|  |
| --- |
| 3GPP TS 22.156 V1.0.0 (2023-09) |
| Technical Specification.  |
| 3rd Generation Partnership Project;Technical Specification Group Services and System Aspects;Mobile Metaverse Services;Stage 1 (Release 19) |
|   |
|  |  |
| The present document has been developed within the 3rd Generation Partnership Project (3GPP TM) and may be further elaborated for the purposes of 3GPP.The present document has not been subject to any approval process by the 3GPPOrganizational Partners and shall not be implemented.This Specification is provided for future development work within 3GPPonly. The Organizational Partners accept no liability for any use of this Specification.Specifications and Reports for implementation of the 3GPP TM system should be obtained via the 3GPP Organizational Partners' Publications Offices. |

|  |
| --- |
|  |
| ***3GPP***Postal address3GPP support office address650 Route des Lucioles - Sophia AntipolisValbonne - FRANCETel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16Internethttp://www.3gpp.org |
| ***Copyright Notification***No part may be reproduced except as authorized by written permission.The copyright and the foregoing restriction extend to reproduction in all media.© 2023, 3GPP Organizational Partners (ARIB, ATIS, CCSA, ETSI, TSDSI, TTA, TTC).All rights reserved.UMTS™ is a Trade Mark of ETSI registered for the benefit of its members3GPP™ is a Trade Mark of ETSI registered for the benefit of its Members and of the 3GPP Organizational PartnersLTE™ is a Trade Mark of ETSI registered for the benefit of its Members and of the 3GPP Organizational PartnersGSM® and the GSM logo are registered and owned by the GSM Association |

Contents

Foreword 4

1 Scope 5

2 References 5

3 Definitions of terms, symbols and abbreviations 6

3.1 Terms 6

3.2 Abbreviations 7

4 Overview 8

5. Functional service requirements 8

5.1 General requirements 8

5.1.1 Operational efficiency, exposure, and coordination 8

5.1.1.1 Description 8

5.1.1.2 Requirements 8

5.2 Specific functional areas 9

5.2.1 Localized mobile metaverse service 9

5.2.1.1 Description 9

5.2.1.2 Requirements 10

5.2.2 Avatar-based real-time communication 11

5.2.2.1 Description 11

5.2.2.2 Requirements 11

5.2.3 Digital asset management 12

5.2.3.1 Description 12

5.2.3.2 Requirements 12

6 Performance requirements 12

6.1 Description 12

6.2 Performance requirements 14

7 Security, authorization and privacy 16

7.1 Description 16

7.2 Requirements 16

7.2.1 General 16

7.2.2 Localized mobile metaverse service 16

7.2.3 Avatar-based real-time communication 16

7.2.4 Digital asset management 16

8 Charging aspects 16

8.1 Description 16

8.2 Requirements 17

8.2.1 Localized mobile metaverse service 17

8.2.2 Avatar-based real-time communication 17

8.2.3 Digital asset management 17

Annex A (informative): Mobile metaverse services 17

A.0 Introduction 17

A.1 Consumer mobile metaverse services 18

A.2 Enterprise mobile metaverse services 18

A.3 Industrial mobile metaverse services 19

A.4 Common aspects 19

Annex <B> (informative): Bibliography 20

Annex <C> (informative): Change history 21

# Foreword

This Technical Specification has been produced by the 3rd Generation Partnership Project (3GPP).

The contents of the present document are subject to continuing work within the TSG and may change following formal TSG approval. Should the TSG modify the contents of the present document, it will be re-released by the TSG with an identifying change of release date and an increase in version number as follows:

Version x.y.z

where:

x the first digit:

1 presented to TSG for information;

2 presented to TSG for approval;

3 or greater indicates TSG approved document under change control.

y the second digit is incremented for all changes of substance, i.e. technical enhancements, corrections, updates, etc.

z the third digit is incremented when editorial only changes have been incorporated in the document.

In the present document, modal verbs have the following meanings:

**shall** indicates a mandatory requirement to do something

**shall not** indicates an interdiction (prohibition) to do something

The constructions "shall" and "shall not" are confined to the context of normative provisions, and do not appear in Technical Reports.

The constructions "must" and "must not" are not used as substitutes for "shall" and "shall not". Their use is avoided insofar as possible, and they are not used in a normative context except in a direct citation from an external, referenced, non-3GPP document, or so as to maintain continuity of style when extending or modifying the provisions of such a referenced document.

**should** indicates a recommendation to do something

**should not** indicates a recommendation not to do something

**may** indicates permission to do something

**need not** indicates permission not to do something

The construction "may not" is ambiguous and is not used in normative elements. The unambiguous constructions "might not" or "shall not" are used instead, depending upon the meaning intended.

**can** indicates that something is possible

**cannot** indicates that something is impossible

The constructions "can" and "cannot" are not substitutes for "may" and "need not".

**will** indicates that something is certain or expected to happen as a result of action taken by an agency the behaviour of which is outside the scope of the present document

**will not** indicates that something is certain or expected not to happen as a result of action taken by an agency the behaviour of which is outside the scope of the present document

**might** indicates a likelihood that something will happen as a result of action taken by some agency the behaviour of which is outside the scope of the present document

**might not** indicates a likelihood that something will not happen as a result of action taken by some agency the behaviour of which is outside the scope of the present document

In addition:

**is** (or any other verb in the indicative mood) indicates a statement of fact

**is not** (or any other negative verb in the indicative mood) indicates a statement of fact

The constructions "is" and "is not" do not indicate requirements.

# 1 Scope

The present document provides Stage 1 normative service and performance requirements for diverse service enablers to enhance XR-based services. The term 'metaverse' in the title of the present document embraces the broader implications of AR and VR.

Service enablers considered in this document include:

- Localized mobile metaverse service functionality;

- Avatar-based real-time communication functionality;

- Digital asset management functionality;

- Operation efficiency, exposure, coordination of mobile metaverse services.

# 2 References

The following documents contain provisions which, through reference in this text, constitute provisions of the present document.

- References are either specific (identified by date of publication, edition number, version number, etc.) or non‑specific.

- For a specific reference, subsequent revisions do not apply.

- For a non-specific reference, the latest version applies. In the case of a reference to a 3GPP document (including a GSM document), a non-specific reference implicitly refers to the latest version of that document *in the same Release as the present document*.

[1] 3GPP TR 21.905: "Vocabulary for 3GPP Specifications".

[2] 3GPP TS 22.228: "Service requirements for the Internet Protocol (IP) Multimedia core network Subsystem (IMS)".

[3] ITU-T Recommendation Y.3090 (02/22): "Digital twin network - Requirements and architecture" (https://www.itu.int/rec/T-REC-Y.3090-202202-I).

[4] 3GPP TS 22.101: "Service principles".

[5] ITU-T Recommendation F.703 (11/00): "Multimedia conversational services".

[6] European Commission: "Shaping Europe's digital future", <https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/shaping-europe-digital-future_en>.

[7] 3GPP TS 22.261: "Service requirements for the 5G system".

[8] 5GAA: "C-V2X Use Cases Volume II: Examples and Service Level Requirements", 5G Automobile Association White Paper, https://5gaa.org/wp-content/uploads/2020/10/5GAA\_White-Paper\_C-V2X-Use-Cases-Volume-II.pdf <accessed 02.09.22>.

[9] O. Holland et al.: "The IEEE 1918.1 "Tactile Internet" Standards Working Group and its Standards," Proceedings of the IEEE, vol. 107, no. 2, Feb. 2019."

[10] A. Ebrahimzadeh, M. Maier and R. H. Glitho: "Trace-Driven Haptic Traffic Characterization for Tactile Internet Performance Evaluation," 2021 International Conference on Engineering and Emerging Technologies (ICEET), 2021, pp. 1-6.

# 3 Definitions of terms, symbols and abbreviations

## 3.1 Terms

For the purposes of the present document, the terms given in 3GPP TR 21.905 [1] and the following apply. A term defined in the present document takes precedence over the definition of the same term, if any, in 3GPP TR 21.905 [1].

**avatar:** a digital representation specific to media that encodes facial (possibly body) position, motions and expressions of a person or some software generated entity.

**Conference**: An IP multimedia session with two or more participants. Each conference has a "conference focus". A conference can be uniquely identified by a user. Examples for a conference could be a Telepresence or a multimedia game, in which the conference focus is located in a game server.

NOTE 1: This definition was taken from 3GPP TS 22.228 [2].

**Conference Focus**: The conference focus is an entity which has abilities to host conferences including their creation, maintenance, and manipulation of the media. A conference focus implements the conference policy (e.g. rules for talk burst control, assign priorities and participant’s rights).

NOTE 2: This definition was taken from 3GPP TS 22.228 [2].

**digital asset**: digitally stored information that is uniquely identifiable and can be used to realize value according to their licensing conditions and applicable regulations. Examples of digital assets include digital representation (avatar), software licenses, gift certificates, tokens and files (e.g. music files) that have been purchased. This is not an exhaustive list of examples.

**digital representation:** the mobile metaverse media associated with the presentation of a particular virtual or physical object. The digital representation could present the current state of the object. One example of a digital representation is an avatar, see Annex A.

**digital twin:** A real-time representation of physical assets in a digital world.

NOTE 3: This definition was taken from ITU-T Recommendation Y.3090 [3].

**digital wallet**: one type of digital asset container, also known as e-wallet or mobile wallet. It is a software application that securely stores digital credentials typically part of personal data, such as payment information, loyalty cards, tickets, and other digital assets. It allows users to make electronic transactions, such as payments and transfers, conveniently and securely using their digital credentials.

NOTE 4: Digital wallets typically employ encryption and authentication mechanisms to protect the stored information and ensure the security of transactions.

**gesture:** a change in the pose that is considered significant, i.e. as a discriminated interaction with a mobile metaverse service.

**immersive:** a characteristic of a service experience or AR/MR/VR media, seeming to surround the user, so that they feel completely involved.

**localization**: A known location in 3 dimensional space, including an orientation, e.g. defined as pitch, yaw and roll.

**location related service experience:** user interaction and information provided by a service to a user that is relevant to the physical location in which the user accesses the service.

**location agnostic service experience:** user interaction and information provided by a service to a user that has little or no relation to the physical location in which the user accesses the service. Rather the service provides interaction and information concerning either a distant or a non-existent physical location.

**mobile metaverse media:** media communicated or enabled using the 5G system including audio, video, XR (including haptic) media, and data from which media can be constructed (e.g., a 'point cloud' that could be used to generate XR media.)

**mobile metaverse:** the user experience enabled by the 5G system of interactive and/or immersive XR media, including haptic media.

**mobile metaverse server:** an application server that supports one or more mobile metaverse services to a user access by means of the 5G system.

**mobile metaverse service:** the service that provides a mobile metaverse experience to a user by means of the 5G system.

**pose:** the relative location, orientation and direction of the parts of a whole. The pose can refer the user, specifically used in terms of identifying the position of a user's body. The pose can also also refer to an entity or object (whose parts can adopt different locations, orientations, etc.) that the user interacts with by means of mobile metaverse services.

**service information**: this information is out of scope of standardization but could contain, e.g., a URL, media data, media access information, etc. This information is used by an application to access a service.

**spatial anchor**: an association between a location in space (three dimensions) and service information that can be used to identify and access services, e.g. information to access AR media content.

**spatial map**: A collection of information that corresponds to space, including information gathered from sensors concerning characteristics of the forms in that space, especially appearance information.

**spatial mapping service:** A service offered by a mobile network operator that gathers sensor data in order to create and maintain a Spatial Map that can be used to offer customers Spatial Localization Service.

**spatial localization service:** A service offered by a mobile network operator that can provide customers with Localization.

**User Identifier:** a piece of information used to identify one specific User Identity in one or more systems.

NOTE 5: This definition was taken from 3GPP TS 22.101 [4].

**User Identity**: information representing a user in a specific context. A user can have several user identities, e.g. a User Identity in the context of his profession, or a private User Identity for some aspects of private life.

NOTE 6: This definition was taken from 3GPP TS 22.101 [4].

**User Identity Profile:** A collection of information associated with the User Identities of a user.

NOTE 7: This definition was taken from 3GPP TS 22.101 [4].

## 3.2 Abbreviations

For the purposes of the present document, the abbreviations given in 3GPP TR 21.905 [1] and the following apply. An abbreviation defined in the present document takes precedence over the definition of the same abbreviation, if any, in 3GPP TR 21.905 [1].

AI Artificial Intelligence

CCTV Closed Circuit TeleVision

DoF Degrees of Freedom

DVE Distributed Virtual Environment

FACS Facial Action Coding System

FOV Field Of View

LiDAR Light Detection And Ranging

MR Mixed Reality

VRU Vulnerable Road User

XR eXtended Reality

# 4 Overview

The term metaverse has been used in various ways to refer to the broader implications of AR and VR. Metaverse in diverse sectors evokes a number of possible immersive user experiences. Products and services can emerge once virtual reality and augmented reality become commonly available and find application in our work, leisure and other activities.

In addition to services that offer location-independent user experiences, this feature also considers content and services that are associated or applicable only in a particular location. These metaverse services are mobile in the sense that mobile users are able to interact with services anywhere and in particular when located where specific services are offered.

 Requirements for diverse service enablers are introduced to the 5G system to support these services, including avatar call functionality, coordination of mobile metaverse services, digital asset management and support for spatial anchors.

# 5. Functional service requirements

## 5.1 General requirements

### 5.1.1 Operational efficiency, exposure, and coordination

#### 5.1.1.1 Description

These capabilities whose service requirements are defined in clause 5.1.1.2 enable diverse mobile metaverse services.

One important class of services involves several users who take part in mobile metaverse services simultaneously, for example, to support a 'virtual sport event' where some of the environment or objects in the match are virtual, that is, they are produced by an application that provides the user with XR media. Users could be local (in the same location) or remote and have a service experience that is immersive and meets the expectations set by the interactive activity.

Another important class of services are those that require coordination of diverse service data flows of sensor data and media in order to satisfy the needs of a digital twin or situational awareness service.

The service requirements in this clause correspond to means by which the 5G system provides access to digital assets and communication services for mobile metaverse services so that:

- the service experience of users of the same service are compatible and consistent;

- the services can operate over a sufficient duration for devices with constrained energy storage;

- the services can communicate efficiently to a large number of authorized users;

- the communication performance for specific mobile metaverse services to specific users can be monitored and exposed to third parties.

#### 5.1.1.2 Requirements

Subject to operator policy, the 5G system shall support a mechanism that enables flexible adjustment of communication services based on e.g., the type of devices (e.g., wearables), or communication duration (e.g., more than one hour), such that the services can be operated with reduced energy utilization.

NOTE 1: Metaverse service experience over an extended period of time (e.g., 2h) requires significant power consumption by the UE. In some cases, a device with no external power supply cannot sustain downloading and rendering of media over a long interval, e.g., for the duration of an entire feature film or athletic event.

The 5G system shall provide a means to associate and coordinate data flows related to one or multiple UEs e.g., associated with the same object in digital twin applications provided by the mobile metaverse service.

Subject to operator policy, regulatory requirements and user consent, the 5G system (including IMS) shall be able to expose network performance information (e.g., observed or predicted bitrate, latency or packet loss) related to one or more users to an authorized third party metaverse application.

NOTE 2: The network performance information can be per UE and can take into account all available access network types, i.e., 3GPP and non-3GPP.

Subject to operator policy, the 5G system (including IMS) shall support a mechanism, including enabling one or more authorized third party(ies) to coordinate multiple service data flows of a single mobile metaverse service delivered to/from one or more UE(s). Multiple UEs may be associated with one user/location or different users at different locations potentially using different access networks, i.e., 3GPP and non-3GPP.

NOTE 3: Coordination refers to the ability to provide an acceptable level of user experience for a given service, e.g., based on latency and synchronization constraints (due to multiple sources or long distance between UEs/users). This can be based on a quantitative bound.

NOTE 4: It is not assumed that it is always possible to coordinate and provide the same capabilities regardless of whether 3GPP or non-3GPP access is used.

The 5G system shall enable the coordination of diverse media, transmitted to a UE from one or more mobile metaverse services associated with a physical location, to be combined to form a localized service experience.

Subject to operator policy, the 5G system shall support exposure mechanisms enabling an authorized third party to determine one or more subscribers to whom mobile metaverse media can be distributed in a resource efficient manner.

Subject to operator policy and user consent, the 5G system shall support a means to provide resource efficient communication of third party mobile metaverse media to one or more subscribers.

## 5.2 Specific functional areas

### 5.2.1 Localized mobile metaverse service

#### 5.2.1.1 Description

Localized mobile metaverse services are immersive and integrated into a user's ordinary experiences. Such service experiences are location-related and can include presentation of AR, MR media.

Localized experiences are effectively present in the user's environment, so that the mobile metaverse media provided for a given mobile metaverse service is both appropriate to and integrated with both the physical world and with mobile metaverse media content displayed. Localized mobile metaverse services can be associated with specific places (3D locations in the physical world). The association between these places and service information is termed a spatial anchor.

Spatial anchors enable mobile metaverse services to be discovered and accessed, if the user is authorized. For example, the service information can convey the mobile metaverse server access information. When the user's application accesses the mobile metaverse service, the media associated with the service can be obtained by the user.



Figure 5.2.1.1-1: Services offering relevant information are anchored in space

Spatial anchors can associate diverse information with spatial location, beyond access control and access information of mobile metaverse services. Type of service information can also allow a user to discover appropriate spatial anchors, e.g., when the user seeks restaurants.

Spatial anchors can be defined by third parties, e.g., service providers, to offer relevant localized services, e.g., associated with specific items or features in their place of business. This information and its associated authorization information, determining who can discover the spatial anchor, can be managed - created, deleted and modified.

Users' localization, that is their precise location and orientation, is important in order to discover spatial anchors. The 5G system offers a spatial localization service to determine this information. Using sensor data related to the user's location, the 5G system can identify where the user is. This is possible by means of processing the sensor data as well as a spatial map.

The spatial map is created using processed sensor data. The 5G system supports a spatial mapping service to customers that, for example, want to offer mobile metaverse services associated with spatial anchors on their premises. Creation of a spatial map for a location makes localization there possible, as well as assignment of spatial anchors in that location.

#### 5.2.1.2 Requirements

Subject to operator policy, the 5G system shall provide a means to define and expose to an authorized third party a spatial anchor, i.e., an association between a physical location (a point or volume in three-dimensional space) and service information.

NOTE 1: Service information can include information to enable users to discover and access services, e.g., type of service, URLs, configuration data, the distance between the user and the spatial anchor, etc.

Subject to operator policy, the 5G system shall enable an authorized third party to request the information associated with a specific spatial anchor.

NOTE 2: How the service and location information is used by the third party to access a mobile metaverse server and the AR media itself is out of scope of this requirement.

Subject to operator policy, regulatory requirements and user consent, the 5G system shall provide a means for a UE to provide sensor data, (e.g., from UE sensors, cameras, etc.) to the network in order to derive localization information, e.g., to produce or modify a spatial map or discover or find spatial anchors. The 5G system shall enable an authorized third party to obtain all the spatial anchors in a given three-dimensional area.

NOTE 3: How an authorized third party identifies which three-dimensional area to request spatial anchors in is not in scope of the 3GPP standard. Spatial localization and mapping information could be used to identify areas of interest.

Subject to operator policy and regulatory requirements, the 5G system shall support mechanisms to expose a spatial map or derived localization information to authorized third parties.

### 5.2.2 Avatar-based real-time communication

#### 5.2.2.1 Description

A user can take part in mobile metaverse services that provide digital representations of several other users simultaneously, for example, to support a Conference using XR media. As these services are interactive and immersive, the 5G system provides a means so that the experience of each user of the same service is compatible and consistent. Users can participate together in this way, whether some of those in the Conference service are located in the same place: these users can experience remote users as AR or MR media. This media is conversational - meaning that all parties can participate, and real time - meaning that all users perceive each other's actions effectively simultaneously.

The 5G system supports a means by which user's pose, gestures and expressions are captured as input for the conversational mobile metaverse service. Devices (e.g., UEs) can capture this information in a standardized form, which is used for the creation of an animated digital representation of the user (e.g., 3D avatar) that can be presented to other users. The system supports communication of this animated digital representation as "avatar media" as well as audio and other media as needed. The 5G system supports the creation of avatar media. Privacy and user consent is needed for sensor data used to capture a user's pose, gestures and (facial) expressions is sent from the UE to the network encoded as avatar media to be rendered.

Capabilities of UEs differ. For example: some UEs can render avatar media, others video, others only text. To support interactive avatar communication, media transcoding can occur in the 5G network. Similarly, the display capabilities of UEs also differ. Avatar, video or text data can be transcoded as appropriate to be displayed to the user.

Capabilities of users also vary. To support accessibility for those with disability due to physical (e.g., impaired hearing, sight, etc.), environmental (e.g., in a noisy environment), conversational media can be transcoded. This is consistent with the objective for Total Conversation [5], clause 4.5. Avatar functionality provides new options, as media can be transcoded to and from an avatar call.

#### 5.2.2.2 Requirements

The 5G system shall support 5G CN to provide real-time feedback in support of conversational XR communication among multiple users simultaneously.

NOTE 1: The feedback can include information such as network condition, achieved QoS. Such information can be used by the IMS, for example, to trigger the codec negotiation.

Subject to operator policy and user consent, the 5G system (including IMS) shall support multimedia conversational communications between two or more users including transfer of real time avatar media and audio media.

NOTE 2: Avatar media can be transmitted on both uplink and downlink.

NOTE 3: Confidentiality of the data used to produce the avatar (e.g., from the UE cameras, etc.) is assumed.

Subject to operator policy and user consent, the 5G system (including IMS) shall support change of media types between video and avatar media for parties of a multimedia conversational communication.

Subject to operator policy, the 5G system (including IMS) shall support transcoding between media such as text, video and avatar media in multimedia conversational communications.

NOTE 4: Text, video or other media could allow a party to control the appearance of its avatar, e.g. to express behaviour, movement, affect, emotions, etc.

NOTE 5: The transcoding of media enables avatar communication, e.g., in scenarios in which UE participating in an IMS call or other service does not support e.g., FACS, encoding avatar media, generating avatar media, etc.

Subject to operator policy, regulatory requirements and user consent, the 5G system (including IMS) shall support the capabilities of rendering the avatar based on the body movement information (e.g., body motion or facial expression) of a human user.

Subject to operator policy, regulatory requirements and user consent, the 5G system (including IMS) shall support the encoding of sensor data capturing the facial expression and movement and gestures of a person, in a standard form.

### 5.2.3 Digital asset management

#### 5.2.3.1 Description

Mobile metaverse services can depend upon information that is associated with the user, e.g., User Identifiers and personal data that are commonly required and represented in a machine-readable format. The requirements as described in 3GPP TS 22.101 [4] clause 26a apply for identification of users. These can be used to provide proof for regulatory constrained service, e.g., proof of residential address for services that are restricted to local residents.

Further, the services can benefit from common information, such as avatar parameters and configuration information, so that a user's digital representation is consistent across different applications.

Finally, some more specific information used by different services can also be shared in different mobile metaverse services and be considered 'digital assets' in that the user needs or could benefit from having this information available when access mobile metaverse services. According to regulation, this information can be considered as personal data.

An example of such service is the EU digital wallet initiative [6]. Both the digital wallet and the digital asset management functionality described in the present document emphasize the need for security, privacy and control over access to authorized parties.

#### 5.2.3.2 Requirements

Subject to operator policy, regulatory requirements and user consent, the 5G system shall be able to provide functionality to store digital assets associated with a user, and to remove such digital assets associated with a user.

Subject to operator policy, regulatory requirements and user consent, the 5G system shall provide a means to allow a user to securely access and update their digital assets.

Subject to operator policy and user consent, the 5G system shall be able to allow an authorized third party to retrieve the digital asset(s) associated with a user, e.g., when the user accesses a specific application.

NOTE: When a user accesses an immersive mobile metaverse service, the authorized third party (service provider) could obtain relevant digital assets of a user associated with that service.

# 6 Performance requirements

## 6.1 Description

The performance requirements shown in table 6.2-1 feature exemplary use cases of mobile metaverse services that require communication services with specific performance levels.

5G-enabled Traffic Flow Simulation and Situation Awareness is a use case in which the real conditions of a road including vehicles and other factors are captured with sensors, modelled in a simulation and used to provide guidance for vehicles and users for efficiency and safety. This is a specific example of a broad category of 'situational awareness' services that capture 'virtual representations' of the real world to then advise or control actions taken in the real world.

Collaborative and concurrent engineering is a form of Conference in which multiple users participate, both together at the same site and remotely, to interact with virtual and physical objects collectively. The use case considers audio, video and haptic interaction.

Metaverse-based Tele-Operated Driving is a use case that enables remote user actuation of equipment, specifically remote-controlled driving in a hazardous environment. The interaction of the user and the remote equipment is facilitated by a digital twin representing the vehicle and the environment it operates in. The status of the digital twin is determined by sensors in the vehicle's vicinity and carried by the vehicle.

The performance requirements for other AR/VR/XR services are included in 3GPP TS 22.261 [7] including:

- clause 7.6.1 AR/VR;

- clause 7.11 KPIs for tactile and multi-modal communication service.

## 6.2 Performance requirements

The 5G system shall support various mobile metaverse services with the following KPIs.

NOTE: Unless stated otherwise, the "Max allowed end-to-end latency" refers to the maximum transmission delay expected between a UE and the mobile metaverse server or vice-versa.

Table 6.2-1: Performance requirements for Mobile Metaverse Services

|  |  |  |  |
| --- | --- | --- | --- |
| Use Cases | Characteristic parameter (KPI) | Influence quantity | Remarks |
| Max allowed end-to-end latency | Service bit rate: user-experienced data rate | Reliability | Area Traffic capacity | Message size (byte) | Transfer Interval | Positioning accuracy | UE Speed | Service Area |
| 5G-enabled Traffic Flow Simulation and Situational Awareness(NOTE 2) | [5-20] ms (NOTE 1) | [10~100] Mbit/s [8](NOTE 6) | > 99.9% | [TBD] Tbit/s/km2 (NOTE 5) | - | 20~100 ms(NOTE 3) | - | < 250 km/h | City or Country wide(NOTE 4) | UL |
| Collaborative and concurrent engineering | [≤10] ms[9](NOTE 7) | [1-100] Mbit/s[9] | [> 99.9%] [9] | [1.55] Tbit/s/km2 (NOTE 8) | Video: 1500Audio: 100[9] | - | - | Stationary or Pedestrian | typically < 100 km2(NOTE 9) | UL and DL audio/video |
| [5] ms UL [1-50] ms DL[9](NOTE 7) | [<1] Mbit/s [9] | [> 99.9%] (without compression)[> 99.999%] (with compression (NOTE 10))[10] | [2.25] Tbit/s/km2 (NOTE 8) | 1 DoF: 2-8 3 DoFs: 6-24 6 DoFs: 12-48 [9] | 0.25-10 ms [9] |  |  |  | UL and DL haptic feedback |
| Metaverse-based Tele-Operated Driving(NOTE 16) | [100] ms [8] (NOTE 11) | [10~50] Mbit/s [8]  | 99% [8] | [~360] Mbit/s/km2 (NOTE 14) | - | 20~100 ms [8](NOTE 12) | [10] cm [8] | [10-50] km/h (vehicle) [8]Stationary or Pedestrian | Up to 10 km radius [8](NOTE 13) | UL real-time vehicle data (video streaming and/or sensor data) [8] |
| [20] ms [8] | [0.1~0.4] Mbit/s [8] | 99,999% [8] | [~4] Mbit/s/km2 (NOTE 14) | Up to 8Kb [8] | 20 ms [8](NOTE 12) | [10] cm [8] | [10-50] km/h (vehicle) [8]Stationary or Pedestrian | Up to 10 km radius [8](NOTE 13) | DL control traffic (commands from the remote driver) [8]. |
| 1-20 ms(NOTE 15) | 16 kbit/s -2 Mbit/s(without haptic compression encoding);0.8 - 200 kbit/s (with haptic compression encoding)(NOTE 15) | 99.999%(NOTE 15) | [~20] Mbit/s/km2 (NOTE 14) | 2-8 (1 DoF) (NOTE 15) |  |  | Stationary or Pedestrian | Up to 10 km radius [8](NOTE 13) | Haptic feedback |
| NOTE 1: The mobile metaverse server receives the data from various sensors, performs data processing, rendering and provide feedback to the vehicles and users.NOTE 2: Examples of typical data volume including 1) camera: 10 Mbit/s per sensor (unstructured), 2) LiDAR: 90 Mbit/s per sensor (unstructured), 3) radar: 10 Mbit/s per sensor (unstructured), and 4) real-time Status information including Telemetry data: [< 50 kbit/s] per sensor/vehicle/VRU (structured). This is to support at least 80 vehicles and 1600 users present at the same location (e.g. in an area of 40m\*250m) to actively enjoy immersive metaverse services for traffic simulation and traffic awareness, the area traffic capacity is calculated considering 2 cameras, 2 Radars, 2 LiDARs on road side, 1600 user’s smart phones and 80 vehicles with 7 cameras, 4 radar and 2 LiDAR for each vehicle.NOTE 3: The frequency considers different sensor types such as Radar/LiDAR (10Hz) and camera (10~50Hz).NOTE 4: The service area for traffic flow simulation and situational awareness depends on the actual deployment, for example, it can be deployed for a city or a district within a city or even countrywide. In some cases a local approach (e.g. the application servers are hosted at the network edge) is preferred in order to satisfy the requirements of low latency and high reliability.NOTE 5: The calculation is this table is done per one 5G network, in case of N 5G networks to be involved for such use case in the same area, this value can be divided by N.NOTE 6: User experienced data rate refers to the data rate needed for the vehicle or human, the value is observed from industrial practice. NOTE 7: The network based conference focus is assumed, which receives data from all the participants, performs rendering (image synthesis), and then distributes the results to all participants. As rendering and hardware introduce some delay, the communication delay for haptic feedback is typically less than 5ms.NOTE 8: To support at least 15 users present at the same location (e.g. in an area of 20m\*20m) to actively enjoy immersive Metaverse service concurrently, the area traffic capacity is calculated considering per user consuming non-haptic XR media (e.g. for video per stream up to 40000 kbit/s) and concurrently 60 haptic sensors (per haptic sensor generates data up to 1024 kbit/s).NOTE 9: In practice, the service area depends on the actual deployment. In some cases a local approach (e.g. the application servers are hosted at the network edge) is preferred in order to satisfy the requirements of low latency and high reliability.NOTE 10: The arrival interval of compressed haptic data usually follow some statistical distributions, such as generalized Pareto distribution, and Exponential distribution [10].NOTE 11: The end-to-end latency does not include sensor acquisition or actuator control on the vehicle side, processing, and rendering on the user side (estimated additional 100ms total). Target e2e user experienced max delay depends on reaction time of the remote driver (e.g. at 50km/h, 20ms means 27cm of remote vehicle movement).NOTE 12: UL data transfer interval around 20ms (video) to 100ms (sensor), DL data transfer interval (commands) around 20ms.NOTE 13: The service area for teleoperation depends on the actual deployment; for example, it can be deployed for a warehouse, a factory, a transportation hub (seaport, airport etc.), or even a city district or city. In some cases, a local approach (e.g., the application servers are hosted at the network edge) is preferred to satisfy low latency and high-reliability requirements.NOTE 14: The area traffic capacity is calculated for one 5G network, considering 4 cameras + sensors on each vehicle. Density is estimated to 10 vehicles/km2, each of the vehicles with one user controlling them. [8]NOTE 15: KPI comes from 3GPP TS 22.261 [7] clause 7.11 “remote control robot” use case.NOTE 16: Examples of typical data volume including 1) ~8Mbps video stream. Four cameras per vehicle (one for each side): 4\*8=32Mbps. 2) sensor data (interpreted objects), assuming 1 kB/object/100 ms and 50 objects: 4 Mbps [8]. |

# 7 Security, authorization and privacy

## 7.1 Description

Security and privacy requirements are important to consider in the context of the present document. Regulatory requirements and user consent are mentioned throughout, emphasizing the importance of data confidentiality. The requirements listed below identify specific capabilities needed for authorization to support functionality described in other clauses of the present document. These requirements supplement the general security requirements for the 5G system defined in 3GPP TS 22.261 [7].

This clause includes requirements that provide functionality to define and enforce authorization policies.

## 7.2 Requirements

### 7.2.1 General

Subject to operator policy, regulatory requirements and user consent, the 5G system shall be able to support mechanisms to expose to a trusted third party the result of the UE authenticating the user.

NOTE: How a UE authenticates the user's identity at the terminal equipment, e.g., using biometrics, is out of the scope of the present document.

### 7.2.2 Localized mobile metaverse service

Subject to operator policy, regulatory requirements and user consent, the 5G system shall support mechanisms to authorize Spatial Localization Service.

Subject to operator policy, the 5G system shall provide an authorized third party a means to define authorization to access spatial anchor information and to manage the spatial anchor(s), e.g., add, remove or modify spatial anchors.

### 7.2.3 Avatar-based real-time communication

Subject to operator policy, regulatory requirements and user consent, the 5G system shall be able to authorize the avatar to be used in mobile metaverse services.

Subject to operator policy, regulatory requirements and user consent, the 5G system shall provide time-bound authorization for specified subscribers to use an avatar in mobile metaverse services.

### 7.2.4 Digital asset management

Subject to operator policy, regulatory requirements and user consent, the 5G system shall provide secure means to authorize the use of digital assets associated with a user (e.g., digital assets belonging to a third party customer).

The 5G system shall provide mechanisms to certify the authenticity of digital assets associated with a user.

# 8 Charging aspects

## 8.1 Description

This clause gathers all charging requirements that apply to functionalities whose requirements are specified in other clauses of the present document.

## 8.2 Requirements

### 8.2.1 Localized mobile metaverse service

The 5G system shall be able to collect charging information for the actions related to spatial anchors, where a third party creates, deletes or modifies a spatial anchor or associated service information.

NOTE: It is assumed that exposure of network anchors and associated service information can be a service provided by a network operator to third parties.

The 5G system shall support the collection of charging information associated with the exposure of a spatial map or derived localization information to authorized third parties.

The 5G system shall support the collection of charging information associated with the production or modification of a spatial map on behalf of an authorized third party.

The 5G system shall support the collection of charging information associated with exposing spatial location service information to authorized third parties.

The 5G system shall be able to collect charging information associated with distribution of third party mobile metaverse media to one or more subscribers.

### 8.2.2 Avatar-based real-time communication

The 5G system shall support collection of charging information associated with initiating and terminating avatar call.

The 5G system shall be able to collect charging information for transcoding services associated with avatar call.

### 8.2.3 Digital asset management

The 5G system shall be able to collect charging information per UE or per application, related to the use of digital assets associated with a user (e.g., typically a human user with a certain subscription).

The 5G system shall be able to collect charging information per UE for managing the digital assets associated with a user (e.g., typically a human user with a certain subscription) or a third party.

NOTE: A third party who has digital assets could be an enterprise customer having service level agreement with the operator.

Annex A (informative):
Mobile metaverse services

## A.0 Introduction

This document defines the "mobile metaverse" as the user experience enabled by the 5G system of interactive and/or immersive XR media, including haptic media. For clarity and unambiguity with respect to many divergent technical and commercial developments, this document avoids the use of the term "metaverse" without qualification with the term "mobile". This annex provides background information for the present document and further explanation of the term mobile metaverse, and the services that it enables.

The term mobile metaverse implies the combination of various technologies to fuse physical and digital worlds, widely impacting society and the economy. Users will access these mobile metaverse services with devices for interaction with XR media and sensors, enabled by mobile telecommunication standards.

Mobile metaverse services are expected to be provided via many and diverse service providers, each catering to different customers, companies and communities. An entertainment company might provide one mobile metaverse service for consumers such as a virtual theme park, while offering a separate mobile metaverse service to its employees. Some providers may specialize in interactive tools to create content, while others present that content. In general, mobile metaverse services can be divided into three general categories: industrial, enterprise and consumer mobile metaverse services.

There are commonalities among these categories. These mobile metaverse services will apply across these domains to varying degrees, sharing technologies, devices and interfaces, and functionalities described in this document. Ultimately, mobile metaverse services will be defined by the applications they enable and the business models they adopt.

Besides "experiencing" virtual world and/or augmented real world media as a passive consumer (where the media is read-only), mobile metaverse services also can enable interaction, the user can "create" and even "control" elements of the media. Depending on the mobile metaverse service, the consequence of user interaction could be experienced by other users. As discussed below, in some cases, user interaction may result even in changes to the real world, through 'actuation' (remote control mechanisms) as discussed in clause A.3.

Mobile metaverse services face technical challenges specific to the media and interactions they offer to users, in the context of mobile networks. Some of these technical challenges are addressed by mobile telecommunications standards, of which this document is a contribution. Another set of challenges are not technical in nature, such as regulatory and safety implications, only some of which are discussed in the present document. Other technical and non-technical challenges are expected to be addressed separately depending on the target environment/vertical.

## A.1 Consumer mobile metaverse services

There has been some initial development of commercial consumer-oriented mobile metaverse services, specifically in metaverse gaming and VR social media 'virtual worlds'.

Mobile metaverse consumer services considered during the study of mobile metaverse services relate to the following use cases:

- Attending (live) VR events (sports, gaming, concerts, etc.) either as spectator or performer, including while moving / commuting;

- Virtual shopping or visit experience (tourism, real estate, etc.);

- Presentation of AR content on a virtual screen, e.g., a feature length movie;

- Interaction with AR content in a location-aware manner, offering spatial localization and mapping to support applications generating AR content for enhanced localized experiences (e.g., in museums, shopping malls);

- Situation awareness about the user's physical surroundings, while walking or driving;

- Experience immersive communications with other entities - digital representations of users or application-generated content, including customer support services, by leveraging avatars, digital assets and wallets.

## A.2 Enterprise mobile metaverse services

Remote working has driven demand for better collaboration and communication tools, and thanks to extended reality (XR) and virtual reality (VR) technologies, many of those tools contribute to the enterprise metaverse.

The enterprise mobile metaverse will eventually envelop the core productivity applications that make business function, e.g., as a digital drafting tool for architecture and engineering, and training sessions via AR and/or VR. Eventually, the enterprise and industrial mobile metaverses will interlink, merging the IT systems in office branches with the Operational Technology (OT) systems of the shop floor.

Main mobile metaverse enterprise services considered during the study of mobile metaverse services relate to the following use cases:

- XR-enabled collaborative and concurrent engineering based on geospatial digital twins (and avatars), for research & prototyping, visual testing & simulation, planning & optimisation, which can further be used for training

- Virtual showrooms, products or stores;

- Interaction with AR content in a location-aware manner, e.g., creating spatial anchors and discovery of them for AR content enhanced in-store shopping experiences;

## A.3 Industrial mobile metaverse services

Industrial mobile metaverse services are expected to provide cost, productivity, safety and flexibility gains. In particular, industrial OT systems begin to support mobile services that include XR media for monitoring and analysis, and also for control of operations through the digital orchestration of robot fleets or user-guided remote operations. Industrial users will gain the ability to visualize and reconfigure their operations, e.g., in response to changing supply and demand or disruptions.

By fusing digital and physical realities, the digital twin has begun to transform many industries. Seaports have begun using digital twins to track every container on their docks. Aerospace companies design and build engines and fuselages virtually to simulate how an aircraft will fly and perform. Many new factories exist just as much digitally, as data and virtual representation, as they do in the physical world. This kind of mobile metaverse service aims at control and awareness of operations down to the smallest detail. Some of the mobile metaverse service enabler requirements support digital twin applications.

Main mobile metaverse industrial services considered during the study of mobile metaverse services relate to the following use cases:

- Remote critical health care, including surgery and treatment;

- AR/VR based tele-operation of a remote device or vehicle (e.g., driving).

## A.4 Common aspects

Common enablers associated with mobile metaverse services are virtual, augmented, mixed and other media (expressed collectively as extended reality, XR.) Whilst these are essential to the realization of the immersive experience of mobile metaverse services, they are not the only enabling functionalities.

This document does not focus on XR media and media services per se, but rather identifies requirements for complementary enablers common to all mobile metaverses services, including:

- Enabling localized (i.e., location-aware) mobile metaverse services by considering the essential mobility aspect of 5G connectivity, and collection and exposure of sensor data or 'results of processed sensor data' as needed by mobile metaverse services to integrate media into an immersive user's experience and expectations;

- Enhancing real-time communications, including over IMS, with avatar-based capabilities, thus allowing more immersiveness in virtual reality, augmented reality, etc. user experiences;

- Enhancing XR media delivery to make it possible to support user experience of multiple services simultaneously;

- Securely storing, exposing and managing digital assets.

Other key enablers of the metaverse implied by some immersive experiences relate to digital twins, defined in this document as "a real-time representation of physical assets in a digital world".

With respect to the metaverse, “geospatial” digital twins relate to the concept of adding geospatial information to a digital twin, such as 3D models, precise location and temporal information – enabling for example holographic-type of object animation. The value proposition of geospatial Digital twins is about enabling immersive user experience, collaboration and simulations that are very close to what to expect in the real world by all 3 categories of mobile metaverse services.

It is foreseen that additional functionalities will be defined in the future to provide capabilities to better offer and operate mobile metaverse services.

Annex <B> (informative):
Bibliography

[Use style "Heading 8" in TSs and "Heading 9" in TRs. Do not use "informative" in the title in TRs.

The Bibliography is optional. If it exists, it shall follow the last technical annex in the document.

The following material, though not specifically referenced in the body of the present document (or not publicly available), gives supporting information.

Bibliography format

<Publication>: "<Title>".

Editor's Note: For references used in the TS, let's separate those that are merely informative from those that are needed for normative requirements. All informative references can be added to this appendix and not clause 2.

Annex <C> (informative):
Change history

|  |
| --- |
| **Change history** |
| **Date** | **Meeting** | **TDoc** | **CR** | **Rev** | **Cat** | **Subject/Comment** | **New version** |
| 08.2023 | SA1 103 | S1-232605 |  |  |  | TS skeleton | 0.0.0 |
| 08.2023 | SA1 103 | S1-232074S1-232406S1-232407S1-232076S1-232623S1-232419S1-232620S1-232412S1-232421S1-232080S1-232422 |  |  |  | ScopeReferencesTerminology & AcronymsOverviewLocalized mobile metaverse servicesAvatar-based real-time communicationDigital asset managementPerformance requirementsSecurity, authorization, privacyCharging aspectsAnnex - Metaverse | 0.1.0 |
| 09.2023 | SA#101 | SP-231018 |  |  |  | MCC lean-up | 1.0.0 |