IOWN GLOBAL FORUM

AI-Integrated Communications Use Case Interim Report

Classification: APPROVED REFERENCE DOCUMENT

Confidentiality: PUBLIC

Version 1.0

[AIC Use Case]

Legal

Contents

List of Figures

List of Tables

0. Preface

The document is provided as an initial work on prospective use cases being considered at the Innovative Optical and Wireless Network Global Forum (IOWN GF). Purpose of the release is to indicate the progress being made and show the technical feasibility study initiated within the GF.

The document is an interim version and will be updated through the GF activities. Final and complete version is to be released at the end of December 2021.

NOTE: The adopters of the systems described in this report should operate the systems cautiously so as not to infringe any human rights. IOWN GF recognizes that these systems should not be operated without technologies that support human rights.

1. Introduction

The world today has experienced faster than ever growth, thanks to advancements in communication and computing technologies. Moving forward, another quantum leap in computing and communication capabilities is expected to empower the world toward a new era of growth. The mission of IOWN GF is to develop fundamental technologies on communication, computing, data, and energy efficiency that would bring in quantum leap performance improvement and enable a much smarter world with advanced applications, including those with digital twin computing.

The main purpose of this document is to provide prominent future-looking use cases to identify application specific service requirements benefiting users from different vertical industries. The collection of requirements will be used to break down into technical requirements for feasibility study.

Use cases are categorized into human centric applications that deals to assist and enhance means for communication (AI Integrated Communications: AIC), and smart city applications that aims for autonomy that goes beyond human capability (Cyber Physical Systems: CPS). Two documents are created to capture the different objectives for each category, and to define distinct set of requirements [IOWN GF, 2020].

2. Scope

The document intends to provide key feature set, and requirements for use cases that can be categorized into AI-Integrated Communication.

Initial focus will be put into the following categories

- Entertainment
- Remote Operation
- Navigation
- Human Augmentation

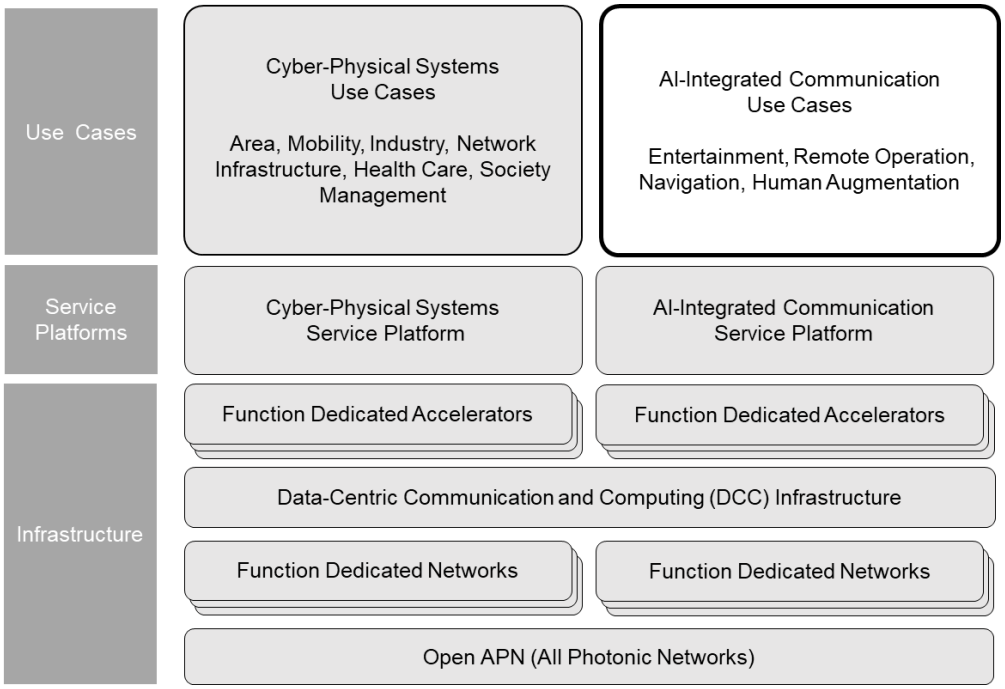

Figure 1 IOWN GF – Work Item Structure

3. Use Case Details

3.1. Entertainment

IOWN GF will pioneer new ways to deliver immersive entertainment regardless of the physical location of the user. In this section, new use cases on the future of music, sports and game entertainment are described.

3.1.1. Interactive Live Music

3.1.1.1. Description

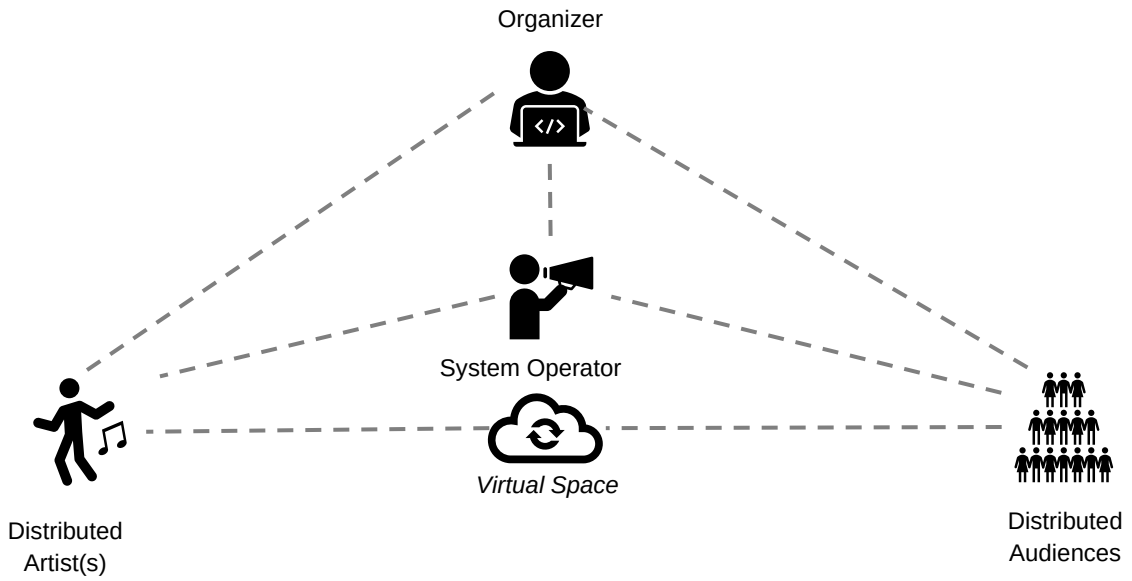

Figure 2 Overview of Interactive Live Music

IOWN will revolutionize how audiences enjoy live music performances. The audiences in their own homes or at a karaoke box will have an immersive experience as if they were at a concert stadium/hall, viewing the artist in high quality audio and video, from any seat of their liking. The distributed audiences will experience a sense of unity with the other audiences, by cheering or singing along together to their favorite artist. The distributed artists will also have an unprecedented experience in interacting more intimately with individual fans or even collaborate live with a remote artist at a distant location.

IOWN GF, with its AI Integrated Network, will also change how the live entertainment is conducted. The system operator will be able to analyze real-time the reaction of the distributed audience and create a virtual stadium/hall with customized sound and visual effects that fits the atmosphere of the crowd. The system operator will also be able to keep track of the QoS delivered to each audience, and dynamically allocate resource as needed. The organizer will be able to collect/analyze the big data gathered during the event to efficiently plan for future events.

The following Persona appear in this use case;

- Audience: People who are remotely watching and participating a live music performance. Up to around 20,000,000 people.

- Artist(s): Person or group performing at the live music performance. Distributed Artists may be remotely located from one another.
- System Operator: Person or group operating the system of the live music performance.
- Organizer: Person or group organizing the live music performance.

3.1.1.2. Key Feature Set

3.1.1.2.1. Persona #1: Audience

As an Audience, I want to;

- Interact with the Artist
    1. View the live performance with high fidelity360-degree free perspective video and audio
    2. Sing, sway, and dance along with the artist
    3. Respond to the artist to root and cheer

so that I can have an immersive experience as if I were at the concert stadium/hall.

- Interact with other Audiences
    1. Cheer together and share excitement with the other Audiences and monitor how I am seen from the other audience to blend in as crowd
    2. View the performance anywhere with multiple devices-e.g., TV, Screen, tablet, smartphone, HMD as that I can change location and switch devices accordingly

so that I can share exciting experiences with other audiences.

3.1.1.2.2. Persona #2: Artist

As an Artist, I want to;

- Interact with the Audience
    1. See, hear, and feel the Audience behavior
    2. Request for cheer, dance, sing

so that I can be as if I am at the stage at crowd filled stadium/hall.

- Interact with other Artist(s)
    1. Sing and play along with other members/performers
    2. Dance and sway along with other members/performers

so that I can feel as if we are on the same stage.

- Interact with the virtual space
- Interact with the System Operator
    1. Adjust feedback from the audience-e.g., volume, view size, etc.
    2. Pick selected audience to be highlighted

so that I can give a feel of exclusiveness to audiences

3.1.1.2.3. Persona #3: System Operator

As a System Operator, I want to;

- Monitor the state of the Artist
  1. Location and direction in respect to camera/mic position to select and create effects—e.g., ROI(Region Of Interest), lightning, etc. both manually and automatically.
  2. Video/Audio quality provided toward the artist to adjust and adapt operation mode
- Monitor the state of the Audiences
  1. Engagement state of audience to select and provide toward the artist the feedback of excitement at the audience
  2. Video/Audio/Data quality provided to audience to adapt operation mode
  3. Video/Audio/Data quality provided toward the artist to select appropriate feed back toward the artist

  so that I can provide the best performance to audiences within the range of affordable costs.
- Select the Artist scene to be used
- Select the Audiences' feedback to be used
- Create a virtual space where artists and audience co-exist and interact with overlaying real and artificial images so that I can give an immersive experience to both artists and audiences.
- Cancel the Artists' and the Audiences' echo so that I can highlight artists and certain audiences.

3.1.1.2.4. Persona #4: Organizer

As an Organizer, I want to;

- Collect behavior of the Artist during the event
  1. Interaction toward the audiences
  2. Interaction toward the other artists
- Collect the behavior of the Audiences during the event
  1. Interaction toward the artist
  2. Interaction toward the other audiences
  3. Other Interaction over the network
- Analyze the collected data

  so that I can enhance and reflect toward next events

3.1.1.3. Service Gap/Requirements

- Immersive virtual space creation: Bandwidth/latency not enough to collect images, sound, and sensor data from artist and audience. Computational power limited to create immersive virtual environment in real-time fashion.
- Immersive virtual space distribution: Bandwidth/latency not enough to distribute personalized virtual space data toward large scale distributed audiences. Also, means to synchronize images, sound and sensor data from Artist and Audience within the required latency lacking.
- Data Analysis platform: No sufficient means to collect data in consolidated manner that including Artist to Audience, Audience to Artist, Audience to Audience. Data transaction toward cloud (virtual space) and data transaction made in P2P both need to be scoped.

Service requirement can be summarized as follows;

- Number of Audiences: ~20,000,000
- Raw Data Rate at the Artist: 90 ~ 230 Gbps
  - Note: the data rate depends on the emulated distance from the eye to the object (1m@120fps: 230 Gbps, 3m@90fps: 90 Gbps)
- Latency
  - Interactive/synchronous components: ~10msec
  - Synchronous Sound: ~20msec
  - Non-verbal communication (e.g. waving hands): ~100msec
  - Verbal Communication: ~300msec
- End-to-end data delivery QoS (includes wireless network)
- Data synchronization within the required latency

3.1.2. Interactive Live Sports

3.1.2.1. Description

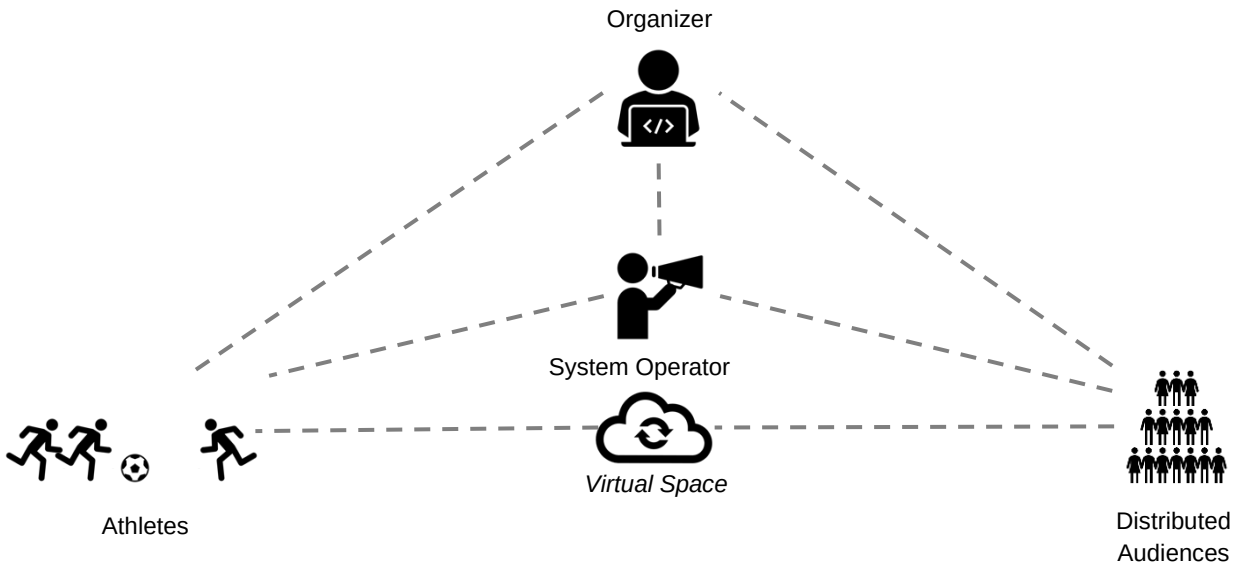

Figure 3 Overview of Interactive Live Sports

IOWN GF will change how distributed audiences view live sports events. The distributed audiences in their own homes, at a sports bar, or at a public screening venue will have an immersive experience as if they were at a stadium/arena, viewing the athletes' performance in high quality video and audio, from any seat of their liking. The distributed audiences will experience a sense of unity with the other audiences, by chanting together for their favorite team or cheer together in joy at a goal scene. The athletes will also have an unprecedented experience in interacting more intimately with individual audience. IOWN GF, with its AI Integrated Network, will also change how live sports event is

conducted. The system operator will be able to analyze real-time the reaction of the distributed audience and create a virtual stadium/arena with customized visual and sound effects that fits the atmosphere of the crowd. The system operator will also be able to keep track of the QoS delivered to each audience, and dynamically allocate resource as needed. The organizer will be able to collect/analyze the big data gathered during the event to efficiently plan for future events.

The following Persona appear in this use case;

- Audience: People who are remotely watching and participating a live sports event.
- Athlete(s): Person or group performing at the live sports event.
- System Operator: Person or group operating the system of the live sports event.
- Organizer: Person or group organizing the live sports event.

3.1.2.2. Key Feature Set

3.1.2.2.1. Persona #1: Audience

As an Audience, I want to;

- Interact with the Athlete(s)
    1. View the Athlete(s) with high fedelity360-degree free perspective video and audio
    2. Playback a highlight (e.g. a goal scene) of the Athletes' performance and Athletes' view
    3. Send chant or cheer to the Athlete(s)
    4. Applaud to the athlete(s) good performance (e.g. at a goal scene)

so that I can have an immersive user experience to be provided as if I am at the stadium/arena.

- Interact with other Audiences
    1. Cheer together and share excitement with other Audiences and monitor how I am seen from the other audience to blend in as crowd
    2. View the performance anywhere with multiple devices-e.g., TV, Screen, tablet, smartphone, HMD as that I can change location and switch devices accordingly

so that I can share exciting experiences with other Audiences.

3.1.2.2.2. Persona #2: Athlete(s)

As an Athlete(s), I want to;

- Interact with the Audience
    1. See, hear, and feel the Audiences' cheer
    2. Request for chant, cheer

so that I can hear audiences' cheering and booing me to feel Audiences' enthusiasm.

3.1.2.2.3. Persona #3: System Operator

As an System Operator, I want to;

- Monitor the state of the Athlete(s)
  1. Location and direction in respect to camera/mic position to select and create effects—e.g., ROI, lightning, etc.
  2. Video/Audio quality provided toward the Athlete(s) to adjust and adapt operation mode

  so that I can provide the best performance to Audiences within the range of affordable costs.
- Monitor the state of Audiences
  1. Engagement state of the Audiences to select and provide toward the Athlete(s) the feedback of excitement at the Audience
  2. Video/Audio/Data quality provided to Audience to adapt operation mode
  3. Video/Audio/Data quality provided toward the Athlete(s) to select appropriate feed back toward the Athletes(s)
- Select the Athlete'(s) highlight scene to be used
- Select the Audiences' feedback to be used
- Create a virtual space where the Athlete(s) and Audience co-exist and interact with overlaying real and artificial images

so that I can give an immersive experience to both athletes and audiences.

3.1.2.2.4. Persona #4: Organizer

As an Organizer, I want to;

- Collect behavior of the Athlete(s) during the sports event
- Collect the behavior of the Audience(s) during the event
  1. Interaction toward the Athlete(s)
  2. Interaction toward the other Audiences
  3. Other Interaction over the network
- Analyze the collected data

so that I can enhance and reflect toward next events

3.1.2.3. Service Gap/Requirements

- Immersive virtual space creation: Bandwidth/latency not enough to collect images, sound, and sensor data from Athlete(s) and Audiences. Computational power limited to create immersive virtual environment sin real-time fashion.
- Immersive virtual space distribution: Bandwidth/latency not enough to distribute personalized virtual space data toward large scale distributed Audiences. Also, means to synchronize images, sound and sensor data from Athlete(s) and Audience within the required latency lacking.
- Data Analysis platform: No sufficient means to collect data in consolidated manner that including Athlete(s) to Audience, Audience to Athlete(s), Audience to Audience. Data transaction toward cloud (virtual space) and data transaction made in P2P both need to be scoped.

Service requirement can be summarized as follows;

- Number of Audience: ~1,000,000,000 [FIFA, 2018].
- Raw Data Rate at Athlete(s): 90 ~ 230 Gbps
  - Note: the data rate depends on the emulated distance from the eye to the object (1m@120fps: 230 Gbps, 3m@90fps: 90 Gbps)
- Latency
  - Interactive/synchronous components: ~10msec
  - Synchronous Sound: ~20msec
  - Non-verbal communication (e.g. waving hands): ~100msec
  - Verbal Communication: ~300msec
- End-to-end data delivery QoS (includes wireless network)

3.1.3. Cloud Gaming

3.1.3.1. Description

In the video game industry, cloud gaming is a game-provision model in which a network and server generate game images and distribute them to user terminals. Cloud gaming has the following advantages: i) the users can play high-spec games on inexpensive terminals, ii) distribution costs including game software updates can be reduced, and iii) it is easy to manage the game data of multiple terminals belonging to a user. However, cloud gaming incurs additional delays (denoted as 1) to 4) in *Figure 4* compared to the terminal-processing game model that generates game video on user equipment such as a game console, PC, or smartphone. For this reason, it is currently difficult to play all games on cloud gaming system with the same quality as consoles.

IOWN will greatly improve the performance of cloud gaming and dramatically change the user experience and delivery model of the game. Game users will be able to play all games immediately without an expensive console or PC. Furthermore, eSports organizers will be able to easily hold fair eSports events at multiple locations with IOWN cloud gaming that is sufficiently low latency for eSports players.

The following Persona appear in this use case as illustrated in *Figure 5*:

- Game User: Person who plays video games
- Game Platformer: Person or group who provides game platforms for users to play games
- Game Contents Provider: Person or group who provides game contents with game platforms
- eSports Organizer: Person or group who holds eSports events with game platforms and contents

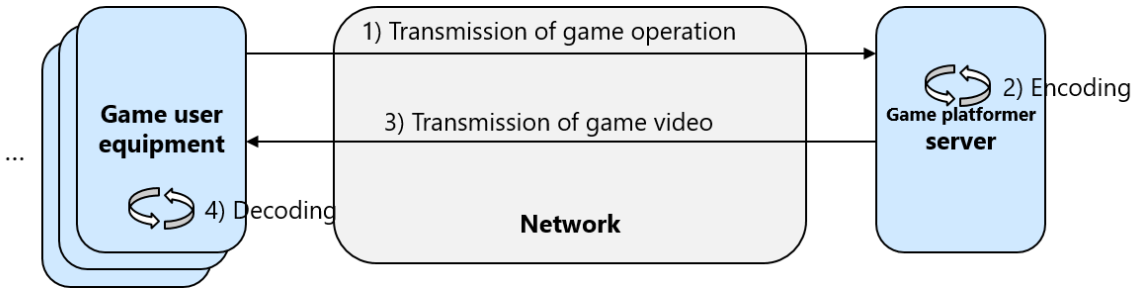

Figure 4 Additional Delay in Cloud Gaming

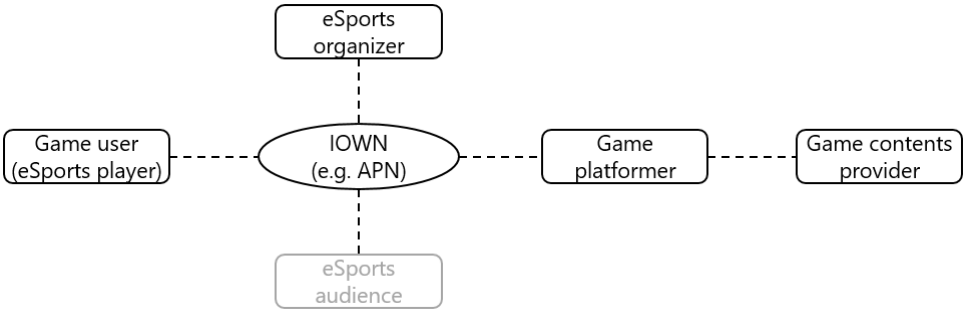

Figure 5 Overview of Cloud Gaming

3.1.3.2. Key Feature Set

3.1.3.2.1. Persona #1: Game User

As a game user, I want to;

- Play all low latency games on inexpensive terminals or without terminals
  1. Play my favorite game contents without having to buy a compatible console first
  2. Play various games without having consoles from multiple manufacturers and multiple generations at home
- Start the game quickly without waiting for the download time of the game content
- Use the game data of my multiple terminals easily

3.1.3.2.2. Persona #2: Game Platformer

As a game platformer, I want to;

- Orchestrate network functions and services to provide the best performance/lowest latency to Game User/eSports Players
- Gain and keep game users on my platform as game hardware technology evolves
- Reduce the development cost of consoles
- Gain new contents provider

3.1.3.2.3. Persona #3: Game Contents Provider

As a game contents provider, I want to;

- Update my game contents frequently and quickly to get users to continue playing the game contents
- Provide my game contents to many users

3.1.3.2.4. Persona #4: eSports Organizer

As an eSports organizer, I want to;

- Hold fair eSports events easily

1. Reduce the amount of game equipment needed at the venue when hosting eSports
2. Provide eSports players with a fair gaming environment

- Broadcast eSports competitions easily without having to capture video at the venue

3.1.3.3. Service Gap/Requirements

As shown Table 1, some games require a delay of 10 ms or less from user operation to screen display (including delays 1) to 4) in Figure 4 Additional Delay in Cloud Gaming. To enable access to all game content in cloud gaming, low-latency and low-jitter network transfer and server processing technology are required. Synchronization of games between multi-users is also required, within the required low-latency.

Table 1: Guidelines for Acceptable Delays by Game Category [Nakajima, 2018]

Game category	Delay
FPS (FP/First-Person shooter、PVP/Player versus Player) Official game	5–15 ms
FPS (PVP general)	10–30 ms
Battle fight (1-to-1)	5–30 ms
Racing game (PvP)	10–50 ms
Super Smash Bros. PvP	10–50 ms
2D side view jump game	10–50 ms
VR space sharing	10–100 ms
FPS (PvE/Player versus Environment)	30–100 ms
RTS (Real-time Strategy) / MOBA (Multiplayer Online Battle Arena)	20–100 ms
2D looking down action	50–100 ms
TPS (Third Person Shooter)	50–100 ms
Smartphone real-time action	50–200 ms
Voice chat	50–200 ms
Massively Multiplayer Online	100–300 ms
MMORPG (Massively Multiplayer Online Role-Playing Game)	100–500 ms
Smartphone turn-based games	200–1000 ms

In the 2020s, it is expected that game consoles with 8K image quality and 120-fps performance will become common. Therefore, to provide a cloud game with the same quality as a game console, IOWN technology, which enables high-quality and low-latency games, is required.

Service requirement can be summarized as follows;

- Number of the game users: ~ 100,000
- Data rate:
  - Raw data rate at the game user: 72 ~ 144 Gbps (8K, 120-fps)
  - Compressed data rate at the game user: 200 ~ 400 Mbps (8K, 120fps)

    Note: There is a trade-off between latency and data rate, so it is worth considering an appropriate solution using IOWN technology in the IOWN Global Forum.
- Latency: 5 ~ 1,000 ms
- Game synchronization within the required latency

3.2. Remote Operation

IOWN GF will expand the scope of various remote operation. In this section, how IOWN GF intends to change remote learning is described.

3.2.1. Professional Learning

3.2.1.1. Description

As Covid-19 spreads throughout the world, almost all cities or states/provinces enter shelter-in-place or completely lockdown stage. While normal stay-at-home life has not been entirely disrupted thanks for the current technologies in communications, there are still some restrictions.

For instance, if someone needs to do remote learning which involves a hand-on training session like disassemble and reassemble a car engine, it is not possible without invoking Augmented Reality and Virtual Reality technologies. In order to fully realize an immersive remote experience, all five human senses need to be digitized and resulted data to be transferred across networks. Streaming holographic videos over network requires high data bandwidth and low latency to accommodate large amount of hologram data and to increase fidelity.

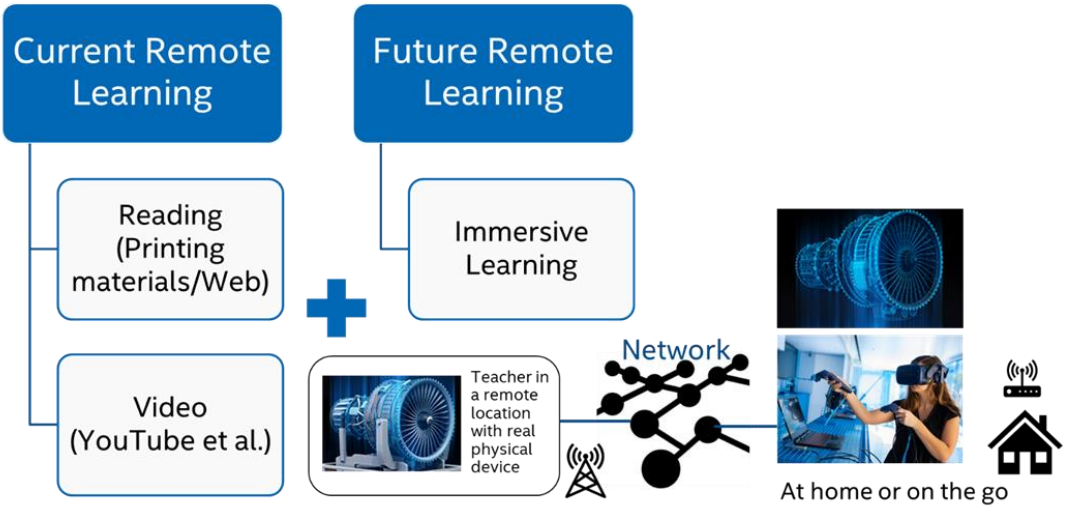

Figure 6 XR Application for Remote Learning

Remote expert support is another use case when shelter-in-place order or lockdown stage is in force. Issues on site such as plumbing problem, heater/air condition break-down, computer problem, or more severe issues, needed to be remedied with professional or expert assistance. However, the travel between the expert's location and the customer site is not possible due to lockdown.

Armed by Augmented Reality and Virtual Reality technologies, the remote expert can have an immersive surrounding experience, sees and comprehends exactly what the customer sees on site.

This allows the expert to better understand the overall context and provide specific troubleshooting and repairing instructions. The real-time haptic feedback enables the customer who may have minimum skill and no prior experience

to execute troubleshooting or repairing simultaneously with closely guided and real-time assistance as illustrated in *Figure 6*.

3.2.1.2. Key Feature Set

3.2.1.2.1. Persona #1: An Educator (for Remote Learning)

As an educator, I want to;

- able to give a remote hand-on teaching session via a virtual reality system
- deliver teaching session via wireless or wired network
- students to follow me INSTANTLY to disassemble/assemble complex engine using hologram model

so that my student(s) can learn effectively as if they were in the same physical classroom

3.2.1.2.2. Persona #2: A Consumer (for Remote Support)

As a consumer without any prior knowledge/experience in plumbing, I want to;

- fix a leak in kitchen with remote support from a plumber
- get support via mobile network
- take high definition photo/image of problem area to generate a 3-D hologram image
- able to stream high definition photo/video around the problem area to the plumber
- mimic in real-time how plumber interacts with digital model simultaneously

so that I can fix plumbing problem without waiting for plumber to come which may be restricted due to various reasons (lock-down, weather/road conditions et al.)

3.2.1.3. Service Gap/Requirements

Currently, remote learning involves students learning through reading of text materials in printing forms or on internet, or through watching video on YouTube or other video platforms. This is acceptable to gain general knowledge. However, if learning involves very complicated objects and requires hand-on experience, for example, to assemble or disassemble a turbo engine, then watching video and reading are not effective as they provide only two-dimensional aspects of object. This is where immersive learning utilizing AR/VR coming in. With AR/VR, a teacher in a remote location working with real physical device, whereas students are at home or on the go following teacher's instruction to work with a 3-D image of the exact model as the teacher has.

Today's technologies cannot deliver this kind of learning because of extremely high bandwidth requirement. For example, a raw hologram, without any compression, with colors, full parallax, and 30 fps will need 4.32 Tbps data rate. In addition to bandwidth requirement, haptic response time needs to be in the order of sub *ms* range [M. Giordani, 2020]. 5G wireless network typically targets for 1 Gbps data rate and 1 ms latency. There are many factors which affect whether these 5G performance goals can be achieved. 5G spectrum allocated, cell traffic level, backhaul capacity, among others, may become roadblock to reach 1 Gbps for average user. 5G numerology (slot duration), edger server location, network configuration (Non-standalone vs. standalone), backhaul configuration, traffic loading impact network latency. Currently, network's E2E network latency in a general commercial 5G network can at best achieve >= 8 ms latency. Network slicing can allow some users to achieve better performance, but latency with 1 ms is yet to be

achieved in commercial 5G network. In summary, current 5G network work cannot meet stringent performance requirements of 4.32 Tbps and less than 1 ms latency.

AR/VR devices based on stereoscopic principle utilize only human's binocular depth perception to create 3D image. Glasses-free truly holographic 3D Displays in the future is expected to provide all depth cues and eliminate eye fatigue or visual discomfort.

- Cameras in mobile device can take 8k HD photo/video
- 12 tiles for VR/AR
- Glasses-free truly holographic 3D displays
- High data rate (4.32 Tbps)
    - A raw hologram, without any compression, with colors, full parallax, and 30 fps, would require 4.32 Tbps
- Ultra-low network latency (haptic response time sub ms to 5.5 ms)
    - The latency requirement for Holographic Telepresence will be in the range of sub-ms, and thousands of synchronized view angles will be necessary.
- Continuous connection reliability (99.9999%)
    - In some remote learning scenarios, where fast haptic response is a critical element of the learning, isochronous operation adds tight constraints on jitter and the communication service reliability as high as 99.9999%.

3.3. Navigation

IOWN GF aims to change how people navigate in the physical world. In this section, new use case on navigation using XR technology is described.

3.3.1. Ultra XR Navigation

3.3.1.1. Description

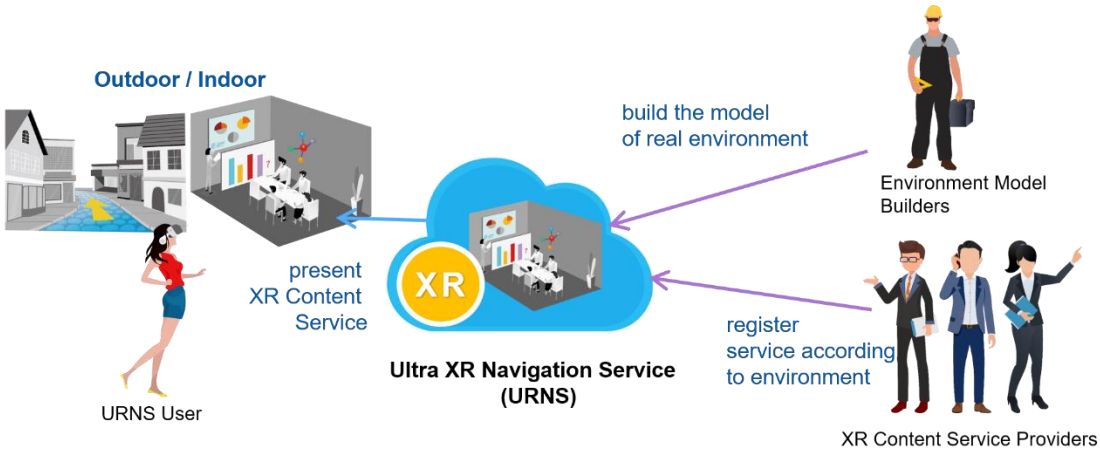

Figure 7 Overview of Ultra XR Navigation Services

IOWN GF will revolutionize how people experience the world and how the services/applications are properly provided to users through Ultra XR Navigation Service platform as illustrated in *Figure 7*. The benefit of XR different from other technologies is to combine the information or the simulation with the people's surrounding environment making data real and vivid. In addition, the worlds changes with the passage of time, and we make various decision by considering past, current and future. Therefore, IOWN GF, with its AI Photonic Network, will fulfill the idea in Time-travel XR Navigation Service platform.

For instance, the following cases could be there;

- At the restaurants or handcraft shops, you could feel the texture, size, weight, smell, and various additional specifications like a finished product by XR device, as if it was made already.
- You can see the past and future of the object that you watch through XR device.
- You can get the experience of traveling while staying at home, when your friends went to tourist site with XR device that interact you.

The following Persona appear in this use case;

- User: People who uses XR devices experience the past, current and future world.
- Environment Model Builder: Person or group model the real-world environment to digital data helping the platform to recognize user's environment.
- XR Content Service Provider: Person or group design the XR Content Service according to a specific environment.
- Platform Provider: A platform that provide XR content services according to user's location.

3.3.1.2. Key Feature Set

3.3.1.2.1. Persona #1: User

As an User, I want to;

- Interact with the platform
    1. See a well-designed XR Content (text, picture, video, 3D Model, etc.) showing up when I walk into an registered location.
    2. See a past and future of XR Content.
    3. Feel XR Content with my body and make a feedback to the platform.
    4. Get a recommendation and advice yield the best result.
    5. Make another User see what I see and manipulate my AR Content.

3.3.1.2.2. Persona #2: Environment Model Builder

As an Environment Model Builder, I want to;

- Interact with the platform
    1. Create an object with huge amount of data, ease and exactly.
    2. Record all the digital features (coordinates, magnet, 3D point cloud, etc.) of a specific environment/location precisely.
    3. Register a recorded environment/location

3.3.1.2.3. Persona #3: AR/VR Content Service Provider

As an XR Content Service Provider, I want to;

- Interact with the platform
    1. Design and register my XR Content Service on a specific environment to express my world precisely.
    2. Register my XR Content Service on a specific environment/location, so that user could be well-served when he walks in this environment.
    3. Make the platform know how to render my content and connect relevant services such as other XR Content Service Provider behind to expand my world.

3.3.1.2.4. Persona #4: Platform Provider

As Platform Provider, I want to;

- Interact with the User
    1. Recognize where the User is precisely.
    2. Show the User XR content according to user's vision.
    3. Recognize User's action such as pointing by finger, nodding, grabbing, walking, and so on for the action from the feedback.

3.3.1.3. Service Gap/Requirements

In order to create an XR world that combines virtual and physical aspects, it is necessary to accurately capture the state of the real world and objects, convert them into data, and link them with the virtual space in real time.

For service users to experience the XR world without feeling uncomfortable;

- A precise space locating technology: With more and more environment/location modeled on cloud and huge space could be possible, the accuracy of locating user's vision could be hard. We need to get an accurate picture of the real world. It is necessary to specify exact locations, both outdoors and indoors, and in some cases to accurately capture the latitude, longitude, and height of objects at the inch level.

  Real-time huge data transmission: Dozens of to hundreds of megabytes of data per object, such as 3D CAD data, need to be instantly downloaded and displayed on a user's device. If the object become more precise, the amount of data is even larger.

From the perspective of the service provider:

- An instant cloud rendering technology: According to use's vision input, a cloud rendering tech should render XR Content then send back to user. The bandwidth/latency and the computing time would be an obstacle of a good user experience.
- Edge AI technology: In order to recognize user's position and user's action, an edge AI technology is needed to shorten the response time of XR content rendering result and the feedback of user's action.

3.4. Human Augmentation

IOWN GF aims to realize technology that would expand human capability. In this section, new use cases on mind-to-mind communication and Another Me application are described.

3.4.1 Mind-to-Mind Communications

3.4.1.1. Description

Human civilization has evolved through communication and interactions with other cultures. Conflicts often arises from misunderstanding and even if the words are understood, it does not mean that background thoughts are comprehended. In the end, this will result in not being able to think and view from the other's standpoint. The aim of mind-to-mind communication is to realize a new form of communication that absorbs differences in language, culture and individual characteristics such as experiences and sensibilities to enable direct understanding of each other's perceptions and emotions.

We have seen the emergence of AI technology that realizes simultaneous interpretations and fills in the vocabulary gap. However, translating and presenting emotions or views that play crucial role in decision making are yet to be achieved.

Mind-to-Mind communications will bring about a new stage of communications that will overcome differences in language, culture, experience, values and sensibilities to gain a real and direct understanding of how the other people perceive and feel things in their minds through the transmission of expressed words and expressions.

This aims to reduce and eliminate communication discrepancies, increase psychological safety, and promote mutual understanding in our fragmented world to create an inclusive and harmonized approaches to problems where people with diverse characteristics can work together, stimulate each other and grow together.

The ultimate goal of this use case is to realize mind-to-mind communications as described in *Figure 8*.

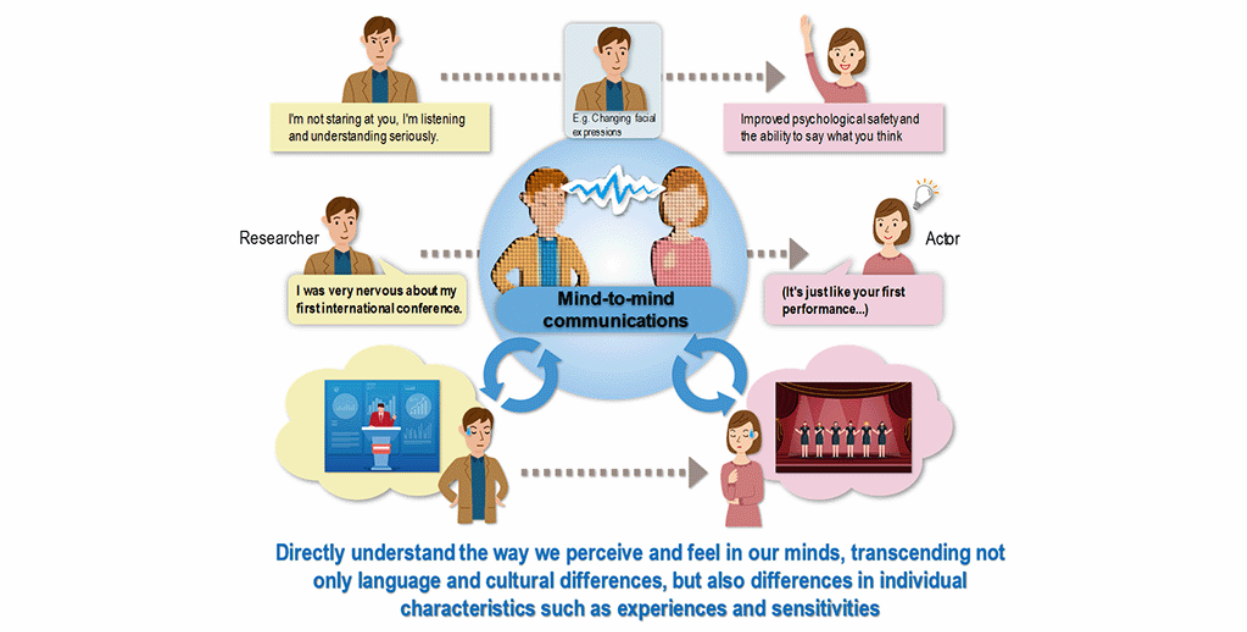

Figure 8 Mind-to-Mind Communications

3.4.1.2. Key Features

3.4.1.2.1 Persona #1: An individual who wants to communicate with others

As an individual, I want to;

- smoothly communicate with others who originates from other countries, background, e.g., education, professions, gender.
- easily understand what the others mean, translated into comprehensible language, expression with no lag.

3.4.1.2.2 Persona #2: Service Provider

As a service provider, I want to;

- calculate psychological state of individuals involved in the mind-to-mind communication. This will be based on the collected and accumulated information of an individual, i.e., context, including backgrounds, behavioral log, situations and cultural values.
- analyze psychological state from dialog text, voice, facial expression and any other physiological signs on real-time basis through wearable devices. In the receiver's end, this conveyed information must be converted appropriately in terms of the receiver's context and be projected graphically in the display of a wearable device in means of AR/VR.

3.4.1.3 Service Gap/Requirements

- In early stage, the differences in communication characteristics such as superficial attitudes and expressions must be transcended to achieve smooth, discrepancy–free face–to–face communications in lag. Which does not interfere with smooth human communications/dialogs. To achieve this, communication and psychological characteristics must be modelled and classified by this time. Also comfortable wearable devices that does not distract an ordinary conversation need to be developed.
- Late, the conversations will be assisted with how the talker perceive and feel things viscerally through messages transcending differences in experiences and sensibilities.
- Ultimately, communications that enable receivers to directly understand how senders perceive and feel things in their minds and thus absorbs the differences in experiences and sensibilities will have to be realized.
- Defining concepts and technologies that can be used as a reference for building human digital twins. In other words, the reference model of what the digitalized human data or 'human digital twin' is will be needed.

3.4.2. Another Me

3.4.2.1. Description

We all use or know about technologies to represent us or act on our behalf, e.g. an answering machine, a health profile, an avatar in a game. What if we combine all these use cases into one model, a digital instance that is controlled by us, that represents us in certain aspects in certain situations, and that accumulates knowledge from all the interactions it is involved in and can instantly make this knowledge available to all use cases?

Such a digital instance of ourselves would have to be controlled by us in a way that would allow us to enrich the model with physical and mental properties of ourselves as we see fit, and to make use of them only in those situations we approve. Individuals could also make use of the "experience" the digital instance would make in its interactions.

Such an alter ego entities in digital twin space, Another Me, can expand our lives and opportunities. It would also be ideal if those entities also grow with us, as we do in real life.

These can compensate for the loss of various opportunities in life and brings us more enriched life such as balancing work with childcare and nursing care, which will allow us to participate and contribute in society.

The ultimate goal of this use case is Another Me as summarized in *Figure 9*.

By the year 2035, Another Me, an alter ego that coexists and grows with itself in digital twin space will address these issues. All the experience and knowledge that are performed either in real life and digital twin spaces will be shared between Another Me and one's real-self, and thus creates more than double opportunities to play multiple roles simultaneously. Another Me aims to improve emotional well–being, health and life satisfaction by enabling users to perceive many opportunities in life such as balancing work and family and participating in many communities at the same time while enabling many growth opportunities in personal experience and growth with Another Me to accelerate self–realization.

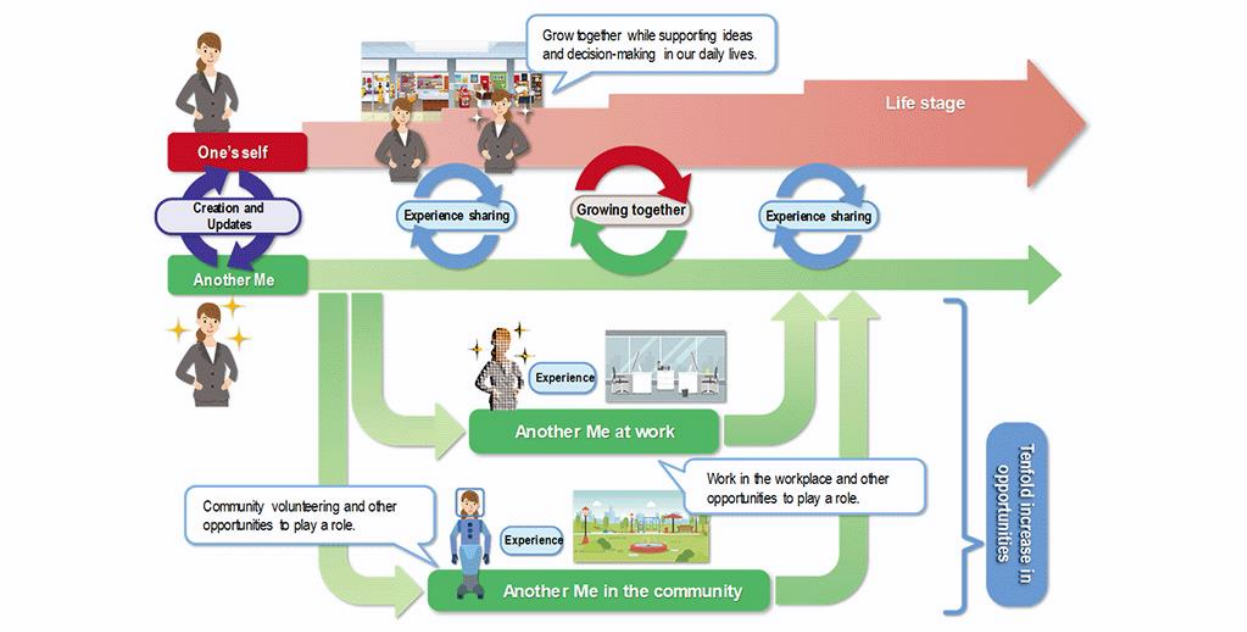

Figure 9 Another Me

3.4.2.2. Key Feature

3.4.2.2.1. Persona #1: Another Me User

As an Another Me User, I want to;

- take a leave from the work and let Another Me do the tasks instead, e.g. not only email replies and meeting arrange but also meeting attendance, discussion, writing various documents. This Another Me is an avatar of me in digital twin space and will behave just like myself. Another Me may not only be confined in a digital twin world but also reach real world through robotics technologies.

3.4.1.2.2 Persona #2: Another Me Service Provider

As an Another Me Service Provider, I want to

- provide storage and computational power to maintain more than one Another Me entities in digital space and provide means to communicate and share the experience of one's real-self.
- make all the thoughts, experiences and memories of the entities (amongst both digital twin and real world) to be shared instantaneously.

3.4.2.3. Service Gap/Requirements

- To accurately reproduce a human being in digital space, the required computational power is said to be in the order of some hundred peta FLOPS. This computational power can only be reached in 2020 by a super-computer. For multiple instances of humans to interact in the digital space, the figure is not achievable by today's technology.

- All transaction between a person and computed result in digital space must be made instantaneously in order to facilitate smooth communication. (The tolerated lags for auditory, visual, and tactile sensing are reported to be 100, 10, and 1 msec, respectively.) Digital twins must also be able to sense everything that a person experiences in real life and be able make a concise feed-backs to real person on what it experienced in the digital world.
- It is not yet realistic to collect all comprehensive personal data, so the technology to reconstruct character of a person must be from a selected and limited data must be established. This information must also be stored securely and retrieved appropriately.
- Defining concepts and technologies that can be used as a reference for building human digital twins. In other words, the reference model of what the digitalized human data or 'human digital twin' is will be needed.
- Human Digital Twin (HDT) controlled robotics would increase the networking and real-time demands of what is discussed today as "cloud robotics", where the concept of an "extended brain" exists to offload heavy computation to data center servers, and to use a "shared brain" to build a common database to collect and organize information about the world and learned skills/behaviors. HDI controlled robotics would have to maintain the differences of the individuals.

4. Requirements

4.1. Functional Requirements

IOWN shall meet the following functional requirements to support AI-Integrated Communication use cases.

Table 2: Functional Requirements for AI-Integrated Communication Use Cases

4.2. Performance Requirements

IOWN shall meet the following performance requirements to support AI-Integrated Communication use cases.

Table 3: Performance Requirements for AI-Integrated Communication Use Cases

5. References

IOWN GF. (2021). IOWN GF CPS Use Case Interim Report.

FIFA. (2018). 2018 FIFA World Cup Russia Global Broadcast and Audience Summary. Retrieved from https://resources.fifa.com/image/upload/2018-fifa-world-cup-russia-global-broadcast-and-audience-executive-summary.pdf?cloudid=njqsntrvdvqv8ho1dag5

M. Giordani, M. P. (2020). Toward 6G Networks: Use Cases and Technologies. IEEE Communications Magazine, vol. 58, no. 3, 55-61.

Nakajima, K. (2018). Technology for Creating Cloud Game. In Gijutsu Hyoron.

History

Revision	Release Date	Summary of Changes
1.0	January 13, 2021	Initial Release