

Integrating Embodied AI with Short Range Communication

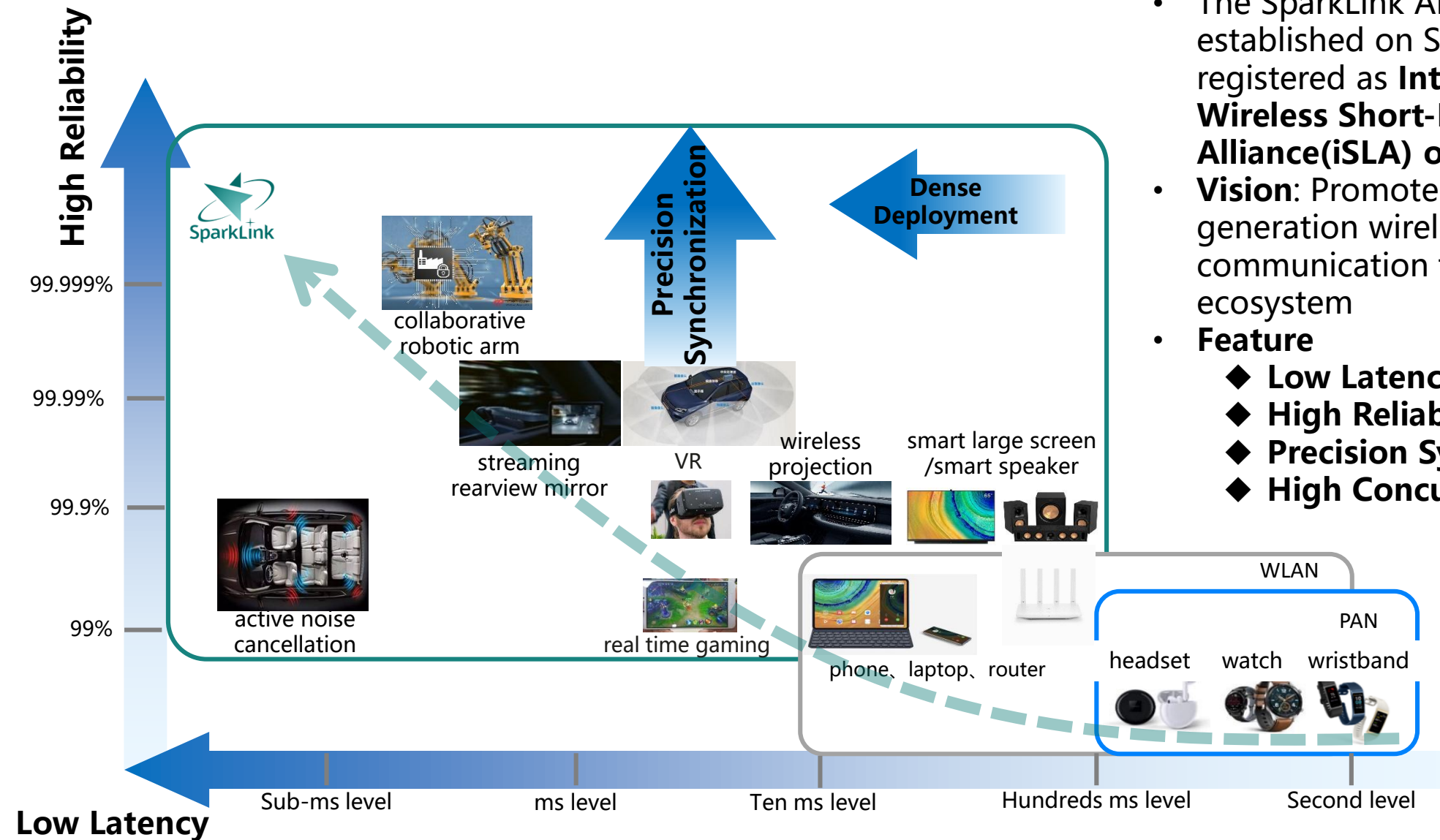
-Challenges and Standardization Opportunities

Hang Liu, on behalf of iSLA

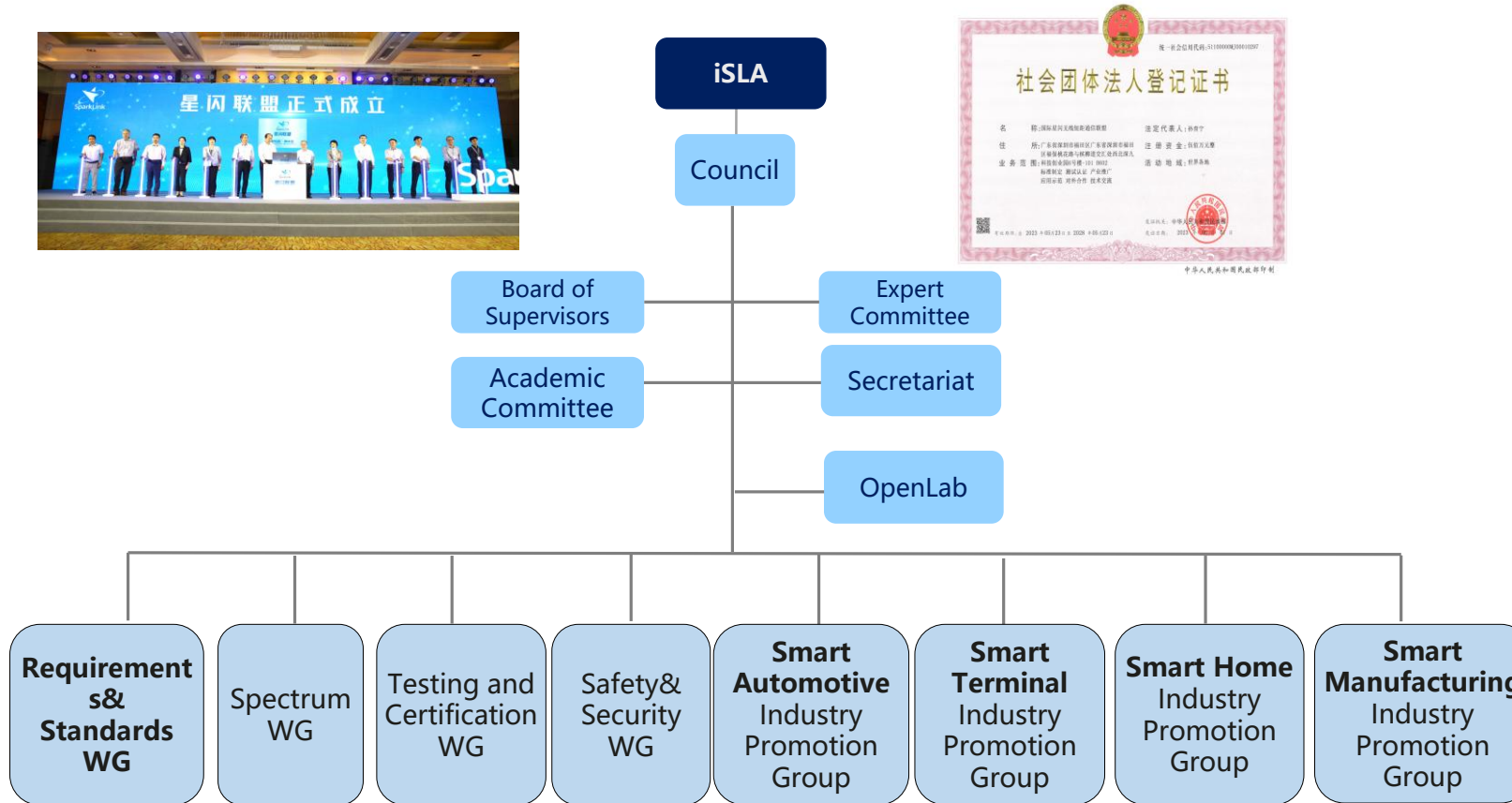
10th Oct, 2025

// Introduction of iSLA

- The SparkLink Alliance was officially established on Sep,2020, and further registered as **International SparkLink Wireless Short-Range Communication Alliance(iSLA) on 2023 in China**
- Vision:** Promote innovation in next generation wireless short-range communication technology and industrial ecosystem
- Feature**
 - ◆ **Low Latency**
 - ◆ **High Reliability**
 - ◆ **Precision Synchronization**
 - ◆ **High Concurrency**



// Organizational Structure



- Requirements & Standards WG
 - ◆ responsible for the formulation of standards
- Spectrum WG
 - ◆ addresses the spectrum requirements , promoting spectrum testing and allocation
- Testing and Certification WG
 - ◆ in charge of testing laboratories, authorization, and certification
- Safety & Security WG
 - ◆ Safety and Security Scenarios and Requirements for SparkLink applications
- Four Industry Promotion Group
 - ◆ Smart Automotive
 - ◆ Smart Terminal
 - ◆ Smart Home
 - ◆ Smart Manufacturing

// Integration embodied AI with short range communications



AI for Network

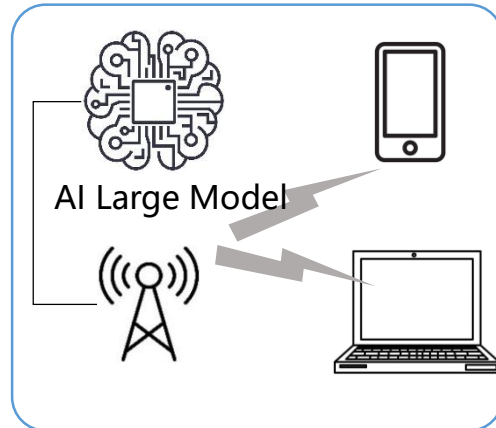
- Enhance communication systems and network intelligence via AI
- Engineering Issues



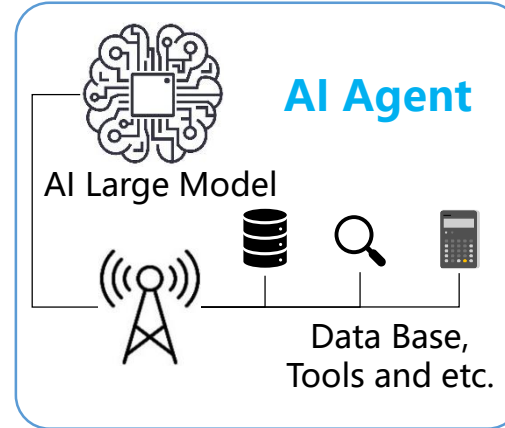
Network for AI

- Improve communication technologies and network architecture to meet the diverse technical demands of AI services

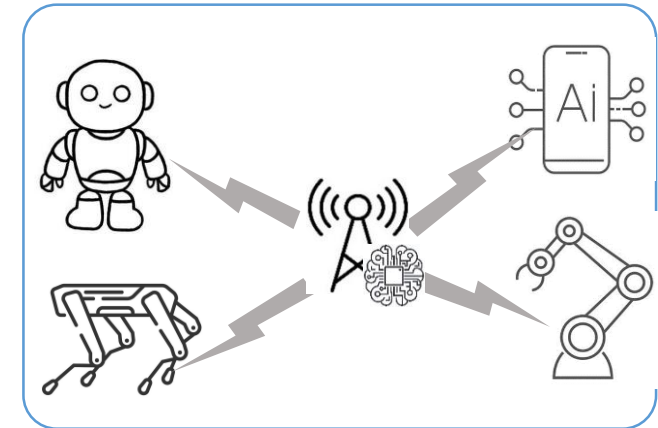
Human-AI Large Model Interaction



Internal Communication of AI Agents



Collaborative Communication among AI Agents



Communication Scenarios

Challenges of multimodal information interaction of embodied AI



Smart Terminal

Google Pixel 9

- Real-time voice interaction: voice chat and Video chat
- Call assistant: Direct My Call, Hold for Me, Call Screen;
- **Multimodal Perception: Gemini Live: Simultaneous input of images, voice, and text; supports screen sharing and video input**

Apple iOS 18.1

- Real-time voice interaction
- Writing assistance: Text style definition, text polishing, native app invocation (such as **email, memo, message**), cross-device capability
- Image retrieval and editing: Semantic image retrieval, editing

Smart Glass

Android XR

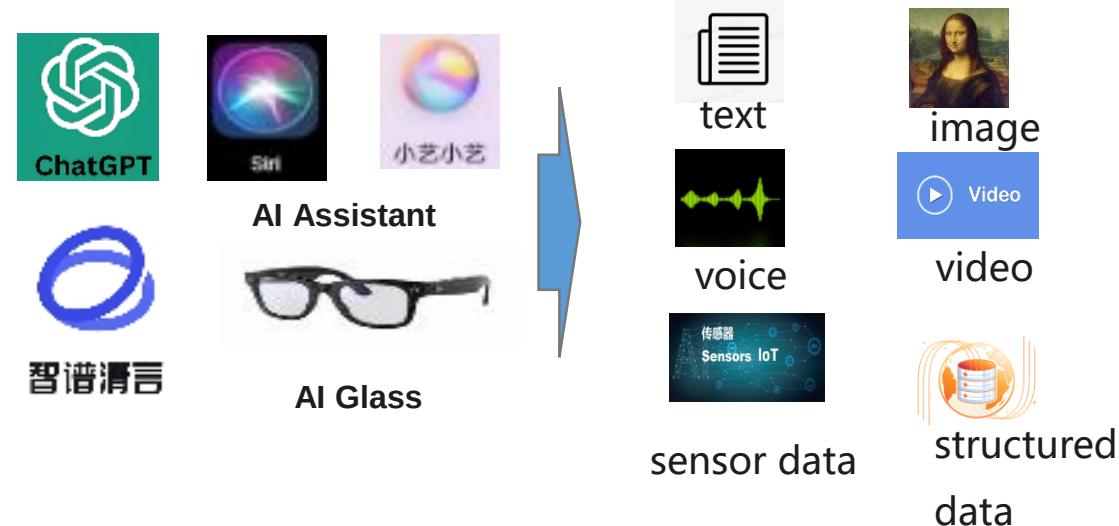
- Visual understanding, knowledge search, agent execution; photo taking, information inquiry and sending, schedule management, real-time translation;

AliQuark

- Alipay Scan-to-Pay: Scan QR codes with glasses, voiceprint authentication for payment;
- Real-time updates on flight and train schedules; Near-eye display navigation.

Source from internet

From text-based conversations to multimodal information interaction



Source from internet

- **How can we achieve unified and flexible encapsulation?**
 - Traditional streaming media encapsulation cannot meet the demands of AI services
 - **more diverse data types**
 - **mixed transmission** of burst and periodic data
 - dynamic changes in services.

// Challenges of multimodal information interaction of embodied AI



- Perception is **task-oriented perception**, not merely simple information transmission;
- Example: Traditional **streaming media** transmission protocols **are not suitable for voice transmission of embodied AI** (which requires high reliability since it is voice command) and also struggle to accommodate burst transmission demands.



The perception module of embodied AI

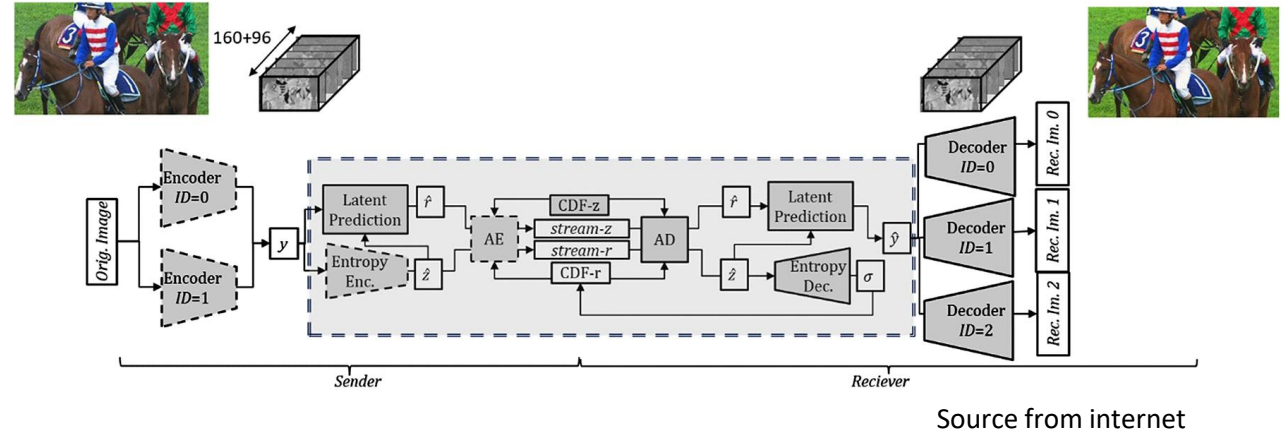
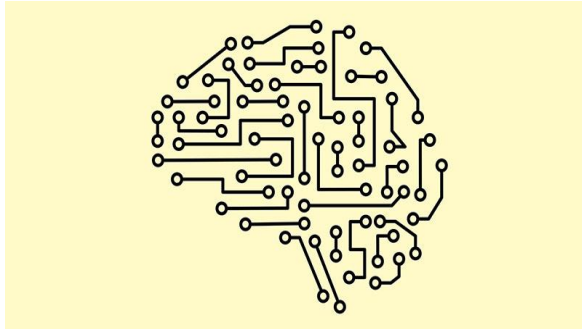
- responsible for gathering information from the external world
- converting it into internal representations that can be **understood and processed**.

Importance:

- Understanding the environment: Identifying **objects, events, and states**.
- Understanding **instructions**: deciphering **user intentions**.
- Driving decision-making: providing the data necessary for **decision-making**.

- Data from different modalities exhibit complex relationships such as **dependency, substitution, and complementarity**.
- How to achieve flexible and unified encapsulation to **meet task requirements**?
- How to define the **Qos of each modalities** considering the task requirements?

Challenges of multimodal information interaction of embodied AI



- There is also a vast amount of user-defined AI-based data compression and representation methods.
- This shall also be supported by a unified capsule framework.



ICS 35.030
CS 1.80



中华人民共和国国家标准

GB 45438—2025

网络安全技术
人工智能生成内容标识方法

Cybersecurity technology—
Labeling method for content generated by artificial intelligence

- Comply with national regulatory requirements for AI-generated content
 - Labelling requirements

Project	Content
Name	《AI Multimodal Information encapsulation and Transmission Optimization》
Period	Aug,2025~Dec,2026
Scope	• Define an AI multimodal information encapsulation protocol and optimize the transmission of multimodal information; • Specifically, it supports flexible and unified encapsulation of various types of modal information (text, audio, images, video, sensor data and etc.); • it supports common multimodal information compression formats as well as AI-based multimodal information compression formats; • it meets regulatory requirements for the identification of AI-generated content; • furthermore, it optimizes the transmission of multimodal information (especially text information), effectively reducing transmission overhead and improving transmission efficiency while ensuring service quality.

感谢聆听!
THANKS!

2025/10/8



国际星闪无线短距通信联盟
International SparkLink Alliance