# “Intelligent Multimedia”

## WHAT’S MISSING?

*WANG Yuntao （Mr.）*
*Rapporteur of Q5/16 AI enabled-multimedia applications*
*Vice Chair of ITU-T Focus Group on metaverse (FG-MV)*
*Artificial Intelligence Research Institute*
*China Academy of Information and Communication Technology*
*No.52 Huayuanbei Road, Beijing, P.R.China*
*Mobile: +86-18611547086*
*Email: wangyuntao@caict.ac.cn*

# Traditional Multimedia

- Definitions

Multimedia is content that uses a combination of different content forms such as text, audio, images, animations, video and interactive content.

--Wikipedia

## Intelligent Multimedia

No Clear Descriptions yet, but

- Easy access to extensive, searchable archives of mixed text, graphics, sounds, narrations, and video footage
- More Human-friendly interactions
- More than Just content consumers, deep mining of multimedia data
- Not only human, but we want machines to understand multimedia as well

……

# Identifying Missing link…


Natural Language Processing
- Machine Translation
- Automatic Abstracting
- Automatic Generation

……

Computer and Machine Vision
- Face recognition
- Object detection

……

Human-Machine Interface
- Speech Interaction
- Brain-computer interface

……

Intelligent Speech
- Speech recognition
- Speech Synthesis
- Question Answering

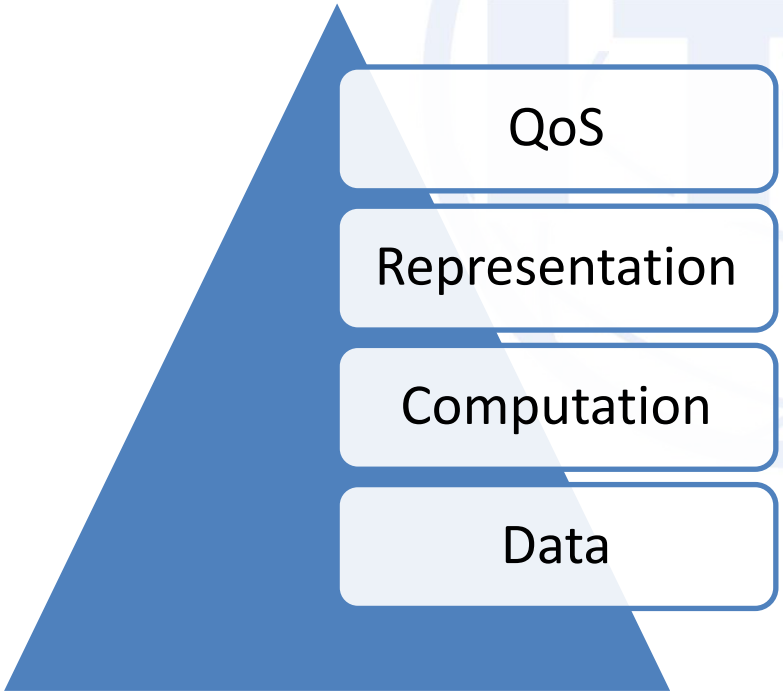
……

Computer and Machine Vision
- Content Audition
- Automatic Pilot

……

We should we do?
Figure out the framework

Applications are booming, we need to identify the common technical barriers behind all these applications, and figure out the Intelligence Enablers.




- New Intelligent QoS requirements
- More Human-friendly requirements
- More mining and analyzing tasks
- New requirements of data preparation

We should we do?
Data: Data preparation

As the gasoline of modern AI industry, data labelling has brought new requirements and challenges.


DATA LABELLING

| Data collection | Data labelling | Data Delivery | Data quality control |
|---|---|---|---|

We should we do?
Computation: System impact

In-depth data mining and analyzing tasks brings new technical demands.
Deep learning is transforming how we design computers -- Jeff Dean

Multimedia Architecture → Intelligent Multimedia Architecture

System Point of View

Representation | Network design | Algorithm | Optimization

We should we do?
Representation: Coding

To facilitate intelligent data mining, new frame structures are proposed.


**Example:**

*SVAC*

Surveillance video and audio coding

Defines new data analysis descriptions:

- Rules for image analysis;
- Object detection;
- Feature analysis;
- Object/Behavior recognition;
- Statistics for objects counting

We should we do?
QoS

New QoS metrics and assessment methodology are required to evaluate the intelligent part.

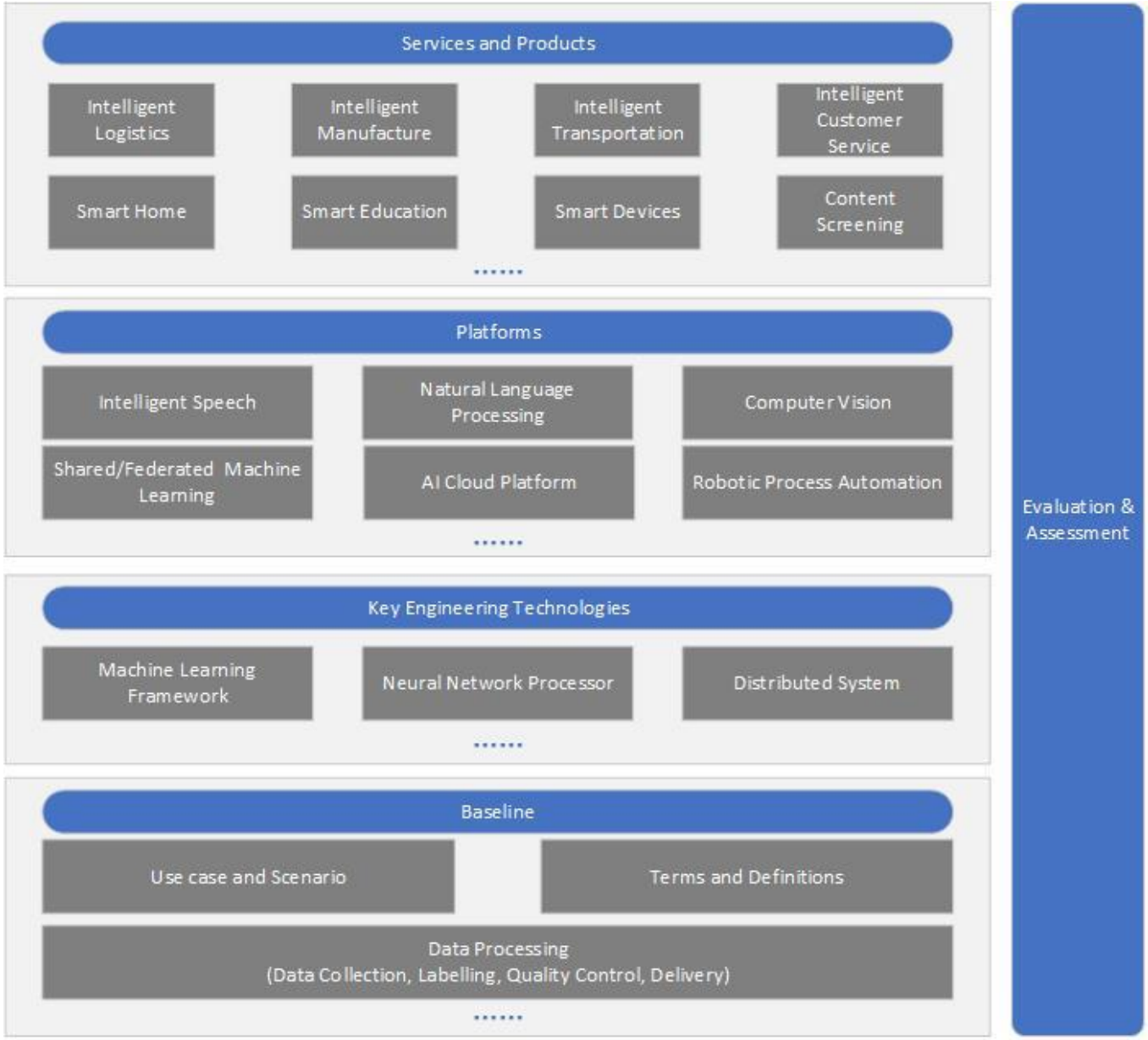
Multimedia QoS → Intelligent Multimedia QoS

- How good is the video quality?
- How good is the compression ratio?

……

- How intelligent is the robot?
- How good is the speech recognition?

……


Services and Products
Intelligent Logistics
Intelligent Manufacture
Intelligent Transportation
Intelligent Customer Service
Smart Home
Smart Education
Smart Devices
Content Screening
.....
Platforms
Intelligent Speech
Natural Language Processing
Computer Vision
Shared/Federated Machine Learning
AI Cloud Platform
Robotic Process Automation
.....
Key Engineering Technologies
Machine Learning Framework
Neural Network Processor
Distributed System
.....
Baseline
Use case and Scenario
Terms and Definitions
Data Processing
(Data Collection, Labelling, Quality Control, Delivery)
.....
Evaluation & Assessment

- Briefing of Q5/16 AI enabled-multimedia applications
  - 56 ongoing work items
  - 28 published recommendations

| Work item | Question | Status | Timing | Approval process | Subject / Title |
|---|---|---|---|---|---|
| F.742.1 (ex F.SCAI) | Q5/16 | Approved 2022-12-14 | 2022 | AAP | Requirements for smart class based on artificial intelligence |
| F.746.13 (ex F.IMCS) | Q5/16 | Approved 2022-03-16 | 2022 | AAP | Requirements for smart speaker based intelligent multimedia communication |
| F.746.15 (ex F.SBNG) | Q5/16 | Approved 2022-12-14 | 2022 | AAP | Requirements for smart broadband network gateway in multimedia content |
| F.746.16 (ex F.AI-ILICSS) | Q5/16 | Approved 2022-12-14 | 2022 | AAP | Technical requirements and evaluation methods of intelligent levels of intelligent customer service systems |
| F.747.11 (ex F.AI-ISD) | Q5/16 | Approved 2022-12-14 | 2022 | AAP | Requirements for intelligent surface-defect detection service in industrial |
| F.747.12 (ex F.AI-MVSLWS) | Q5/16 | Approved 2022-12-14 | 2022 | AAP | Requirements for artificial intelligence based machine vision system in smart logistics warehouse |
| F.748.14 (ex F.DH-2D) | Q5/16 | Approved 2022-03-16 | Jan-22 | AAP | Requirements and evaluation methods of non-interactive 2D real-person digital human application systems |
| F.748.15 (ex F.DH-FM) | Q5/16 | Approved 2022-03-16 | Jan-22 | AAP | Framework and metrics for digital human application systems |
| F.748.17 (ex F.AICP-MD) | Q5/16 | Approved 2022-12-14 | 2022 | AAP | Technical specification for artificial intelligence cloud platform: AI model development |
| F.748.18 (ex F.AI-DLEMT) | Q5/16 | Approved 2022-12-14 | 2022 | AAP | Metric and evaluation methods for AI-enabled multimedia application computing power benchmark |
| F.748.19 (ex F.AI-FASD) | Q5/16 | Approved 2022-12-14 | 2022 | AAP | Framework for audio structuralizing based on deep neural network |
| F.748.20 (ex F.AI-DMPC) | Q5/16 | Approved 2022-12-14 | 2022 | AAP | Technical framework for deep neural network model partition and collaborative execution |
| F.748.21 (ex F.FDIS) | Q5/16 | Approved 2022-12-14 | 2022 | AAP | Requirements and framework for feature-based distributed intelligent |
| F.748.22 (ex H.FDISarch) | Q5/16 | Approved 2023-09-13 | Jul-23 | AAP | Functional architecture for feature-based distributed intelligent systems |
| F.748.23 (ex F.ML-ICSMIReqs) | Q5/16 | Approved 2024-04-26 | Feb-24 | TAP | Requirements and framework for intelligent crowdsensing multimedia interaction based on deep learning |
| F.748.24 (ex F.TCEF-FML) | Q5/16 | Approved 2024-04-15 | Feb-24 | TAP | Trusted contribution evaluation framework on federated machine |
| F.748.25 (ex F.AI-SCS) | Q5/16 | Approved 2023-09-13 | Jul-23 | AAP | Requirements for speech interaction of intelligent customer services |
| F.748.26 (ex F.AI-CPP) | Q5/16 | Approved 2024-02-13 | Dec-23 | AAP | Technical specification for artificial intelligence cloud platforms: Performance evaluation |
| F.748.34 (ex F.AI-MKGDS) | Q5/16 | Approved 2024-06-13 | 2024 | AAP | Requirements for the construction of multimedia knowledge graph database structure based on artificial intelligence |
| F.748.35 (ex F.FML-TS-FR) | Q5/16 | Approved 2024-06-13 | 2024 | AAP | Requirement and framework of trustworthy federated machine learning |
| F.748.36 (ex F.MAS) | Q5/16 | Approved 2024-06-13 | 2024 | AAP | Requirements and framework of multi-algorithm scheduling systems |
| F.748.37 (ex F.JSQSUDAC) | Q5/16 | Approved 2024-06-13 | 2024 | AAP | Requirements and functional framework of joint semantic query system of unstructured data across clusters |
| F.748.38 (ex F.AICP-GA) | Q5/16 | Approved 2024-06-13 | 2024 | AAP | Technical specification for artificial intelligence cloud platform: General architecture |
| F.748.39 (ex F.AICP-FRRC) | Q5/16 | Determined 2024-04-26 | 2024 | TAP | Functional requirements and reference architecture of artificial intelligence cloud platform for smart grid operation and maintenance |
| F.748.11 (ex F.AI-DLPB) | Q5/16 | Approved 2020-08-13 | 2021 | AAP | Metrics and evaluation methods for deep neural network processor |
| F.748.12 (ex F.AI-DLFE) | Q5/16 | Approved 2021-06-13 | 2021 | AAP | Deep learning software framework evaluation methodology |
| F.748.13 (ex F.AI-MLTF) | Q5/16 | Approved 2021-06-13 | 2021 | AAP | Technical framework for shared machine learning system |
| F.Sup4 (ex F.Supp-OCAIB) | Q5/16 | Agreed 2021-04-30 | 2021 | Agreement | Overview of convergence of artificial intelligence and blockchain |

# ITU-T AI related work

## SG13

- address the requirements, architectures, functional capabilities and application programming interfaces of converged future networks. Key areas of focus include network softwarization and orchestration, information-centric networking, content-centric networking, and the application of machine learning technologies.

## SG5

- lead study group on electromagnetic field (EMF), environment, climate action, sustainable digitalization, and the circular economy.
- Guidelines on the environmental efficiency of machine learning processes in supply chain management

## SG12

- expert group responsible for the development of international standards (ITU-T Recommendations) on performance, quality of service (QoS) and quality of experience (QoE). This work spans the full spectrum of terminals, networks and services, ranging from speech over fixed circuit-switched networks to multimedia applications over mobile and packet-based networks.

## SG9

- carries out studies on the use of telecommunication systems in the distribution of television and sound programs supporting advanced capabilities such as ultra-high definition and 3D TV. This work also covers the use of cable and hybrid networks – primarily designed for the distribution of television and sound programs to the home – as integrated broadband networks to provide interactive voice, video and data services, including Internet access.

## SG17

- work to build confidence and security in the use of information and communication technologies (ICTs) continues to intensify in order to facilitate more secure network infrastructure, services and applications. The ITU-T SG17 coordinates security-related work across all ITU-T Study Groups, often working in cooperation with other standards development organizations (SDOs) and various ICT industry consortia.

## SG20

- develops international standards (ITU-T Recommendations) providing commonly agreed guidance for implementing the Internet of Things (IoT) and its applications, as well as smart cities and communities. Its work supports digital transformation in both urban and rural areas enabled by solutions in fields such as IoT, digital twins, and artificial intelligence.

## SG11

- responsible for signalling, producing international standards (ITU T Recommendations) that define how telephone calls and other ICT services are handled in the network.

# Thank you!

**WANG Yuntao**（Mr.）
Deputy Chief Engineer
Artificial Intelligence Research Institute
China Academy of Information and Communication Technology
No.52 Huayuan Bei Road, Beijing, P.R.China
Mobile: +86-18611547086
Email: wangyuntao@caict.ac.cn, wangyuntao812@gmail.com