지털 환경에서 스스로 상황을 인지하고 목표를 달성하기 위해 행동하는 활용 영역

❍ **디지털 워크플로우 자동화 분야**

- **Agentic AI / Autonomous AI:** 복잡한 ERP 시스템 관리, 자율적인 투자 포트폴리오 조정 등 B2B 소프트웨어 환경에서 독립적인 업무 프로세스 자동화를 수행
- **LAM:** 사용자가 "비행기 표를 예매해 줘"라고 말하면 AI가 직접 웹브라우저나 앱의 UI를 클릭하여 예매를 완료하는 등 디지털 비서의 핵심 제어 기술로 활용

❍ **현실 세계인지 및 로보틱스 분야**

- **World Models / 공간 지능:** 자율주행차나 드론이 주행 경로의 물리적 변화를 예측하고 3D 공간의 장애물을 인식하는 내비게이션 및 시뮬레이션 시스템에 활용
- **Physical AI / VLA:** 공장의 제조 로봇, 물류 창고의 분류 로봇이 시각(Vision)으로 물체를 인식하고 언어(Language) 지시를 받아 로봇 팔(Action)을 정밀하게 움직이는 데 적용
- **RL (강화학습):** 위 모든 에이전트와 로봇이 수많은 시행착오 시뮬레이션을 통해 최적의 행동 패턴을 학습하는 데 사용되는 필수 훈련 방법



(그림 4) 활용성 측면의 AI 핵심 기술

4) 엔터프라이즈 신뢰 및 최적화 배포 (Enterprise Trust & Optimized Deployment)

엔터프라이즈 신뢰 및 최적화 배포 분야는 AI를 기업의 보안 규정, 비용 한도, 국가의 규제, 디바이스의 한계에 맞춰 안전하고 실용적으로 안착시키는 인프라 및 거버넌스

❍ **정확성 및 투명성 확보 분야**

- **RAG:** 사내 보안 문서나 최신 매뉴얼을 AI와 연동하여, 할루시네이션(환각) 없이 정확한 사내 헬프데스크나 법률/의료 자문 시스템을 구축할 때 필수적 적용
- **XAI:** 금융권의 대출 심사 거절 사유, 의료 AI의 암 진단 근거 등 AI의 결정이 법적/윤리적 감사를 통과해야 하는 규제 산업 등에서 활용

❍ **효율적 구동 및 보안 배포 분야**

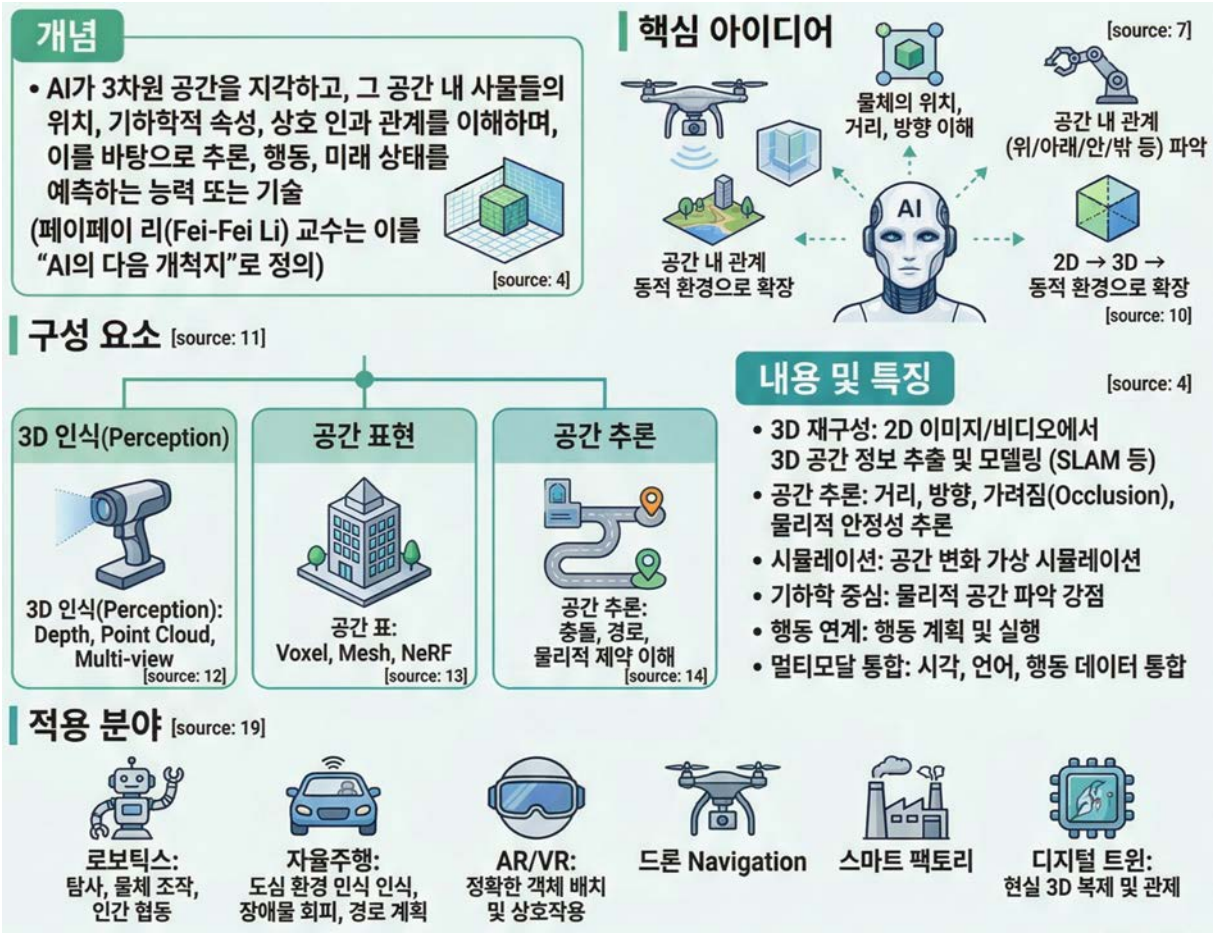
- **SLM / 온디바이스 AI:** 스마트폰의 실시간 오프라인 통역, PC의 로컬 AI 비서 등 인터넷 연결 없이 민감한 개인정보를 보호하며 저전력으로 AI를 구동하는 데 활용
- **엣지 AI:** 자율주행차나 스마트 팩토리 센서 등 수 밀리초(ms) 단위의 빠른 반응 속도가 생명인 현장에서 클라우드 지연 없이 즉각적인 AI 연산을 수행
- **소버린 AI:** 국방, 공공기관, 국가 인프라 등 데이터의 해외 반출을 엄격히 금지하는 환경에서 독자적인 국가/기업 전용 AI 클라우드를 구축하는 전략적 활용

3 공간 지능 및 월드 모델의 개념과 피지컬 AI와의 관계

가. 공간 지능의 개념과 특징

구분	공간 지능(Spatial Intelligence) 설명
개념	• AI가 3차원 공간을 지각하고, 그 공간 내 사물들의 위치, 기하학적 속성, 상호 인과 관계를 이해하며, 이를 바탕으로 추론, 행동, 미래 상태를 예측하는 능력 또는 기술 (페이페이 리(Fei-Fei Li) 교수는 이를 "AI의 다음 개척지"로 정의)
내용	• **3D 재구성:** 2D 이미지/비디오에서 3D 공간 정보를 추출 및 모델링 • **공간 추론:** 사물 간의 거리, 방향, 가려짐(Occlusion), 물리적 안정성 추론 • **시뮬레이션:** 공간 변화에 따른 결과 가상 시뮬레이션 수행
특징	• **기하학 중심:** 물리적 공간의 형태와 배치를 정확히 파악하는 데 강점 • **행동 연계:** 단순히 보는 것을 넘어, 행동을 계획하고 실행하기 위한 지능 • **멀티모달 통합:** 시각, 언어, 행동 데이터를 통합하여 공간을 이해
적용 분야	• **로보틱스:** 로봇의 탐사, 물체 조작, 인간과의 협동 • **자율주행:** 복잡한 도심 환경 인식, 장애물 회피, 경로 계획 • **AR/VR:** 현실 공간에 디지털 객체의 정확한 배치 및 상호작용 • **디지털 트윈:** 현실 도심이나 공장을 3D로 완벽히 복제하여 관제

(그림 5) 공간 지능의 개념과 특징



Dr. Fei-Fei Li

페이페이 리(Fei-Fei Li)는 미국 스탠포드 대학 컴퓨터공학과 교수

인공지능(AI) 분야, 특히 컴퓨터 비전과 공간 지능 분야를 선도하는 세계적인 석학이고, 스타트업 '월드 랩스(World Labs)' 공동 창업

From Words to Worlds: Spatial Intelligence is AI's Next Frontier

FEI-FEI LI
NOV 10, 2025

2,501 411

[AI는 지금] 'AI 대모' 페이페이 리, 월드 랩스로 7조원 몸값 노린다

| 블룸버그 "수억 달러 투자 논의 중"…기업가치 1년 만에 5배 급등

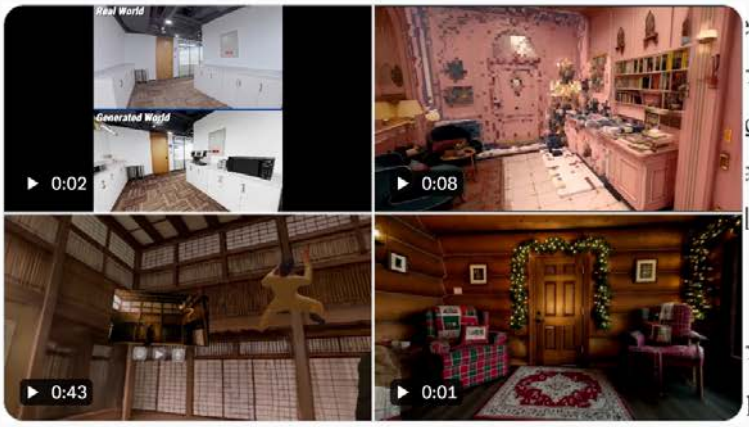
컴퓨팅 | 입력 :2026/01/23 11:25 수정: 2026/01/23 13:25

omated arithmetic and simple logic,
reverberates today: can machines think? It took
imagination to see what he saw: that intelligence might someday be built rather than born. That insight later launched a relentless scientific quest called Artificial Intelligence (AI). Twenty-five years into my own career in AI, I still find myself inspired by Turing's vision. But how close are we? The answer isn't simple.

Today, leading AI technology such as large language models (LLMs) have begun to transform how we access and work with abstract knowledge. Yet they remain wordsmiths in the dark; eloquent but inexperienced, knowledgeable but ungrounded. **Spatial intelligence will transform how we create and interact with real and virtual worlds—revolutionizing storytelling, creativity, robotics, scientific discovery, and beyond. This is AI's next frontier.**

공간 지능은 우리가 현실 세계와 가상 세계를 창조하고 상호작용하는 방식을 혁신적으로 변화시켜 스토리텔링, 창의성, 로봇 공학, 과학적 발견 등 모든 분야에 혁명을 일으킬 것이다. 이것이 바로 AI의 다음 개척지이다.

The pursuit of visual and spatial in
I entered the field. It's why I spent years building ImageNet, the first large-scale visual



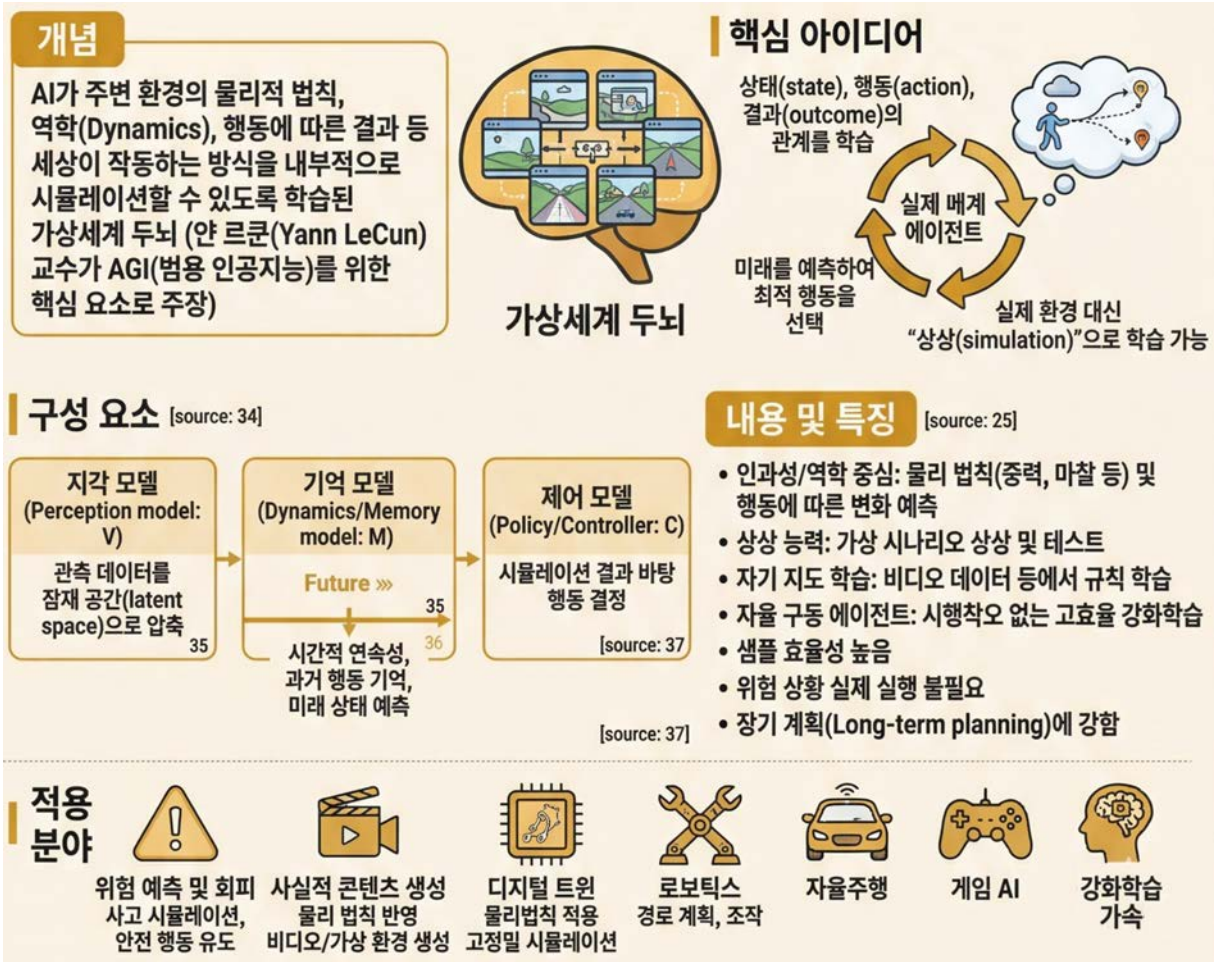
y elements enabling the birth
nd modern compute like
c lab at Stanford has spent the
arning. And it's why my
ldenhall, and I created World
in full, for the first time.

ny it matters, and how we're
pact that will reshape creativity, embodied intelligence, and human progress.

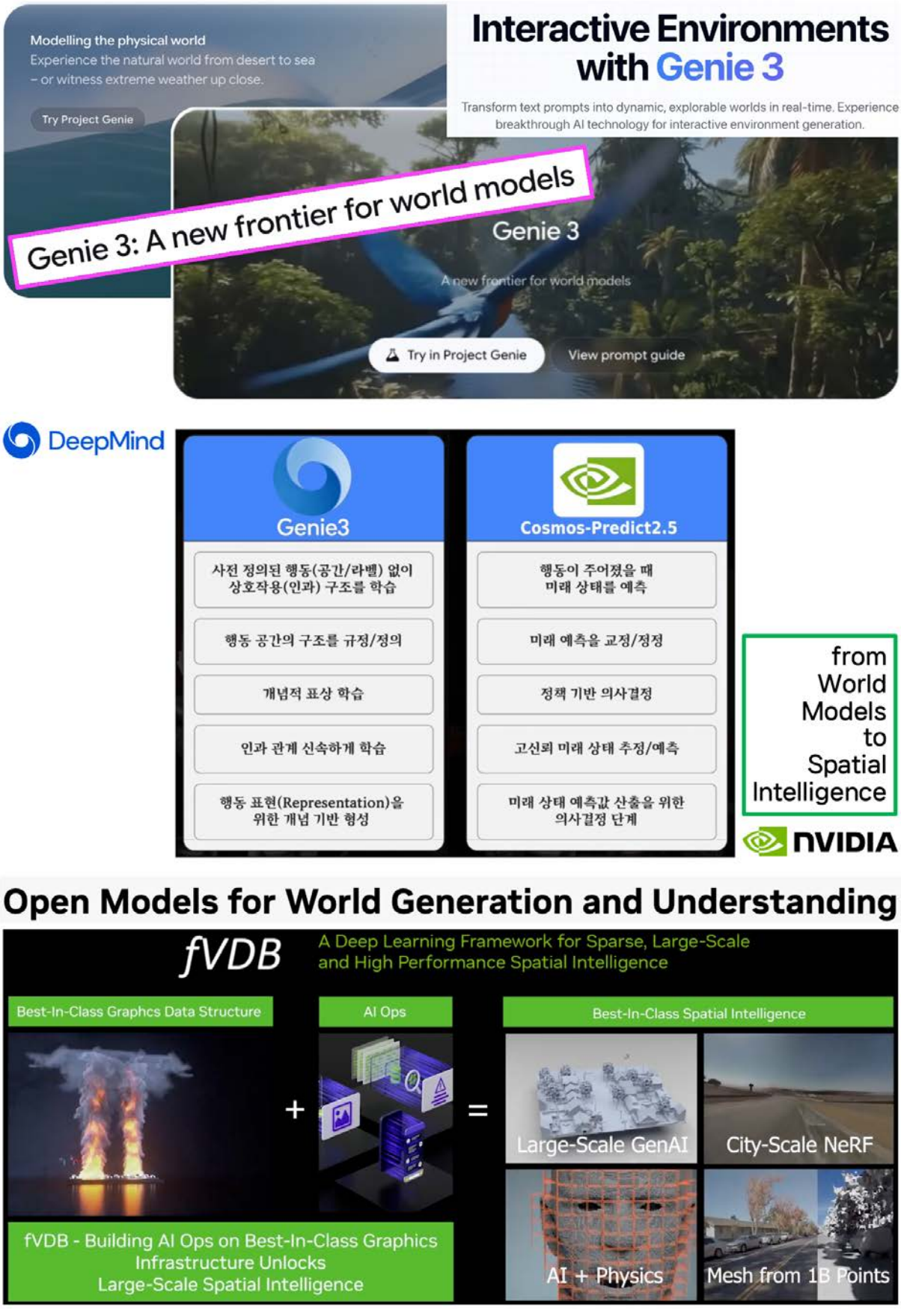
(그림 6) 공간 지능 : Fei Fei Li의 차세대 AI

나. 월드 모델의 개념과 특징

구분	월드 모델(World Models) 설명
개념	• AI가 주변 환경의 물리적 법칙, 역학(Dynamics), 행동에 따른 결과 등 세상이 작동하는 방식을 내부적으로 시뮬레이션할 수 있도록 학습된 가상세계 두뇌 (얀 르쿤(Yann LeCun) 교수가 AGI(범용 인공지능)를 위한 핵심 요소로 주장)
내용	• 지각 모델(V): 고차원 입력을 압축된 표현(잠재 공간)으로 변환 • 기억 모델(M): 시간적 연속성과 과거 행동을 기억하고 미래 상태 예측 • 제어 모델(C): 월드 모델의 시뮬레이션 결과를 바탕으로 행동 결정
특징	• 인과성/역학 중심: 물리 법칙(중력, 마찰 등)과 행동에 따른 세상의 변화를 예측 • 상상 능력: 실제 행동 전 가상으로 여러 시나리오를 '상상'하고 테스트 • 자기 지도 학습: 방대한 비디오 데이터 등에서 스스로 세상의 규칙을 학습
적용 분야	• 자율 구동 에이전트: 시행착오 없는 고효율 강화학습(가상 환경 내 학습) • 위험 예측 및 회피: 사고 시나리오를 가상으로 시뮬레이션하여 안전 행동 유도 • 사실적 콘텐츠 생성: 물리 법칙이 반영된 고품질 비디오 및 가상 환경 생성 • 디지털 트윈: 물리법칙을 적용한 고정밀 시뮬레이션으로 자율 시스템 통제



(그림 7) 월드 모델의 개념과 특징



(그림 8) 월드 모델 : 구글 Genie3 vs. 엔비디아 Cosmos

다. 공간 지능 및 월드 모델과 피지컬 AI와의 관계

1) 공간 지능과 월드 모델의 차이

○ AI 생태계의 여러 구성 요소 중에, 피지컬 AI(자율제조, 로봇 및 자율주행)의 핵심 요소로 급부상하고 있는 **월드 모델(World Models)과 공간 지능(Spatial Intelligence)은 '물리적 세계의 이해'라는 공통 목표**를 가지며, **"보완 관계(상호 의존적)"로서 둘은 완전히 별개도 아니고, 종속도 아님**

구분	공간 지능	월드 모델
본질	공간 기반 "이해 능력" **공간(Spatial intelligence)**	시간 기반 "예측 모델" **시간(Temporal intelligence)**
초점	공간 구조 이해	미래 상태 예측
데이터	3D/Geometry	상태 전이(Time series)
역할	인식 + 해석	시뮬레이션
포함 범위	주로 환경 표현	동적 + 정책

○ 전통적인 공간 인식(SLAM, Simultaneous Localization and Mapping) 기술은 기하학적의 정적인 3D 지도 작성에 치우쳤다면, **공간 지능은 의미와 맥락을 이해하고, 월드 모델은 그 지도 위에서 사물들이 시간에 따라 어떻게 움직이고 행동에 의해 어떻게 변하는지(예측)를 더해 'Spatially Intelligent World Model'**을 완성하는 수준으로 고도화

2) 공간 지능과 월드 모델의 관계 및 피지컬 AI 연계

○ 피지컬 AI란, 일반적으로 물리 세계에서 실제 행동하는 AI(로봇, 자율주행, 드론 등)를 의미하며, **공간 지능 측면에서 "세상이 어떻게 생겼는가?", 월드 모델 측면에서 "앞으로 어떻게 변할 것인가?", 행동 시스템(RL/Agent) 측면에서: "그래서 무엇을 할 것인가?"**를 적용

- 구조적 프로세스: [센서 입력] → **(공간 지능)** → [환경 이해] → **(월드 모델)** → [미래 예측] → **(RL/Agent)** → [행동]

○ **공간 지능과 월드 모델은** 별개가 아니며, **피지컬 AI를 구현하기 위해 서로 필수적으로 '보완'하고 '통합'되는 관계**로서 특히, 최근 기술 트렌드에서는 월드 모델이 공간 지능을 구현하는 핵심 아키텍처(엔진)로 자리 잡고 있음

- **피지컬 AI를 구현에서 공간 지능은 '목표/능력', 월드 모델은 '수단/엔진'**
- **공간 지능은** "로봇이 이 방의 구조를 이해하고 물건을 옮길 수 있게 하자"는 목표 능력
- **월드 모델은** 그 로봇의 두뇌 속에 "방의 3차원 지도"를 그리고, "물건을 들면 중력에 의해 아래로 힘을 받는다"는 법칙을 시뮬레이션해 주는 기술 엔진
- **즉, 피지컬 AI는 공간 지능의 '공간과 인지'에 월드 모델의 '시간과 역학'을 추가**

3) 공간 지능과 월드 모델의 피지컬 AI 적용

로봇과 공간 컴퓨팅이 3차원 공간을 이해하는 지능과 물리법칙 학습을 내재화하면 피지컬 AI가 제대로 작동될 수 있으며, 이에 AI 기반 물리법칙을 내재한 공간구조 솔루션인 **구글 딥마인드의 Genie는 LWM(물리법칙+공간구조 상호작용 학습)**, **엔비디아의 Omniverse 및 COSMOS 플랫폼은 WFM(물리 기반 가상 3D영상+시뮬레이션 학습)** 방식으로 접근하고 있음

❍ **로봇이 컵을 집는 상황(예시)**

- **공간 지능(Spatial Intelligence)**: 컵의 위치, 거리, 자세뿐 아니라 주변 환경과의 3차원 관계를 인식하고 상태를 추정
- **월드 모델(World Model)**: 로봇 팔을 움직였을 때 컵과의 접촉, 힘, 마찰 등 물리적 상호작용을 고려하여 결과 상태를 예측
- **피지컬 AI(Physical AI)**: 정책(Policy), 경로 계획(Planning), 제어(Control)를 통합하여 가장 안정적이고 성공 확률이 높은 grasp 전략을 선택 및 실행

❍ **자율주행 자동차 운행(예시)**

- **공간 지능(Spatial Intelligence)**: 차선, 차량, 보행자뿐 아니라 속도, 거리, 상대 위치 등 동적 환경을 포함한 3차원 상황을 인식 및 추적
- **월드 모델(World Model)**: 보행자, 차량 등의 미래 행동과 상호작용을 시간에 따라 경로 예측(Trajectory Prediction)
- **피지컬 AI(Physical AI)**: 인식과 예측 결과를 기반으로 경로 계획 및 의사결정을 수행하고 제어 시스템을 통해 감속, 회피, 차선 변경 등의 행동을 실행

❍ **피지컬 AI(로봇/자율주행)와의 연계: 두뇌와 신체의 조화**

- **공간 지능(두뇌의 인지 품질):** 센서(카메라, LiDAR 등)를 통해 획득한 정보를 바탕으로 객체의 위치, 형태, 자세뿐 아니라 객체 간 관계와 맥락을 포함한 3차원 환경을 이해하고 상태를 추정(물리적 사실적인 시뮬레이션을 할 수 있을 때 "공간 지능이 높다"고 평가)
- **월드 모델(두뇌의 시뮬레이터):** 로봇이 수집한 시청각 데이터를 바탕으로 내부적으로 가상의 3D 공간을 구축하고, 자신의 행동에 따른 세상의 변화를 물리적 상황을 고려하여 매우 빠른 속도로 시뮬레이션 수행(예: 팔을 이렇게 뻗으면 컵에 닿을 것이고, 특정 조건에서 컵은 넘어지지 않을 것이다")
- **피지컬 AI(신체/로봇):** 인지(공간 지능)와 예측(월드 모델)을 기반으로 경로 계획(Planning), 정책(Policy), 제어(Control)를 수행하고, 실제 환경에서 행동을 실행하며 지속적으로 데이터를 수집·학습하는 하드웨어-소프트웨어 통합 시스템

❍ **결론적으로, 피지컬 AI는 고도화된 '공간 지능'을 가진 '월드 모델'을 내장함으로써, 실제 환경에서 시행착오를 최소화하고 안전하며 영리하게 행동할 수 있음**

4 산업 분야의 최적 AI 시스템 구현 방향

가. 산업 분야의 AI 시스템 구성

AI 기반의 제조 등 산업 분야 시스템의 구성 요소로 **디지털 트윈(Digital Twin), 공간 지능(Spatial Intelligence), 월드 모델(World Models), 피지컬 AI(Physical AI)**를 들 수 있으며, 이들 영역의 조합을 통해 **스스로 물리적 환경을 이해하고 최적의 공정을 추론하여 제어하는 '인지 자율형 제조' 프레임워크**를 추구

글로벌 선도 기업들이 지향하는 **가장 이상적인 자율 제조(Autonomous Manufacturing)** 아키텍처는 **"디지털 트윈 + 공간 지능 + 월드 모델 + 피지컬 AI"을 모두 수용**하는 조합으로 이 4개 영역은 파편화된 솔루션이 아니라, 현장을 인지하고, 예측하며, 실제 행동으로 옮기는 **'완전 자율 폐쇄 루프(Autonomous Closed-loop)'**를 형성

1) 산업 분야에 필요한 AI 시스템

❍ **디지털 트윈 (Digital Twin) : 데이터 통합 및 가상 샌드박스**

- 현실의 공장 설비, 라인 상태, 물류 흐름을 실시간 센서 데이터로 연동하여 가상 공간에 똑같이 복제한 인프라를 조성
- AI가 현실 세계에 직접 개입하기 전에, 새로운 공정이나 로봇의 움직임을 수만 번 시뮬레이션할 수 있는 안전하고 통제된 환경을 제공

❍ **공간 지능 (Spatial Intelligence) : 3D 기하학적 지각**

- 디지털 트윈 내의 가상 공장과 실제 물리적 공장의 3차원 구조, 설비 간의 틈새, 부품의 형태 등을 정확히 인지하는 시각적 두뇌
- 단순한 2D 이미지 인식을 넘어, 사물의 깊이, 겹침, 형태적 특징을 파악하여 피지컬 AI가 조작해야 할 대상의 정확한 위치와 맥락을 모델링

❍ **월드 모델 (World Models) : 물리 법칙 기반의 예측 엔진**

- 공간 지능이 파악한 3D 상태에서, 특정 행동을 취했을 때 물리 법칙(중력, 마찰, 가속도 등)에 의해 어떤 결과가 나타날지 가상으로 예측
- 현실에서 발생할 수 있는 충돌이나 불량 발생 등의 시행착오를 가상 세계에서 미리 겪어보고 최적의 움직임 궤적과 제어 값을 도출하는 핵심 추론 엔진

❍ **피지컬 AI (Physical AI) : 현실 세계의 자율 실행**

- 월드 모델이 검증한 최적의 행동 지침을 바탕으로 실제 공장에서 움직이는 자율주행 물류 로봇(AMR), 다관절 로봇 팔, 정밀 제어 설비 등 시스템
- 이들이 현실에서 작업을 수행하며 수집한 새로운 센서와 공간 데이터는 다시 디지털 트윈으로 피드백되어 전체 시스템을 끊임없이 진화

2) 4대 영역과 프로세스 관계

❍ **디지털 트윈 → 데이터 수집 및 환경 동기화:** 현실의 센서, 설비 데이터가 모이는 백지 상태의 '도화지'이며 단순한 점(Point)과 선(Line), 혹은 수치 데이터의 집합

❍ **공간 지능 → 3D 형태 및 의미 인지:** 도화지 위의 데이터에 '의미론적 입체감'을 부여하여 "이 좌표의 데이터 뭉치는 항공우주 엔진 부품이고, 저 데이터는 조립용 로봇 팔이며, 두 객체 사이의 거리는 1.5m이다"라고 3D 기하학과 사물의 쓰임새(Affordance)를 해석

❍ **월드 모델 → 시공간 예측 및 계획:** 공간 지능이 파악한 3D 맥락 위에서 시간의 흐름(Dynamics)과 물리 법칙(중력, 조립 마찰력 등)을 더해 시뮬레이션을 수행하고, "로봇 팔이 부품을 잡으면 3초 뒤 어떤 물리적 저항이 생길까?"를 계산하여 최적의 제어 정책 수립

❍ **피지컬 AI → 물리적 제어 및 실행:** 앞선 3단계의 완벽한 계산을 바탕으로 현실 세계의 로봇이나 AMR/AGV가 오차 없이 작업을 수행하고, 그 결과를 다시 디지털 트윈으로 전송하여 폐쇄 루프(Closed-Loop)를 완성

3) 4대 영역의 역할 및 관계

❍ **디지털 트윈 (Digital Twin) ← "현실의 동기화된 복제본(가상 환경 인프라)"**

- **(역할)** 현실 세계의 객체, 시스템, 환경을 실시간 센서 데이터를 통해 가상 공간에 동일하게 구현한 3D 모델 및 데이터 아키텍처
- **(관계)** 디지털 트윈 자체는 데이터의 거울(Mirror)일 뿐, 스스로 생각하거나 예측하는 지능은 부족하며, 여기에 지능을 부여하기 위해 '공간 지능'과 '월드 모델'이 투입

❍ **공간 지능 (Spatial Intelligence) ← "3D 공간의 의미론적 이해(인지 능력)"**

- **(역할)** AI가 3D 공간의 기하학적 구조(Geometry)와 객체 간의 관계, 깊이(Depth), 그리고 사물의 쓰임새(Affordance)를 이해하는 능력
- **(관계)** 디지털 트윈이 만들어 놓은 3D 가상 환경을 단순 픽셀이나 좌표계가 아닌, "이것은 문이고, 저것은 장애물이며, 여기는 통과할 수 있는 공간이다"라고 AI가 논리적으로 인식

❍ **월드 모델 (World Model) ← "물리 법칙 기반의 미래 시뮬레이터(예측 및 계획 두뇌)"**

- **(역할)** 공간 지능을 통해 파악된 환경에서 에이전트가 특정 행동을 취했을 때 물리적으로 어떤 결과가 발생할지 인과관계와 동역학(Dynamics)을 시뮬레이션
- **(관계)** 디지털 트윈에서 공간 지능으로 상황을 파악한 뒤, 월드 모델을 가동하여 " 이 밸브를 열거나 이동하면 10초 뒤 환경이 어떻게 변할까?"를 가상으로 사전 테스트(Planning) 수행

❍ **피지컬 AI (Physical AI) ← "물리적 현실에서의 실행자(행동 및 피드백)"**

- **(역할)** 로봇, 자율주행차, 드론 등 물리적인 몸(Embodiment)을 가지고 현실 세계에서 직접 힘(Force)을 가하고 움직이는 주체
- **(관계)** 월드 모델이 도출한 최적의 행동 계획(Action Plan)을 현실에서 실행, 오차나 새로운 변수는 실시간으로 센싱되어 다시 디지털 트윈을 업데이트하는 무한 피드백 루프

나. 디지털 트윈의 고도화 - 일반적 DT vs 생성적 DT

1) 일반적 디지털 트윈과 생성적 디지털 트윈의 개요

일반적 디지털 트윈은 특정 지역의 실제 사진, 이미지, 동영상 등을 활용해 3D 모델링/렌더링 등 기법으로 현실 세계와 똑같은 미러 월드를 구축, 반면에 '공간 지능 + 월드 모델' 기반의 생성적 디지털 트윈은 '생성형 AI' 기법 등을 적용해 가상 세계 구축
제조 현장이나 스마트 시티에서 말하는 '일반적 디지털 트윈'과, 공간 지능 및 월드 모델이 결합된 최적 AI 시스템에서의 '생성적 디지털 트윈'은 그 **본질적인 정체성부터 기술적 근간까지 현격히 다른 세대(Generation)의 시스템**으로 구분

❍ AI 기술이 적용되면서 일반적 디지털 트윈이 현실을 비추는 '정적인 거울(Static Mirror)'이라면, **최적 AI 시스템의 생성적 디지털 트윈은 스스로 생각하고 미래를 생성해 내는 '인지 시뮬레이터(Cognitive Simulator)'로서 피지컬 AI와 연동 강화**

< 표 1 > 일반적 디지털 트윈 vs 생성적 디지털 트윈 비교

비교 구분	일반적 디지털 트윈 (3D 모델링 및 렌더링 기반)	생성적 디지털 트윈 (공간 지능 + 월드 모델 기반 AI)
개념 (Concept)	현실 세계의 시각적 및 물리적 '복제본(Replica)'으로 데이터 모니터링 위주의 가상 공간과 대시보드 운영	물리 법칙과 공간 맥락을 이해하고, 발생하지 않은 미래까지 생성하는 **인지 예측 공간(Cognitive Sandbox)** 구축
기술 (Technology)	3D 모델링/렌더링, CAD/BIM, 유니티/언리얼 엔진 렌더링, 사진 측량(Photogrammetry), 규칙 기반(Rule-based) IoT 연동	**생성형 AI**(NeRF, 3D Gaussian Splatting, Diffusion 등), **공간 지능**(VLM, Semantic SLAM), **월드 모델**(PINN, Transformer) 기술 적용
구축 (Construction)	전문가가 실제 사진/도면을 바탕으로 수동/반자동으로 3D 환경을 디자인하고 렌더링(정적, 노동 집약)	생성형 AI가 2D 사진/센서 데이터만으로 **실시간 3D 공간을 자동 생성 및 의미론적 라벨링(동적, 자가 진화)**
실제/가상 (Real/Virtual)	실제→가상(단방향 위주), 현실 변화를 가상세계 화면에 '지연 반영(Delayed Mirroring)'	실제↔가상(폐쇄 루프), 가상이 미래를 생성하여 **현실의 피지컬 AI를 통제(Sim-to-Real Control)'**
활용 (Utilization)	원격 관제, 센서 임계치 알람, 시각적 상태 확인, 규칙 기반의 제한적 시뮬레이션 및 관제	자율 행동 모델 훈련, 물리적 충돌 예측 및 회피, 극한/재난 시나리오 생성(Synthetic Data) 및 대응

❍ **개념(Concept): 복제 거울세계 vs 인지 가상세계**

- **(일반적 디지털 트윈)** 현실 세계에 존재하는 설비의 외형과 현재 상태(온도, 압력 등)를 화면상에 똑같이 보여주는 데 목적이 있으며 즉, '현재 어떻게 생겼고 어떤 상태인가'를 사람이 보기 편하게 만든 3D 관제용 인터페이스 중심의 시뮬레이션

- **(생성적 디지털 트윈)** 기계가 공간의 기하학적 의미와 인과관계를 이해하는 '뇌(Brain)' 역할을 하며, 겉 모습만 복제하는 것이 아니라, "이 객체는 어떤 물리적 속성을 가졌으며, 힘을 가했을 때 미래에 어떻게 변할 것인가"를 계산하고 생성해 내는 지능형 시뮬레이터

❍ **기술 (Technology): 시각 모델링 및 렌더링 엔진 vs 생성형 AI 및 물리 모델**

- **(일반적 디지털 트윈)** 게임 엔진(Unity, Unreal 등)이나 3D 모델링 소프트웨어를 통해 시각적인 '껍데기'를 렌더링하며, 물리적 움직임은 엔지니어가 사전에 입력한 스크립트(Script)나 수학적 규칙에 의존

- **(생성적 디지털 트윈)**
  - **생성형 3D AI(NeRF 등):** 복잡한 3D 모델링 작업 없이, 로봇이 찍은 몇 장의 2D 사진만으로 빛의 반사까지 계산된 초정밀 3D 공간을 인공지능이 즉각적으로 '생성(Generation)'
  - **공간 지능(VLM 등):** 생성된 공간 내의 객체들이 무엇인지(기둥인지, 부품인지, 사람인지) 의미를 스스로 파악
  - **월드 모델(PINN 등):** 외형뿐만 아니라 질량, 유체역학 등 눈에 보이지 않는 물리 법칙 신경망을 3D 객체에 부여

❍ **구축 (Construction): 수동적/정적 맵핑 vs 생성적/동적 진화**

- **(일반적 디지털 트윈)** 공장을 지을 때 도면(CAD)을 바탕으로 3D 에셋을 일일이 만들고, 센서 데이터를 API로 하나씩 연결해야 하며, 공장 구조가 바뀌면 개발자가 다시 수동으로 3D 모델을 수정해야 하는 **정적인 한계**

- **(생성적 디지털 트윈)** 자율 이동 로봇이나 드론이 현장을 돌아다니면, 생성형 AI가 실시간으로 3D 공간 지도를 업데이트(Semantic SLAM)하며, 새로운 장비가 들어오면 VLM이 스스로 장비의 종류와 기능을 파악하여 가상 환경의 맥락(Context)을 **동적으로 진화**

❍ **실제/가상 (Real/Virtual): 단방향 투영 vs Sim-to-Real 폐쇄 루프**

- **(일반적 디지털 트윈)** 현실의 센서값이 변하면 가상 세계의 그래프나 색상이 변하는데 현실이 원본이고 가상은 그림자라 볼 수 있으며, 가상의 화면을 보고 최종 판단을 내리고 현실의 시스템(기계/장치)을 조작할 때 사람이 관여

- **(생성적 디지털 트윈)** 가상과 현실이 동등한 주체로 상호작용하며, 가상 환경(월드 모델)에서 현실보다 빠른 속도로 미래를 생성하고 테스트한 뒤에, 그 결과값(최적의 제어 명령)을 역으로 현실의 피지컬 AI에게 내려보내는 즉, **가상의 연산이 현실의 물리적 행동을 지배하는 'Sim-to-Real'의 완벽한 폐쇄 루프(Closed-Loop)** 구조

❍ **활용 (Utilization): 사후 모니터링 및 관제 vs 극한 시나리오 생성 및 자율 제어**

- **(일반적 디지털 트윈)** 현실의 물리적 설비를 가상 공간에 3D로 똑같이 구현하여, 설비에 빨간불이 들어왔을 때 알람을 울려주거나, 장비의 과거 가동 기록을 3D로 다시 돌려보는 사후 분석 및 관제의 용도로 활용

- (생성적 디지털 트윈)
  - **희귀 데이터 생성(Synthetic Data Generation):** 생성형 AI 모델(Diffusion 등)을 통해 현실에서는 실험하기 위험한 상황(예: 공장 대형 화재, 주요 배관 연쇄 폭발, 기상 악화시 항만 크레인 붕괴 등)을 가상 환경에서 무한히 생성
  - **제로샷(Zero-shot) 자율 학습:** 피지컬 AI(강화학습 에이전트)는 이렇게 생성된 극한의 가상 시나리오 속에서 훈련을 거치며, 현실에서 단 한 번도 겪어보지 않은 위험 상황에서도 스스로 대처할 수 있는 자율 제어 능력을 확보

2) 일반적 디지털 트윈과 생성적 디지털 트윈의 비교

일반적 **디지털 트윈이 '가상 현실(VR)' 기술의 연장선에서 '현실을 얼마나 똑같이 묘사하고 잘 보여줄 것인가'에 집중했**다면, **생성적 디지털 트윈은 '이 복잡한 물리적 환경을 어떻게 수학적으로 이해하고, 미래를 통제하여 최적의 결정을 내릴 것인가'**에 집중하는 지능형 통제라 할 수 있으며, 고도의 전략과 물리적 정밀성이 요구되는 첨단 자율시스템에서 가치 극대화

❍ 실시간 모니터링 (Real-time Monitoring) 관점

- **(공통점)** 두 시스템 모두 현실 세계의 센서, IoT 기기, 카메라 등으로부터 데이터를 수집하여 가상 공간에서 실시간으로 시각화하고 현장 상태를 파악하는 '데이터 수집 창구' 역할을 수행

- (특 징)
  - **(일반적 디지털 트윈의 특징)** 수치와 시각적 표상(Representation)에 집중하며, "현재 크레인의 모터 온도는 85도이다", "선박의 현재 좌표는 X, Y이다"라는 1차원적 팩트를 대시보드와 3D 화면에 띄움
  - **(생성적 디지털 트윈의 특징)** 시각-언어 모델(VLM)과 공간 지능을 통해 모니터링되는 데이터의 의미론적 맥락(Semantic Context)을 파악하며, 단순히 온도가 85도인 것을 넘어, "이 온도는 인접한 유압 밸브의 한계치를 위협하는 위험 상태이며, 3D 공간상 작업자의 동선과 겹쳐 있다"고 상황 자체를 인지

- **(차이점)** 일반적 디지털 트윈이 데이터의 '단순 시각화(Human-readable)'에 머문다면, 생성적 디지털 트윈은 데이터의 '기계적 이해 및 인지(Machine-understandable)' 수행

❍ 통제 및 동기화 (Control & Synchronization) 관점

- **(공통점)** 물리적 현실 세계와 가상의 디지털 세계 간의 상태 불일치를 최소화하고, 특정 조건 발생시 기계나 설비를 제어하여 시스템의 안정성 유지

- (특 징)
  - **(일반적 디지털 트윈의 특징)** 단방향 동기화(물리 → 가상)가 중심으로 구축되며, 제어는 사전에 설정된 임계치(Rule-based)를 넘었을 때 경고 알람을 울리거나 기계를 정지시키는 수준, 복잡한 의사결정과 최종 통제권은 '인간 관리자(Human-in-the-loop)' 권한

• **(생성적 디지털 트윈의 특징)** 양방향 폐쇄 루프 동기화(물리 → 가상)가 구축되고 월드 모델이 도출한 최적의 해결책을 대형행동모델(LAM)을 통해 피지컬 AI(로봇, 무인 선박 등)의 제어 명령어로 즉각 번역하여 하달되어, 인간의 개입 없이 시스템이 스스로 통제하는 '자율 제어(Autonomous Control)' 가능

- **(차이점)** 일반적 디지털 트윈은 인간의 통제를 돕는 '수동적 보조 도구(Passive Tool)'이며, 생성적 디지털 트윈은 스스로 최적화 정책을 현실에 실행하는 '주도적 에이전트(Active Agent)' 역할

❍ **시뮬레이션 (Simulation) 관점**

- **(공통점)** 현실에서 직접 테스트하기 위험하거나 비용이 많이 드는 상황을 가상의 환경에서 미리 구동해 보아, 안전성과 경제성을 도모

- **(특 징)**

  • **(일반적 디지털 트윈의 특징)** 정해진 규칙과 기존의 물리 엔진(예: Unity, Unreal)에 의존하는 **결정론적(Deterministic) 시뮬레이션**으로, 엔지니어가 사전에 스크립트를 짜놓은 제한된 시나리오 내에서만 작동하며, 예상을 벗어난 변수(바람의 미세한 변화, 비정형 장애물 등)가 개입되면 시뮬레이션 한계

  • **(생성적 디지털 트윈의 특징)** 물리 정보 신경망(PINN)과 생성형 AI(Diffusion)를 결합한 **확률론적, 생성적(Generative) 시뮬레이션**이며, 과거에 한 번도 발생하지 않았던 극단적 재난 시나리오(예: 폭우 속 다중 충돌)를 무한히 생성해 내고, 그 안에서 강화학습(RL)을 통해 피지컬 AI가 스스로 생존 및 회피 궤적을 훈련(Sim-to-Real)하여 작동

- **(차이점)** 일반 디지털 트윈의 시뮬레이션이 정해진 길을 따라가는 '기차(Scripted)'라면, 생성적 디지털 트윈은 무한한 가능성의 공간을 탐험하고 스스로 길을 만드는 '탐험선(Generative & Exploratory)' 수준

❍ **예측 (Prediction) 관점**

- **(공통점)** 과거와 현재의 데이터를 분석하여 향후 발생할 수 있는 장비 고장, 공정 지연 등의 문제를 사전에 파악

- **(특 징)**

  • **(일반적 디지털 트윈의 특징)** 주로 통계적 기법이나 단순 머신러닝에 기반한 선형적 예측(Linear Prediction)을 수행하며, "A 부품의 진동 데이터 패턴을 보니 3일 뒤 고장 날 확률이 80%이다"라는 식의 개별 설비 단위의 예측 유지보수(PdM)에 집중

  • **(생성적 디지털 트윈의 특징)** 월드 모델(Transformer, SSM)을 활용하여 시스템 전체의 인과관계를 계산하는 동역학적, 복잡계 예측(Dynamic Complex Prediction)을 수행하며, "A 장비가 멈추면, B 장비의 대기 시간이 길어지고, 이는 전체 물류 흐름에 병목을 유발하여 2시간 뒤 최종 조립 라인이 중단된다"는 식의 연쇄적 시공간적 나비효과를 계산

- **(차이점)** 일반 디지털 트윈은 개별 객체의 '상태 타임라인(Statistical Trend)'을 예측하지만, 생성적 디지털 트윈은 거대한 3D 환경 안에서 모든 객체가 상호작용하는 **'물리적 인과관계의 흐름(Causal Dynamics)'** 전체를 통찰

3) 일반적 디지털 트윈과 생성적 디지털 트윈의 사례

전통적 방식의 **일반적 디지털 트윈**과 공간 지능 및 월드 모델이 결합된 **생성적 디지털 트윈은 '구축(Construction)' 과정에서 근본적인 패러다임의 차이**가 있으며, 이에 따라 전자는 스마트시티, 시설 제어 등 비교적 모니터링 중심의 관제 시뮬레이션에 적용되고, 후자는 물리 법칙 등이 요구되는 복잡한 공간이나 정교한 제조 공정 등의 자율 시스템에 적용

❍ 일반적 디지털 트윈이 시각적 복제를 위한 **'수동적 그래픽 렌더링'**에 가깝다면, AI가 내재된 생성적 디지털 트윈은 물리적 맥락과 상호작용을 스스로 이해하는 **'자동화된 인지 공간 생성'**에 초점을 두어 구축 방식, 공간 범위, 적용 기술에 현저한 차이

< 표 1 > 일반적 디지털 트윈 vs 생성적 디지털 트윈 3대 요소 비교

비교 구분	일반적 디지털 트윈 (정적/시각적 복제)	생성적 디지털 트윈 (동적/인지적 생성)
구축 방식	**수동/반자동 모델링** **(Manual/Rule-based)** • CAD 도면, 정적 사진, 레이저 스캐닝 기반 • 인간 엔지니어가 3D 에셋을 직접 디자인 및 배치	**실시간 동적 생성** **(Generative/Semantic)** • 자율 드론/로봇의 2D 영상, 라이다(LiDAR) 실시간 주입 • AI가 3D 공간을 자동 렌더링 및 의미론적 라벨링(VLM)
공간 범위	**단절된 스케일** **(Segmented Scale)** • 폴리곤(Polygon) 한계로 거시(도시)와 미시(부품) 통합 곤란 • GIS(도시)와 CAD(부품)가 별도 플랫폼에서 구동됨	**다중 스케일 연속 통합** **(Seamless Multi-scale)** • 잠재 공간(Latent Space) 압축을 통해 도시 인프라부터 기계 부품의 물리적 마찰까지 하나의 환경에서 연속 확대/축소 가능
적용 기술	**렌더링 엔진 & 관계형 DB** • Unity, Unreal Engine(시각 엔진) • RDBMS(단순 센서 로그 저장) • RESTful API(주기적 데이터 호출)	**생성형 공간 AI & 물리 신경망** • NeRF, 3D Gaussian Splatting (초정밀 공간 생성) • PINN(물리 엔진), Semantic SLAM(공간 지능) • Vector DB(맥락 저장), 5G 특화망/Edge AI(초저지연)

❍ **구축 방식 (Method): 정적 모델링 vs 동적 라벨링**

- **일반적 디지털 트윈의 방식과 사례**

  • **(방식)** 건설 도면(CAD/BIM)이나 고정된 사진 측량(Photogrammetry)을 바탕으로 3D 디자이너와 엔지니어가 수작업으로 가상 환경을 구축하며, 공장 라인이 변경되면 도면을 다시 그리고 3D 모델을 수동으로 업데이트

- **(사례: 항만 운영)** 항만 터미널 구축시 설계 도면을 바탕으로 크레인과 야드의 3D 모델을 엔진(Unreal, Unity)에서 배치하며, 실제 야드에 컨테이너가 무작위로 쌓이더라도, 화면은 센서 좌표에 따라 미리 만들어진 '컨테이너 그래픽 덩어리'를 이동시키는 것에 불과

- 생성적 디지털 트윈의 방식과 사례

- **(방식)** 영상(Video)과 라이다(LiDAR) 센서 데이터를 실시간으로 받아 생성형 AI(예: NeRF)가 3D 공간을 스스로 렌더링하며, VLM(시각-언어 모델)이 결합되어 구축과 동시에 "이것은 H빔", "저것은 용접 로봇"이라고 의미를 스스로 라벨링
- **(사례: 그린 스마트 해양도시 플랫폼)** 해양도시나 대규모 항만에 드론이나 무인 선박이 순찰을 돌면, 실시간으로 전송되는 영상 데이터를 바탕으로 파도의 높이, 방파제의 미세한 균열, 정박 중인 AI 선박의 위치가 동적으로 3D 렌더링되어 트윈 환경을 실시간으로 자가 구축(Self-constructing)

❍ 3D 공간 범위 (Spatial Scope): 단절 vs 연속

- 일반적 디지털 트윈의 범위와 사례

- **(범위)** 시각적 렌더링 부하(Polygon Count) 때문에 거시적 범위(도시, 도로)와 미시적 범위(기계 밸브, 부품 내부)를 하나의 시스템에 담지 못하며, 거시 공간은 GIS(지리정보 시스템)로, 미시 공간은 CAD 전용 뷰어로 분리하여 구축
- **(사례: 스마트 시티 및 시설 제어)** 도시 관제 센터 화면에서 도로망과 빌딩의 에너지 사용량은 볼 수 있지만, 특정 빌딩 지하에 있는 펌프의 내부 임펠러(회전날개) 마모 상태나 유체 흐름을 보려면 다른 시설물관리시스템(FMS) 프로그램을 작동시켜야 가능

- 생성적 디지털 트윈의 범위와 사례

- **(범위)** VAE(변분 자기부호화기)를 통한 데이터 압축과 Transformer(트랜스포머)의 주의 집중(Attention) 기술을 통해, 데이터 부하 없이 거시 환경부터 나노 단위의 미시 환경까지 하나의 연속된 차원에서 구축
- **(사례: AI 선박 및 기계 정비)** 항해 중인 거대한 선박의 위치와 해상 날씨(거시 스케일)를 관제하던 중, 선박 엔진 룸 내부의 특정 실린더 헤드(미시 스케일)로 화면을 줌인(Zoom-in)하며, 이때 월드 모델 내에 구축된 PINN(물리 정보 신경망)을 통해 실린더 내부의 열역학적 팽창 상태와 균열 진행 범위를 이음새 없이(Seamless) 동일한 플랫폼 내에서 실시간으로 확인하고 제어

❍ 적용 기술 (Technology): 시각 그래픽 엔진 vs 인지·물리 AI

- 일반적 디지털 트윈의 기술과 사례

- **(기술)** 핵심 기술은 눈에 보이기 위한 시각 엔진(Unity, Unreal)이며, 데이터베이스는 주로 센서의 시계열 로그를 저장하는 관계형 DB(RDBMS)를 사용하고, 현실 세계와의 통신은 초당 몇 회 수준의 RESTful API 호출에 의존
- **(사례: 제조 공장 관제)** 컨베이어 벨트 모터의 온도가 90도를 넘으면, API를 통해 데이터를 받은 엔진이 3D 화면의 모터 색상을 붉은색으로 바꾸고 경고창을 띄우는 규칙 기반(Rule-based) 아키텍처로 구축

- 생성적 디지털 트윈의 기술과 사례

  • **(기술)** 시각을 넘어선 공간 인지(Semantic SLAM)와 물리 법칙(PINN)이 핵심이며, 객체 간의 인과관계와 맥락을 저장하기 위해 벡터 DB(Vector DB)나 그래프 DB(Graph DB)를 구축하며, 피지컬 AI와의 즉각적인 피드백 루프를 위해 통신 지연이 0에 가까운 5G 특화망과 엣지 컴퓨팅(Edge AI) 환경이 필수적으로 요구

  • **(사례: 제조 공장 복잡계 제어)** 용접 로봇(피지컬 AI) 주변에 예기치 않은 장애물이 떨어지면, 엣지 AI 기반의 Semantic SLAM이 즉각 장애물의 3D 형상과 의미(철제 부품)를 인식하여 트윈에 업데이트하며, 동시에 월드 모델이 PINN을 가동해 회피 시 조인트에 가해지는 물리적 토크(Torque)를 밀리초(ms) 단위로 계산하고, 강화학습(RL) 정책에 따라 완전히 새로운 우회 궤적을 렌더링하여 로봇에게 실행 명령을 하달하는 복합 지능 아키텍처로 구축

다. 최적 AI 시스템의 대안 분석

일반적인 제조 산업 분야에 **효율성, 생산성, 경제성, 안전성을 위해 AI를 도입/적용**할 경우,
1) **(피지컬 AI + 디지털 트윈)** 피지컬 AI와 디지털 트윈만 있어도 되는지?
2) **(피지컬 AI + 공간 지능)** 피지컬 AI와 공간 지능만 있어도 되는지?
3) **(피지컬 AI + 월드 모델)** 피지컬 AI와 월드 모델만 있어도 되는지?
4) **(피지컬 AI + 디지털 트윈 + 공간 지능)** 피지컬 AI와 디지털 트윈, 공간 지능, 필요한지?
5) **(피지컬 AI + 디지털 트윈 + 월드 모델)** 피지컬 AI와 디지털 트윈, 월드 모델, 필요한지?
6) **(피지컬 AI + 디지털 트윈 + 공간지능 + 월드 모델)** 피지컬 AI와 디지털 트윈, 공간 지능, 월드 모델 4가지 모두 필요한지?

산업 분야의 **최적 AI 시스템은 "디지털 트윈(검증) + 공간지능(인지) + 월드 모델(예측) + 피지컬 AI(실행)"으로 4개 영역을 모두 결합하여 구축**하는 것이 가장 이상적이라고 평가

1) 피지컬 AI + 디지털 트윈 ⇒ 불충분

❍ **분석 결과**

- **(평가) 부분적 자동화 및 관제용으로는 가능하나, 지능적 최적화 불가**
- **(결론) "전통 스마트 팩토리" 수준에서는 충분히 가능(가장 현실적 조합)**
- **(이유)** 디지털 트윈은 공정 최적화, 병목 분석, 비용 절감, 반면에 피지컬 AI는 실제 실행
- **(적용)** 반도체, 자동차, 배터리처럼 **정형화된 공정**

❍ **장점 한계**

- **(장점)** ROI 명확(이미 산업에서 검증)하고 구현 난이도 상대적으로 낮으며, 생산성/경제성 개선에 매우 효과적
- **(한계)** 비정형 작업(랜덤 부품, 작업 변화)에 약하며, 환경변화 대응 느리고 자율성 제한

❍ **평가 근거**

- 디지털 트윈은 공장의 현재 상태 데이터를 모아주는 '거울'이며, 피지컬 AI는 지정된 동작을 수행하는 '몸'으로 과거와 현재의 데이터는 완벽히 알 수 있지만, 미래를 시뮬레이션하는 '뇌(WM)'와 3D 공간을 의미론적으로 이해하는 '눈(SI)'이 없음

❍ **영향 효과**

- **안전성(보통)** 센서 임계치 초과 시 기계를 멈추는 사후 대응적(Reactive) 안전망만 가동
- **생산성/효율성(낮음)** 설비 고장이나 병목 현상을 미리 예측하여 동선을 수정할 수 없으므로 다운타임(Downtime)이 발생

2) 피지컬 AI + 공간 지능 ⇒ 불충분

❍ **분석 결과**

- **(평가) 단일 로봇 셀(Cell) 단위의 작업에는 탁월하나, 공장 단위 제어 불가**
- **(결론) "유연 자동화"에는 가능하지만 전체 최적화는 부족**
- **(이유)** 공간 지능은 비정형 환경에서 물체 인식·조작 가능하지만, 디지털 트윈이 없어 공정 전체 최적화/예측에 부족
- **(적용)** 창고, 물류, 소량 다품종 생산

❍ **장점 한계**

- **(장점)** 물류, 피킹, 협동로봇에 강하고, 작업 유연성 매우 높음
- **(한계)** 공정 전체 효율 개선에 약하며, 병목, 비용 최적화 못함

❍ **평가 근거**

- 로봇이 3D 비전(SI)을 통해 비정형 부품의 형태를 완벽히 이해하고 조립할 수는 있지만, 공장 전체의 물류 흐름 데이터(DT)가 없고, 조립 후 물리적 변화를 예측(WM)하지 못함

❍ **영향 효과**

- **생산성(국소적 높음)** 해당 로봇의 작업 속도는 빠름
- **경제성/효율성(최하)** 로봇 시스템 혼자만 똑똑할 뿐, 전후 공정과의 조율이 불가능하여 공장 전체의 자원 할당이 낭비

3) 피지컬 AI + 월드 모델 ⇒ 부적절

❍ **분석 결과**

- **(평가) 기술적으로 성립 불가능하며 도입 시 시스템 붕괴 위험**
- **(결론) "연구/고급 자율화" 수준에서는 강력하지만 산업 적용은 아직 제한적**
- **(이유)** 월드 모델은 미래 상태 예측 + 강화학습 기반 의사결정이만 현실 공정은 제약조건 많음(안전, 품질, 규제)
- **(적용)** 고급 로봇 연구, 실험적 공정

❍ **장점 한계**

- **(장점)** 자율 최적화 가능성 매우 높으며, 복잡한 상황 대응 가능
- **(한계)** 데이터 요구량이 크고 안정성 검증 어려우며, 산업 적용 사례는 아직 제한적

❍ **평가 근거**

- 월드 모델은 예측을 위해 '현재의 완벽한 상태 데이터'를 필요로 하지만, 현실을 복제한 환경(DT)과 시각적 인지(SI) 없이 뇌(WM)만 가동하면, AI는 허구의 상황을 예측하는 환각(Hallucination)에 빠질 수 있음

❍ **영향 효과**

- **안전성/경제성(치명적 실패)** 잘못된 물리적 예측을 현실의 기계(피지컬 AI)에 명령하게 되어 막대한 장비 파손과 인명 사고를 유발할 수도 있음

4) 피지컬 AI + 디지털 트윈 + 공간 지능 ⇒ 고도화 한계

❍ **분석 결과**

- **(평가) 현실 인식과 실행력은 완벽하나, '미래'에 대한 계획 능력 부재**
- **(결론) "현재 산업에서 가장 이상적인 균형 조합"**
- **(이유)** 디지털 트윈은 전체 최적화, 공간 지능은 현장 유연성, 피지컬 AI는 실행
- **(적용)** 스마트팩토리 고도화, 인간-로봇 협업 환경

❍ **장점 한계**

- **(장점)** 효율성과 유연성을 동시 확보할 수 있으며, 현실 제조업의 요구사항을 대부분 충족하고 안전성 검증도 가능(시뮬레이션 기반)
- **(한계)** 구축 비용 증가하고 시스템 통합 복잡

❍ **평가 근거**

- 공장 전체의 데이터(DT)를 알고, 3D 공간의 맥락(SI)도 이해하여 복잡한 로봇 제어가 가능하지만, 시간의 흐름에 따른 물리 법칙(동역학, 열역학 등)을 계산하는 월드 모델이 없음

❍ **영향 효과**

- **생산성/안전성(우수)** 현재 주어진 상황에서의 대처는 훌륭함
- **경제성(불완전)** 새로운 공정을 설계하거나 변경할 때, 가상 공간에서 미리 시뮬레이션(Sim-to-Real)해 볼 수 없어, 결국 현실 제조 라인에서 직접 테스트하며 막대한 '물리적 시행착오 비용'을 지불

5) 피지컬 AI + 디지털 트윈 + 월드 모델 ⇒ 구조적 한계

❍ **분석 결과**

- **(평가) 거시적 물류 최적화는 가능하나, 미시적 물리 조작 불가**
- **(결론) "차세대 자동화 구조(고급 최적화)"**
- **(이유)** 디지털 트윈은 검증된 시뮬레이션, 월드 모델은 학습 기반 의사결정을 하지만 3차원 공간 인식의 지능이 부족
- **(적용)** 고난도 공정 최적화(예: 화학, 에너지)

❍ **장점 한계**

- **(장점)** 공정 자동 최적화 가능하며, 예측 기반으로 유지보수 및 생산 계획 가능
- **(한계)** 모델 불확실성과 설명 가능성 부족(Black-box 문제)

❍ **평가 근거**

- 전체 공장의 병목을 예측(WM)하고 데이터를 통합(DT)하는 데는 완벽하지만 3D 비정형 객체를 이해하는 공간 지능(SI)이 누락되어 있음

❍ **영향 효과**

- **효율성(우수)** AMR/AGV 배차 등 2D 평면 데이터 기반의 최적화는 완벽
- **안전성/생산성(저하)** 부품이 컨베이어 벨트에서 이탈하거나, 로봇 팔이 3D 공간에서 복잡한 각도로 부품을 끼워 맞춰야 하는 등 '입체적이고 의미론적인 물리 제어'가 필요한 순간 오작동 가능

6) 피지컬 AI + DT + SI + WM (4가지 통합) ⇒ 최적 완성형

❍ **분석 결과**

- **(평가) 완전 자율 제조(Autonomous Manufacturing)를 위한 객관적 표준 아키텍처**
- **(결론) "궁극적 최적 구조(과잉 설계 부담)"**
- **(이유)** 디지털 트윈은 시뮬레이션, 공간 지능은 인식, 월드 모델은 예측/학습, 피지컬 AI는 실행하여 최적의 역할 분담
- **(적용)** 제조공정 모든 분야에 적용 가능

❍ **장점 한계**

- **(장점)** 완전 자율공장 구현하며, 모든 상황 대응 가능
- **(한계)** 비용이 매우 크고 시스템 복잡도 폭증하며, 유지보수 곤란 및 실제 ROI 불확실

❍ **평가 근거**

- 인지-예측-실행의 사이버-물리 시스템(CPS) 밸류체인이 완결되며, **DT가 데이터를 모으고, SI가 3D 기하학을 해석하며, WM이 물리 법칙 기반으로 미래를 수만 번 시뮬레이션한 뒤, 피지컬 AI가 오차 없이 실행 가능**

❍ **영향 효과**

- **효율성(우수)** 모든 장비의 대기 시간을 수학적 한계까지 줄임
- **생산성(우수)** 장비 고장이나 병목을 사전 예측하여 가동 중단(Downtime)을 0으로 수렴
- **경제성(우수)** 현실의 물리적 시운전 및 불량품 폐기 비용을 가상의 완벽한 시뮬레이션으로 대체하여 매몰 비용을 차단
- **안전성(우수)** 3D 공간 내 작업자의 돌발 행동까지 예측하여 충돌 궤적을 사전에 차단

라. 자율 시스템의 핵심 AI 기술 적용

산업 분야의 AI 시스템에서 **디지털 트윈**은 물리적 현실의 데이터를 기반으로 가상 환경을 구축하며, AI는 **공간 지능**을 발휘해 그 가상 환경 내의 구조적, 의미론적 맥락을 파악하고, 이를 바탕으로 **월드 모델**이 물리 법칙에 기반한 수만 번의 동역학 시뮬레이션을 돌려 최적의 미래 행동 궤적을 계획하고, 이 계획된 궤적은 현실 세계의 **피지컬 AI**로 전송되어 정밀하고 안전한 물리적 제어로 발현되어 작동

이와 같은 최적 AI 시스템을 구현하기 위해서는 **핵심 AI 기술**의 투입이 요구되며, 기존에 도출된 AI 기술과 4대 영역을 매트릭스로 분석한 결과 **"VLM, VAE, RNN/Transformer/SSM, PINN, RL, LAM, Diffusion/GenAI, XAI" 등이 공통적으로 필요**

1) VLM (Vision Language Model, 시각-언어 모델)

❍ **필요 이유**

- 기계가 수집한 3D 시각 데이터를 인간이 이해하는 '의미(Semantics)'와 '목적'으로 연결하는 통역사가 필요하기 때문

❍ **기술 속성**

- 이미지/영상(Vision)과 자연어(Text)를 동시에 처리하여, 시각적 맥락을 언어로 추론하는 멀티모달(Multimodal) 및 제로샷(Zero-shot) 추론 능력

❍ **핵심 기능**

- 3D 공간 내의 객체에 의미론적 라벨을 부여하고(예: "잡을 수 있는 밸브"), 인간 작업자의 자연어 지시("엔진 옆 붉은 부품을 조립해")를 기계의 시각적 목표로 매핑

❍ **영향 분석**

- **효율성:** 복잡한 코딩 없이 자연어로 전체 공정을 통제할 수 있어 직관적인 인간-기계 상호작용(HMI)을 구현
- **안전성:** 시각적으로 위험한 상황(예: "작업자가 크레인 반경에 진입함")을 언어적 맥락으로 파악하여 즉각적인 경고를 보냄

2) VAE (Variational Autoencoder, 변분 자기부호화기)

❍ **필요 이유**

- 제조 현장의 라이다(LiDAR) 및 고해상도 카메라 데이터는 너무 방대하여, 이를 그대로 시뮬레이션에 넣으면 실시간(Real-time) 연산이 불가능하기 때문

❍ **기술 속성**

- 고차원의 원시 데이터를 정보 손실을 최소화하면서 저차원의 잠재 공간(Latent Space)으로 압축하는 차원 축소 및 특징 추출 알고리즘

❍ **핵심 기능**

- 디지털 트윈에 모인 무거운 시각/공간 데이터를 가벼운 수치 벡터로 압축하여 월드 모델의 예측 엔진에 입력값으로 제공

❍ **영향 분석**

- **효율성:** 월드 모델의 시뮬레이션 연산 속도를 기하급수적으로 끌어올려 밀리초(ms) 단위의 실시간 반응을 가능
- **경제성:** 방대한 데이터 처리에 필요한 GPU 컴퓨팅 서버 및 전력 유지 비용을 대폭 절감

3) RNN / Transformer / SSM (시계열 및 상태 공간 예측 모델)

❍ **필요 이유**

- 제조 공정은 정지된 화면이 아니라 시간의 흐름에 따라 연속적으로 변하는 동적(Dynamic) 시스템이므로, 미래의 연속적 변화를 추론해야 하기 때문

❍ **기술 속성**

- 과거부터 현재까지의 데이터 시퀀스에 주의 집중(Attention) 메커니즘을 적용하여, 장기적인(Long-horizon) 시간적 인과관계를 계산하는 동역학 추론 모델

❍ **핵심 기능**

- 현재의 환경 상태와 특정 행동을 입력받아, 시간 경과에 따른 물리적 상태 변화와 장비들의 이동 궤적(Trajectory)을 예측

❍ **영향 분석**

- **생산성:** 컨베이어 벨트나 물류 라인에서 얼마 뒤에 발생할 병목 현상을 미리 계산하여 사전에 경로를 수정함으로써 가동 중단(Downtime)을 방지
- **안전성:** AMR/AGV와 로봇 간의 미래 이동 궤적을 겹쳐보고 충돌 가능성을 예측하여 완벽하게 회피

4) PINN (Physics-Informed Neural Network, 물리 정보 신경망)

❍ **필요 이유**

- AI가 단순히 통계적 확률로만 미래를 예측하면 '질량 없는 물체'나 '에너지 보존 법칙 위배'와 같은 치명적인 물리적 환각(Hallucination)이 발생하기 때문

❍ **기술 속성**

- AI의 학습 손실 함수(Loss Function)에 뉴턴 역학, 유체 역학(나비에-스토크스 방정식) 등의 수학적 미분 방정식을 강제로 주입한 하이브리드 신경망

❍ **핵심 기능**

- 월드 모델이 가상 공간에서 돌리는 수만 번의 시뮬레이션 결과가 현실 세계의 절대적인 물리 법칙을 위배하지 않도록 통제

❍ **영향 분석**

- **경제성:** 현실의 중력, 마찰력을 가상에 완벽히 구현하여 현실에서의 물리적 테스트(시행착오 및 파손) 비용을 최소화(Sim-to-Real 격차 해소)
- **안전성:** 항공우주 부품 조립 등 극도의 정밀도가 요구되는 환경에서, 물리적으로 불가능하거나 위험한 제어 명령이 내려지는 것을 원천 차단

5) RL (Reinforcement Learning, 강화학습)

❍ **필요 이유**

- 예측된 미래 환경 속에서, 수많은 선택지 중 목표를 달성하기 위한 '가장 완벽한 정답(행동)'을 기계가 스스로 찾아내야 하기 때문

❍ **기술 속성**

- 보상(Reward)과 페널티(Penalty)를 기반으로 에이전트가 가상 환경과 상호작용하며 최적의 제어 정책(Policy)을 학습하는 마르코프 결정 과정(MDP) 알고리즘

❍ **핵심 기능**

- 월드 모델이 만든 가상 시뮬레이터 안에서 수만 번의 충돌과 실패를 경험하며, 로봇 관절의 최적 토크 값이나 조립 궤적을 찾아냄

❍ **영향 분석**

- **효율성:** 인간 엔지니어가 수학적으로 일일이 계산할 수 없는 복잡한 다관절 로봇의 최적 동선을 AI가 스스로 도출
- **생산성:** 가장 짧은 시간 내, 최소한의 에너지를 사용하여 조립을 완료하는 극한의 최적화 정책을 공정에 이식

6) LAM (Large Action Model, 대형 행동 모델)

❍ **필요 이유**

- 강화학습으로 찾은 정책은 추상적이므로, 이를 다수의 기계 장비들이 알아들을 수 있는 순차적이고 구체적인 '행동 명령'으로 번역해야 하기 때문

❍ **기술 속성**

- 고차원의 작업 목표를 인식하고, 이를 달성하기 위해 필요한 하위 API 호출 및 기계 제어 시퀀스를 생성하는 거시적 오케스트레이션(Orchestration) 모델

❍ **핵심 기능**

- 시스템으로부터 "엔진을 조립하라"는 지시를 받아, "1단계 크레인 이동 → 2단계 부품 파지 → 3단계 AMR/AGV 이송 → 4단계 로봇 팔 조립" 등 다단계(Step-by-step)의 구체적인 물리 제어 신호로 분해하여 하달

❍ **영향 분석**

- **생산성:** 여러 대의 피지컬 AI가 톱니바퀴처럼 맞물려 작업하도록 거시적으로 조율하여 시스템 전체의 단위 시간당 산출량을 극대화
- **효율성:** 작업 간의 대기 시간(Idle time)을 최소화하고, 공정 간의 매끄러운 연결을 보장

7) Diffusion / GenAI (생성형 AI)

❍ **필요 이유**

- 월드 모델이 예측한 복잡한 수학적 결과값을 인간 관리자가 시각적으로 확인하고, 기계학습을 위한 '가상의 희귀 데이터'를 무한히 확보해야 하기 때문

❍ **기술 속성**

- 잠재 공간의 노이즈를 점진적으로 제거(Denoising)하여 고해상도의 3D 객체, 이미지, 비디오 시퀀스를 합성해 내는 확률적 생성 모델

❍ **핵심 기능**

- 예측된 미래 물리 상태를 3D 비디오로 렌더링하고, 극단적 불량품이나 재난 상황(화재, 폭우) 등 현실에서 구하기 힘든 합성 데이터(Synthetic Data)를 생성

❍ **영향 분석**

- **경제성:** 희귀한 불량품 데이터를 수집하기 위해 공정을 고의로 망가뜨릴 필요 없이, AI가 생성한 데이터로 피지컬 AI를 학습시켜 막대한 데이터 수집 비용을 절감
- **안전성:** 인간 관리자가 AI가 계획한 위험 시나리오를 고화질 3D 시뮬레이션 화면으로 직관적으로 사전 검토(Review) 가능

8) XAI (Explainable AI, 설명 가능한 AI)

❍ **필요 이유**

- 제조, 항만, 에너지 등 치명적인 산업 인프라에서는 AI가 아무리 똑똑해도 그 결정을 '인간이 논리적으로 납득'할 수 없으면 현장에 도입할 수 없기 때문

❍ **기술 속성**

- 블랙박스 신경망 내부의 가중치와 주의집중 영역(Attention Map)을 추적하여, 의사결정의 인과적 근거를 시각화 및 정량화하는 분석 알고리즘

❍ **핵심 기능**

- 공정과정에서 "로봇이 왜 갑자기 작동을 멈췄는가?", "왜 이 부품을 불량으로 예측했는가?"에 대한 정확한 물리적, 공간적, 수치적 근거(Log)를 제공

❍ **영향 분석**

- **안전성:** 공정 내 사고 발생 시 AI의 판단 근거를 투명하게 감사(Auditing)할 수 있어, 치명적인 오류가 반복되는 것을 완벽히 통제
- **효율성:** 시스템 오작동 시 블랙박스 모델을 맹목적으로 뜯어고칠 필요 없이, XAI가 지목한 근거 변수(예: 센서 노이즈)만 빠르게 디버깅(Debugging) 가능

< 표 3 > 핵심 AI 기술별 디지털 트윈, 공간 지능, World Model, 피지컬 AI 간의 상호 연계

핵심 AI 기술 (속성 및 기능)	디지털 트윈 (데이터, 시뮬레이션)	공간 지능 (3D 구조 의미 파악)	월드 모델 (학습 및 미래 예측)	피지컬 AI (물리적 제어 및 실행)
VLM (시각-언어 모델) 단순 픽셀 데이터를 공간적 맥락으로 변환, 논리적 인과관계 연결	트윈에 수집된 시각/센서 데이터의 "A구역 장비, 장치", "장애물" 등의 **의미론적 텍스트 라벨(Labelling) 부여**	3D 공간 내 객체의 쓰임새를 이해하여, **"저 손잡이는 잡고 돌릴 수 있다"는 기하학적·논리적 맥락 해석**	시각적 미래 시나리오를 논리적 규칙과 결합하여, **물리적 변화에 대한 인과적 추론을 언어적 맥락으로 변환**	작업자의 자연어 명령("엔진 옆 붉은색 볼트를 조립해")을 **제로샷(Zero-shot)으로 이해하고 시각 정보와 매핑하여 실행**
VAE (변형 오토인코더) 고차원 데이터의 저차원 잠재 공간(Latent Space) 압축	방대한 실시간 센서 및 3D 원시 데이터의 연산 효율성을 위해 **저차원 잠재 공간으로 압축 및 동기화**	복잡한 3D 포인트 클라우드에서 불필요한 노이즈를 제거하고 **공간의 핵심 기하학적 특징만 추출**	무거운 원시 데이터가 아닌, 압축된 상태 변수를 사용하여 **동역학 시뮬레이션의 연산 속도 극대화**	로봇의 비전 센서로 들어오는 방대한 실시간 프레임을 즉각적으로 압축하여 **인지 반응속도 최적화**
RNN/ Transformer/SSM 시계열 데이터 처리 및 장기(Long-horizon) 인과관계 추론	가상 환경 내 객체와 센서의 과거 상태변화 이력을 모델링하여 추적 가능한 **동적 인과관계 및 맥락 파악하는 추론 엔진화**	시야에서 일시적으로 가려진 객체(Occlusion)의 이동경로를 기억하고 **시공간적(Spatio- temporal) 맥락 추적**	현재 상태와 행동을 기반으로 시간의 흐름에 따른 장기(Long-horizon) 인과관계를 계산하는 **핵심 예측 엔진**	예측된 미래 상태를 바탕으로 로봇 관절(Actuator) 등에 끊김 없이 **매끄럽고 연속적인 물리적 제어 신호(Trajectory) 제공**
PINN (물리 정보 신경망) AI 손실 함수(Loss Function)에 질량, 유체, 역학 등 물리법칙 구현	질량, 유체, 열역학 등 현실의 물리적 수치가 완벽히 동기화된 **오차 없는 정밀 가상 환경(도화지) 유지**	공간 지능이 파악한 3D 형태를 바탕으로 무게 중심, 표면 마찰력 등 **보이지 않는 물리적 속성 추론**	AI의 비논리적 환각을 차단하여, 시뮬레이션의 모든 프레임이 **실제 물리 법칙(에너지 보존 등)을 엄격하게 따르도록 강제**	가상에서 최적화된 제어 정책을 현실의 질량과 중력을 가진 기계에 적용할 때 발생하는 **Sim-to-Real 격차 원천 차단**
RL (강화학습) 보상 체계 기반의 최적화 및 제어 정책(Policy) 학습	현실 세계를 복제한 트윈 환경을 에이전트가 탐색할 수 있는 **거대한 상태 공간(State Space)으로 활용**	복잡한 3D 장애물 사이를 충돌 없이 통과하거나 정확한 파지점(Grasping point)을 찾았을 때 **보상(Reward) 부여**	수만 번의 극단적 충돌과 실패를 안전하게 테스트하며, 주어진 목표를 달성하는 **가장 효율적인 정책(Policy) 훈련**	검증 완료된 최적 행동정책을 이식받아, 불확실한 현실의 노이즈 속에서도 **강건한 (Robust) 자율 제어 수행**
LAM (대형 행동 모델) 고차원적 목표를 물리적 기기의 단계별 제어 명령으로 적용	공장이나 항만 전체의 설비 가동 상태와 자원 흐름을 거시적으로 모니터링하여 **작업 할당 조율**	공간 지능의 3D 맥락(장애물 위치, 목적지 거리 등)을 바탕으로 목표 달성을 위한 **다단계 이동 및 조작 경로(Path) 기획**	월드 모델의 예측 결과를 바탕으로 "어떤 순서로 개입해야 가장 안전한가"를 계산하여 **최적 행동 시퀀스 생성**	생성된 거시적 행동 시퀀스를 로봇의 작업 단위로 번역 및 오케스트레이션 기능, **직접적인 물리 기계 언어로 번역 및 하달**
Diffusion/GenAI (생성형 AI) 잠재 공간의 데이터를 고해상도 이미지, 3D 렌더링, 영상으로 복원	센서 사각지대나 해상도가 낮은 구역의 데이터를 보완하여 **고정밀 3D 가상 환경 복원(예: 3D Gaussian Splatting)**	2D 카메라 이미지 몇 장만으로 객체의 뒷면이나 내부의 3D 입체 구조를 **초정밀 해상도로 렌더링**	수학적으로 존재하는 미래의 물리적 상태 예측값을 인간 관리자가 직관적으로 볼 수 있는 **가상 비디오로 생성**	피지컬 AI의 비전 모델 훈련을 위해, 폭우, 조명 고장 등 현실에서 수집하기 힘든 **희귀 합성 데이터 무한 생성**
XAI (설명 가능한 AI) 블랙박스 모델의 예측 및 의사결정 과정에 대한 인과적 근거 제공	가상 환경 데이터 베이스 내에서 병목이나 통신 지연이 발생하고 있는 **구조적 원인 노드(Node) 시각화**	시스템이 왜 특정 3D 객체를 장애물이나 조립 대상으로 인식했는지 등 **시각적 근거(Attention Map) 제시**	월드 모델이 특정한 충돌이나 파손 미래를 예측하게 된 결정적인 **물리적 파라미터 (마찰계수 저하 등) 추적 및 설명**	로봇이 왜 작업을 중단했거나, 설계된 궤도를 이탈하여 회피 기동을 했는지에 대한 **논리적 의사결정 로그 제공**

마. 기능별 핵심 AI 기술의 매핑

최적 AI 시스템의 **핵심 AI 기술(VLM, VAE, RNN/Transformer/SSM, PINN, RL, LAM, Diffusion/GenAI, XAI)**은 목표(효율성, 생산성, 경제성, 안정성)를 달성하기 위해 "**공간 및 상황 인지, 물리적 사실주의 및 가상 시뮬레이션, 자율 제어 및 실행, 시스템 통제 및 거버넌스**"의 기능 측면에서 요소 기술로 적용

1) 공간 및 상황 인지 → 공간 지능 및 디지털 트윈의 눈

❍ **VLM (Vision-Language Model):** 현장의 카메라와 센서가 수집한 시각 정보를 언어로 변환하여, 관리자가 "현재 3번 라인의 조립 상태가 어때?"라고 물었을 때 AI가 상황을 정확히 이해하고 브리핑 제공

❍ **VAE (Variational Autoencoders):** 정상적인 공정 데이터의 '표준 분포(잠재 공간)'를 학습하여, 미세한 진동이나 소음 변화만으로도 설비의 이상 징후(Anomaly)를 초기에 탐지하는 예지 보전에 탁월(안정성 확보)

2) 물리적 사실주의 및 가상 시뮬레이션 → 월드 모델 및 디지털 트윈의 뇌

❍ **PINN (Physics-Informed Neural Networks):** 제조 현장의 핵심으로 AI 시뮬레이션에 열역학, 유체역학 등의 물리 법칙을 강제하여, 가상 테스트 결과가 실제 현장과 100% 일치하도록 보장(경제성 및 효율성 극대화)

❍ **RNN / Transformer / SSM:** 시계열 데이터(온도, 압력, 로봇 관절의 궤적 등)의 흐름과 문맥을 분석하여, 다음 순간에 벌어질 물리적 상태 변화를 예측

❍ **Diffusion / GenAI:** 실제 현장에서 구하기 힘든 '희귀 불량 데이터'나 '극단적인 상황'을 합성(생성)하여 월드 모델을 훈련시키는 데 사용되며 데이터 수집 비용을 혁신적으로 낮춤

3) 자율 제어 및 실행 → 피지컬 AI의 행동과 손발

❍ **RL (Reinforcement Learning):** 가상의 월드 모델 안에서 피지컬 로봇이 수만 번의 시행착오를 거쳐 가장 빠르고 안전한 조립 궤적이나 주행 경로를 스스로 학습(생산성 기여)

❍ **LAM (Large Action Model):** 물리적인 로봇 제어를 넘어, 제조실행시스템(MES)이나 ERP 시스템의 UI를 AI가 직접 조작하여 발주, 작업 지시서 발행 등의 디지털 워크플로우를 자율적으로 처리

4) 시스템 통제 및 거버넌스 → 전사적 안정성

❍ **XAI (Explainable AI):** AI가 특정 부품을 불량으로 판정하거나 라인 가동을 중단시켰을 때, 그 판단 근거(예: 히트맵 시각화 등)를 엔지니어/관리자에게 명확히 설명하여 현장의 수용성과 신뢰도를 확보

바. 단계별 최적 AI 시스템의 구현 방향

제조 현장에서 필요한 **AI 기반 시스템(플랫폼/솔루션)**은 “디지털 트윈 ⇄ 공간 지능 ⇄ 월드 모델 ⇄ 피지컬 AI”가 계층적으로 연결된 루프 구조로 상호 연계되어야 효율성을 극대화

- **디지털 트윈** : 현실 세계 동기화한 외부에서 **정밀하게 시뮬레이션·검증·최적화하는 플랫폼**
- **공간 지능** : 디지털 트윈내 3D 지형, 설비, 부품의 **형태적 맥락을 인지 해석하는 솔루션**
- **월드 모델** : 공간 지능이 파악한 내부에서 물리법칙의 **미래를 예측하고 의사결정하는 시스템**
- **피지컬 AI** : 공간 지능, 월드 모델과 연계한 현실에서 **행동하고 데이터를 만드는 실행 시스템**

1) 1단계: 디지털 트윈 (Digital Twin) → 데이터 수집 및 가상화

❍ **단계 개요:** 현실 세계의 제조 현장을 디지털 공간에 오차 없이 실시간으로 복제하는 진입점으로 데이터를 수집하고 가상화하는 단계

❍ **핵심 역할:** 공장 내의 모든 물리적 자산과 운영 상태를 데이터화하여 가상 공간에 매핑

❍ **하드웨어 및 인프라:** LiDAR 등 IoT 센서를 통해 온도, 압력, 진동 등의 수치 데이터를 수집하고, 카메라를 통해 현장의 영상 데이터를 실시간으로 확보

❍ **적용 AI 기술**

- **VAE (변형 오토인코더):** 센서 및 비전 카메라로 들어오는 방대한 고차원 원시 데이터를 연산 효율성을 위해 저차원 잠재 공간으로 압축하며, 이를 통해 지연 없는 실시간 가상 공장 구축과 초기 이상 탐지(예지 보전)를 수행

2) 2단계: 공간 지능 (Spatial Intelligence) → 지각 및 맥락 이해

❍ **단계 개요:** 디지털 트윈으로 구축된 3D 가상 환경 내에서 사물 간의 기하학적, 의미론적 관계를 AI가 인지하는 단계

❍ **핵심 역할:** 단순한 좌표 데이터가 아니라, "저 장비 옆에 이동 중인 물류 로봇이 있고, 그 앞에는 조립 대기 중인 부품이 있다"는 식의 입체적 맥락을 파악

❍ **적용 AI 기술**

- **VLM (시각-언어 모델):** 카메라 등 시각 데이터를 기반으로 현장의 상황을 언어적 의미로 해석(예: "현재 A 라인 3번 컨베이어 벨트에 장애물이 있음")
- **VAE 연계:** 1단계에서 압축된 데이터로부터 공간의 핵심적인 기하학적 특징만 빠르고 정확하게 추출하여 복잡한 환경을 모델링

3) 3단계: 월드 모델 (World Models) → 예측 및 시뮬레이션

❍ **단계 개요:** 공간 지능이 파악한 현재 상태를 바탕으로, 미래에 일어날 상황을 예측하고 행동의 결과를 가상으로 테스트하는 '두뇌' 역할의 단계

❍ **핵심 역할:** 현실 세계의 물리 법칙을 가상 세계에 동일하게 적용하여, 로봇이 실제로 움직이기 전에 수많은 시나리오를 미리 검증

❍ **적용 AI 기술**

- **PINN (물리 정보 신경망):** 시뮬레이션에 중력, 마찰력, 유체 역학 등 실제 물리 법칙을 강제하여, 가상 테스트 결과가 현실과 완벽히 일치하도록(Sim-to-Real 격차 해소) 구현
- **RNN/Transformer:** 과거부터 현재까지의 시계열 데이터를 분석하여 시간 흐름에 따른 기계의 마모도나 다음 공정의 병목 현상 등 미래 상태를 논리적으로 추론
- **Diffusion/GenAI (생성형 AI):** 현실에서 수집하기 힘든 극단적인 고장 상황이나 희귀 불량 시나리오를 가상으로 생성하여 시스템이 미리 대비하고 학습할 수 있도록 도움

4) 4단계: 피지컬 AI (Physical AI) → 실행 및 제어

❍ **단계 개요:** 월드 모델에서 검증이 끝난 최적의 행동 시나리오를 현실 세계의 기계 장치들을 통해 직접 실행하는 최종 단계

❍ **핵심 역할:** 디지털 세계의 지능을 물리적 세계의 역학적 움직임으로 변환하여 실제 제조 공정을 수행

❍ **적용 AI 기술**

- **RL (강화학습):** 월드 모델의 가상 환경에서 수만 번의 사전 훈련을 통해 도출된 보상 체계를 바탕으로, AMR/AGV, 다관절 로봇 팔 등의 궤적을 가장 빠르고 안전하게 제어
- **LAM (대형 행동 모델):** "A구역의 부품을 B조립 라인으로 이송하라"와 같은 관리자의 거시적 자연어 명령이나 시스템의 목표를, 로봇이나 생산관리시스템(MES/ERP)이 이해할 수 있는 다단계 세부 작업 명령으로 변환하여 하달

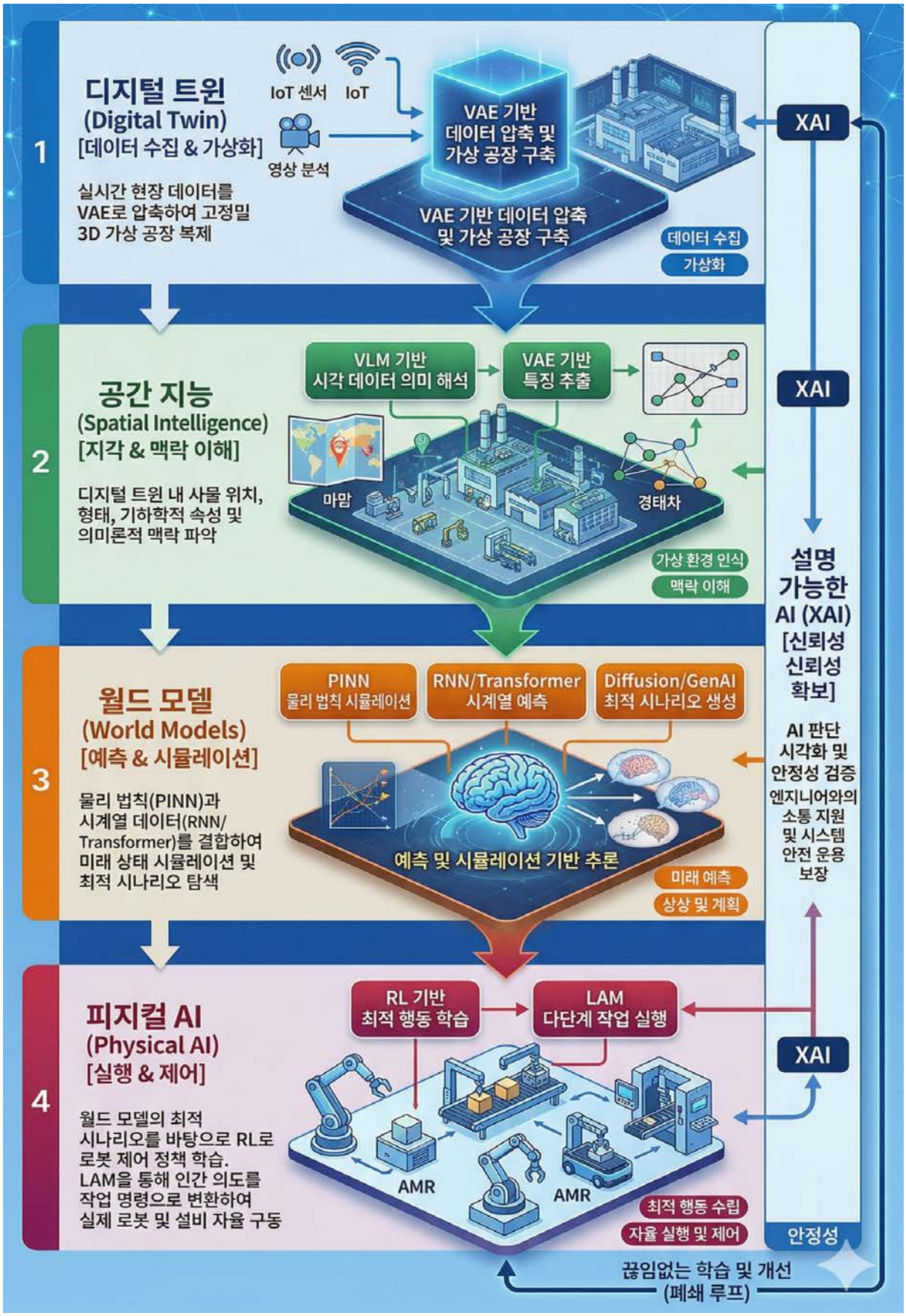
5) 통합 신뢰 기반 및 폐쇄 루프

❍ **단계 개요:** 4단계에서 피지컬 AI가 실제 현장에서 작동하며 발생시킨 새로운 결과와 변수들은 다시 1단계의 디지털 트윈(IoT/카메라)으로 실시간 수집

❍ **핵심 역할:** 끊임없는 학습 및 개선(폐쇄 루프)의 순환 고리를 통해 시스템은 멈추지 않고 진화하며 제조 공정의 자율 최적화를 완성

❍ **적용 AI 기술**

- **XAI(설명 가능한 AI):** 1~4단계 전체를 아우르는 수직 기둥으로, AI가 특정 설비의 가동을 멈추거나 로봇의 경로를 변경했을 때, 그 판단 근거(시각적 Attention Map, 인과관계 로그 등)를 엔지니어에게 투명하게 제공하며, 이는 규제 대응과 현장 작업자의 시스템 수용성을 위해 신뢰성 측면에서 필수적 사항



(그림 9) 최적 AI 시스템의 영역별/단계별 연계

5 AI 기반 자율 시스템의 산업 적용

가. 전기차 모터 부품의 자율 제조(예시)

1) 개발 및 설계 (Development & Design): 가상 엔지니어링과 물리적 검증

❍ **기능 구현**

- 물리적인 시제품을 만들기 전, 디지털 공간에서 최적의 설계를 도출하고 검증
- 개발자가 목표 성능(예: 출력 200kW, 발열 최소화)을 입력하면, AI가 수만 가지 설계안을 생성하고 가상 테스트를 마친 뒤 최적안을 제시

❍ **적용 기술 및 시스템**

- **디지털 트윈 & Diffusion/GenAI:** 과거의 모터 설계 도면과 데이터를 학습한 GenAI가 새로운 형태의 모터 하우징 및 코일 권선 패턴 시안 등을 무수히 생성하여 **디지털 트윈** 공간에 제공
- **월드 모델 & PINN:** 생성된 설계안이 실제 주행 환경에서 어떻게 작동할지 테스트하는 것으로 이때 PINN(물리 정보 신경망)이 적용되어, 가상 공간 내에서 전자기장 분포, 회전 시의 유체 저항, 열역학적 발열 등 실제 물리 법칙이 엄격하게 적용된 시뮬레이션을 수행하며, 이를 통해 막대한 비용이 드는 물리적 풍동 실험과 파괴 테스트를 생략 가능

2) 부품 가공 및 제조 (Manufacturing & Processing): 무중단 예지 보전

❍ **기능 구현**

- 설계도가 현장으로 넘어오면 원자재를 깎고 가공하는 공정이 시작되며, 장비의 고장을 사전에 막고 수율을 극대화하는 것이 핵심
- CNC(Computer Numerical Control) 밀링 머신이 모터 케이스를 가공하는 동안, 미세한 진동 변화를 AI가 감지하여 "예를 들어, 2시간 뒤 절삭 공구 파손 예상, 즉시 교체 요망"이라는 알림을 띄우고 라인 속도를 스스로 조절

❍ **적용 기술 및 시스템**

- **IoT 센서 & 카메라:** 가공 장비 내부에 부착된 고정밀 IoT 센서가 분당 수만 건의 온도/진동 데이터를 수집하고, 카메라는 칩(쇳가루)의 배출 형태를 실시간으로 촬영
- **디지털 트윈 & VAE:** 방대한 센서 원시 데이터를 VAE가 저차원 잠재 공간으로 압축하여 통신 지연을 없애고, 동시에 정상 가동 상태의 데이터 분포를 학습하여, 아주 미세한 이상 파형(Anomaly)을 즉각 탐지
- **RNN/Transformer:** 압축된 시계열 데이터를 분석하여, 설비의 마모도 진행 상황을 장기적 관점(Long-horizon)에서 추론하고 고장 시점을 정확히 예측

3) 조립 (Assembly): 고정밀 자율 협동 제어

❍ **기능 구현**

- 가공된 모터 코어, 스테이터(고정자), 로터(회전자) 등을 결합하는 가장 복잡한 단계
- 다관절 로봇 팔 여러 대가 부품이 담긴 박스의 위치가 매번 조금씩 바뀌어도 스스로 부품을 찾아 정확한 각도로 집어 들고, 부딪힘 없이 모터를 조립

❍ **적용 기술 및 시스템**

- **공간 지능 & VLM:** 조립 셀에 설치된 다중 카메라의 시각 정보를 **VLM**이 해석하며, 즉 "현재 스테이터가 3도 기울어져 있음", "로봇 팔의 궤적에 도구가 놓여 있음" 등 3D 공간의 기하학적 형태와 의미론적 맥락을 파악
- **월드 모델 & RL(강화학습):** 조립 전, 월드 모델에서 로봇이 부품을 끼워 넣을 때 발생하는 마찰력과 저항을 계산하며, **RL**을 통해 가상 환경에서 수만 번의 조립 연습을 거친 로봇은 부품이 꽉 끼거나 저항이 발생했을 때 힘을 조절하는 최적의 제어 정책(Policy)을 확보
- **피지컬 AI:** 학습이 완료된 로봇 팔이 현장에서 0.01mm의 오차도 없이 모터를 조립

4) 품질 검수 (Quality Inspection): 설명 가능한 자율 검수

❍ **기능 구현**

- 완성된 모터의 결함을 찾아내고, 결함의 원인을 추적
- 컨베이어 벨트를 지나는 모터를 비전 시스템이 스캔한 뒤, 불량품을 솎아내고 엔지니어의 모니터에 결함 위치와 불량 판정 이유를 정확히 시각화하여 보고

❍ **적용 기술 및 시스템**

- **공간 지능 & 카메라:** 초고해상도 비전 카메라와 3D 스캐너가 모터 외관의 스크래치, 볼트 체결 불량 등을 스캔
- **XAI (설명 가능한 AI):** AI가 특정 모터를 불량으로 판정했을 때, 블랙박스처럼 결과만 내놓는 것이 아니라, **XAI**가 히트맵을 생성하여 "3번 하우징 볼트의 체결 깊이가 기준치보다 0.2mm 부족하여 불량으로 판정함"과 같이 명확한 인과적 근거를 엔지니어에게 제공

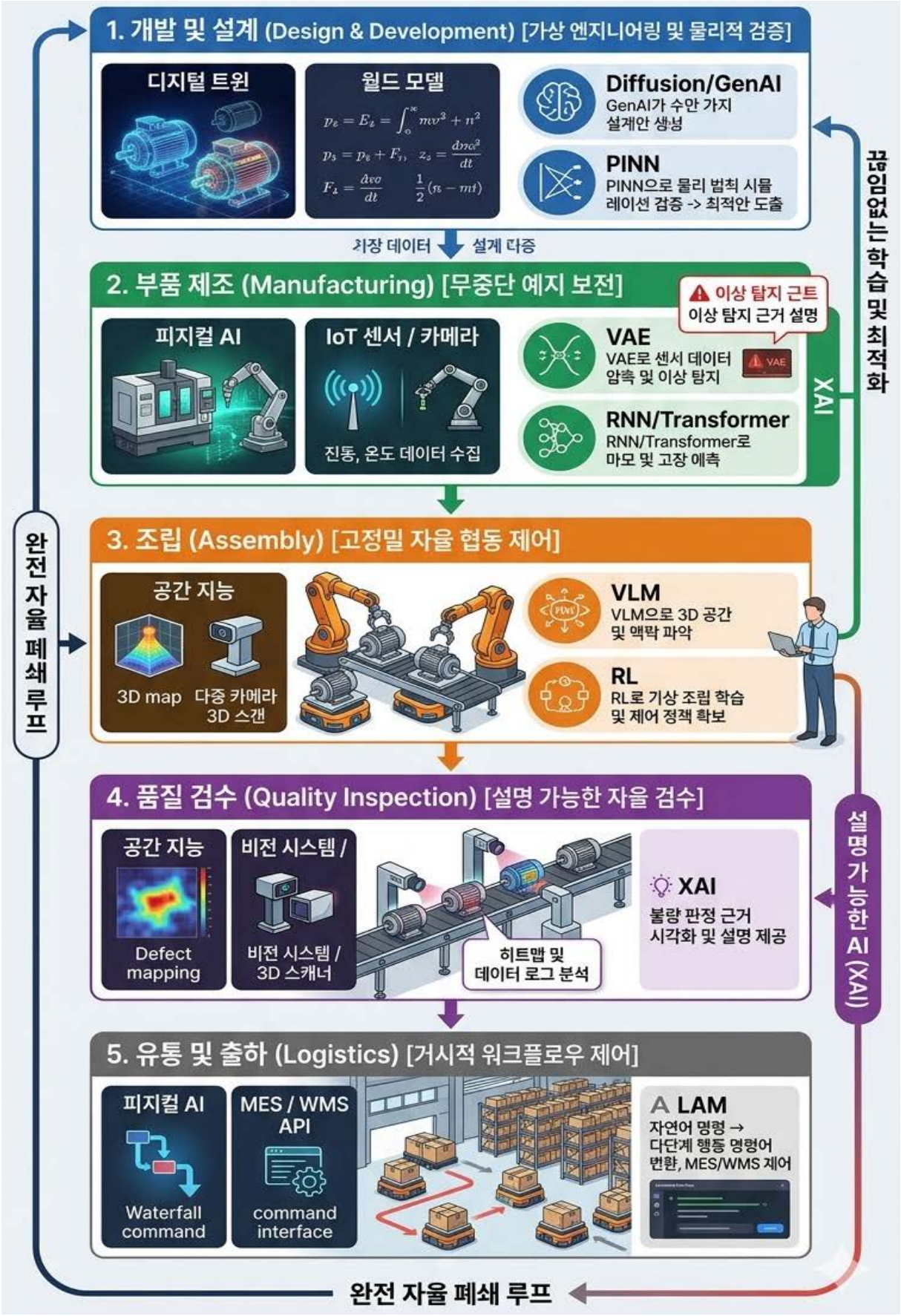
5) 유통 및 출하 (Distribution & Logistics): 거시적 워크플로우 제어

❍ **기능 구현**

- 합격 판정을 받은 모터들을 포장하여 물류 창고로 이동시키고 출하 일정을 맞춤
- 직원이 "오늘 완성된 EV 모터 A형 500대를 2번 게이트로 출하해"라고 음성이나 텍스트로 지시하면, 공장 내 물류 로봇들이 스스로 대형을 짜서 제품을 이송

❍ **적용 기술 및 시스템**

- **피지컬 AI & LAM:** 사용자의 자연어 명령을 **LAM**이 인식하여 공장의 MES 및 WMS의 API를 직접 제어하며, "A구역 재고 확인 → AMR 10대 배차 → 2번 게이트로 이동 경로 할당"이라는 다단계 물리 기계 명령어로 변환하여 물류 로봇들에게 작업 하달 및 실행



(그림 10) AI 기반 시스템: 전기차 부품의 적용(예시)

나. 항만 컨테이너터미널의 자율 항만(예시)

1) 게이트(Gate) 구역: 무정차 반출입 및 항만 혼잡도 최소화

❍ **기능 구현**

- 외부 트럭이 터미널에 진입하여 화물을 인계하는 첫 관문으로, 병목 현상 해소와 화물 상태의 즉각적인 인지가 목표

❍ **연계 장비 및 인프라**

- 외부 트럭, 컨테이너, 게이트 아치(다중 채널 카메라, 3D LiDAR, IoT 센서) 등

❍ **적용 기술 및 운영**

- **디지털 트윈 기반 실시간 수집:** 트럭이 무정차로 게이트를 통과할 때, 아치형 센서 프레임이 컨테이너의 6면 외관과 트럭 번호를 초고속으로 스캔하여 가상 공간에 데이터를 복제
- **VAE (데이터 압축 및 예지):** 여러 대의 카메라에서 쏟아지는 고용량 영상 데이터를 **VAE**가 지연 없이 저차원 잠재 공간으로 압축하여 AI Port Simulator로 전송하며, 동시에 정상 컨테이너의 표준 분포와 다른 미세한 변형(위험물 누출, 구조적 파손)을 0.1초 이내에 탐지하여 사고를 예방
- **VLM (시각 데이터의 언어화):** 스캔된 영상에서 ISO Code(컨테이너 번호), 씰(Seal) 파손 여부를 판독하고, "우측 측면 20cm 스크래치 발생"과 같이 상태를 즉각적인 텍스트 묘사로 변환하여 시스템에 기록

2) 야드(Yard) 구역: 지능형 보관, 리핸들링 최소화 및 배치 최적화

❍ **기능 구현**

- 수만 개의 컨테이너가 적재되는 공간으로, 불필요한 장비 움직임을 줄이고 최적의 반출입 순서를 꿰뚫어 보는 것이 핵심

❍ **연계 장비 및 인프라**

- ARMGC(자동화 야드 크레인), AGV(무인이송장비), 리치스태커, 스마트 센서 등

❍ **적용 기술 및 운영**

- **공간 지능 (3D 지각):** 야드의 수많은 블록(Block), 베이(Bay), 티어(Tier) 구조와 이동 중인 AGV, 리치스태커의 위치를 3차원 입체 공간으로 모델링하며, 사각지대 없이 장비 간의 기하학적 맥락과 거리를 실시간으로 파악
- **RNN/Transformer (시계열 예측):** 선박 입출항 스케줄과 과거 반출입 이력 데이터를 시계열로 분석하며, "이 컨테이너는 내일 오후 B 선박에 실릴 확률이 92%이므로 하단이 아닌 최상단에 적부해야 한다"는 판단을 내려, 컨테이너 화물을 몇 번씩 옮기는 **리핸들링 (Re-handling) 최소화**
- **RL (강화학습) 기반 피지컬 AI:** 가상 야드 환경에서 수만 번 훈련된 **RL 정책(Policy)**을

바탕으로, 수십 대의 AGV가 교차로에서 충돌 없이 최단 거리로 이동하도록 라우팅하며, 또한 ARMGC 트롤리의 최적 이동 궤적을 제어하여 작업 시간을 단축

- **XAI (설명 가능한 판단):** 시스템이 야드 공간 효율을 위해 특정 컨테이너를 C블록으로 임시 이동(Remarshalling)시키는 결정을 내렸을 때, 관제사나 리치스태커 작업자에게 "익일 새벽 양하 물량과 동선이 겹쳐 선제적으로 병목을 해소하기 위함"이라는 인과적 근거를 대시보드에 명확히 제공하여 현장의 운영 신뢰성 제고

3) 선석(Berth) 구역: 하역 한계 극복 및 크레인 생산성(Gross Crane Rate) 극대화

❍ **기능 구현**

- 초대형 선박이 접안하여 화물을 싣고 내리는 구역으로, 해풍과 파도 등 통제 불가한 자연 변수 속에서 거대 장비를 안전하고 신속하게 운영

❍ **연계 장비 및 인프라**

- STS(안벽 크레인), 스프레더 센서, 선박 등

❍ **적용 기술 및 운영**

- **월드 모델 & PINN (물리 기반 정밀 제어):** 수십 톤의 화물을 매단 STS 크레인은 바닷바람에 의해 심한 스웨이(Sway, 흔들림) 현상을 겪는데, **PINN**은 진자 운동, 와이어 장력, 유체 역학(풍압) 등 실제 물리 법칙을 AI 시뮬레이션에 강제 적용하며, **월드 모델**은 이를 바탕으로 스프레더가 흔들릴 궤적을 0.1초 앞서 예측하고, 흔들림을 상쇄하는 역방향 모터 토크 값을 STS 제어기(피지컬 AI)에 하달하여 **크레인 운영 생산성을 극대화**
- **Diffusion/GenAI (극단적 시나리오 학습):** 초속 30m 이상의 돌풍, 폭우 등 현실에서 데이터를 수집하기 힘든 극한의 기상 악화 시나리오나 스프레더 결함 상황을 **AI**가 가상으로 합성(생성)하며, 시스템은 이 희귀 데이터를 미리 학습하여, 미증유의 사태에서도 안전한 하역 제어 정책을 유지

4) 관제 및 오케스트레이션: 거시적 행동 제어 및 전사적 경제성 확보

❍ **기능 구현**

- 게이트-야드-선석을 잇는 전체 항만의 물류 흐름을 조율하는 두뇌 역할을 담당

❍ **연계 장비 및 인프라**

- TOS, AI Port Simulator 등

❍ **적용 기술 및 운영**

- **LAM (대형 행동 모델):** 항만 운영 관제사가 시스템에 "A 선박 양하 작업을 개시하고, 냉동 화물은 리퍼(Reefer) 야드로 최우선 이송해"라고 일상 언어로 지시하며, LAM은 이 고차원적 목표를 해석하여, "STS 3기 작업 할당 → 하부 AGV 15대 배차 → 목적지 리퍼 야드 ARMGC 대기 지시"라는 수백 개의 세부 장비 API 제어 명령으로 자율 변환을 수행함으로써 모든 피지컬 AI가 하나의 유기체처럼 움직이게 하여 대기 시간과 에너지 낭비를 없애고 전체 운영의 **경제성 달성**

(그림 11) AI 기반 시스템: 항만 물류의 적용(예시)

[참고 문헌]

1) 공간 지능(Spatial Intelligence) 및 월드 모델(World Models)

- Bruce, J., Dennis, M., Edwards, A., Parker-Holder, J., Shi, Y., et al., "Genie: Generative Interactive Environments", *International Conference on Machine Learning (ICML)*, arXiv:2402.15391, 2024/02/23, pp.1-27.
- Ken Museth, Robert Jensen, Jonathan Swartz, Francis Williams and Jiahui Huang, "Building Spatial Intelligence from Real-World 3D Data Using Deep-Learning Framework fVDB", *NVIDIA Technical Blog,* Jul 29, 2024.
- Fei-Fei Li, From Words to Worlds: Spatial Intelligence is AI's Next Frontier, *World Labs Research & Insights*, 2025.11.10.
- Graison Thomas, "The World Model Inflection: 2025 Made It Real", *Medium,* Dec 15, 2025.
- NVIDIA Cosmos Team, "Cosmos World Foundation Model Platform for Physical AI", *arXiv (NVIDIA Research)*, arXiv:2501.03575, 2025/07/09, pp.1-75.
- NVIDIA, "NVIDIA and Global Robotics Leaders Take Physical AI to the Real World", *NVIDIA Press Release*, March 16, 2026.
- Prachi Mishra, "Cosmos Predict 2.5 & Transfer 2.5: Evolving the World Foundation Models for Physical AI", *Huggingface,* October 28, 2025.
- Yongji Wang, Mehdi Bennani, et al., "Discovery of Unstable Singularities", *Cornell University*, arXiv:2509.14185, September 17. 2025.
- Yue Liao, Pengfei Zhou, Siyuan Huang, Donglin Yang, Shengcong Chen, Yuxin Jiang, Yue Hu, Jingbin Cai, Si Liu, Jianlan Luo, Liliang Chen, Shuicheng Yan, Maoqing Yao, Guanghui Ren, "Genie Envisioner: A Unified World Foundation Platform for Robotic Manipulation", arXiv:2508.05635, 2025/11/04, pp.1-26.
- Yuncong Yang, Jiageng Liu, Zheyuan Zhang, Siyuan Zhou, Reuben Tan, Jianwei Yang, Yilun Du, Chuang Gan, "MindJourney: Test-Time Scaling with World Models for Spatial Reasoning", arXiv:2507.12508, 2025/11/01, pp.1-22.

2) 비전-언어-행동 모델 (LAM & VLA)

- Ding, Z., Tong, Y., Yue, Y., Zhang, J., et al., "VLAS: Vision-Language-Action Model with Speech Instructions for Customized Robot Manipulation", *ICLR Proceedings*, 2025/02/21, pp.1-18.
- Ghosh, D., Walke, H., Pertsch, K., Finn, C., Levine, S., "Octo: An Open-Source Generalist Robot Policy", *Robotics: Science and Systems(RSS)*, Vol.20, No.1, 2024/07/15, pp.1-13.

- Hoang, T., Huang, K., Kokane, S., Zhang, J., Liu, Z., "LAM SIMULATOR: Advancing Data Generation for Large Action Model Training via Online Exploration and Trajectory Feedback", *Findings of the Association for Computational Linguistics (ACL)*, ACL 2025, No.1, 2025/08/01, pp.12921-12934.
- Jiang, Y., Yong, S., Huang, S., et al., "An Embodied Generalist Agent in 3D World", *International Conference on Machine Learning(ICML)*, 2024/05/09, pp.1-39.
- Moo Jin Kim, Karl Pertsch, Siddharth Karamcheti, Ted Xiao, et.al, "OpenVLA: An Open-Source Vision-Language-Action Model", arXiv:2406.09246, 2024/09/05, pp.1-37.

3) 물리 정보 신경망(PINN) 및 AI 기반 제어

- Barry-Straume, J., Verulkar, A. D., Sarshar, A., Popov, A. A., Sandu, A., "Ensemble based Closed-Loop Optimal Control using Physics-Informed Neural Networks" Virginia Tech Computational Science Laboratory, arXiv:2510.18195, 2025/10/22, pp.1-35.
- Farea, A., Yli-Harja, O., Emmert-Streib, F., "Understanding Physics-Informed Neural Networks: Techniques, Applications, Trends, and Challenges", *AI (MDPI)*, Vol.5, No.3, 2024/08/29, pp.1534-1557.
- Lauten, S., Otten, M., "Physics-Informed Neural Networks for Gate Design using Quantum Optimal Control", arXiv:2511.09463, 2025/11/12, pp.1-11.
- Seo, Jaemin, "Past rewinding of fluid dynamics from noisy observation via physics-informed neural computing", *Physical Review E*, Vol.110, No.2, 2024/08/06, pp. 025302(1-12).
- Socaciu, T., Paşcu, P., "Physics-Informed Neural Networks in Pricing Financial Options", *BRAIN(Broad Research in Artificial Intelligence and Neuroscience)*, Vol.16, No.2, 2025/07/29, pp.474-488.
- Suleiman Tawil, Cosmos-Predict2.5: Inside NVIDIA's Breakthrough in Physical World Simulation, *Medium,* Nov 6, 2025.
- Yongji Wang, Sam Blackwell, "Discovering new solutions to century-old problems in fluid dynamics", *Google DeepMind,* September 18, 2025.

4) 디지털 트윈 및 스마트 제조 (Digital Twin & Manufacturing RL)

- Donggun Lee, Yong-Shin Kang, Sang Do Noh, "Digital twin-driven deep reinforcement learning for real-time optimisation in dynamic AGV systems", *International Journal of Production Research*, Vol. 64, Issue 1, 2025/08/25, pp.106-124.
- Khdoudi, A., Masrour, T., El Hassani, I., El Mazgualdi, C., "A Deep-Reinforcement-Learning-Based Digital Twin for Manufacturing Process Optimization", *Systems (MDPI)*, Vol.12, No.2, 2024/01/24, pp.38-55.

- Lang Dai, "Intelligent manufacturing digital twin creation based on BIM and reinforcement learning", *Proceedings of SPIE*, Vol.13691, No.1, 2025/07/08, pp.1369134(1-8).
- Lee, D., Kang, Y.-S., Noh, S. D., "Digital twin-driven deep reinforcement learning for real-time optimisation in dynamic AGV systems", International Journal of Production Research (Taylor & Francis), Vol.64, No.1, 2025/08/25, pp.106-124.
- Tsai, Y.-K., Karkaria, V., Chen, Y.-P., Chen, W., "Digital Twin-enabled Multi-generation Control Co-Design with Deep Reinforcement Learning", Northwestern University, arXiv:2510.10694, 2025/10/12, pp.1-13.
- Ying-Kuan Tsai, Vispi Karkaria, Yi-Ping Chen, Wei Chen, "Multi-Generation Control Co-Design for Digital Twin-Enabled Systems With Deep Reinforcement Learning", *ASME International Design Engineering Technical Conferences*, Vol.3, No. V03AT03A006, 2025/10/27, pp.1-15.
- Zheng, H., Xie, W., Ryzhov, I. O., Choy, K., "Digital Twin Calibration with Model-Based Reinforcement Learning", arXiv:2501.02205, 2025/01/04, pp.1-28.

5) 항만 물류 및 AGV (Port Logistics & AGV)

- Lianhua An, Junqi Li, Hongchen Li, Zixuan Yang, Mi Chao & Jia Hu, "Energy-aware dispatching for automated container terminals", *Transportmetrica A: Transport Science*, Vo. 23, No.1, 2025/10/17, pp.1-20.
- Lu Zhen, Qian Zhang, Zheyi Tan & Shuaian Wang, "Scheduling AGVs in ports with battery charging and swapping", *IISE Transactions*, Vol.57, Issue 10, 2025/04/07, pp.1245-1262.
- Songle Leng, Yan Wang, Yanjie Zhou, Xuehao Feng, Paul T.W. Lee, Ruibin Bai, "The growing role of artificial intelligence in smart container ports: its application and future research directions", *Journal of International Logistics and Trade (Emerald)*, Vol.24, No.1, 2026/01/06, pp.1-35.
- Xurui Yang, Hongtao Hu, Yuren Wang & Chen Cheng, "AGV scheduling in automated container terminals considering multi-load strategy and charging requirements", *International Journal of Production Research*, Vol. 63, Issue 23, 2025/07/22, pp.9269-9297.
- Yuxuan Zhanga, Liang Chena, Moshi Zhoub, Xiangyu Baoa, Funing Jiac, Changhui Liu, Lei Zhangc, Yu Zhenga, "Deep reinforcement learning-based dynamic integrated scheduling of automated guided vehicles and yard cranes for container terminal loading operations", *Engineering Applications of Artificial Intelligence*, Vol.163, Part 1, 2026/01/11, pp.1-22.

6) 공간 지능 및 월드 모델 관련 영상 (링크)

- Fei Fei Li SBS 포럼: https://www.youtube.com/watch?v=bqeTK1J7UXE
- World Labs: https://x.com/lucas_flatwhite/status/2006033891828850921
- Marble Labs: https://marble.worldlabs.ai/
- NVIDA 공간지능: https://www.youtube.com/watch?v=6JCp0hXb3S4
- Genie3: https://deepmind.google/models/genie/
- Genie3 공간 만들기: https://www.youtube.com/watch?v=PDKhUknuQDg&t=59s
- Genie3 설명: https://www.youtube.com/watch?v=mMKIyxw3H9Q
- NVIDIA fVDB: https://blogs.nvidia.co.kr/blog/fvdb-bigger-digital-models/
- Cosmos-Predict2.5: https://research.nvidia.com/labs/dir/cosmos-predict2.5/
- Genie3, Cosmos-Predict2.5: https://www.youtube.com/watch?v=Ymj4c0qzZHA

[부록 : AI 용어]

AI 전체 기술을 다음 기준(①**지능을 만드는 기술,** ②**지능을 사용하는 기술**)으로 분류하여 정리
① **AI Core Stack(AI 자체/직접 기술 계층)**: 모델 중심 + 지능 자체를 구성하는 기술
② **AI System & Ecosystem Stack(AI 활용/연계 기술 계층)**: 서비스·인프라·운영·외부 시스템

1. AI Core Stack (AI 자체 기술)

1) AI 패러다임 / 알고리즘 (AI Paradigms & Algorithm)

- **ANI (Artificial Narrow Intelligence, 약한 인공지능)**: 특정 작업(예: 번역, 추천 등)에 특화된 AI로, 제한된 범위 내에서만 지능을 발휘함
- **AGI (Artificial General Intelligence, 인간 수준 범용 지능)**: 인간처럼 다양한 문제를 이해·학습·추론할 수 있는 범용 인공지능
- **ASI (Artificial Superintelligence, 초지능)**: 인간의 모든 지능을 초월하는 수준의 이론적 인공지능
- **Generative AI (생성형 인공지능)**: 텍스트, 이미지, 음성 등 새로운 데이터를 생성하는 AI 기술
- **Multimodal AI (멀티모달 인공지능)**: 텍스트, 이미지, 음성 등 다양한 형태의 데이터를 동시에 이해·처리하는 AI
- **Agentic AI (에이전트형 인공지능)**: 목표를 기반으로 계획을 세우고 도구를 활용해 자율적으로 작업을 수행하는 AI
- **Autonomous AI (자율형 인공지능)**: 인간의 개입 없이 스스로 판단하고 행동하는 AI 시스템
- **Embodied AI (체화된 인공지능)**: 로봇 등 물리적 몸을 통해 환경과 상호작용하며 학습하는 AI
- **Physical AI (피지컬 AI)**: 로봇, 드론, 자율주행 차량 등 현실 세계에서 물리적 환경과 직접 상호작용하는 AI. 센서 데이터와 물리적 제약을 고려하여 행동을 계획하고 수행하는 시스템
- **Spatial Intelligence (공간 지능)**: 3차원 환경 내에서 객체 위치, 거리, 움직임 등 공간 정보를 이해하고 활용하는 능력. 로봇 내비게이션, AR/VR, 3D 시뮬레이션 등에 필수적인 인지 능력
- **World Models (세계 모델)**: 에이전트가 환경의 동작 원리와 상태 변화를 내재적으로 학습하여 시뮬레이션하거나 예측할 수 있는 모델로 강화학습 및 자율 시스템에서 미래 상태 예측과 의사결정에 활용
- **Foundation Models (기반 모델)**: 대규모 데이터로 사전 학습된 범용 AI 모델로, 다양한 다운스트림 작업에 적용, LLM, 멀티모달 모델 등이 대표적이며, 전이학습과 생성형 AI에 활용

2) 모델 아키텍처 (Neural Architectures)

▸ **기본 구조**

- **DNN (Deep Neural Network, 심층 신경망)**: 여러 층의 신경망을 통해 복잡한 패턴을 학습하는 기본적인 딥러닝 구조

- **CNN (Convolutional Neural Network, 합성곱 신경망)**: 이미지와 같은 공간적 데이터를 처리하는 데 특화된 신경망
- **RNN (Recurrent Neural Network, 순환 신경망)**: 순차 데이터(문장, 시계열 등)를 처리하며 이전 정보를 기억하는 구조
- **LSTM (Long Short-Term Memory, 장단기 기억 네트워크)**: 장기 의존성 문제를 해결한 RNN 구조
- **GRU (Gated Recurrent Unit, 게이트 순환 유닛)**: LSTM을 단순화한 효율적인 순환 구조

▸ Transformer 계열

- **Transformer (트랜스포머)**: Self-Attention 메커니즘을 통해 입력 간 관계를 병렬적으로 학습하는 모델 구조
- **Encoder-Decoder (인코더-디코더 구조)**: 입력을 이해(Encoder)하고 출력으로 생성(Decoder)하는 구조
- **BERT (Bidirectional Encoder Representations from Transformers, 양방향 인코더 표현 모델)**: 문맥을 양방향으로 이해하는 인코더 기반 언어 모델
- **GPT (Generative Pre-trained Transformer, 생성형 사전학습 트랜스포머)**: 다음 단어를 예측하며 텍스트를 생성하는 디코더 기반 모델
- **T5 (Text-To-Text Transfer Transformer, 텍스트-투-텍스트 트랜스포머)**: 모든 NLP 문제를 텍스트 입력-출력 형태로 통합한 모델
- **Vision Transformer (ViT, 비전 트랜스포머)**: 이미지를 패치 단위로 나누어 처리하는 트랜스포머 기반 비전 모델
- **Diffusion Transformer (DiT, 확산 트랜스포머)**: 확산 모델과 트랜스포머를 결합한 생성 모델 구조

▸ 생성 모델

- **GAN (Generative Adversarial Network, 생성적 적대 신경망)**: 생성자와 판별자가 경쟁하며 데이터를 생성하는 모델
- **VAE (Variational Autoencoder, 변분 오토인코더)**: 확률 분포를 학습해 새로운 데이터를 생성하는 인코더-디코더 모델
- **Diffusion Models (확산 모델)**: 노이즈를 점진적으로 제거하며 데이터를 생성하는 확률 기반 모델
- **Normalizing Flow (정규화 흐름 모델)**: 가역 변환을 통해 확률 분포를 정확히 모델링하는 생성 방식

▸ 고급 구조

- **Graph Neural Network (GNN, 그래프 신경망)**: 노드와 엣지로 구성된 그래프 데이터를 학습하는 모델
- **Neural Radiance Fields (NeRF, 신경 방사장)**: 여러 각도의 2D 이미지를 이용해 복잡한 3D 장면을 재구성하고 새로운 시점의 이미지를 고품질로 만드는 딥러닝 기반 기술
- **Physics-Informed Neural Networks (PINN, 물리 기반 신경망)**: 뉴럴 네트워크 학습 과정에서 물리 법칙이나 수학적 제약을 통합하여 모델을 학습. 과학 시뮬레이션, 유체역학, 구조 해석 등 실제 물리 문제 해결에 사용

- Neural Operator (**신경 연산자**): 함수 간의 매핑을 학습하여 물리/수치 문제를 해결하는 모델

▸ **확장 구조**

- Sparse Models (**희소 모델**): 일부 파라미터만 활성화하여 효율성과 확장성을 높인 모델
- Mixture of Experts (MoE, **전문가 혼합 모델**): 여러 전문가 모델 중 일부만 선택적으로 활성화하는 구조
- Memory-Augmented Networks (**메모리 확장 신경망**): 외부 메모리를 활용하여 장기 정보 저장 및 활용이 가능한 모델

3) 모델 유형 (Model Classes)

- Language Model (LM, **언어 모델**): 텍스트 데이터의 확률 분포를 학습하여 문장을 이해하거나 생성하는 기본 모델
- Large Language Model (LLM, **대규모 언어 모델**): 대량의 데이터와 파라미터로 학습되어 고급 추론 및 생성 능력을 갖춘 언어 모델
- Small Language Model (SLM, **소형 언어 모델**): 경량화된 구조로 빠르고 효율적인 추론에 최적화된 언어 모델
- Large Multimodal Model (LMM, **대규모 멀티모달 모델**): 텍스트, 이미지, 음성 등 다양한 입력을 동시에 이해하고 생성하는 대형 모델
- Vision-Language Model (VLM, **시각-언어 모델**): 이미지와 텍스트를 함께 처리하여 설명, 질의응답 등을 수행하는 모델
- Vision-Language-Action (VLA, **시각-언어-행동 모델**): 시각과 언어 정보를 기반으로 실제 행동까지 수행하는 모델
- Large Action Model (LAM, **대규모 행동 모델**): 복잡한 작업 수행을 위해 행동 계획 및 실행 능력을 학습한 모델
- Speech/Audio Models (AudioLM **계열, 음성/오디오 모델**): 음성 인식, 생성, 변환 등을 수행하는 오디오 기반 모델
- Code Models (**코드 모델**): 프로그래밍 코드 생성, 이해, 수정 등을 수행하는 특화 모델
- Decision Transformer (**결정 트랜스포머**): 트랜스포머 구조를 활용하여 시퀀스 기반 의사결정 문제를 해결하는 모델
- Diffusion Policy Models (**확산 정책 모델**): 확산 모델을 활용해 로봇 제어 등 연속적인 행동 정책을 학습하는 모델
- Embedding Models (**임베딩 모델**): 데이터를 벡터 공간으로 변환하여 의미적 유사도를 표현하는 모델
- Retrieval Models (**검색 모델**): 대규모 데이터에서 관련 정보를 검색하고 선택하는 데 최적화된 모델

4) 학습 패러다임 (Learning Paradigms)

▸ **기본**

- **Supervised Learning (지도 학습)**: 입력과 정답(label)을 함께 사용하여 모델을 학습하는 방식
- **Unsupervised Learning (비지도 학습)**: 정답 없이 데이터의 구조나 패턴을 스스로 학습하는 방식
- **Self-Supervised Learning (자기지도 학습)**: 데이터 자체에서 정답을 생성하여 학습하는 방식
- **Semi-Supervised Learning (준지도 학습)**: 소량의 라벨 데이터와 다량의 비라벨 데이터를 함께 활용하는 방식

▸ 일반화 / 적응

- **Transfer Learning (전이 학습)**: 사전학습된 모델을 다른 작업에 재사용하여 학습 효율을 높이는 방법
- **Meta Learning (메타 학습)**: "학습하는 방법" 자체를 학습하여 새로운 작업에 빠르게 적응하는 기법
- **Continual Learning (지속 학습 / 평생 학습)**: 기존 지식을 유지하면서 새로운 정보를 계속 학습하는 방식

▸ 저데이터 학습

- **Few-shot Learning (소량 샘플 학습)**: 적은 수의 예시만으로 새로운 작업을 수행하는 능력
- **Zero-shot Learning (제로샷 학습)**: 학습하지 않은 작업도 설명만으로 수행하는 능력
- **One-shot Learning (원샷 학습)**: 단 하나의 예시로 개념을 학습하는 방식

▸ 강화학습

- **Reinforcement Learning (RL, 강화학습)**: 에이전트가 환경과 상호작용하며 행동을 선택하고 보상을 최대화하도록 학습하는 알고리즘. 정책 학습, 가치 함수, 보상 설계 등을 통해 최적 행동을 탐색
- **Deep RL (Deep Reinforcement Learning, 심층 강화학습)**: 딥러닝과 강화학습을 결합한 방식
- **RLHF (Reinforcement Learning from Human Feedback, 인간 피드백 기반 강화학습)**: 인간 평가를 보상으로 활용해 모델을 개선하는 방법
- **RLAIF (Reinforcement Learning from AI Feedback, AI 피드백 기반 강화학습)**: AI가 생성한 피드백을 활용해 모델을 학습하는 방식
- **Offline RL (오프라인 강화학습)**: 환경과 상호작용 없이 기존 데이터로만 학습하는 강화학습 방식
- **Multi-Agent RL (다중 에이전트 강화학습)**: 여러 에이전트가 상호작용하며 학습하는 강화학습 방식

▸ 데이터 중심

- **Data-centric AI (데이터 중심 인공지능)**: 모델보다 데이터 품질 개선에 초점을 맞춘 AI 접근 방식
- **Synthetic Data (합성 데이터)**: 실제 데이터 대신 인위적으로 생성된 학습용 데이터
- **Data Augmentation (데이터 증강)**: 기존 데이터를 변형하여 학습 데이터를 늘리는 기법
- **Active Learning (능동 학습)**: 모델이 중요한 데이터를 선택해 라벨링을 요청하며 학습하는 방식

5) 추론 / 인지 / 사고 메커니즘 (Reasoning Mechanisms)

- Prompt Engineering (**프롬프트 엔지니어링**): 원하는 결과를 얻기 위해 입력 문장을 설계·최적화하는 기법
- Prompt Tuning (**프롬프트 튜닝**): 프롬프트를 학습 가능한 형태로 만들어 모델 성능을 개선하는 방법
- Fine-tuning (**파인튜닝**): 사전학습된 모델을 특정 작업에 맞게 추가 학습하는 과정
- Instruction Tuning (**지시문 튜닝**): 다양한 지시문 데이터를 통해 모델이 명령을 잘 따르도록 학습하는 방법
- Context Engineering (**컨텍스트 엔지니어링**): 모델 입력에 포함되는 정보(문맥)를 설계하여 성능을 높이는 기법
- Chain-of-Thought (CoT, **사고 연쇄 추론**): 중간 추론 과정을 단계적으로 생성하여 문제 해결 능력을 향상시키는 방식
- Tree-of-Thought (ToT, **사고 트리 추론**): 여러 추론 경로를 탐색하며 최적의 답을 찾는 방식
- Graph-of-Thought (**그래프 기반 사고**): 복잡한 관계를 그래프 형태로 확장하여 추론하는 방법
- Self-Consistency (**자기 일관성**): 여러 추론 결과를 생성한 뒤 가장 일관된 답을 선택하는 기법
- Planning (**계획 수립**): 주어진 목표를 달성하기 위해 행동 순서와 전략을 사전에 설계하는 과정
- Reasoning (**추론**): 주어진 정보나 지식을 기반으로 논리적 결론을 도출하는 사고 과정
- Tool Use (**도구 사용**): AI가 외부 소프트웨어, API, 장치 등 기능적 도구를 활용해 문제 해결이나 작업 수행을 하는 능력
- Function Calling (**함수 호출**): AI가 특정 작업을 수행하기 위해 프로그래밍 함수나 API를 실행하는 기능적 인터페이스 사용 방법
- Long-context Modeling (**장문맥 처리 모델링**): 긴 입력 문맥을 유지하며 이해·추론할 수 있는 능력

6) 지식 & 검색 통합 구조

- Retrieval-Augmented Generation (RAG, **검색 증강 생성**): 외부 데이터베이스에서 정보를 검색한 후 이를 기반으로 더 정확한 답변을 생성하는 방식
- Semantic Search (**의미 기반 검색**): 키워드가 아닌 의미 유사도를 기반으로 관련 정보를 찾는 검색 기술
- Vector Representation Learning (**벡터 표현 학습**): 데이터를 벡터 공간에 임베딩하여 의미적 관계를 학습하는 기법
- Knowledge Graph Integration (**지식 그래프 통합**): 개체와 관계로 구성된 지식 그래프를 활용해 구조적 추론을 강화하는 방식

7) 에이전트 및 자율성

- AI Agents (**인공지능 에이전트**): 목표를 달성하기 위해 환경을 인식하고 행동을 선택하는 자율 시스템

- **Autonomous Agents (자율 에이전트)**: 인간 개입 없이 스스로 판단하고 작업을 수행하는 에이전트
- **Multi-Agent Systems (MAS, 다중 에이전트 시스템)**: 여러 에이전트가 협력 또는 경쟁하며 문제를 해결하는 구조
- **Swarm Intelligence (군집 지능)**: 다수의 단순 에이전트가 집단적으로 복잡한 행동을 만들어 내는 방식
- **ReAct Framework (Reasoning + Acting, 추론-행동 프레임워크)**: 추론과 행동을 반복하며 문제를 해결하는 에이전트 구조
- **Digital Agents (디지털 에이전트)**: 소프트웨어 환경에서 사용자 작업을 대신 수행하는 가상 에이전트
- **Copilot / AI Assistant (코파일럿 / 인공지능 비서)**: 인간의 작업을 보조하며 생산성을 향상시키는 AI 시스템
- **Synthetic Agents (합성 에이전트)**: 시뮬레이션 환경에서 생성된 인공지능 에이전트
- **Program Synthesis (프로그램 합성)**: 자연어 또는 명세로부터 자동으로 코드나 프로그램을 생성하는 기술

8) 확률 / 인과 / 의사결정

- **Causal AI (인과 인공지능)**: 변수 간의 원인과 결과 관계를 학습하여 더 정확한 의사결정을 수행하는 AI
- **Bayesian Networks (베이지안 네트워크)**: 확률적 의존 관계를 그래프로 표현하여 추론하는 모델
- **Probabilistic Programming (확률 프로그래밍)**: 확률 모델을 코드 형태로 정의하고 자동으로 추론하는 접근 방식
- **Markov Decision Process (MDP, 마르코프 의사결정 과정)**: 상태, 행동, 보상을 기반으로 최적 정책을 찾는 수학적 모델
- **Hidden Markov Model (HMM, 은닉 마르코프 모델)**: 관측되지 않는 상태를 확률적으로 추정하는 시계열 모델

9) 신뢰성 / 정렬 / 안전성

- **Explainable AI (XAI, 설명 가능한 인공지능)**: 모델의 예측 결과를 사람이 이해할 수 있도록 설명하는 기술
- **Interpretable AI (해석 가능한 인공지능)**: 모델 내부 구조와 작동 원리를 직접 이해할 수 있는 AI
- **Trustworthy AI (신뢰할 수 있는 인공지능)**: 공정성, 안전성, 투명성 등을 충족하는 신뢰 가능한 AI
- **Responsible AI (책임 있는 인공지능)**: 윤리적·사회적 책임을 고려하여 설계되고 운영되는 AI
- **AI Safety (인공지능 안전성)**: AI가 의도하지 않은 위험한 행동을 하지 않도록 보장하는 기술과 연구 분야

- Alignment (**정렬, 인간 가치 정렬 문제 포함**): AI의 목표와 인간의 가치 및 의도를 일치시키는 과정 및 문제
- Robustness (**강건성**): 노이즈, 공격, 데이터 변화에도 안정적으로 작동하는 능력
- Generalization (**일반화**): 학습하지 않은 새로운 데이터에도 잘 대응하는 능력
- Hallucination (**환각 현상**): 사실과 다른 정보를 그럴듯하게 생성하는 생성형 AI의 문제

10) AI 기본 응용 능력 (Core Capabilities)

▸ Vision

- Object Detection (**객체 탐지**): 이미지나 영상에서 특정 객체의 위치와 종류를 식별하는 기술
- Image Segmentation (**이미지 분할**): 이미지의 각 픽셀을 분류하여 객체 영역을 정밀하게 구분하는 기술
- OCR (Optical Character Recognition, **광학 문자 인식**): 이미지 속 텍스트를 추출하여 디지털 문자로 변환하는 기술
- 3D Vision (**3차원 비전**): 깊이와 공간 정보를 포함한 3D 환경을 인식하고 이해하는 기술

▸ Speech

- ASR (Automatic Speech Recognition, **자동 음성 인식**): 음성을 텍스트로 변환하는 기술
- TTS (Text-To-Speech, **텍스트 음성 변환**): 텍스트를 자연스러운 음성으로 생성하는 기술
- Voice Cloning (**음성 복제**): 특정 사람의 목소리를 학습하여 유사한 음성을 생성하는 기술

▸ Physical Intelligence

- SLAM (Simultaneous Localization and Mapping, **동시 위치 추정 및 지도 작성**): 환경 지도를 만들면서 동시에 자신의 위치를 추정하는 기술
- Motion Planning (**경로 계획**): 목표 지점까지 최적의 이동 경로를 계산하는 기술
- Manipulation Learning (**조작 학습**): 로봇이 물체를 잡고 이동시키는 등의 물리적 작업을 학습하는 기술

2. AI System & Ecosystem Stack (연계/활용 기술)

1) 데이터 엔지니어링 & 표현

- **Token (토큰)**: 거대 언어 모델(LLM)이 텍스트를 이해하고 생성하는가장 작은 데이터 처리 단위(글자, 단어, 부분 단어 등)
- **Tokenization (토큰화)**: 텍스트를 단어, 서브워드, 문자 등의 단위(토큰)로 분할하는 과정
- **Chunking (청킹)**: 긴 문서나 데이터를 일정 크기의 의미 단위로 나누는 처리 방식
- **Labeling (라벨링)**: 데이터에 정답(클래스, 태그 등)을 부여하여 학습에 활용하는 과정
- **Embedding Pipeline (임베딩 파이프라인)**: 데이터를 벡터로 변환하고 저장·활용하는 전체 처리 흐름
- **Data Pipelines (데이터 파이프라인)**: 데이터 수집, 정제, 변환, 저장까지 자동화된 처리 과정

2) 모델 개발 & 학습 운영

- **Training (학습)**: 데이터를 이용해 모델의 파라미터를 최적화하는 과정
- **Fine-tuning Infrastructure (파인튜닝 인프라)**: 사전학습 모델을 특정 목적에 맞게 재학습하기 위한 시스템 환경
- **LoRA (Low-Rank Adaptation, 저랭크 적응)**: 사전학습 모델의 일부 파라미터만 조정하여 효율적으로 모델을 튜닝하는 방법
- **PEFT (Parameter-Efficient Fine-Tuning, 파라미터 효율적 파인튜닝)**: 적은 수의 파라미터만 업데이트하여 학습 효율과 메모리 사용을 최적화하는 파인튜닝 기법
- **Hyperparameter Optimization (하이퍼파라미터 최적화)**: 학습률, 배치 크기 등 설정값을 조정해 모델 성능을 개선하는 과정

3) 추론 & 서비스 레이어

- **Inference Serving (추론 서빙)**: 학습된 모델을 통해 실시간 또는 배치 방식으로 예측 결과를 제공하는 시스템
- **Model Serving (모델 서빙)**: 모델을 API 형태로 배포하여 외부 서비스에서 사용할 수 있게 하는 기술
- **API Deployment (API 배포)**: 모델 기능을 REST/HTTP 등의 인터페이스로 서비스에 연결하는 과정
- **Batching (배치 처리)**: 여러 입력 데이터를 한 번에 묶어 처리하여 연산 효율과 처리량을 높이는 기법
- **Caching (캐싱)**: 이미 계산된 결과를 저장해 재사용함으로써 속도와 자원 효율을 개선하는 기술

4) 예측 모델

- **Decision Tree (의사결정나무)**: 데이터의 조건 분기를 기반으로 결과를 예측하는 직관적이고 해석 가능한 모델

- Random Forest (랜덤 포레스트): 여러 의사결정나무를 결합하여 예측 정확도를 높이고 과적합을 줄이는 앙상블 모델
- GBM (Gradient Boosting Machine, 그래디언트 부스팅 머신): 이전 모델의 오차를 순차적으로 보완하며 예측 성능을 향상시키는 부스팅 방식의 모델
- XGBoost (Extreme Gradient Boosting, 익스트림 그래디언트 부스팅): 병렬 처리와 정규화를 통해 속도와 예측 성능을 크게 개선한 GBM 기반 모델
- LightGBM (Light Gradient Boosting Machine, 라이트 그래디언트 부스팅 머신): 대규모 데이터에서도 빠르고 효율적인 학습이 가능한 경량 부스팅 모델
- CatBoost (Categorical Boosting, 범주형 부스팅): 범주형 데이터를 효과적으로 처리하여 높은 예측 성능을 제공하는 부스팅 모델

5) 검색 & 메모리 인프라

- Vector Database (벡터 데이터베이스): 임베딩 벡터 형태로 변환된 데이터를 효율적으로 저장하고 검색하는 특화된 데이터베이스(DB)이며, 대규모 AI 모델에서 생성된 텍스트, 이미지, 오디오 등의 벡터를 기반으로 유사도 검색(Similarity Search)을 빠르게 수행
- Indexing Systems (인덱싱 시스템): 빠른 검색을 위해 데이터를 구조화하고 색인을 생성하는 시스템
- Retrieval Infrastructure (검색 인프라): 대규모 데이터에서 관련 정보를 효율적으로 검색하기 위한 전체 시스템 구성

6) AI 애플리케이션 프레임워크

- LangChain (랭체인): LLM 기반 애플리케이션을 구축하기 위한 체인 및 에이전트 프레임워크
- LlamaIndex (라마인덱스): 외부 데이터와 LLM을 연결하여 검색 및 질의응답을 구현하는 프레임워크
- Orchestration Systems (오케스트레이션 시스템): 여러 AI 컴포넌트와 서비스 흐름을 조정·관리하는 시스템
- Workflow Engines (워크플로우 엔진): 작업 순서와 실행 로직을 정의하고 자동으로 처리하는 시스템

7) 에이전트 시스템 구현

- Agent Frameworks (에이전트 프레임워크): AI 에이전트를 설계·구현하기 위한 개발 도구 및 구조
- Tool Integration Systems (도구 통합 시스템): 외부 API, 데이터베이스, 소프트웨어 등을 AI와 연결하는 시스템
- Task Execution Engines (작업 실행 엔진): 계획된 작업을 실제로 실행하고 결과를 관리하는 실행 시스템

8) 성능 최적화 & 경량화

- **Quantization (양자화)**: 모델의 수치 정밀도를 낮춰(예: FP32 → INT8) 연산 속도와 메모리 효율을 개선하는 기법
- **Pruning (프루닝)**: 중요하지 않은 가중치나 뉴런을 제거하여 모델 크기와 연산량을 줄이는 방법
- **Distillation (지식 증류)**: 큰 모델(교사)의 지식을 작은 모델(학생)에 전달해 성능을 유지하면서 경량화하는 기법
- **Model Compression (모델 압축)**: 다양한 기법을 활용해 모델 크기와 연산 비용을 줄이는 기술의 총칭
- **Compilation (AI Compiler, AI 컴파일러)**: 모델을 하드웨어에 최적화된 형태로 변환하여 실행 성능을 높이는 기술

9) 인프라 & 하드웨어

- **GPU (Graphics Processing Unit, 그래픽 처리 장치)**: 대규모 병렬 연산에 강점을 가진 AI 학습 및 추론용 핵심 하드웨어
- **TPU (Tensor Processing Unit, 텐서 처리 장치)**: 텐서 연산에 특화된 AI 가속기로, 고속 딥러닝 연산을 지원
- **NPU (Neural Processing Unit, 신경망 처리 장치)**: AI 연산을 위해 설계된 전용 칩으로 모바일·엣지 환경에서 활용됨
- **Edge Computing (엣지 컴퓨팅)**: 데이터를 클라우드가 아닌 디바이스 근처에서 처리하는 분산 컴퓨팅 방식
- **Cloud AI Infrastructure (클라우드 AI 인프라)**: AI 모델 학습·배포를 위한 클라우드 기반 컴퓨팅 자원과 플랫폼
- **Distributed Training (분산 학습)**: 여러 장치나 서버를 활용해 모델 학습을 병렬로 수행하는 방식
- **Model Parallelism (모델 병렬화)**: 하나의 큰 모델을 여러 장치에 분할하여 병렬로 학습하거나 추론하는 방식
- **Data Parallelism (데이터 병렬화)**: 동일 모델을 여러 장치에서 데이터의 서로 다른 배치를 동시에 학습하는 방식

10) 운영 / 관측 / 품질관리

- **MLOps (Machine Learning Operations, 머신러닝 운영)**: 모델 개발부터 배포·운영까지 전체 생명주기를 관리하는 방법론
- **LLMOps (Large Language Model Operations, 대규모 언어 모델 운영)**: LLM 특화된 배포, 모니터링, 최적화 운영 체계
- **AIOps (Artificial Intelligence Operations, 인공지능 운영)**: IT 운영에 AI를 적용해 자동화 및 이상 탐지를 수행하는 기술
- **DataOps (Data Operations, 데이터 운영)**: 데이터 파이프라인과 품질을 지속적으로 관리하는 운영 방식

- Model Monitoring (모델 모니터링): 운영 중인 모델의 성능과 상태를 지속적으로 추적·관리하는 활동
- Drift Detection (드리프트 탐지): 데이터나 모델 성능이 시간에 따라 변화하는 현상을 감지하는 기술
- Evaluation Systems (평가 시스템): 모델의 성능과 품질을 체계적으로 측정하고 검증하는 시스템
- Benchmarking (벤치마킹): 표준 지표와 데이터셋을 통해 모델 성능을 비교·평가하는 방법

11) 거버넌스 / 정책 / 보안

- AI Governance (인공지능 거버넌스): AI 개발과 활용에 대한 정책, 기준, 관리 체계를 수립하는 활동
- AI Regulation (인공지능 규제): 법과 제도를 통해 AI의 안전성과 책임성을 확보하는 정책적 장치
- AI Ethics (인공지능 윤리): 공정성, 투명성, 책임성 등 윤리적 기준을 AI에 적용하는 원칙
- Privacy-preserving AI (프라이버시 보호 인공지능): 개인정보를 보호하면서 데이터를 활용하는 기술(예: 연합학습, 암호화)
- Sovereign AI (주권형 인공지능): 국가나 조직이 데이터와 AI 기술을 독립적으로 통제·운영하는 개념

12) 산업 및 시스템 통합

- Autonomous Systems (자율 시스템): 외부 개입 없이 환경을 인식하고 스스로 판단·행동하는 시스템
- Sensor Fusion (센서 융합): 다양한 센서 데이터를 결합해 더 정확한 인식과 판단을 수행하는 기술
- Digital Twin (디지털 트윈): 현실 시스템을 가상 환경에 복제하여 분석·예측·최적화를 수행하는 기술
- Edge AI Applications (엣지 AI 응용): 디바이스에서 직접 AI를 실행하는 응용 서비스 (예: 스마트폰, IoT)
- AI SaaS (AI Software as a Service, AI 서비스형 소프트웨어): 클라우드를 통해 AI 기능을 서비스 형태로 제공하여 사용자가 설치 없이 활용할 수 있는 플랫폼
- AI Platforms (AI 플랫폼): 데이터 수집, 모델 학습, 배포, 모니터링 등 AI 생애주기 전체를 지원하는 통합 환경 또는 시스템

산업 분야의 AI 시스템 구현 방안

발 행 처	(주)심스리얼리티
작 성 자	AI융합원, 임명환
발 행 일	2026. 3. 31.

※ 본 자료는 사전 허가없이 무단 배포 및 복제를 금지합니다.