ITU Focus Group Technical Report

(12/2023)

ITU Focus Group on metaverse
(FG-MV)

FGMV-17

Guidelines and requirements on interpreting in the metaverse

Working Group 8: Sustainability, Accessibility & Inclusion

Technical Report ITU FGMV-17

Guidelines and requirements on interpreting in the metaverse

Summary

This document provides guidelines and requirements on interpreting in the metaverse. It summarises typical use settings that require interpreting in the metaverse, including conference interpreting, public service interpreting and sign-language interpreting. It describes technical requirements for interpreting in the metaverse. It also provides advice for all parties in interpreted events in the metaverse, including organizers, speakers, interpreters and audience in interpreting-facilitated events in the metaverse.

Note

This Technical Report is an informative ITU-T publication. Mandatory provisions such as those found in ITU-T Recommendations are outside the scope of this publication. This publication should only be referenced bibliographically in ITU-T Recommendations.

Change Log

This document contains Version 1.0 of the ITU Technical Report on "*Guidelines and requirements on interpreting in the metaverse*" approved at the 4th meeting of the ITU Focus Group on metaverse (FG-MV), held on 4-7 December 2023 in Geneva, Switzerland.

Editor:	Binhua Wang University of Leeds	E-mail: B.H.W.wang@leeds.ac.uk
Editor & Task Group Co-Chair:	Yong Jick Lee Center for Accessible ICT, Korea (Rep. of)	E-mail: ylee@caict.re.kr
Editor:	Fei Gao Chongqing University of Posts and Communications	E-mail: fei_gao1@hotmail.com

Editor:	Lihong Pan University of Leeds	E-mail: L.Pan1@leeds.ac.uk
WG8 Co-Chair:	Nevine Tewfik MCIT Egypt	E-mail: ntewfik@mcit.gov.eg
WG8 Co-Chair:	Pilar Orero UAB Spain	E-mail: pilar.orero@uab.cat
Task Group Co-Chair:	Paola Cecchi-Dimeglio Harvard University	E-mail:pcecchidimeglio@law.harvard.edu

Table of contents

Technical Report ITU FGMV-17

Guidelines and requirements on interpreting in the metaverse

1   Scope

This document provides guidelines and requirements on interpreting in the metaverse. It summarises typical use settings including conference interpreting, public service interpreting and sign-language interpreting in the metaverse. It describes technical requirements for conference interpreting, public service interpreting and sign-language interpreting in the metaverse. It also provides advice for all parties in interpreted events in the metaverse, including organizers, speakers, interpreters and audience in interpreting-facilitated events in the metaverse.

2   References

The following ITU-T Recommendations and other references contain provisions which, through reference in this text, constitute provisions of this Technical Report. At the time of publication, the editions indicated were valid. All Recommendations and other references are subject to revision; users of these Technical Report are therefore encouraged to investigate the possibility of applying the most recent edition of the Recommendations and other references listed below. A list of the currently valid ITU-T Recommendations is regularly published. The reference to a document within this Technical Report does not give it, as a stand-alone document, the status of a Recommendation.

[ITU-T F.791]   Recommendation ITU-T F.791 (2015), *Accessibility terms and definitions.*

[ITU-T J.301]   Recommendation ITU-T J.301 (2014), *Requirements for augmented reality smart television systems*.

3   Terms and definitions

3.1   Terms defined elsewhere

The Technical Report uses the following terms defined elsewhere:

3.1.1   artificial intelligence [b-ISO/IEC 2382]: An interdisciplinary field, usually regarded as a branch of computer science, dealing with models and systems for the performance of functions generally associated with human intelligence, such as reasoning and learning.

3.1.2   augmented reality [b-ITU-T J.301]: A type of mixed reality where graphical elements are integrated into the real world in order to enhance user experience and enrich information.

3.1.3   diverse users [b-ISO/IEC 71]: Individuals with differing abilities and characteristics or accessibility needs.

3.1.4   service [b-ISO/IEC 9241-11]: Means of delivering value for the customer by facilitating results the customer wants to achieve.

3.1.5   system [b-ISO/IEC 9241-11]: Combination of interacting elements organized to achieve one or more stated purposes.
3.1.6   task [b-ISO/IEC 9241-11]: Set of activities undertaken in order to achieve a specific goal.

3.1.7   user interface [b-ISO/IEC 9241-11]: All components of an interactive system (software or hardware) that provide information and/or controls for the user to accomplish specific tasks with the interactive system.

3.1.8   virtual reality [b-ISO 9241-394]: Set of artificial conditions created by computer and dedicated electronic devices that simulate visual images and possibly other sensory information of a user's surrounding with which the user is allowed to interact.

3.2 Terms defined in this Technical Report

This Technical Report defines the following terms:

None

4 Abbreviations and acronyms

AI   Artificial Intelligence

GDPR   General Data Protection Regulation

NLP   Natural Language Processing

PSI   Public Service Interpreting

VR   Virtual Reality

5 Conventions

In this Technical Report:

The expression "**is required to**" indicates a requirement that must be strictly followed and from which no deviation is permitted if conformance to this Technical Report is to be claimed.

The expression "**is recommended**" indicates a requirement that is recommended but which is not absolutely required. Thus, this requirement need not be present to claim conformance.

The expression "**can optionally**" and "**may**" indicates an optional requirement that is permissible, without implying any sense of being recommended. This term is not intended to imply that the vendor's implementation must provide the option and the feature can be optionally enabled by the network operator/service provider. Rather, it means the vendor may optionally provide the feature and still claim conformance with this Technical Report.

6 Guidelines on conference interpreting in the metaverse

Metaverse conference interpreting is required to facilitate communication in multilingual conferences in the metaverse, enabling participants with different languages and cultural backgrounds to understand and interact with each other in the remote multilingual conferences.

6.1 Typical use settings

Typical use settings where conference interpreting in the metaverse is required include:

International Conferences
Large-scale international conferences in the metaverse, which are typically organised by international institutions and between nations and participated by delegates from different countries, often require metaverse conference interpreting to facilitate multilingual communication.

Corporate Meetings
Corporate meetings in multinational companies and organisations, which are held in the metaverse, often require metaverse conference interpreting for their global team meetings, board meetings, or training sessions for staff from different countries.

Inter-governmental and Diplomatic Meetings
Bilateral and multilateral meetings, and other diplomatic meetings, which are held in the metaverse, involving representatives from various countries often require metaverse conference interpreting.

Press Conferences
Metaverse conference interpreting is required to enable media outlets and journalists to access information during press conferences and media briefings, which are held in the metaverse.

Academic Conferences

Metaverse conference interpreting is required to facilitate multilingual discussions and presentations during academic conferences with international participants, which are held in the metaverse.

International Training and Workshops
Metaverse conference interpreting is required for international training sessions and workshops involving multilingual participants, which are held in the metaverse.

International Business Meetings and Negotiations
Metaverse conference interpreting is required for international business meetings and negotiations, which are held in the metaverse, to facilitate multilingual communication between different enterprises from different countries.

6.2 Equipment and technical requirements

6.2.1 Conference organisers

Organisers of multilingual conferences in the metaverse is recommended to use a platform to organise the interpreting-facilitated virtual meetings. The platform is required to have real-time audio and video streaming capabilities, affording a comprehensive online system that facilitates and manages conference interpreting services for such events. The platform is required to be designed to meet the needs of multilingual conference interpreting (in most cases simultaneous interpreting) provided by interpreters to event organisers, speakers and the audience from different language backgrounds. The equipment and technical requirements are required to ensure the following functionalities that are essential to conference interpreting in the metaverse:

- Event setup by organisers

  The facilities are required to ensure that organisers can set up their conferences on the platform, specifying the date, time, agenda and other details. They can also define the language options available and the directionalities for the interpreting services during each session.

- Pre-event accessibility for interpreters

  The facilities are required to ensure that interpreters can access the event agenda, presentation slides, and relevant conference documents in advance to prepare for the interpreting task.

- Management of interpreters

  The facilities are required to enable management of interpreters, including scheduling, assigning specific language pairs, coordinating their availability and monitoring their performance.

- Communication interface among relevant actors

  The facilities are required to ensure communication between the organiser, interpreters, and speakers of the event, particularly between: event moderators and interpreters, interpreter and technicians; interpreters using the same outgoing language channel, all interpreters in the same virtual event.

- The user interface for speakers and audience

  The platform is recommended to offer a user-friendly interface for both speakers and the audience, making it easy to navigate and access conference interpreting services in the metaverse.

- Technical support

  The platform is recommended to include technical support services to assist all the event actors (organisers, interpreters, speakers, and audiences) prior to, during, and after the event. Such support includes technical instructions and troubleshooting mechanisms in case of issues that may arise during the event.

- Recording and transcription

    The platform is recommended to offer the option to record interpreted content if it is agreed by the speakers and the interpreters, and provide transcriptions, allowing participants to access the interpreted content after the event.

6.2.2 Conference interpreters

The conference interpreters offering conference interpreting services in the metaverse are required to be equipped with the following equipment:

- Computer

    A powerful and reliable computer connected to the internet, which can run the metaverse conference interpreting platform and handle audio and video streaming feeds.
- Stable Internet Connection

    Stable and high-speed internet connection is essential for smooth communication in the metaverse. A sufficient amount of bandwidth is crucial for metaverse conference interpreting. For video and audio communication, a recommended minimum of 1 Mbps upload and 1 Mbps download speed is suitable for conducting interpreting. However, higher speeds, such as 3 Mbps or more, are preferred for better performance, especially in large-scale conferences.
- High-Quality Headset

    A professional-grade headset with excellent audio quality and noise-cancelling features for clear remote communication, which is an essential tool for interpreters in metaverse conference interpreting. It enables interpreters to hear the source language clearly and to deliver interpretations clearly to the participants in the target language.
- The most critical aspect of a headset for interpreters is the sound quality. It is required to have clear audio reproduction with a wide frequency response, allowing interpreters to hear speech nuances and subtle intonations accurately in remote communication. Noise-cancellation technology is also vital since it helps block out background noise, ensuring that interpreters can focus attention on the speaker's voice and eliminate distractions during the interpreting process.
    A high-quality microphone is essential for interpreters as it ensures that their voice is transmitted clearly to the participants. The microphone is required to have noise-cancelling capabilities to eliminate ambient noise.
- Dual Monitors

    Dual monitors are not compulsory but can be useful if interpreters need to work on the metaverse conference interpreting platform on one screen and at the same time view conference documents on the other screen.

6.2.3 Speakers

- Computers, headset, and internet connection

    Speakers of the online conferences in the metaverse are recommended to be equipped with computers, headsets, and internet connections as suggested for interpreters.
- Presentation software

    Speakers are recommended to use software to support any slides or visual aids used during the presentation, such as but not limited to *Microsoft PowerPoint, Google Slides, Keynote, Prezi*.

6.3   Advice for metaverse conference organisers, speakers and interpreters

6.3.1   For metaverse conference organisers

Organising an online virtual event with metaverse conference interpreting requires meticulous planning and execution to ensure a seamless cross-language and cross-cultural communicative experience for the speakers and other participants. The following advice can help event organisers to use metaverse conference interpreting effectively and deliver a successful virtual event:

- Identify language combinations

  Determine the language combinations and interpreting directionalities that are required for the event based on the languages used by the speakers and language profiles of the participants.

- Book a reliable metaverse conference interpreting platform

  Select a reliable metaverse conference interpreting platform that offers high-quality audio and video streaming, a user-friendly interface, and reliable technical support. Set up the language channels in the metaverse conference interpreting platform. Test the platform prior to the event to ensure its effectiveness.

- Book interpreters

  If the event involves more than one language pair, ensure that different interpreters are booked for the different language pairs, and provide them with clear instructions on when to join the virtual event. Make sure that two interpreters work in pairs in simultaneous interpreting so that they can take turns to interpret in every 15-20 minutes.

- Provide access of conference documents to interpreters

  Provide the interpreters who will deliver metaverse conference interpreting with the conference agenda, Presentation slides, and other conference documents in good advance. This allows them to do task-based preparation, which helps to ensure quality interpreting during the event.

- Briefing for interpreters

  If time allows, conduct a briefing session with the interpreters to give them context of the event and to discuss rundown of the event and user expectations.

- Technical rehearsal

  If the situation allows, conduct a technical rehearsal with the interpreters and speakers to familiarise them with the facilities and workflow, and to resolve any potential technical issues.

- Inform speakers and participants about interpreting

  At the beginning of the conference, inform speakers and participants about the availability of the metaverse conference interpreting services, and the language channels that they can select.

- Offer 'standing-by' technical support

  Provide interpreters, speakers, and participants with instant access to technical support during the event in case that they encounter any issues with the metaverse conference interpreting setup. Monitor the status of interpreting during the event. Be vigilant for any audio, video, or connectivity issues and get technical support to address them promptly.

- Recording and transcription

  If it is agreed by the speakers and the interpreters, consider recording the event and providing transcriptions of interpreted content afterwards for participants' reference.

- Gather feedback

After the virtual event, gather feedback from participants, speakers, and interpreters to identify areas of improvement for future events.

6.3.2 For metaverse conference interpreters

The following advice is provided to conference interpreter for best practices in metaverse conference interpreting to ensure effective communication and collaboration with other interpreters as well as speakers.

- Pre-task preparation

  Be familiarised with the event agenda, speech topics, presentation slides and other conference documents beforehand. Participate in pre-event briefings and coordination meetings with other interpreters and event organisers to discuss logistics, interpreting modes, and potential challenges.

- Technical setup

  Ensure a stable internet connection with sufficient bandwidth to allow smooth video and audio streaming. Use a high-quality headset with noise-cancellation features and a reliable microphone to deliver clear interpretations. Dual monitors are recommended, with one used for the interface of the metaverse conference interpreting platform and the other for viewing conference documents.

- Professionalism

  Conduct metaverse conference interpreting in a professional manner during the event, maintaining impartiality and neutrality in interpreting. Dress appropriately as they may be visible to the speakers and participants in the virtual events.

- Collaboration with other interpreters

  Coordinate closely with other interpreters to ensure smooth transitions during relay or team interpreting setups. Use pre-arranged visual cues or signals to indicate when you need to switch or take over interpreting duties. The typical length of a shift for simultaneous interpreting ranges from 15 to 30 minutes. Interpreters need to work in pairs and take turns to deliver the rendition.

- Interaction with speakers

  If necessary, request speakers politely and discreetly to slow down or repeat information for accurate interpretation. If necessary, interrupt speakers briefly to ask for clarification on ambiguous terms or jargon.

- Confidentiality and privacy

  Adhere to strict confidentiality and data privacy guidelines, especially when interpreting sensitive or confidential events.

6.3.3 For metaverse conference speakers

Working remotely with interpreters as a speaker requires collaboration with them to ensure effective communication during virtual events in the metaverse. Here are some essential guidelines for speakers working with conference interpreters in metaverse conference interpreting:

- Pre-event preparation

  Provide speech topics, presentation slides, and other prepared materials to interpreters in good advance to help them prepare.

- Clarify key terminology, abbreviations and acronyms

  Identify any specialised and technical terms that might be used during the presentation and provide necessary translations or explanations to the remote interpreters. Explain to the interpreters the abbreviations and acronyms that are not commonly known by people who are

not in your field or subject area.

- The use of scripts

  If speakers plan to read a written script, try to do so at a moderate speed which is similar to naturally spoken speech. Also, be sure to give the interpreters as early as possible a copy of the script that will be read.

- Speaking pace and clarity

  Speak clearly and at a moderate pace to allow interpreters to convey your message to the audience. Avoid speaking too fast or using long and complex sentences that may be too difficult to render.

- Communication with interpreters

  If possible, establish direct communication with interpreters prior to the event to clarify any doubts and to explain technical jargons. Be open to responding to any requests or questions from the interpreters in their briefing session.

- Visual aids and gestures

  If possible, use visual aids, such as presentation slides, to complement the presentation or speech and provide additional context for the interpreters and audience. Consider using gestures or visual cues to enhance non-verbal communication.

- The font size of the words or characters on the slides

  The font size of the words or characters on the presentation slides in a presentation are recommended to be easily readable by the interpreters and audience through the remote screen. The title on each slide are recommended to be larger and more prominent than the other text. The main content, such as bullet points or paragraphs, are recommended to be clear and easily readable.

7   Guidelines on public service interpreting in the metaverse

In almost every nation, there are groups of people who are language minorities, such as migrants and ethnic minorities, who do not speak the language of the majority. Whether and how they access public services constitute a significant issue for both themselves and the society not only because of the language barriers that they have to face with, but also because the language in healthcare, legal and social services is often highly specialised, which usually takes several years of learning to attain native-speaker fluency. Therefore, the language minorities often need to rely on public service interpreters to access public policies and public services, such as healthcare, legal, education and social services (e.g., housing, social security, welfare, etc.). Public service interpreting (PSI) is essential for members of language minorities to access public services and to communicate their needs and voices, which plays a crucial role in bridging linguistic barriers and ensuring accessibility and inclusivity in the metaverse. Within the metaverse, PSI involves interpreters offer interpreting remotely, enabling individuals to access interpreting services remotely.

7.1   Typical use settings

The typical use settings of public service interpreting (PSI) in the metaverse include but is not limited to:

1) Police services: PSI is required in police investigations, interviews, and questioning with the language minorities in the metaverse, who do not speak the language of the majority or speak it fluently.

2) Court and legal proceedings: PSI is required in courtrooms in the metaverse, where interpreters facilitate communication between language minorities and court and legal professionals, ensuring fair trials and proper understanding of legal matters.

3) Immigration services: PSI is required in communication between language minorities and immigration officers and applicants during interviews, visa applications, and other immigration-related procedures in the metaverse.

4) Legal consultations with solicitors: PSI is required by language minorities in legal consultations and meetings with solicitors in the metaverse.

5) Government public services: PSI is required by language minorities in their communication with government offices in order to access government public services in the metaverse.

6) Hospitals and healthcare services: PSI is required to facilitate communication between medical professionals and language minority patients in the metaverse, ensuring their access to healthcare services.

7) Social services and support programmes: PSI is required to assist language minorities in various social services settings, such as welfare offices, counselling centres, and community support programmes, fostering accessibility and inclusivity.

7.2 Technical requirements

In the metaverse, there are two primary forms of public service interpreting: real-time public service interpreting by human interpreters and automatic public service interpreting by the machine.

7.2.1 Real-time public service interpreting by human interpreters

This section is about the technical prerequisites that enable real-time interpreting by human interpreters within the metaverse. These requirements encompass communication platforms, language processing tools, security measures, and adaptability to ensure seamless interpretation during live interactions. Here are the technical requirements for real-time interpreting by human interpreters:

1) Communication platforms

To enable real-time interpreting within the metaverse, the platform itself is required to support robust and low-latency communication. Whether it is a virtual meeting space, a virtual meeting platform, or a social hub, the communication platform in the metaverse is recommended to offer:

High-quality audio and video: Clear and high-fidelity audio and video transmission is essential for interpreters to accurately hear and observe the participants' interactions. A reliable and high-bandwidth connection is crucial to minimise audio and video lags.

Multiparty support: The platform is recommended to accommodate multiparty interactions, allowing multiple participants to communicate simultaneously. This capability ensures that interpreters can listen to one participant and interpret for others as needed, promoting seamless communication.

Multilingual support: The platform is recommended to be capable of handling multiple languages and language pairs. It should allow interpreters to select their language of expertise and seamlessly switch between languages during real-time interpreting.

2) Language processing tools

Human interpreters need support from advanced language processing tools to enhance their efficiency and accuracy during real-time interpretation. These tools may include:

Real-time translation aids: Artificial Intelligence (AI)-powered real-time translation tools can assist interpreters by providing quick access to relevant terminology and context-specific language support. These aids act as valuable references during high-demand situations.

Gesture and emotion recognition: Integrating gesture and emotion recognition technologies can provide interpreters with non-verbal cues that aid their understanding of the participants' emotions and intentions. This additional context enhances the accuracy of interpreting.

Voice recognition and speech synthesis: Voice recognition tools can transcribe spoken language into text, allowing interpreters to access written transcripts, while speech synthesis facilitates the generation of spoken translations.

3) Security and privacy measures

In the metaverse, real-time interpreting sessions may involve sensitive information exchange during virtual meetings, healthcare consultations, or immigration consultations. Ensuring the security and privacy of such interactions is critical, and the technical requirements are recommended to include:

End-to-end encryption: Implementing robust encryption protocols ensures that all communication between participants and interpreters remains confidential and protected from unauthorised access.

Secure user authentication: The platform is recommended to require secure user authentication to verify the identity of participants and prevent unauthorised users from joining interpreting sessions.

Data protection compliance: Adhering to relevant data protection regulations, such as the General Data Protection Regulation (GDPR), ensures that personal data is handled responsibly and ethically.

4) Adaptability and customisation

Flexibility and adaptability are essential technical requirements for real-time interpreting in the metaverse. The platform is recommended to allow interpreters to customise their workspace and settings to suit their preferences and work style. Additionally, providing options to fine-tune language models or adapt to domain-specific jargon ensures higher accuracy and context-aware delivery in interpreting.

5) Accessibility and cross-platform compatibility

To maximise the accessibility of real-time interpreting services, the platform is recommended to be compatible with various devices commonly used in the metaverse. Whether accessed through desktop computers, laptops, VR headsets, or mobile devices, the platform is recommended to offer a seamless experience across different platforms.

Furthermore, audio or video recordings could be contemplated for individuals who either prefer or necessitate spoken information, especially in scenarios where proficiency in a native language might be limited or for languages with a significant emphasis on oral tradition.

6) Training and support

A user-friendly and intuitive interface is essential for interpreters to navigate the platform easily. Additionally, comprehensive training and support is recommended to be provided to interpreters to familiarise them with the platform's features and functionalities, ensuring they can utilise the technology effectively.

7.2.2 Automatic public service interpreting by the machine

Automatic public service interpreting by the machine offers a promising solution by leveraging artificial intelligence and language processing technologies. This section outlines the technical requirements for automatic public service interpreting in the metaverse.

1) Natural Language Processing (NLP) algorithms

At the core of automatic interpreting lies sophisticated natural language processing algorithms. These algorithms are designed to analyse, understand, and interpret human language in a contextually appropriate manner. The technical requirements for NLP algorithms in the metaverse include:

Speech recognition: Automatic interpreting often begins with speech recognition, where the system transcribes spoken language into text format.

Multilingual support: The algorithms is recommended to be capable of handling multiple languages and language pairs, accommodating the diverse linguistic landscape of the metaverse.

Contextual understanding: NLP algorithms is recommended to be able to grasp the context of the conversation, as the meaning of words and phrases often depends on the surrounding dialogue.

Sentiment analysis: Integrating sentiment analysis allows the machine to comprehend the emotional undertones of conversations, providing more nuanced interpretations.

2) Large multilingual datasets

Machine translation systems rely heavily on vast and diverse multilingual datasets for training and fine-tuning their models. The technical requirements for these datasets in the metaverse include:

Multimodal data: Incorporating multimodal data (e.g., audio, video, text) enables the machine to consider non-verbal cues and context for improved translations.

Domain-specific data: Specialised datasets for various industries or domains help create more accurate and context-aware interpretations in specific contexts.

3) Real-time translation engine

A real-time translation engine is an essential technical requirement to provide instantaneous and fluid translations during live interactions within the metaverse. This engine is recommended to offer:

Low latency: Minimal delays between spoken or written input and the translated output are critical for smooth communication.

Multilingual capability: The translation engine is recommended to be capable of supporting multiple languages and language pairs in real-time.

Continuous updates: Regular updates to the translation engine ensure that it stays up to date with language changes and evolving linguistic trends.

4) Data privacy and security

Automatic interpreting in the metaverse involves real-time exchange of conversations, making data privacy and security paramount. The technical requirements for data privacy and security include:

End-to-end encryption: Implementing robust encryption protocols ensures the protection of all communication between users and the machine translation system.

Data anonymisation: Anonymising user data prevents any personal information from being linked to specific conversations, preserving user privacy.

5) Continuous Learning and Adaptability

To enhance the accuracy and effectiveness of automatic interpreting, the machine translation system is recommended to continuously learn and adapt. The technical requirements for continuous learning include:

Feedback mechanism: A feedback loop allows users to provide input on the translation quality, enabling the system to learn from its mistakes and improve over time.

Recordings: The interpreting session is recommended to be recorded for subsequent analysis following the meeting. This recording serves the purpose of assessing accuracy and facilitating enhancements in the performance of the automated interpreting system.

Domain-specific adaptation: The system is recommended to be able to adapt to domain-specific jargon and terminology, ensuring contextually relevant translations.

6) Cross-platform compatibility

The machine translation system is recommended to be compatible with various platforms and devices commonly used in the metaverse, such as virtual meeting spaces, social platforms, and gaming environments. Cross-platform compatibility ensures that users can access interpreting services seamlessly from different devices.

7.3 Advice for meeting organisers, speakers, interpreters and audience

This section presents guidelines for meeting organisers, speakers, interpreters and the audience to foster a positive and engaging experience within the metaverse. The guidelines cover pre-meeting planning, speaker preparations, interpreting best practices, and strategies for audience involvement. By adhering to these guidelines, all stakeholders can contribute to creating memorable and accessible virtual meetings that connect people in the global metaverse.

7.3.1 Advice for meeting organisers

Multilingual considerations: Assess the linguistic diversity of the audience and potential participants. Provide language options for meeting materials, registration forms, and promotional content to accommodate various language preferences.

Interpreting services: Arrange for professional interpreting services to facilitate real-time interpretation during the meeting. Plan for the required number of interpreters based on language needs and the number of participants.

Technical setup: Ensure that the metaverse platform supports multilingual meetings, offers real-time translation features, and allows for easy switching between languages during the meeting.

Cultural sensitivity: Promote cultural sensitivity throughout the meeting planning process, including respecting various time zones and avoiding scheduling conflicts with significant holidays or cultural meetings.

7.3.2 Advice for speakers

Language:

- Speakers is recommended to maintain a moderate pace while speaking to allow interpreters to provide accurate interpreting in real-time.
- Speakers is recommended to be aware of their audience's language preferences and prepare to deliver their content clearly and concisely for diverse language groups.
- Speakers is recommended to avoid complex jargon or idiomatic expressions that may be difficult for non-native speakers to understand. Choose language that is accessible and easily translatable.

Collaboration:

- Work closely with interpreters during the meeting to ensure smooth communication.
- Provide materials or key points in advance to help interpreters prepare.

7.3.3 Advice for interpreters

Research the topic: Familiarise yourself with the meeting's subject matter to provide contextually relevant interpretations.

Coordinate with speakers: Reach out to speakers in advance to discuss terminology and clarify any potential language challenges.

Remain neutral during interpreting: Interpreters are required to aim for accuracy and impartiality, refraining from injecting personal opinions or biases into their interpretations.

Multi-tasking: You will need to be able to manage the devices and technical settings while doing interpreting in the metaverse.

Time management: Ensure that interpreting is delivered in a timely manner, allowing for an efficient flow of the meeting.

7.3.4 Advice for the audience

1) Choose the right language: Select the language channel that suits your understanding best for real-time interpreting during the meeting.

2) Allow time for interpreting when speaking: When you participate discussions, Q&A, and other interactions, please allow time for interpreting.

3) Respect language differences: Be patient and understanding of language barriers, recognising that interpreting may be delivered with slight delays due to internet connection and streaming.

8 Sign-language interpreting in the metaverse

Sign language interpreting in the metaverse is necessary to ensure accessibility for those who are deaf or hard of hearing and use sign language as their primary communication medium. Sign language is a linguistic system, as the name implies. Sign language communicates using signs consisting of handshape (or hand-form), orientation (or palm orientation), location (or place of articulation), movement, and non-manual markers (or facial expression). Just as the types of speech languages vary, there are various types of sign language with various vocabulary and grammar systems by region and culture. Sign language is not an alternative expression that responds to voice language one-on-one but is an independent language that reflects the social experience, culture, and value system of the deaf in their own vocabulary and grammar system. There is a signed communication called an international sign (IS), which is not a sign language as it is created on the spot by highly trained sign language interpreters and which changes by virtue of the different sign language requirements of the participants for every single meeting.

The problem of interpreting speech and sign language is not just about converting speech into hand signals or vice versa but also requires a process as complex as translating different speech languages. Therefore, the problem of expressing sign language or translating it into voice language through metaverse is not simple. In particular, sign language recognition technology is still very rudimentary because it requires understanding the vocabulary and grammar system of sign language beyond the scope of simple gesture recognition technology.

8.1 Typical use settings

Depending on the environment, sign language interpreters are provided one-on-one interpreting with one interpreter per deaf person, or several people who are deaf and hard-of-hearing are interpreted by one interpreter simultaneously. In the case of international conferences, sign language interpreters corresponding to each language may be deployed. British and American sign languages are different and require different interpreting even in the same English-speaking world.

In particular, due to the nature of sign language interpreters that require interpreting through hand movements, it is physically difficult for one interpreter to provide interpreting for a long time, so several interpreters may be needed to conduct interpreting alternately.

When a single interpreter interprets multiple people's sign language, or vice versa, it is often difficult to identify who the speaker is in a video conference environment where many people talk simultaneously.

The critical thing in sign language interpreting in the metaverse is that the speaker, deaf and hard-of-hearing people, are required to be shown with focus at the centre. Many video chat apps do not consider sign language interpreting, so deaf and hard-of-hearing people who make actual remarks in sign language do not receive focus but instead sign language interpreters receive screen focus. It is

necessary for the speaker, the deaf, and the interpreter to be coupled and to put the speaker, deaf and hard-of-hearing people, at the centre.

8.2 Technical requirements

One of the main problems with sign language recognition is that sign language does not have a generalised writing system despite various academic attempts. Here, the character system refers to the expression, such as the alphabet of the speech-language as symbols, and the code refers to a system in which a computer can recognise characters, such as Unicode corresponding to characters. Unlike speech languages, sign language does not have a character system or code, so writing and coding a sign language dictionary is difficult. Unlike voice languages, sign language corporations are generally collected as uncoded videos and have interpreting into voice languages, making recognising and analysing sign language challenging.

The technology for expressing sign language has advanced compared to sign language recognition as the technology for automatic sign language generation using avatars has developed. However, many people who are deaf or hard-of-hearing still think that the quality of sign language expressed by avatars does not reach the level that can be applied. In addition, the following advantages and disadvantages exist in expressing sign language with avatars:

Pros:

- The initial cost of developing avatars in expressing sign language is high, but once built, it can be operated at a low cost.
- It can be operated 24 hours a day, 365 days a year, without relying on the availability of human interpreters.
- It solves the problem of personal information infringement that may occur due to the exposure of the face of a human interpreter.
- In particular, at a low cost, sign language may be inserted into pre-recorded voice and video data.

Cons:

- The interpreting quality of avatar interpreters is lower than that of human interpreters (especially, the delivery of facial expressions is not precise)
- Interpreting is difficult to be conducted by avatar interpreters when there is a lot of noise or loud background sound
- Clear interpreting distinguishing each speaker can be difficult if multiple people need to be interpreted at once

In reality, specific quality standards for the avatars' expression ability are required to be prepared to apply the technology of sign language expression using avatars in the metaverse. In particular, depending on the characteristics of sign language, which is highly dependent on facial expressions, there is a big difference in the transmission power of sign language depending on whether the avatar's face can intuitively express a person's expression.

8.3 Advice for event organisers, speakers, interpreters and audience

As of now, the technology to recognise and express sign language has a very high technical difficulty. Of course, sign language recognition and expression technology are gradually developing through the development of artificial intelligence technology and the efforts of many sign language researchers. Still, it is practically difficult to expect an automated sign language communication system using artificial intelligence technology to be commercialised in the near future. Therefore, to support the interpreting of sign language in the metaverse, it is necessary to assume the intervention of human sign language interpreters.

With two methods: providing sign language and subtitling for speech, many people think that these two options satisfy the accessibility requirements of deaf and hard-of-hearing people. However, many

people who are deaf and hard-of-hearing often have difficulty interpreting speech language, even if written in letters. This is because sign language is not an alternative expression of speech language but a separate language system. Also, many acquired people who are deaf and hard-of-hearing do not understand sign language. Therefore, both methods are required to be provided.

Human sign language interpreters are preferred to the use of avatars. Human sign language interpreters can voluntarily offer wider language choices, as every country has at least one sign language. Avatars are suitable for one-way communication and are recommended to be pre-recorded. Avatar doesn't work well in real-time bi-directional situations. Human sign language interpreters work just like voice interpreters. They can interpret for a larger audience, such as individuals, meetings, and one-on-one arrangements. Avatars cannot translate for living humans, such as participants in large gatherings asking questions, in which actual human sign language interpreting is required in the two-way communication.

Bibliography

[b-ITU-T F.791]	Recommendation ITU-T F.791 (2015), *Accessibility terms and definitions.*
[b-ITU-T J.301]	Recommendation ITU-T J.301 (2014), *Requirements for augmented reality smart television systems.*
[b-ISO/IEC 2382]	ISO/IEC 2382:2015, *Information technology – Vocabulary.*
[b-ISO/IEC 71]	ISO/IEC 71:2014, *Guide for addressing accessibility in standards.*
[b-ISO/IEC 9241-11]	ISO/IEC 9241-11:2018, *Ergonomics of human-system interaction – Part 11: Usability: definitions and concepts.*
[b-ISO 9241-394]	ISO 9241-394: 2020, *Ergonomics of human-system interaction – Part 394: ergonomic requirements for reducing undesirable biomedical effects of visual induced motion sickness during watching electronic images.*
[b-Alley]	Alley, E. (2012). Exploring remote interpreting. *International Journal of Interpreter Education*, *4*(1), 10.
[b-Braun 2013]	Braun, S. (2013). Keep your distance? Remote interpreting in legal proceedings: A critical assessment of a growing practice1. *Interpreting*, *15*(2), 200-228.
[b-Braun 2020]	Braun, S. (2020). "Technology in interpreting," in *Routledge Encyclopedia of Translation Studies*. eds. M. Baker and G. Saldanha (London: Routledge), 569–574.
[b-DHACA]	Department of Home Affairs, Commonwealth of Australia. (2019). "Language Services Guidelines: Supporting Access and Equity for People with Limited English".
[b-Fantinuoli 2019]	Fantinuoli, C. (2019). *Interpreting and technology*. Language Science Press.
[b-Fantinuoli 2021]	Fantinuoli, C., & Prandi, B. (2021). Remote simultaneous interpreting and COVID-19: Conference interpreters' experiences and perspectives. In C. Fantinuoli & B. Prandi (Eds.), *Remote simultaneous interpreting: Theory and practice*. Springer Singapore. https://doi.org/10.1007/978-981-19-6680-4_7
[b-Hale]	Hale, S., Goodman-Delahunty, J., Martschuk, N., & Lim, J. (2022). Does interpreter location make a difference? A study of remote vs face-to-face interpreting in simulated police interviews. *Interpreting*, *24*(2), 221-253.
[b-Moser-Mercer]	Moser-Mercer, B. (2005). Remote interpreting: The crucial role of presence. *Bulletin vals-asla*, *81*, 73-97.
[b-Schiele]	Schiele, H., Bos-Nehles, A., Delke, V., Stegmaier, P., & Torn, R. J. (2022). Interpreting the industry 4.0 future: technology, business, society and people. *Journal of business strategy*, *43*(3), 157-167.
[b-Seeber]	Seeber, K. G., Keller, L., Amos, R., & Hengl, S. (2019). Expectations vs. experience: Attitudes towards video remote conference interpreting. *Interpreting*, *21*(2), 270-304.
[b-Wang]	Wang, B.H. & Pan, L.H. (2022). Delivering the first MA/PGDip Business and Public Service Interpreting programme in the UK. In Riccardo Morrato & Defeng Li (eds.) *Global Insights into Public Service Interpreting*. Routledge.

[b-AIIC 2020a]	AIIC (2020a). Covid-19 Distance Interpreting Recommendations for Institutions and DI Hubs. Available at: https://aiic.org/document/4839/AIIC%20Recommendations%20for%20Institutions_27.03.2020.pdf (Accessed Jun. 30, 2023).
[b-AIIC 2020b]	AIIC (2020b). AIIC Guidelines for Distance Interpreting (Version 1.0). Available at: https://aiic.org/document/4418/AIIC%20Guidelines%20for%20Distance%20Interpreting%20Version%201.0 (Accessed Jun. 30, 2023).
[b-AIIC 2021]	AIIC (2021). AIIC and distance interpreting. Available at: https://aiic.org/site/world/about/profession/distanceinterpreting (Accessed Jun. 30, 2023).
[b-APCI]	APCI Code of Conduct. https://apciinterpreters.org.uk/about-us/code-of-conduct/
[b-Ziaul]	Ziaul, M. (2019). Five key requirements of metaverse. https://metaversetroop.com/requirements-of-metaverse
[b-Hyperspace Academy]	Hyperspace Academy (2022). Interprefy real-time interpreting for multilingual metaverse events. https://hyperspace.mv/academy/interprefy-realtime-interpreting-for-multilingual-metaverse-events/?trk=public_post_comment-text
[b-MootUp]	MootUp (2022). MootUp by Hyperspace – Metaverse Events Platform Demo. https://youtu.be/m6NZrorzKrw
[b-CFAINT09]	NRPSI Standard on Undertake remote interpreting assignments (*CFAINT09*). https://www.ukstandards.org.uk/PublishedNos-old/CFAINT09.pdf
[b-Ashwani]	Ashwani K. (2023), What are the technical requirements for accessing the metaverse?. https://www.devopsschool.com/blog/what-are-the-technical-requirements-for-accessing-the-metaverse/#:~:text=The%20first%20thing%20you%20need,card%2C%20and%20a%20fast%20processor

_____