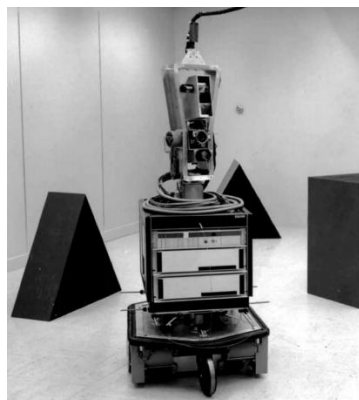


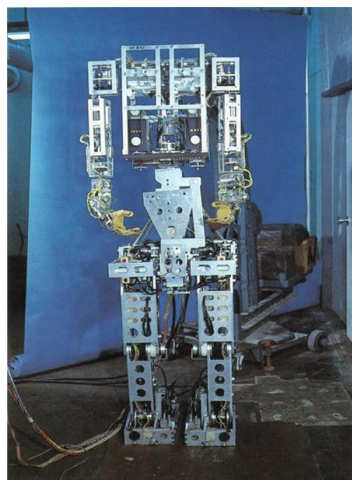
What is embodied artificial intelligence and why it matters to ITU?

Kai WEI,
Director of Artificial Intelligence Institute, CAICT
10 Oct 2025

1. What is embodied AI?
2. Challenges of embodied AI
3. Standard gap and ITU-T's role



1961 Shakey



1973 WABOT-1



1999 AIBO



2000 ASIMO



Embodied AI

- Sensory & Execution
- Cognition
- Collaborating
- Learning

Smart Robot

- Sensory & Execution

Robot

- Programmed

Humanoid Robot

- Robot with body, head and limbs
- looking and moving like a human

Embodied AI is a transformative shift in AI that integrates AI into physical systems, enabling them to interact with and learn from their environment through sensory inputs and actions.

Aspect	(Non-embodied) AI (ChatGPT, Image Recognition program)	Embodied AI (Robots, Self-driving Cars)
Form of Existence	"A Brain in a Vat": Pure software, residing on servers.	"An Embodied Brain": Software integrated with a physical/virtual body.
Primary Goal	Process information in a digital space; passively waits for input (e.g., a prompt).	Take action in a physical environment; acts proactively based on the situation.
Input/Output	Data (text, images, speech)/Data (generated text, classification labels)	Perception (sensor data, cameras, LiDAR, touch)/Action (movement, manipulation)
Learning Paradigm	Learns from static datasets: Trained by analyzing massive existing collections of images and text.	Learns through interaction & experience: Learns like an infant, through trial and error, touching, observing, and failing.

Embodied AI= AI + Physical Body(Action)

Embodied AI enables robots from automation to autonomy, allowing them to performs behaviours or tasks independently, learn, and adapt to changing environments.

Perspective	Robotics	Embodied AI
1. From "Programming" to "Learning"	All behaviors are based on explicit code written by programmers.	Abilities are acquired through real-time environmental interaction (trial-and-error) .
2. From "Perception for Action" to "Perception-Cognition-Action"	Sensory data is used directly to execute predefined actions. e.g., verifying workpiece position, performing force-controlled assembly.	Sensory data is used to construct an internal model of the world, (e.g., "This is a messy room") and object properties, enabling actions driven by cognition .
3. From "Closed World" to "Open World"	The environment and all variables (e.g., object location, lighting, process) must be predictable and pre-defined.e.g., verifying workpiece position, performing force-controlled assembly.	Its purpose is to cope with environments, and must handle unseen objects, adapt to dynamic changes, and deal with ambiguity .

Embodied AI=Robots + Brain(Cognition)

What is embodied AI?

Embodied AI refers to systems that integrate **cognitive abilities with sensory and action capabilities**, enabling machines to perceive, reason, and act autonomously in dynamic environments.

*IEEE Systems, Man, and Cybernetics Society
Announcement: Establishment of the Technical Committee on
Embodied AI Systems*

IEEE SMC

Embodied AI refers to the **integration of artificial intelligence into physical systems**, enabling them to interact with the physical world. These systems can include general-purpose robots, humanoid robots, autonomous vehicles (AVs), and even factories and warehouse facilities.

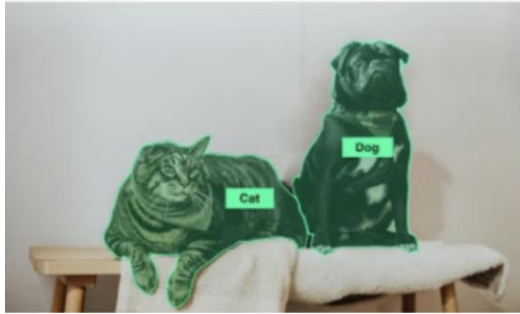
*NVIDIA Glossary
<https://www.nvidia.com/en-us/glossary/embodied-ai/>*

NVIDIA

Robotics for tasks such as **navigation, manipulation**, and instruction following.....**any tangible and intelligent machine** that moves through space is a form of embodied AI, whether it is a self-driving car, a robot vacuum, or a robotic arm in the factory.

*Searching for Computer Vision North Stars
<https://direct.mit.edu/daed/article/151/2/85/110602/Searching-for-Computer-Vision-North-Stars>*

Stanford Institute for HCAI



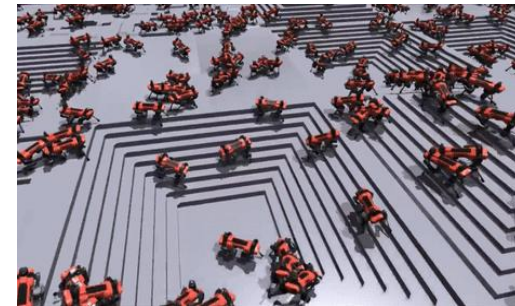
Recognition



1.Cognition



Programming



2.Learning

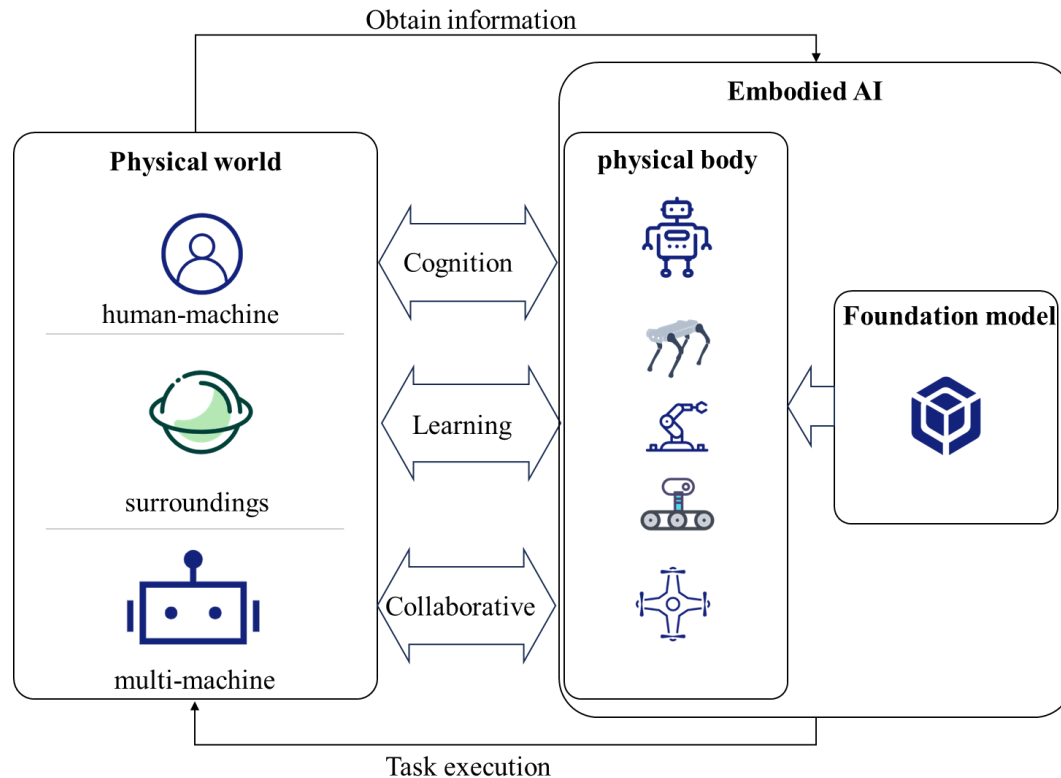


Isolated



3.Collaborative

Conceptual diagram and key features



Definition

Embodied artificial intelligence(EAI): Research, development, and implementation of artificial intelligence integrated into physical systems that interact autonomously with and adapt to the physical world.

ITU-T F.RF-EAI “Requirements and framework for embodied artificial intelligence systems”



Growing Market

\$ 38 Billion

by 2035^[1]

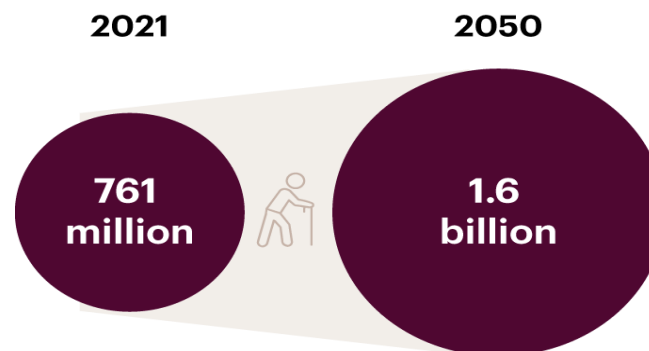
1 Billion in Use

by 2050^[2]

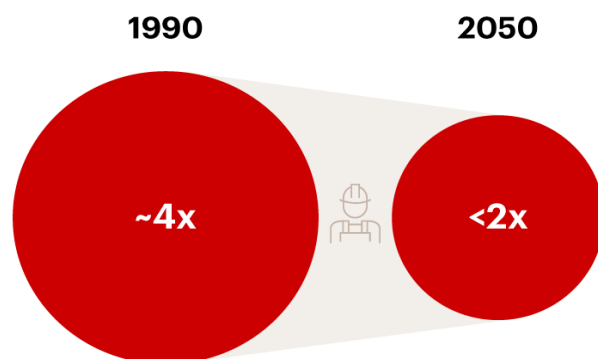
— only for humanoid robots

Societal Needs^[3]

Elderly Care Crisis



Labor Shortage



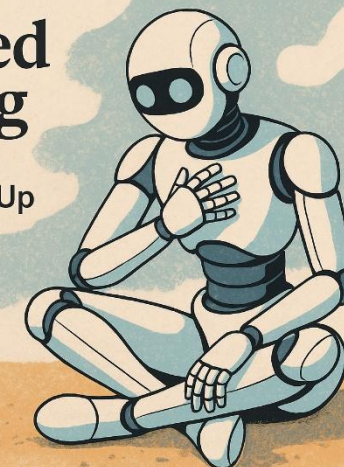
The way towards AGI

读万卷书，行万里路

Read ten thousand books, travel
ten thousand miles.

Embodied Learning

Rethinking AGI
from the Ground Up



[1]Goldman Sachs, Jan, 2024, Global Automation: Humanoid Robots – The AI Accelerant

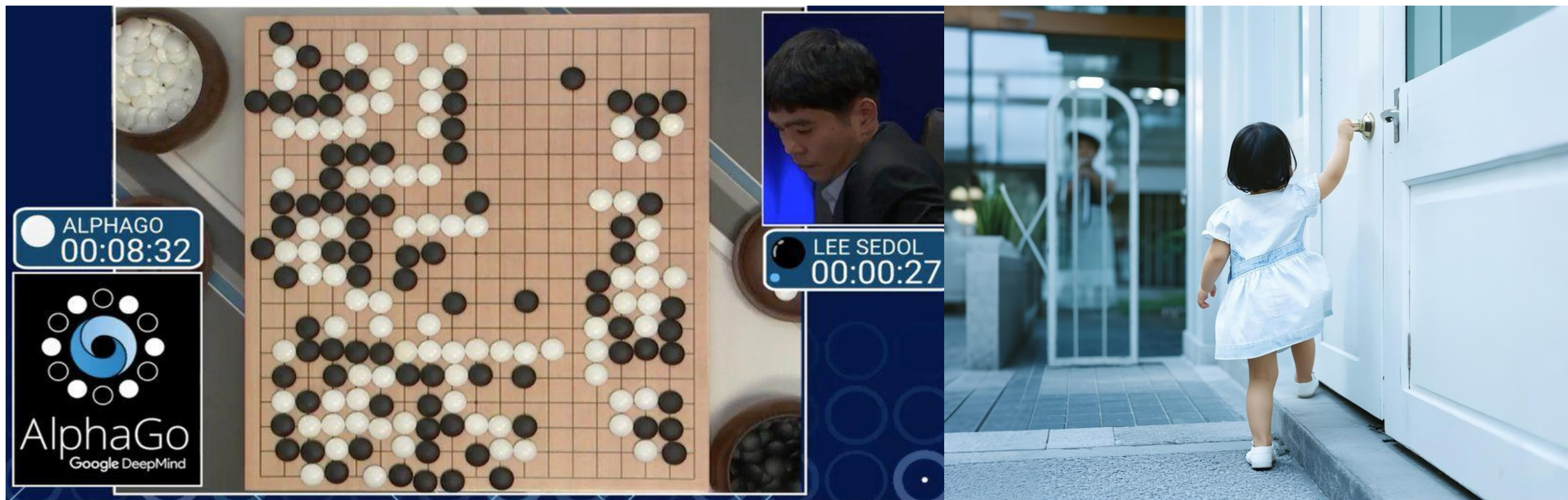
[2]Morgan Stanley, May 2025, Humanoids: A \$5 Trillion Market

[3]UN Department of Economic and Social Affairs & International Labor Organization

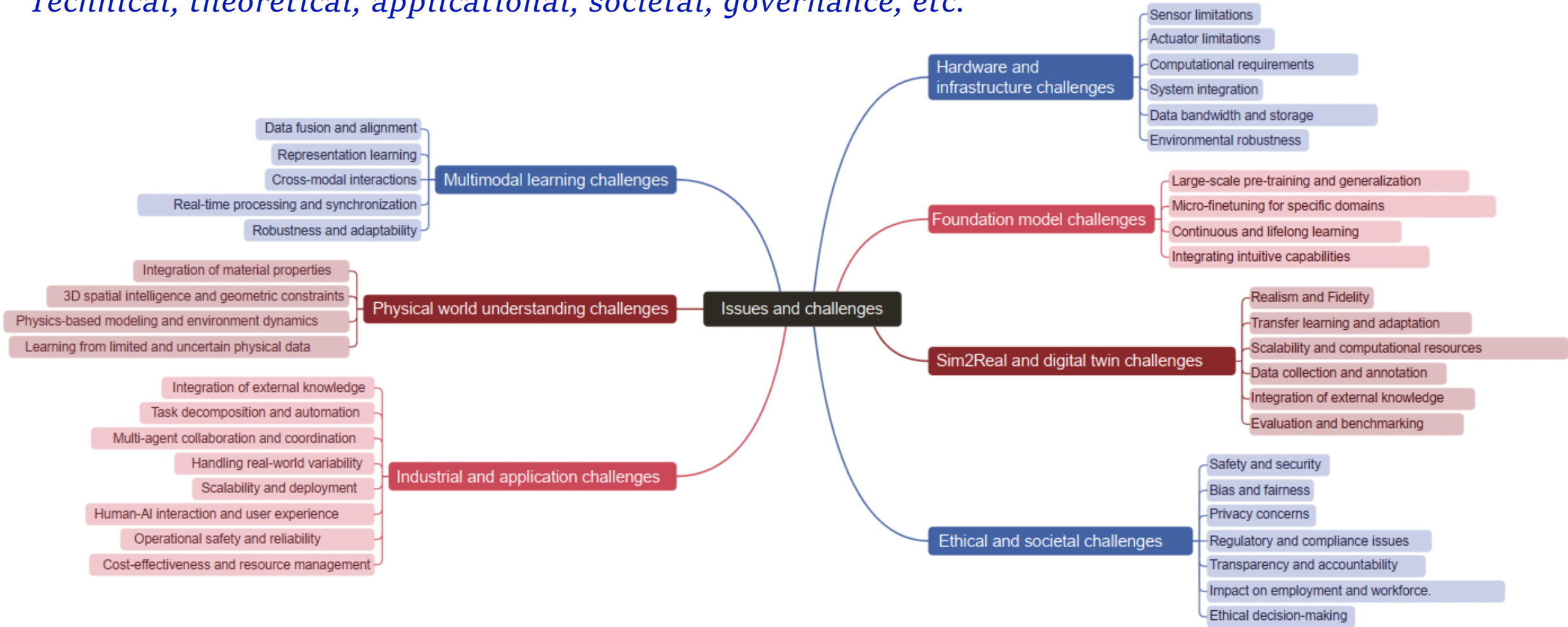
1. What is Embodied AI?
2. Challenges of embodied AI
3. Standard gap and ITU-T's role

“It is comparatively easy to make computers exhibit adult level performance on intelligence tests or playing checkers, and difficult or impossible to give them the skills of a one-year-old when it comes to perception and mobility.”

—Hans Moravec, 1988



Technical, theoretical, applicational, societal, governance, etc.

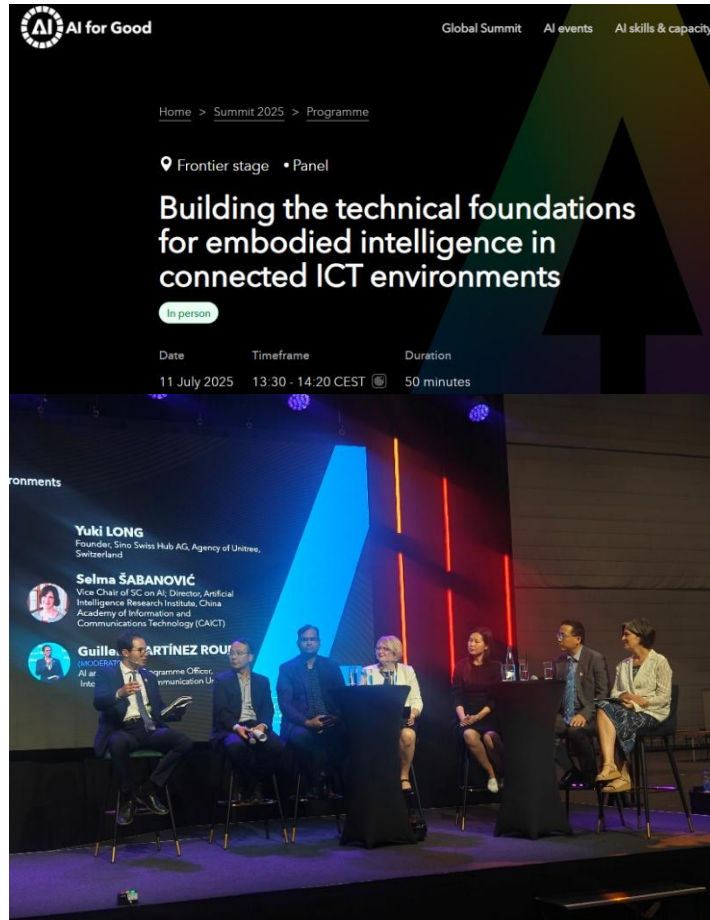


Sources: Sun F, Chen R, Ji T, et al. A comprehensive survey on embodied intelligence: Advancements, challenges, and future perspectives[J]. CAAI Artificial Intelligence Research, 2024, 3: 9150042.

1. What is Embodied AI?
2. Challenges of embodied AI
3. Standard gap and ITU-T's role

Dimension	Key Challenges	Standardization Directions
1. Data & Perception	• Heterogeneous multimodal and multimedia data • Lack of high-quality, task-specific datasets	• Format and quality standard for Multimodal training data • Benchmark dataset & annotation specifications
2. Network & Collaboration	• Complex task allocation across cloud-edge-device • Lack of common multi-agent communication protocols	• Unified interfaces for collaborative computing • Standardized real-time agent interaction protocols
3. Model & Evaluation	• No unified benchmarks for embodied models	• Assessment criteria and metrics • Standardized performance benchmarks
4. System Integration	• High complexity in integrating heterogeneous components • Hard interoperability between vendors	• Modular development frameworks • Plug-and-play component & middleware standards
5. Human-Robot Interaction	• Lack of safety protocols for physical HRI • Inconsistent user interfaces and experiences	• Safety standards for physical interaction • Human-centered design guidelines
6. Safety & Security	• Data privacy and protection risks • Lack of algorithmic transparency and explainability	• End-to-end security & privacy frameworks • Standards for explainable EAI & audit trails
7. Ethics & Governance	• Ambiguity in accountability and liability • Risk of value misalignment	• Accountability & liability frameworks • Ethical guidelines & compliance standards

COMMUNITY



USE CASE STUDY



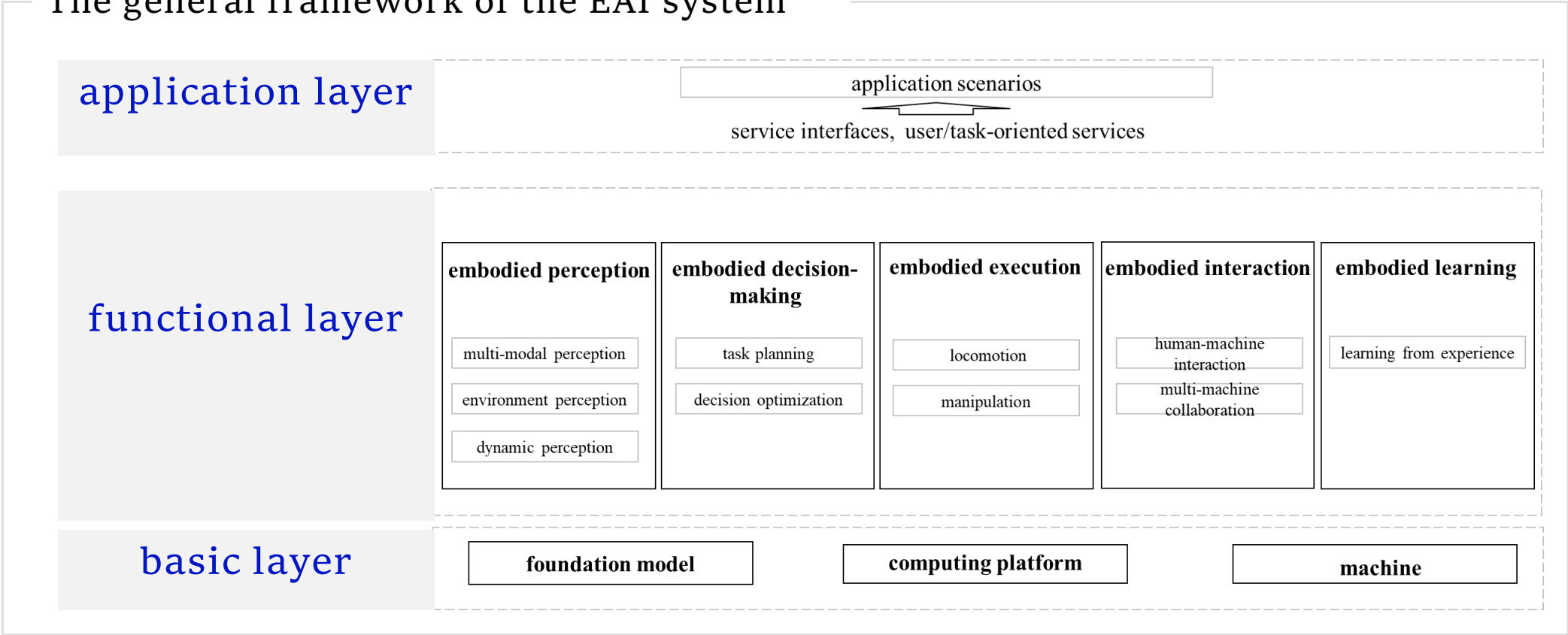
STANDLIZATION

...standards are essential in translating the values we agree on as a global community into real-world systems.

Standards should not be construed as constraints on innovation.

Draf Rec. ITU-T F RF-EAI “Requirements and framework for embodied artificial intelligence systems, provides a unified framework for system capability building.

The general framework of the EAI system



- Embodied AI represents a promising pathway toward AGI that goes beyond the current domains of AI and robotics, with the potential to greatly benefit humanity.
- Moravec's paradox is still validate , EAI faces significant challenges, some of which can be effectively addressed through standardization.
- The ITU plays a key role in this effort, with its Study Group 21 serving as an open, collaborative platform actively advancing the field.

Looking foreword, we recommend continuing workshop discussions or to establishing a Focus Group in the future, to explore the opportunities for "EAI for Good“, and to foster standardization studies, inclusively.



AI for Good

Thanks

weikai@caict.ac.cn