

# Physical AI의 시대

## - 소비자 보호를 위한 시사점

### 목 차

1. 들어가는 말	/ 01
2. Physical AI 활용과 소비자 보호	/ 08
3. AI 기술과 미래	/ 15

## 1. 들어가는 말

### 1.1 최근 AI 기술동향 - 2026 CES

- 엔비디아의 CEO 젠슨 황은 2026 CES에서 자율주행차용 AI 모델과 신형 칩의 출시를 발표<sup>1)</sup>
  - 젠슨 황은 물리적 AI를 위한 ChatGPT의 시대가 도래했음을 선포
  - 기계가 이성적으로 명령을 이해하고 행동하기 시작하는 것이 현실 세계에서 가능해졌음을 시사
  - 메르세데스 벤츠 CLA에 엔비디아의 운전자 지원 소프트웨어가 탑재될 것을 소개하며, 이는 “first entire stack endeavor”<sup>2)</sup>이며, 이 모델은 샌프란시스코에서 보행자를 피하고 회전하는 모습을 시연
  - 젠슨 황이 언급한 “entire stack”은 과거 자동차 회사들이 각 부품들을 업체로부터 제공받아 조립하던 상황에서 벗어나 엔비디아가 칩, 운영체제, 인식 AI, 자율주행 알고리즘, 시뮬레이션 훈련 인프라까지 통합하여 제공할 수 있음을 보여 주는 첫 사례로 지목됨
  - 자율주행이 가능한 것은 NVIDIA Halos를 통하여 자율주행차의 전체 스택에서 차량 안전이 보장되기 때문

1) Flynn, Kerry. 2026. “ChatGPT Moment for Physical AI’: Nvidia CEO Launches New AI Models and Chips.” Axios, January 5, 2026.

2) 일반적으로는 full-stack이라고 정의되며 robot hardware platform과 database, AI models, training pipeline, deployment system을 모두 포함한 구성으로 이해된다. John VerWey et al., “AgiBot World Colosseo: A Large-Scale Manipulation Platform for Scalable and Intelligent Embodied Systems,” arXiv preprint arXiv:2503.06669v3 (2025), <https://arxiv.org/abs/2503.06669> 젠슨 황은 entire stack으로 표현했으나 학술적으로는 full-stack, end-to-end, cross-stack 등으로 표현되기도 한다. 자율주행에서 “인지-예측-제어”의 전 과정의 chain을 하나로 묶는 것이 full-stack인 것이다. 다만, 젠슨 황의 표현을 직접 인용하여 시기적으로 Physical AI의 시대가 도래하였음을 강조하기 위하여 위의 표현을 직접 인용했음을 밝힌다.

- NVIDIA Halos는 차량 아키텍처, AI 모델, 칩, 소프트웨어, 도구 및 서비스가 통합되어 “클라우드에서 자동차까지” 자율주행차의 안전한 개발 및 배포를 보장하는 종합 안전 시스템
- 이러한 시스템은 설계 시점, 배포 시점, 검증 시점의 가드레일을 포함하여 전체 개발 수명 주기를 포함하고 AI 기반 AV 스택의 안전성과 설명가능성을 통합<sup>3)</sup>
- 결과적으로, 차량은 방대한 실제 운전 데이터와 인공 운전 데이터를 학습하고 복잡한 환경과 상황에서 인간과 같은 의사결정으로 안전하게 운전하는 것이 가능함. 그러면서도 시스템과 운전자 간의 협력적 조향은 여전히 유지
- 젠슨 황은 루빈 플랫폼의 출시를 통하여 차세대 AI의 서막을 알림<sup>4)</sup> CES 2026에서 “physical AI의 ChatGPT moment”<sup>5)</sup>라는 표현을 통해 AI와 컴퓨팅의 시대적 전환이 물리적 세계에서 지각-판단-행동이 가능한 자율주행 시스템으로 이동되고 있음을 확인

## □ 2026 CES의 가장 큰 화두였던 Physical AI

- 2026 CES에서는 Physical AI를 로봇의 혁명이라 칭하며, 로봇 공학으로 인한 인류 삶의 질 향상에 대한 기대 확산; 최신 휴머노이드 로봇(아틀라스 프로토타입 등)의 시연 진행<sup>6)</sup>

3) 이는 세 가지 고성능 컴퓨팅 인프라를 기반으로 구현되는데, 인공지능 모델의 학습을 위한 NVIDIA DGX™, 시뮬레이션 환경 구축 및 검증을 위한 NVIDIA Omniverse™, Cosmos™, 최종적으로 실제 현장 배치를 위한 NVIDIA AGX™로 구성된다. NVIDIA, “Autonomous Vehicle Safety: From the Cloud to the Car — NVIDIA Halos,” NVIDIA, accessed Feb. 2026, <https://www.nvidia.com/en-us/self-driving-cars/safety/>

4) NVIDIA. 2026. “NVIDIA Kicks Off the Next Generation of AI With Rubin — Six New Chips, One Incredible AI Supercomputer.” NVIDIA Newsroom, January 5, 2026.

5) NVIDIA. 2026. “NVIDIA Releases New Physical AI Models as Global Partners Unveil Next-Generation Robots.” NVIDIA Newsroom, January 5, 2026.

6) Consumer Technology Association. 2026. “The Future Feels Closer Than Ever: Day 3 of CES 2026.” CES, January 9, 2026.

- 보스턴 다이내믹스(항법과 조작 능력 등이 강화된 휴머노이드 통합), 캐터필러(광산 및 작업 현장에서의 자율성 도입), 프랑카 로보틱스, 휴머노이드, LG 일렉트로닉스(실내 작업용 홈 로봇), NEURA 로보틱스(소형 휴머노이드) 등 글로벌 업계 선두들은 NVIDIA의 로봇 스택을 활용하여 새로운 AI 기반 로봇을 선보임<sup>7)</sup>
- NVIDIA의 이러한 행보는 Physical AI를 생성형 AI 다음의 주요 성장 분야로 확대하며 서버에 머무르던 대규모 모델들이 병원과 산업, 가정 등으로 옮겨 갈 수 있는 기술적 기반을 제공

## 1.2 Physical AI - 인공지능과 로봇 공학의 융합 시대

### □ Physical AI 시대의 도래

- 생성형 AI의 등장과 활용으로 지난 3년간 인공지능 챗봇, 최첨단 AI 연구소 등에서의 새로운 모델의 출시는 Physical AI 시대의 도래를 예언해 왔음<sup>8)</sup>
- Physical AI는 AI 발달의 최종 단계로 기대되어 모건스탠리의 분석에 따르면 휴머노이드 로봇 시장은 2050년까지 5조 달러의 시장을 가질 것으로 전망<sup>9)</sup>

7) NVIDIA는 개발자들이 자원 집약적인 사전 학습을 우회하고, 차세대 AI 로봇과 자율 기계 개발에 집중할 수 있도록 오픈 모델을 구축하고 있음, Hugging Face를 통하여 모두 구매가 가능하고 이러한 모델들은 물리적 기반 합성 데이터 생성과 로봇 정책 평가를 가능하게 하는 모델, 지능형 기계가 인간처럼 물리적 세계에서 이해하고 행동할 수 있는 오픈 추론 비전 언어모델(VLM)등을 포함. NVIDIA. 2026. "NVIDIA Releases New Physical AI Models as Global Partners Unveil Next-Generation Robots." NVIDIA Newsroom, January 5, 2026.

8) 휴머노이드 로봇은 인간을 위해 설계된 환경에서 효율적으로 작동할 수 있으면서 배치 및 운용 시 최소한의 조정만 필요로 하는 이상적인 범용 로봇 구현체로 골드만 삭스에 따르면 휴머노이드 로봇 글로벌 시장은 2035년까지 380억 달러에 이를 것으로 예상 Huang, Madison. 2025. "What Is NVIDIA's Three-Computer Solution for Robotics?" NVIDIA Blog, August 8, 2025.

9) CNBC. 2025. "Morgan Stanley Says Humanoid Robots Will Be a \$5 Trillion Market by 2050 — How to Play It." CNBC, April 29, 2025.

- 인공지능과 로봇틱스의 융합은 아직 역사가 길지 않아 통일된 명칭도 없는 상태, 일부 기업에서는 embodied AI라고 칭하거나, physical AI, embodied machine intelligence, generative physical AI 등으로 부르기도 함<sup>10)</sup>
- 인공지능 분야가 비약적으로 발전하면서 로봇의 학습 방식, 훈련 방법, 현실 세계의 피드백을 해석하는 방식, 로봇이 데이터 학습에 활용할 수 있는 데이터 출처의 변화가 수반됨, 구체적으로는 대규모 언어모델(LLMs), 멀티모달 파운데이션 모델 (multimodal foundation models), 강화 학습(Reinforcement Learning), 시뮬레이션-현실 전이(Simulation-to-Reality) 부분의 발전이 로봇의 추론 능력과 지각, 기술 습득, 훈련 전반에 걸친 범주적 개선을 이끌었음<sup>11)</sup> 결국 이러한 기술이 상호 결합하여 로봇이 인간의 의도를 이해하고(LLMs), 환경의 관련 요소를 지각하고(멀티모달 모델), 과업을 효율적으로 수행하는 방법을 학습하여(강화학습), 고비용의 실제 환경 훈련 대신 비교적 저비용의 시뮬레이션 데이터를 활용하여 학습을 수행하는 것임<sup>12)</sup>

## □ Physical AI 의 활용

- 국제 로봇틱스 협회에서는 로봇을 “의도된 과업을 수행하기 위하여 환경 내에서 이동하며, 일정 수준의 자율성을 갖추고 두 개 이상의 축(AXIS)

10) John VerWey et al., “AgiBot World Colosseo: A Large-Scale Manipulation Platform for Scalable and Intelligent Embodied Systems,” arXiv:2503.06669v3 (2025).

11) VerWey, Supra note 30.; Roya Firoozi, Johnathan Tucker, Mac Schwager et al., “Foundation models in robotics: Applications, challenges, and the future,” The International Journal of Robotics Research 44, no. 5 (2024): 701-739. 특히 의사 결정 분야에서는 LLM과 VLM은 고수준의 계획을 수행하는 로봇을 도우며 맥락적 이해의 능력치를 향상시키고 언어 기반의 보상은 형성된 보상을 제공하기에 강화학습 에이전트를 사용하는 데 활용된다. 로봇의 훈련과 추론을 위한 소프트웨어의 맥락에서는 LLM이, 멀티 모달 모델은 지각 및 객체 이해를 처리하고, 강화 학습은 기술 습득을 가능하게 하고, 시뮬레이션 현실 전이는 대규모 훈련을 가능하게 한다.

12) VerWey, Supra at 7.

에서 프로그래밍이 가능한 구동 매커니즘 (actuated mechanism)으로 정의”하고 로봇에 해당하지 않는 범주는 소프트웨어, 드론, 음성비서, 자율주행 자동차, ATM, 스마트 세탁기 등으로 구분, Physical AI의 맥락에서 로봇은 자율 이동 로봇(AMRs), 매니플레이터 암(manipulator arms), 또는 휴머노이드 로봇을 중심으로 이해되고<sup>13)</sup> 이 외의 수십 종의 로봇들은 4족 보행 로봇(dog-like), 의료 보조 로봇(수술 지원 등), 그리고 인간이 통제 루프안에 개입하는 협동 로봇(cobots) 등이 있으며 산업현장에서 주로 활용<sup>14)</sup>

- Physical AI가 활용되기 위하여 직면한 과제는 로봇 공급망의 일부가 아직 산업의 초기 단계라는 점, 주요 하드웨어의 기술 돌파구를 찾기 어렵다는 점, 확장 가능한 제조시설의 준비가 저조하여 배터리, 모터, 센서, 액추에이터는 알고리즘이나 소프트웨어보다도 느리게 진화되며 상대적으로 마진율이 낮아 투자 자본에 인내심이 요구됨<sup>15)</sup>
- 세계경제포럼(WEC)은 Physical AI가 제조업 및 스마트 공장에서 반복적인 작업과 통제된 환경에서 활용될 것을 예고, 복잡한 환경을 인지하고 학습하여 폭넓게 활용될 것으로 기대<sup>16)</sup>
- 산업현장에서의 자동화는 끝없는 진화를 거듭, ‘규칙 기반’의 산업용 로봇에서 ‘훈련 기반 로봇’인 Physical AI까지 성장,<sup>17)</sup> Physical AI는

13) International Organization for Standardization, ISO 8373: Robots and Robotic Devices - Vocabulary, 2nd ed. (Geneva: ISO, 2012), <https://www.iso.org/obp/ui/-iso:std:iso:8373:ed-2:v1:en>; The IFR also explicitly states what a robot is not: software, drones, voice assistants, autonomous cars, nor ATMs, smart washing machines, etc., See: International Federation of Robotics, “World Robotics Press Conference 2024.”; Madison Huang, “What is NVIDIA’s Three-Computer Solution for Robotics?” NVIDIA (blog), August 8, 2025, <https://blogs.nvidia.com/blog/three-computers-robotics/>.

14) VerWey, *Supra* at 7, note 24.

15) VerWey, *Supra*, at 29.

16) Theresa Wolf and Andrea Willige, “Physical AI Is Changing Manufacturing - Here’s What the Era of Intelligent Robotics Looks Like,” World Economic Forum, September 9, 2025, <https://www.weforum.org/stories/2025/09/what-is-physical-ai-changing-manufacturing/>

17) *Id.*, section “The evolution of industrial robotics.”

특정 프로그램을 엄격히 준수할 수도 있고 일정 수준의 변형된 작업도 수행이 가능하기에 중·대량 생산에서 비반복적 생산 작업에도 적용 가능

- 국내 선행 연구에서의 Physical AI로는 대표적으로 QT-Opt(강화학습 기반 로봇 조작), Dreamer(강화학습 에이전트), Gato(범용 트랜스포머 기반 에이전트), RoboCat(비전 기반 로봇 조작을 위한 자기 개선 범용 에이전트), Optimus(Tesla가 공개한 실제 로봇) 등을 들 수 있음<sup>18)</sup>
- Physical AI를 학습에서 시뮬레이션 영역을 통한 학습을 활용하여 Physical AI의 분야가 확장되는 추세. 물리 환경에서 직접 학습을 위한 시간, 비용, 안전의 문제를 해결하기 위하여 시뮬레이션 기반 학습과 실제 환경으로의 전이(Sim-to-Real)를 위한 기법이 연구됨<sup>19)</sup>
- Physical AI의 핵심 엔진이 되는 World Model은 단순히 언어나 이미지 데이터의 처리에서 나아가 AI가 물체의 움직임, 물리 법칙, 공간과 시간의 변화 등 현실 세계에 대한 구조적 이해와 인과관계를 이해 하도록 설계됨. 생성형 AI 시대의 인공지능 언어를 넘어 현실 세계 자체를 이해하도록 새로운 지능 체계가 주목받기도 함<sup>20)</sup>
- 특히 자율주행차 분야에서는 각종 센서를 통해 도로 상황을 확인하고 차량과 보행자의 움직임을 예측하면서 속도를 줄이거나 차선을 변경하는 등의 판단을 내리는 등의 물리적 행동이 연결될 수 있음. 예컨대 집안의 구조를 파악하여 경로를 스스로 찾거나 사람 대신 위험 지역을 순찰 하는 보안 로봇 등의 활용이 가능<sup>21)</sup>

18) 김중훈, 김의직, 김동완, "Physical AI의 최근 연구 동향과 발전 방향에 대한 연구", 『사물인터넷융복합논문지』 제11권 제4호, 한국사물인터넷학회, 2025, 108-110면.

19) Domain Randomization, SGF World Model 등은 의료 분야에서 적용되어 로봇 보조 수술(Robot-Assisted Surgery, RAS)에서 Sim-to-Real 기반 접근이 제안됨. 김중훈 외, 위의 연구, 111면.

20) IGLOO Corp., AI가 세상에 나온다, 피지컬 AI (Physical AI), IGLOO SECURITY & INTELLIGENCE(Jan.6, 2026), <https://www.igloo.co.kr/security-information/ai%EA%B0%80-%EC%84%B8%EC%83%81%EC%97%90-%EB%82%98%EC%98%A8%EB%8B%A4-%ED%94%BC%EC%A7%80%EC%BB%AC-aiphysical-ai/>

- Google DeepMind의 경우, 텍스트, 오디오, 비디오를 섭렵하여 ‘다중 모달 추론’을 통하여 복잡한 문제를 해결하는 Gemini 모델이 등장. Gemini 2.0을 기반으로 한 Gemini Robotics는 고급 시각 언어 행동 모델로서 Gemini의 체계적 추론 능력을 활용하여 프로그램을 실행할 수 있음. Gemini 2.0가 탑재된 차세대 휴머노이드 로봇은 훈련에서 경험하지 못했던 다양한 과제도 즉석에서 해결이 가능할 뿐 아니라 다양한 지시, 새로운 환경에서도 능숙<sup>22)</sup>

---

21) IGLOO Corp., 위의 기사 참조.

22) Gemini Robotics Team et al., Gemini Robotics: Bringing AI into the Physical World, arXiv:2503.20020v1 [cs.RO] (submitted March 25, 2025), <https://arxiv.org/abs/2503.20020v1> Physical AI에 대한 이론적 배경으로, Vahid Salehi, Fundamentals of Physical AI, arXiv:2511.09497 (submitted November 12, 2025), arXiv preprint, <https://arxiv.org/abs/2511.09497>

## 2. Physical AI 활용과 소비자 보호

### 2.1 기술의 발달 시대의 소비자

#### □ AI 시대의 소비자

- Physical AI의 활용 초기 단계로서 소비자 경험을 구체적으로 구현하는 학술 연구는 매우 제한적이므로 AI를 통한 소비자의 경험이 물리적, 감각적으로 소비자의 실제 행동 및 지각에 영향을 미치는 요소에 대한 연구로 Physical AI 시대의 소비자 보호를 추론하고자 함
- 디지털 플랫폼을 정기적으로 사용하여 관광 서비스를 계획, 탐색 또는 예약하는 18세~45세 사이의 이용자를 대상으로 한 연구에서 실증 분석 결과 AI는 인지적 경험, 정서적 경험, 사회적 경험과 같은 소비자 경험의 다양한 차원으로 통계적으로 유의미한 긍정적 영향을 미치는 것으로 분석됨.<sup>23)</sup>
- 실제로 챗봇은 부분적 매개 효과를 나타내어 정서적 경험과 신체적·감각적 경험을 향상시키는데 중요한 역할을 하고 증강현실은 가상 여행 경험을 제공하고 실제 여행 탐색을 시뮬레이션 하여 정서적 유대감을 강화하는 기능을 수행, 즉 AI의 지능적 기능은 상호작용 기술을 통하여 체험적 가치로 전환되고 그로 인하여 영향력을 발휘<sup>24)</sup>
- 인공지능이 발달하면서 소매업자가 서비스를 제공하는 소비 환경에 변화가 있었음. 첨단 기술 엔터테인먼트와 소매 쇼핑을 결합하여 인공지능 내장 혼합 현실(AI-embedded mixed reality, 이하 'MR') 전시에 대한 소비자 반응을 조사한 결과,

23) Manzar Ahmed & Nermin Gouhar, "The Effect of AI on Consumer Experience, Engagement on Social Media through the Mediating Role of Chatbots, Virtual Influencers, and Augmented Reality," *Scientific Culture*, Vol. 12, No. 2.1 (2026), pp.694-695.

24) Id. at 682.

증강 객체(Augmented Object)와 연관된 음성인식 및 머신러닝을 통한 합성의 인공지능의 품질은 소비자로 하여금 공간 몰입과 MR 즐거움, 소비자의 새로운 경험 인식과 연결된 MR 몰입도 높임을 보임. 이러한 요소들은 소비자의 참여를 높이고 구매 의도와 사회적 집단과 경험을 공유하려는 의도 등의 행동 반응에도 긍정적인 영향을 미치는 것으로 분석되었음<sup>25)</sup>

- 결국 인간과 컴퓨터의 상호작용을 통한 긍정적인 소비자 경험과 참여를 촉진하기 위하여 AI의 품질이 중요한 역할을 함을 밝힘. 이는 기술 강화 소매 환경에서는 AI의 품질이 행동 의도를 증가시키는 데 기여함이 입증됨<sup>26)</sup>

#### □ Physical AI의 활용과 우려

- Physical AI의 사례로서 ‘고객 서비스 로봇’의 인식된 위험; 생성형 AI의 활용례에 비해 Physical AI가 실제 일상생활에서 빈번하게 활용되는 범용 사례는 많지 않아 일부 특정 영역에서의 서비스 로봇의 사용과 그에 따른 위험의 인식, 정보 보안의 영향에 대한 이론적 검증이 존재하므로 이를 분석
- 소비자에게 인지된 위험(perceived risk)에 대해서는 고객 서비스 로봇의 사용 의향에 부정적 영향을 미치나 정보 보안(information security)에 대해서는 긍정적인 영향을 미치는 것으로 연구됨<sup>27)</sup>

25) Ahmed & Gouhar, supra at 678.

26) Eunyong (Christine) Sung, Sujin Bae, Dai-In Danny Han, and Ohbyung Kwon, “Consumer Engagement via Interactive Artificial Intelligence and Mixed Reality,” *International Journal of Information Management* 60 (October 2021): 102382, <https://doi.org/10.1016/j.ijinfomgt.2021.102382>

27) Abraham Pizam et al., “The Role of Perceived Risk and Information Security on Customers’ Acceptance of Service Robots in the Hotel Industry,” *International Journal of Hospitality Management* 117 (February 2024): 103641, <https://doi.org/10.1016/j.ijhm.2023.103641>.

- Covid-19 사태 이후 서비스 로봇은 인간 상호 간의 상호작용을 최소화하거나 대체하여 전염병을 줄일 수 있음이 확인되어 병원, 공항, 호텔 등의 다양한 환경에서 로봇이 채택됨<sup>28)</sup>
- AI 기술은 소비자의 취향과 선호를 매우 정교하게 예측하도록 진화되어 새로운 상품의 효율적 선택을 지원하겠지만 성별, 연령, 소득에 따른 개인화된 차별 문제와 민감정보의 생성 문제도 여전히 해결되지 못하고 있음<sup>29)</sup>
- 기존의 알고리즘 계약에 의한 소비자 권리 침해의 우려와 더불어 일상생활에서 가깝게 활용되고 보다 로봇의 자율성은 높아지며, 사용 부속의 full-stack에 따른 소비자 보호를 위한 법·정책 논의가 필요하므로 “인간 중심적 기술의 개발”과 이에 따른 소비자 위상의 제고가 필요한 시점<sup>30)</sup>

## 2.2 기술 발달과 규제에 대한 인식의 변화

### □ AI로 발생한 손해와 책임 분배 문제

- AI 기술의 발달은 사용자의 편의를 위하는 방법으로 진화할 수밖에 없는데 결국 AI의 자율성이 커지게 됨에 따라 이로 인한 손해의 발생에 대한 책임 공백의 문제가 대두됨<sup>31)</sup>
- 민사 책임의 분배의 측면에서, 그동안 AI는 내부 작동 과정을 설명하기

28) Id., section “Introduction.”

29) 이금노, “인공지능 알고리즘 기반 경제에서의 소비자 문제연구”, 정책연구 18-17, 한국소비자원 정책연구 보고서, 2018, 47면 이하.

30) 조은선, 나중연, “AI 쇼핑 환경에서의 소비자자율성: 소비자 경험과 인식을 중심으로”, 『소비자학연구』 제36권 제4호, 한국소비자학회, 2025, 170면.

31) 이종구, “인공지능과 제조물책임”, 『법학논집』 제49권 제4호, 단국대학교 법학연구소, 2025, 25면.

쉽지 않을 뿐만 아니라 개발자가 설계한 대로 출력되지도 않았기에 사후적으로 AI의 기대하지 않았던 결과값에 대한 원인을 역추적하는 데 대한 문제점이 지적되었음. 이는 손해를 책임질 당사자를 특정할 수 없는 원인으로 지목됨<sup>32)</sup>

- 일반적인 AI 규제 측면에서는 수평적 규제체계와 맥락특유적 규제 체계가 대립하고 있으나 일관성을 갖춘 규제를 위해서는 학습과 작업을 구분하여 용례(use cases)에 따라 맞춤형, 핀셋형으로 유형화하여 대응하는 것이 효율적. 이러한 맥락특유적 규제를 위하여 작업(tasks)의 단계에서 생성모델, 판별모델, 자율모델 등으로 범주화하여 규제 체계를 정비하고 이를 보완할 수 있는 인간 사용자의 판단, 인지, 행동을 정리하는 작업이 요구됨<sup>33)</sup>
- 특히 자율모델의 경우 자율주행차나 로봇 등 안전성이 요구되는 영역으로 이들을 대체하거나 보완하기 위한 인간 사용자(사람)의 활동을 규율하는 기존의 규제를 정비하여 대응하려는 노력이 필요<sup>34)</sup>

## □ 능력 위험과 규제의 접근방식 변화

- AI 기술 발달에 따른 규제의 접근 방식의 변화 - AI의 기술 발달은 인류에게 편의성을 가져옴으로써 전 지구적으로 엄청나게 빠른 확산을

32) AI의 특성에 따른 법적 문제에 대한 선행연구는 이종구, 앞의 연구, 25면; AI의 활용이 가속화 되면서 범용 모델과 생성모델은 2022년 ChatGPT 이후 대중화 되었으나 이에 대한 통제론도 함께 부각시키게 되었다. 박상철, “인공지능의 법적 규율 I: 범용모델과 생성모델”, 『서울대학교 법학』 제65권 제1호, 서울대학교 법학 연구소, 2024, 67면.

33) Park, Sangchul, “Bridging the Global Divide in AI Regulation: A Proposal for a Contextual, Coherent, and Commensurable Framework,” *Washington International Law Journal*, Vol. 33, No. 2 (2024), pp. 216-269. 다만 이 연구에 대한 인용은 SSRN의 출판 예정버전인 Park, Sangchul, Bridging the Global Divide in AI Regulation: A Proposal for a Contextual, Coherent, and Commensurable Framework (August 19, 2024). Available at SSRN: <https://ssrn.com/abstract=4950781> or <http://dx.doi.org/10.2139/ssrn.4950781> pp.36-41.

34) Id., at 8-9, 42-64.

가져옴. 동시에 빠른 기술의 발달로 인하여 초창기 법적 규제 측면에서는 통일된 규율 대상의 정의가 어려웠음.<sup>35)</sup> 인공지능이 다른 기술과는 달리 자율성을 가질 뿐 아니라 설계자의 의도와 다른 결과값을 가져올 수 있는 예측 불가능성, 불투명성 등의 AI로 인한 위험에 대한 규율 방식도 복잡다기하게 발달하였음<sup>36)</sup>

- 최근 발달 된 AI의 능력으로 볼 때, AI의 규제적 측면에서 AI 기술의 발달에 따라 AI의 잠재적 능력을 기반으로 상상할 수 있는 위험으로 파악할 수 있음. 이에 따라 능력 위험에 대한 법적인 규율은 AI의 고성능을 전제로 논의되기에 법제별로도 논의의 선상에 있으나 EU AI Act<sup>37)</sup>나 우리나라의 ‘인공지능기본법’<sup>38)</sup> 등이 그와 관련한 법제로 여겨짐<sup>39)</sup>

## □ 인공지능기본법의 시행과 영향

- EU AI Act의 경우, 고성능 AI 모델을 별도로 규제할 것을 제시, 모델의 평가, 중대 사고 및 시정조치의 보고, 사이버 보안 등의 규제로 구체화<sup>40)</sup>

35) AI로 인한 위험의 규율은 AI의 개념을 정의하는 것으로부터 시작되어야 하는데 AI와 AI 시스템에 대한 개념의 정의에 있어서 AI 시스템에 대한 정의가 보다 일반적이기에 본 글에서도 이를 따르되 이는 일반적인 개념에서의 AI를 지칭하는 것으로 한다. 학계에서 AI 시스템으로 그 개념을 정의하는 것은 OECD, Explanatory Memorandum on the updated OECD Definition of an AI System, 2024에서의 표현이며 선행 연구로는 이상용, “고성능 인공지능의 능력위험 규율-자율규제의 현황과 시사점-”, 『저스티스』 통권 제211호, 한국법학원, 2025, 53면 참조.

36) 이상용, 위의 연구, 53면.

37) 관보 게재 정식명칭은 Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence and amending Regulations (EC) No 300/2008, (EU) No 167/2013, (EU) No 168/2013, (EU) No 2018/858, (EU) 2018/1139 and (EU) 2019/2144 and Directives 2014/90/EU, (EU) 2016/797 and (EU) 2020/1828 (Artificial Intelligence Act) OJ L 1689, 12 July 2024. 이고, 이글에서는 ‘EU AI Act’로 칭한다.

38) 2024년 12월 26일 본회의를 통과하고, 2025년 1월 21일 공포되어 2026년 1월 22일 시행인 “인공지능 발전과 신뢰 기반 조성 등에 관한 기본법”

39) 이상용, 앞의 연구, 55면. 이는 범용 AI에 대한 규율을 포함하고 우리나라의 경우, 기본법의 형태로 만들어졌으나 최근에는 법적 규제와 대별되는 연성규범으로서 자율규제의 사례도 보고되고 있다. 이에 관한 선행 연구로 이승민, “온라인 플랫폼과 자율규제 논의의 기초”, 『경제규제와 법』 제15권 제2호, 서울대학교 공익산업법 센터, 2022.

40) 시스템적 리스크로서 “역내 시장에 중대한 영향을 미치고, 공중보건, 안전, 공공안전, 기본권, 사회 전반에 실제 또는 합리적으로 예견 가능한 가치사슬에 걸쳐 대규모로 전파될 수 있는 부정적 효과가 있는 범용 AI 모델의 고영향 성능에 특유한 리스크”(art. 3(65))로 정의, 박상철, 각주 31)의 연구, 각주 52) 및 88면.

- 인공지능기본법은 EU의 AI Act를 비롯한 비교법적 논의들을 참조하여 적용을 위한 개념과 의무를 규정하였으나 EU의 ‘인공지능 가치사슬 참여자 유형에 따른 차등화된 의무 부과’의 형식이 아닌 포괄적 의무를 규정한 점에서 근본적인 차이가 존재하고 ‘고영향 인공지능’과 관련한 각종 의무를 부과한 점이 특징이나 하위법령에 따라 규제의 정도 및 구체적인 모습이 정해짐<sup>41)</sup>
- EU AI Act를 참조하였음에도 입법과정 당시 AI 환경은 급격한 변화의 시기를 거치며 규제의 수준이 상당 부분 완화되어 진흥이 강조되는 양상을 보이며<sup>42)</sup> EU식 규제모델을 참조하면서도 필요 최소한의 규제 및 AI 진흥의 조화라는 제3의 길을 선택한 것으로 평가. 그럼에도 추상적인 규정과 포괄적인 조항의 존재도 지적됨<sup>43)</sup>
- AI 생태계에서의 한 축을 이루는 스타트업 기업들은 위험 관리 및 이용자 보호 등의 안전성 확보 조치를 마련해야 함, 정부도 AI 생성 영상 등 전반에 워터 마크 등 표시 의무 적용을 검토하였다가 업계 의견을 수용하기도 함<sup>44)</sup>

## □ AI와 책임

- 유럽집행위원회의 ‘인공지능책임지침안’에서의 AI의 도입과 위험에 대한 통제에 대한 제안의 이유는 과실책임주의 아래에서 AI의 특성으로 인한 자율성과 불투명성에 따른 손해배상청구의 어려움을 유럽

41) 최경진, “인공지능 기본법:제정 의의와 과제”, 『미래성장연구』 제11권 제1호, 고려대학교 미래성장연구원, 2025, 89면.

42) 국가 AI 발전과 신뢰 기반 조성을 위한 추진 체계 마련, AI 산업육성 지원, AI에 대한 안전·신뢰 기반 조성 등

43) 학계에서는 수평적·포괄적 규제에 대한 지양, 분야별 후속 법제 정비의 요구, 고영향 인공지능의 개념 정의 및 합리화 방안의 마련 요구 등 실효성 있는 규제 체계를 위한 지적들이 있었다. 한국인공지능법학회, “인공지능기본법 제정에 대한 의견서”, 2024. 참조.

44) 전자신문, “코 앞 다가온 ‘AI 기본법’ 시행 …기대와 우려 공존”, 2026. 1. 18. 입력, 최종 방문일 2026. 2. 20.

연합 차원에서 해결하기 위함. 대상과 범위에 대한 자세한 규정과 적용 배제 영역도 명확화하였음. 그러나 기존의 제조물책임지침과 디지털서비스법을 배제하며 증명수단의 공개와 입증책임을 새로이 규정하여 인공지능책임지침안의 지향점을 명확히 밝힘<sup>45)</sup>

- 우리 법에서도 생성형 인공지능을 활용한 의료기기의 경우 안전성과 유효성 평가에 대한 “생성형 인공지능 의료기기 허가·심사 가이드라인”을 제정·발간, 전(全) 주기 위험 요인 분석 허가심사 시 고려해야 할 사항의 마련<sup>46)</sup>, 각 분야별 구체적 가이드라인 필요
- 인공지능기본법 제2조 제4호 및 제33조에서 “사람의 생명, 신체의 안전 및 기본권에 중대한 영향을 미치거나 위험을 초래할 우려” 및 특정 영역에 대하여 고영향 인공지능의 기준 및 예시에 관한 가이드라인을 수립하고 보급하도록 과기정통부장관에게 역할이 부여되고, 고영향 인공지능사업자에게는 일정한 책무가 부과됨
- 고영향 인공지능사업자의 경우 ‘신뢰성 및 안전성 확보 조치’가 이행 의무로서 부과되는데 이에 대한 구체적인 사항은 고시 및 가이드라인에서 구체화<sup>47)</sup>

45) 특히 EU에서는 디지털 제조물로 발생한 손해와 그에 대한 책임 공백에 대처를 위하여 2024년 10월 23일 제조물책임지침개정안이 통과된 바 있다. 이는 과학기술의 발전에 대응하면서도 기존의 제조물책임지침을 개정하는 방식의 2022년 지침안이 제안되었고 2022년 지침안 제6조에서는 “결함의 존재 유무”의 판단에 대하여 사유를 구체적으로 규정하였다. 2022년 9월 28일 유럽집행위원회는 인공지능법과 함께 인공지능 책임지침안을 제안한 바 있다. 아울러 제조물책임지침에 대한 분석과 인공지능책임지침안에 대한 선행연구로 박신욱, “유럽 제조물책임지침 개정으로부터의 시사”, 『민사법학』 제109호, 한국민사법학회, 2024. 287면, 295면 이하. 아울러 인공지능의 일반적 특징에 대한 입증 곤란에 대한 분석은 이종구, 앞의 연구, 32면 이하.

46) 식품의약품안전처, 2025년 1월 24일.

47) 과학교술정보통신부, 「AI기본법 하위법령집」, 2025.9. 323면.

### 3. AI 기술과 미래

#### □ AI 시대의 정책의 역할

- AI의 발달로 인하여 거버넌스 전략에 따른 정부의 혁신 방식이 AI의 활용에 따른 위험의 선제적 대응을 주도할 가능성이 커짐
- AI 기술 발달에 따른 모니터링, 규제 유연화 등의 전략은 통합적 시행으로 효과를 발휘할 수 있고 신뢰할 수 있는 AI 생태계의 조성이 필요한 시점, 유연한 거버넌스와 더불어 구체적이고 명확한 가이드라인이 있어야 기술 발달의 발목을 잡지 않음<sup>48)</sup>
- 국가 AI 위원회의 법·제도 분과의 하위에 법조계와 학계뿐만 아니라 산업계 목소리의 반영을 제안
- 소비자의 욕구를 반영하되 소비자 중심의 법제를 위하여 기술의 발달에 대하여 ‘소비자의 의사결정’이 반영되도록 해석할 필요<sup>49)</sup>
- 새로운 AI 기술의 발달을 만들어 내는 기업들의 경우에도 자율규제를 통하여 재앙적 위험(catastrophic risks)에 대응하려는 노력 중, Anthropic의 경우 ‘책임 있는 스케일링 정책(responsible scaling policy, RSP)’ 등을 발표<sup>50)</sup> 기업의 자구노력이 의미를 가질 수 있도록 공동체에서도 이를 장려하는 분위기가 조성될 필요

48) NIA 한국지능정보원, 「AI.GOV 해외동향 2025-1호」, 19면.

49) 소비자법의 설계 및 규율 체계의 해석 등에서 소비자 중심의 법 해석이 되도록 소비자법의 기본 전제를 복기할 시점이다. 같은 취지로 정신동, “인공지능을 이용한 소비자계약의 체결”, 『외법논집』 제49권 제3호, 한국외국어대학교 법학연구소, 2025, 234면.

50) 이상용, 각주 35)의 연구, 56면. 이상용 교수의 연구에 따르면 Anthropic의 AI 모델에서 ASL-2 표준이 적용되더라도 위험 수준에 맞추어 안전조치가 강화되고 뿐만 아니라, 배포와 보안이 구별된다. 이상용, 같은 연구 58면 각주 28) 참조.

- 단순히 AI 사업자로 하여금 책무를 부과하고 이행했는지 묻기보다, 소비자 또한 AI의 기술 발달에 따른 편익의 향유에 걸맞는 윤리의식을 가져야 안전한 이용을 담보할 수 있을 것으로 기대
- 소비자 보호를 위해서는 법제에 대한 연구와 함께 ‘소비자 자율성’에 대한 개념과 범위에 대한 연구도 필요, 단순히 AI의 자율성 증대만을 두려워 할 것이 아니라, 소비자의 선택의 궁극적인 자율성과 독립성에 있음을 인지하고 소비자 스스로 필요성을 인식하여 의사결정이 이루어지는 상황이 되도록 AI 사용 환경에 대한 지속적 정책 연구 필요<sup>51)</sup>

---

51) 안전한 AI 사회의 구현을 위해서 ‘인간 중심적 개발’의 중요성 부각, 조윤선 외, 앞의 연구, 170면.; Serena B. Rinaldo, “Consumer Independence, Vulnerability and Public Policy: The Case of Free Matter for the Blind,” *Journal of Consumer Affairs*, Vol. 46, No. 1 (2012), pp. 107-119.

## [ 참고 문헌 ]

- 김종훈· 김의직· 김동완, “Physical AI의 최근 연구 동향과 발전 방향에 대한 연구.” 『사물인터넷융복합논문지』 제11권 제4호, 한국사물인터넷학회, 2025, 108-111면.
- 박상철, 「인공지능의 법적 규율 I: 범용모델과 생성모델」, 『서울대학교 법학』 제65권 제1호, 서울대학교 법학연구소, 2024.
- 이금노, “인공지능 알고리즘 기반 경제에서의 소비자 문제연구.” 정책연구 18-17, 한국소비자원, 2018.
- 이상용, “고성능 인공지능의 능력 위험 규율—자율규제의 현황과 시사점—.” 『저스티스』 통권 제211호, 한국법학원, 2025.
- 이승민, “온라인 플랫폼과 자율규제 논의의 기초.” 『경제규제와 법』 제15권 제2호, 서울대학교 공익산업법센터, 2022.
- 이종구, “인공지능과 제조물책임.” 『법학논집』 제49권 제4호, 단국대학교 법학연구소, 2025.
- 정신동, “인공지능을 이용한 소비자계약의 체결.” 『외법논집』 제49권 제3호, 한국외국어대학교 법학연구소, 2025.
- 조은선· 나중연, “AI 쇼핑 환경에서의 소비자자율성: 소비자 경험과 인식을 중심으로.” 『소비자학연구』 제36권 제4호, 한국소비자학회, 2025.
- 최경진, “인공지능 기본법: 제정 의의와 과제.” 『미래성장연구』 제11권 제1호, 고려대학교 미래성장연구원, 2025.
- 과학기술정보통신부, 「AI기본법 하위법령집」, 2025. 9.
- 식품의약품안전처, 「생성형 인공지능 의료기기 허가·심사 가이드라인」, 2025. 1.

- 한국인공지능법학회, 「인공지능기본법 제정에 대한 의견서」, 2024.
- 한국지능정보사회진흥원(NIA), 「AI.GOV 해외동향 2025-1호」, 2025.
- Ahmed, Manzar, and Nermin Gouhar, “The Effect of AI on Consumer Experience, Engagement on Social Media through the Mediating Role of Chatbots, Virtual Influencers, and Augmented Reality.” *Scientific Culture* 12, no. 2.1 (2026): 678-698.
- Flynn, Kerry, “‘ChatGPT Moment for Physical AI’: Nvidia CEO Launches New AI Models and Chips.” *Axios*, January 5, 2026.
- Firoozi, Roya, Johnathan Tucker, Mac Schwager, et al., “Foundation Models in Robotics: Applications, Challenges, and the Future.” *The International Journal of Robotics Research* 44, no. 5 (2024): 701-739.
- Huang, Madison, “What Is NVIDIA’s Three-Computer Solution for Robotics?” *NVIDIA Blog*, August 8, 2025.
- Park, Sangchul, “Bridging the Global Divide in AI Regulation: A Proposal for Contextual, Coherent, and Commensurable Framework.” *Washington International Law Journal* 33, no. 2 (2024): 216-269.
- Pizam, Abraham, et al., “The Role of Perceived Risk and Information Security on Customers’ Acceptance of Service Robots in the Hotel Industry.” *International Journal of Hospitality Management* 117 (2024): 103641.
- Rinaldo, Serena B., “Consumer Independence, Vulnerability and Public Policy: The Case of Free Matter for the Blind.” *Journal of Consumer Affairs* 46, no. 1 (2012): 107-119.
- Salehi, Vahid, “Fundamentals of Physical AI.” *arXiv:2511.09497*, 2025.

- Sung, Eunyong (Christine), Sujin Bae, Dai-In Danny Han, and Ohbyung Kwon, “Consumer Engagement via Interactive Artificial Intelligence and Mixed Reality.” *International Journal of Information Management* 60 (2021): 102382.
- VerWey, John, et al., “AgiBot World Colosseo: A Large-Scale Manipulation Platform for Scalable and Intelligent Embodied Systems.” arXiv:2503.06669v3, 2025.
- Wolf, Theresa, and Andrea Willige, “Physical AI Is Changing Manufacturing—Here’s What the Era of Intelligent Robotics Looks Like.” World Economic Forum, September 9, 2025.
- Consumer Technology Association, “The Future Feels Closer Than Ever: Day 3 of CES 2026.” CES, January 9, 2026.
- CNBC, “Morgan Stanley Says Humanoid Robots Will Be a \$5 Trillion Market by 2050.” April 29, 2025.
- NVIDIA, “Autonomous Vehicle Safety: From the Cloud to the Car — NVIDIA Halos.” Accessed February 2026.
- NVIDIA, “NVIDIA Kicks Off the Next Generation of AI With Rubin.” NVIDIA Newsroom, January 5, 2026.
- NVIDIA, “NVIDIA Releases New Physical AI Models as Global Partners Unveil Next-Generation Robots.” NVIDIA Newsroom, January 5, 2026.
- International Organization for Standardization, ISO 8373: Robots and Robotic Devices—Vocabulary, 2nd ed. Geneva: ISO, 2012.

\*본 연구의 내용은 연구자 개인의 견해이며 공식 견해가 아닙니다.