ITUPublications

International Telecommunication Union

Standardization Sector

ITU Focus Group Technical Specification

(10/2023)

ITU Focus Group on metaverse

(FG-MV)

FGMV-05

Requirements of accessible products and services in the metaverse: Part II – User perspective

Working Group 8: Sustainability, Accessibility & Inclusion

Technical Specification ITU FGMV-05

Requirements of accessible products and services in the metaverse: Part II – User perspective

Summary

This Technical Specification provides requirements on how to develop an accessible metaverse from a user perspective. This document considers the various metaverse components and the actions that users, regardless of their capabilities, may perform to access the metaverse, create an identity within the metaverse, navigate the metaverse and interact in the metaverse. The document is related to ITU FG-MV Technical Specification on "Requirements of accessible products and services in the metaverse: Part I – System design perspective" and provides requirements on the role of users in creating and assessing accessibility services.

Keywords

Accessibility, accessibility services, accessible products, accessible services, metaverse for all, personalisation, user interaction, user-centric

Note

This 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 Specification on "*Requirements of accessible products and services in the metaverse: Part II - User perspective*" approved at the third meeting of the ITU Focus Group on metaverse (FG-MV), held on 3-5 October 2023 in Geneva, Switzerland.

Acknowledgements

This Technical Specification was researched and written by Estella Oncins (Universitat Autònoma de Barcelona, Spain), Carlo Eugeni (University of Leeds, United Kingdom), Anna Matamala (Universitat Autònoma de Barcelona, Spain), and Paola Cecchi Dimeglio (Harvard University) as a contribution to the ITU Focus Group on metaverse (FG-MV). The development of this document was coordinated by Nevine Tewfik (Egypt) and Pilar Orero (UAB, Spain), as FG-MV Working Group 8 Co-Chairs, and by Yong Jick Lee (Center for Accessible ICT, Korea (Republic of)) and Paola Cecchi-Dimeglio (Harvard University) as Co-Chairs of Task Group on accessibility & inclusion.

Special thanks to Massimiliano Claps, Yong Jick Lee, Hideo Imanaka and Andy Quested, for their helpful reviews and contributions.

Additional information and materials relating to this Technical Specification can be found at: https://www.itu.int/go/fgmv. If you would like to provide any additional information, please contact Cristina Bueti at tsbfgmv@itu.int.

Editor:	Estella Oncins Universitat Autònoma de Barcelona Spain	**Tel:** +34610655149 **E-mail:** estella.oncins@uab.cat
Editor:	Carlo Eugeni University of Leeds United Kingdom	**Tel:** +393397229542 **E-mail:** c.eugeni@leeds.ac.uk

Editor:	Anna Matamala Universitat Autònoma de Barcelona Spain	**Tel:** +34619426305 **E-mail:** anna.matamala@uab.cat
Editor:	Paola Cecchi Dimeglio Harvard University United States	**Tel:** +1 (617) 498-6232 **E-mail:** cecchidimeglio@law.harvard.edu
WG8 Co-Chairs:	Nevine Tewfik MCIT Egypt	**Email**: ntewfik@mcit.gov.eg
	Pilar Orero UAB Spain	**Email**: pilar.orero@uab.cat
Task Group Co-Chair & Editor:	Paola Cecchi-Dimeglio Harvard University United States	**Email**: pcecchidimeglio@law.harvard.edu
Task Group Co-Chair:	Yong Jick Lee Center for Accessible ICT Korea (Republic of)	**Email**: ylee@caict.re.kr

Table of contents

Page

Page

Technical Specification ITU FGMV-05

Requirements of accessible products and services in the metaverse: Part II - User perspective

1 Scope

This Technical Specification provides requirements on how to develop accessible products and services in the metaverse taking a user perspective. The metaverse combines virtual reality, augmented reality, mixed reality, and other artificial intelligence (AI) and blockchain technologies, allowing users to perform a wide variety of tasks such as interact, work, socialise, entertain or transact in the virtual world.

The accessibility requirements provided in this Technical Specification will benefit users with various needs, including, but not limited to, people with learning and reading difficulties, neurodiverse users, people with sight or hearing loss, older people, and non-native language speakers.

This Technical Specification acknowledges the various needs and preferences of users.

Note: This Technical Specification complements ITU FG-MV Technical Specification on "Requirements of accessible products and services in the metaverse: Part I – System design perspective".

2 References

The following ITU-T Recommendations and other references contain provisions which, through reference in this text, constitute provisions of this Technical Specification. At the time of publication, the editions indicated were valid. All Recommendations and other references are subject to revision; users of this Technical Specification 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 published regularly. The reference to a document within this Technical Specification 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*.

3 Definitions

3.1 Terms defined elsewhere

This Technical Specification 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 Avatar [b-ISO/IEC 23005-4]: Entity that can be used as a (visual) representation of the user inside the virtual environments.

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

3.1.5 Easy-to-understand language [b-ISO/IEC 23859]: Any language variety which enhances comprehensibility. Note 1 to entry: Easy-to-understand language includes plain language, easy

language and any intermediate variety. These varieties share many recommendations, but the extent of comprehensibility is different as they address different user needs.

3.1.6 Extended reality: [b-ITU-T P.1320]: An environment containing real or virtual components or a combination thereof, where the variable X serves as a placeholder for any form of new environment (e.g., augmented, assisted, mixed, virtual or diminished reality).

3.1.7 Internet of Things [b-ITU-T Y.4000]: A global infrastructure for the information society, enabling advanced services by interconnecting (physical and virtual) things based on existing and evolving interoperable information and communication technologies.

3.1.8 Mixed reality [b-ISO/IEC 18038]: Merging of real and virtual worlds to generate new environments where physical and synthetic objects co-exist and interact.

3.1.9 Product [b-ISO/IEC 9241-11]: Item that is made or created by a person or machine.

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

3.1.11 System [b-ISO/IEC 9241-11]: Combination of interacting elements organized to achieve one or more stated purpose.

3.1.12 Task [b-ISO/IEC 9241-11]: Set of activities undertaken in order to achieve a specific goal.

3.1.13 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.

3.1.14 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 Specification

None.

4 Abbreviations and acronyms

AI Artificial Intelligence

5 Conventions

In this Technical Specification:

- The keywords "is required to" indicate a requirement which must be strictly followed and from which no deviation is permitted if conformance with this Recommendation is to be claimed.
- The keywords "is recommended" indicate a recommendation which is not absolutely required. Thus, this requirement need not to be fulfilled to claim conformance.
- The keywords “should” or “may” indicate an optional requirement which is permissible. 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 vendor. Rather, it means the vendor may optionally provide the feature and still claim conformance with the specification.

6 Introduction

The metaverse provides multisensory immersive experiences. Users can interact through avatars in a virtual world. Users can access virtual spaces that include content from a wide range of sectors, from education to health care, the arts, and many others.

The metaverse is open to all. However, some users may struggle to access the metaverse and navigate through it because accessibility features and services are not available. Due to its novelty, the metaverse is an opportunity to develop an accessible virtual world from the beginning, overcoming the barriers that are often encountered in the non-virtual world.

6.1 Motivation

The physical world poses barriers to many users. The metaverse offers a unique opportunity to create an accessible virtual world from its early stages of development. To this end, the needs of diverse users must be considered. This includes, but is not limited to, users who may experience challenges accessing audio or visual content, reading and understanding written language, understanding oral language, speaking, touching, using fine motion, and moving the body or parts of it.

6.2 User considerations

Users perform different actions in the metaverse:

- they access the metaverse,
- they create an avatar identity in the metaverse,
- they perceive content and navigate the metaverse, and
- they interact in the metaverse.

The metaverse includes virtual spaces, virtual content, and virtual people in the form of avatars.

EXAMPLE 1 A user accesses the metaverse, and navigates to a virtual cafeteria, where the user interacts with another virtual user in the form of an avatar. Later, the user moves to a virtual space where the user buys a ticket to access an online exhibition. At the end of the exhibition, the user interacts with a satisfaction survey.

7 Metaverse components from a user perspective

Four key components build up the metaverse: hardware, software, avatars, and content.

7.1 Hardware

The metaverse can be accessed through different types of hardware components, such as head-mounted displays, hand-based input devices, non-hand-based input devices, and motion input devices [b-Park].

7.2 Software

Different types of software components and platforms are used in the metaverse. Software is used for scene and object recognition, for sound and speech recognition, for scene and object generation, for sound and speech synthesis, and for motion rendering [b-Park], among other AI and machine learning applications.

7.3 Avatars

Users can take the form of an avatar and navigate through different virtual spaces. An avatar is a computer representation of a user, a user's persona, character or computer (AI).

Users can interact with other users through their avatars.

7.4 Content

The metaverse offers a wide variety of content, including products and services. This content is available in diverse virtual spaces.

Content is related to a wide range of fields, including but not limited to the arts, health, education, marketing, gaming, work, social, and government.

8 Requirements for processes on an accessible metaverse

8.1 General accessibility requirements

It is required that users, regardless of their capabilities, are able to access, perceive, navigate and interact in the metaverse.

It is required that users, regardless of their capabilities, are able to create their own identity in the metaverse.

EXAMPLE 2 A user with sight loss may want to be represented as a human with a cane or with a guide dog. Another blind user may want to be represented with a totally different avatar.

It is required that users are able to provide input and receive response by means of alternative options, such as speech, written text or haptics.

8.2 Accessing the metaverse

To access the metaverse, users need to have stable connectivity. It is recommended that stable connectivity is guaranteed to all users regardless of their situation.

The main hardware component for accessing the metaverse is head-mounted displays. Head-mounted displays can be a technically limiting barrier for some users.

It is required that users are given the possibility to access the metaverse through several types of devices, such as hand-based input devices, hands-free input devices, and motion input devices. It is required that users have alternative options, such as speech, written text or haptics, to authenticate themselves in order to access the metaverse.

It is required that interoperability between hardware components to access the metaverse and user assistive devices is ensured.

It is required that users are able to customize the hardware and software according to their needs.

8.3 Creating an avatar identity in the metaverse

Avatar identities, in the context of the metaverse, are digital self-representations that enable users to express, explore, and experience their identities in new, expansive ways. These identities can range from idealized versions of the user to entirely unique creations, allowing for innovative expressions of personal identity. They encapsulate the evolution of identity over time, mirroring personal growth, shifts in interests, and life events.

Offering a safe space for self-exploration, avatar identities let users experiment with different facets of their identity and understand themselves better. They promote diversity and inclusion by allowing individuals to sidestep societal biases linked to physical, sensory and/or cognitive attributes, such as appearance, age, race, gender, and disability. They form the primary mode of interaction, communication, and socialization within the metaverse.

Avatar identities also provide a platform for users to portray their cultural and socio-economic backgrounds, or to experiment with identities from different contexts. They cater to temporal self-identity, enabling users to display various aspects of their identity at different times, even representing past, present, or future selves. Avatar identities foster a playful environment for experimentation without the constraints of real-world consequences, allowing users to navigate the realms of gender, age, disability, race, and even species.

Avatars serve a crucial role in maintaining user anonymity online, acting as a protective barrier against unwanted disclosure of personal information. They are the main mode of interaction within the metaverse, influencing social communication and facilitating emotional safety and comfort in the virtual world. Lastly, avatar identities offer limitless opportunities for creativity and innovation in

their design, while also allowing contextual flexibility, letting users adapt their avatars to different situations - be it a professional meeting or a casual social interaction.

It is required that users have the choice and option to choose whether or not they wish to represent themselves using an avatar that mirrors their intersectionality, disabilities and diversity. Thereby offering a unique dimension of self-expression.

8.4 Navigating the metaverse

It is required that users are able to perceive content and move in the metaverse and between virtual spaces in the metaverse, regardless of their capabilities. This implies offering alternative navigational options and guaranteeing interoperability with various controllers.

EXAMPLE 3 A user may use a haptic controller that simulates white cane interactions.

EXAMPLE 4 A user may need to navigate without access to the audio, hence needing text-based navigation cues and alternatives to auditory verbal and non-verbal elements, including sound direction.

EXAMPLE 5 A user may need to rotate their view without physically moving their head.

When developing software components, it is required that accessibility features are included so that users can select those suited for their needs. Interoperability between software components and user-assistive devices is required.

EXAMPLE 6 A user may want to turn on a colour-blind mode whereas another one may prefer to increase visual contrast, magnify text, enlarge fonts, or use a screen reader.

8.5 Interacting in the metaverse

Interaction in the metaverse can happen through oral and written language but also through non-verbal communication such as facial expressions, gestures, or tone of voice.

It is required that users are able to provide input by means of alternative options, according to their needs. The options may include speech, keyboards, gestures, and eye-tracking.

It is required that users are able to receive responses by means of alternative options according to their needs.

EXAMPLE 7 A user with hearing loss may receive spoken response which is transferred into an alternative written format.

EXAMPLE 8 A user may give input in a written format which is then converted into an oral format with a personalised voice.

EXAMPLE 9 Users with diverse needs may have alternative options to authenticate themselves to the different available services.

9 Requirements on accessibility and translation services in the metaverse

9.1 General considerations

Users may need accessibility and translation services to fully enjoy the metaverse. These accessibility and translation services may be necessary to navigate through a virtual space, to communicate with another avatar or to access and interact with a virtual product or service.

It is required that personally identifiable information protection and safety are ensured for all users, regardless of their capabilities, when using accessibility and translation services in the metaverse.

9.2 Subtitling /captioning

Subtitles/captions provide written text alternatives to spoken words. Subtitles/captions can be in the same language as the spoken words (intralinguistic) or in a different language (interlinguistic).

Subtitles/captions can also include written alternatives to non-speech information such as character, channel, language identification, paralinguistic elements, sound, silence, and music.

Subtitles/captions follow some presentation restrictions to allow users to adequately read the subtitles/captions and enjoy the visuals. These restrictions refer to various parameters such as number of lines, line-breaks, number of characters per second, number of characters per line, and image synchronisation.

Subtitles/captions benefit users who cannot access the audio content. Subtitles/captions also benefit users who can access the audio content but cannot fully understand the spoken language.

It is required that users are able to activate and deactivate subtitles/captions.

It is required that users are able to select different subtitling features such as language, font size, font type, font colour, contrast, text alignment, and placement.

9.3 Transcripts

Transcripts provide a written, word-by-word alternative to spoken words. Transcripts can also include written alternatives to non-speech audio information.

Transcripts are same-language and do not need to comply with the presentation restrictions followed by subtitles/captions.

Transcripts can be interactive and highlight text phrases as they are spoken.

Transcripts benefit users who cannot access the audio content. Interactive transcripts benefit users with neurodiverse needs, such as reading difficulties.

It is required that users are able to activate and deactivate transcripts.

It is required that users are able to select different transcription features such as placement, font size, font type, font colour, contrast, and alignment.

9.4 Audio description

Audio descriptions provide an auditory spoken alternative to visual content. An alternative name for this accessibility service is descriptive audio.

Audio description may apply to dynamic (i.e., a video) or static (i.e., an object) content.

Audio introductions may be used to complement audio descriptions. Audio introductions are audio texts that provide a summary of the visual content before accessing it.

Audio descriptions and audio introductions benefit users who cannot access the visual content.

It is required that users are able to activate and deactivate audio descriptions and audio introductions.

It is required that users are able to select different audio parameters such as language, volume, and voice type.

Audio description may be combined with audio subtitles.

9.5 Audio subtitles

Audio subtitles are subtitles which are voiced over. Alternative names for this accessibility service are audio captions, spoken subtitles, spoken captions, and audio text.

Audio subtitles provide an alternative to written subtitles and on-screen text and are received auditorily.

Audio subtitles benefit users who cannot see or cannot read the subtitles and cannot understand the source spoken language.

It is required that users are able to activate and deactivate audio subtitles.

Audio subtitles can be generated by a human voice or through text-to-speech technologies.

It is required that users are able to select different audio parameters such as language, volume, and voice type.

9.6 Sign language interpreting

Sign language interpreting provides a visual gestural alternative to oral language.

Sign language interpreting benefits users who cannot access the audio and understand sign language.

It is required that users are able to activate and deactivate sign language interpreting.

It is required that users are able to select different features such as sign language, positioning of the sign language interpreter, choice of avatar.

9.7 Oral language interpreting

Oral language interpreting provides an oral language translation from a source language to a target language.

Oral language interpreting benefits users who cannot understand the source language.

It is required that users are able to activate and deactivate oral language interpreting.

It is required that users are able to select different features of oral language interpretation such as language and volume.

9.8 Easy-to-understand language

Easy-to-understand language refers to different language simplified varieties that enhance comprehensibility. It ranges from Easy Language (or Easy-to-Read), which is the most simplified variety, to Plain or Clear Language, which is closer to standard language.

It is recommended that Plain Language is used in the metaverse. It is recommended that Easy Language alternatives are also provided. Paratextual elements such as images or glosses may be used to enhance comprehension.

Easy Language benefits users who have difficulties reading or understanding language.

It is required that users are able to activate and deactivate Easy Language content.

It is required that users are offered an easy way to go to a “safe place” in the metaverse in case they do not understand or feel overwhelmed.

9.9 Revoicing

Revoicing refers to the translation of voicing of a source language audio track into a target language audio track. This revoicing can be done through dubbing or voice-over.

In dubbing, the original voices are replaced by the target language voices, which are synchronised with the lip movements of visible characters and with the length of the utterances in the source content.

In voice-over, the target language voices overlap with the original voices, which can still be heard with a lower volume. There is no lip synchronisation.

There are differences across countries in terms of dubbing and voice-over implementation and usage.

Dubbing and voice-over benefit users who can access the audio but cannot understand the source language.

It is required that users are able to activate and deactivate dubbing and voice-over tracks.

10 Activating accessibility and translation services in the metaverse

It is required that users are provided with an easy-to-understand way of selecting and changing between different accessibility and translation services.

It is required that users are provided with an easy way of selecting the language of the services.

10.1 Personalisation

It is required that user interfaces in the user device personalise the experience to meet end user's needs and preferences in terms of accessibility and translation services.

10.2 Customisation

It is required that users are able to make choices in the selection and display of accessibility and translation services to tailor the experience to their own needs and preferences.

11 Inclusion of users

11.1 Inclusion of users in creating accessibility services in the metaverse

It is recommended that diverse user profiles are included in the process of creating accessibility services where possible.

EXAMPLE 10 Users with disabilities can be involved in focus group discussions and be employed as consultants.

11.2 Inclusion of users in evaluating accessibility services in the metaverse

It is recommended that users be involved in the evaluation of accessibility services.

EXAMPLE 11 Users with learning disabilities can assess how easy to read and understand is content in the metaverse.

12 Identification of accessibility and translation services

It is recommended that the availability of the different accessibility and translation services is identified by a standardised and easily recognisable logo.

It is required that instructions about how to activate, use, and deactivate one or more accessibility and translation services are available in the end-user's language and in an easy-to-understand format. It is required that an easy way to request assistance is available.

Bibliography

[b-ITU-R BT.2207-6] ITU-R BT.2207-6 (2022), *Accessibility to broadcasting services for persons with disabilities*.

[b-ITU-R BT.2447-2] ITU-R BT.2447-2 (2021), *Artificial intelligence systems for programme production and exchange*.

[b-ITU-R BT.2420-5] ITU-R BT.2420-5 (2022), *Collection of usage scenarios of advanced immersive sensory media systems.*

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

[b-ITU-T Y.4000] Recommendation ITU-T Y.4000/Y.2060 (2016), *Overview of the Internet of things*.

[b-ITU-T P.1320] Recommendation ITU-T P.1320 (2022), *Quality of experience assessment of extended reality meetings*.

[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-ISO/IEC 2382] ISO/IEC 2382:2015, *Information technology – Vocabulary*.

[b-ISO/IEC 9241-11] ISO/IEC 9241-11:2018, *Ergonomics of human-system interaction – Part 11: Usability: definitions and concepts.*

[b-ISO/IEC 71] ISO/IEC 71:2014, *Guide for addressing accessibility in standards*.

[b-ISO/IEC 18038] ISO/IEC 18038: 2020, *Information technology – Computer graphics, image processing and environmental representation – Sensor representation in mixed and augmented reality.*

[b-ISO/IEC 23859] ISO/IEC 23859:2023, *Information technology – User interfaces – Requirements and recommendations on making written text easy to read and understand.*

[b-ISO/IEC 23005-4] ISO/IEC 23005-4:2018, *Information technology — Media context and control — Part 4: Virtual world object characteristics.*

[b-Park] Park, Sang-Min & Young-Gab, Kim (2022), *A Metaverse: Taxonomy, Components, Applications, and Open Challenges*. IEEEAccess, 10, 4209-4251. doi: 10.1109/ACCESS.2021-3150175.

[b-W3C] World Wide Web Consortium (2021), *XR Accessibility User Requirements.* Available [viewed 18 September 2023] at: https://www.w3.org/TR/xaur/