|  |  |
| --- | --- |
| 3GPP TS 22.156 V0.0.0 (2023-07) | |
| Technical Specification. | |
| 3rd Generation Partnership Project;  Technical Specification Group TSG SA;  Services and System Aspects;  Mobile Metaverse Services;  (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 address  3GPP support office address  650 Route des Lucioles - Sophia Antipolis  Valbonne - FRANCE  Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16  Internet  http://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 members  3GPP™ is a Trade Mark of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners LTE™ is a Trade Mark of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners  GSM® and the GSM logo are registered and owned by the GSM Association |

Contents

Foreword 4

Introduction 5

1 Scope 6

2 References 6

3 Definitions of terms, symbols and abbreviations 6

3.1 Terms 6

3.2 Abbreviations 8

4 Overview 8

5. Functional requirements 9

5.1 Description 9

5.2 General requirements 9

5.2.1 General 9

5.2.2 Requirements 9

5.2.3 Security and privacy aspects of mobile metaverse service functionality 9

5.2.4 Charging requirements for mobile metaverse service functionality 9

5.3 Specific functional areas 10

5.3.1 Localized mobile metaverse service functionality 10

5.3.1.1 General 10

5.3.1.2 Requirements 11

5.3.2 Digital representation of users and avatar functionality 11

5.3.2.1 General 11

5.3.2.2 Requirements 12

5.3.3 Operational efficiency, exposure, and coordination of mobile metaverse functionality 12

5.3.3.1 General 12

5.3.3.2 Requirements 13

5.3.4 Digital asset management functionality 14

5.3.4.1 General 14

5.3.4.2 Requirements 14

6 Performance requirements 14

6.1 Description 14

6.2 Performance requirements 16

Annex <A> (informative): Avatars and avatar communication 18

Annex <B> (informative): Bibliography 19

Annex <C> (informative): Change history 20

# 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.

Editor's Note: ------------------------- BEGIN 'Introduction' CONTRIBUTION -------------------------

# Introduction

The present document addresses a diverse range of enablers that provide specific services and capabilities to a range of applications that employ XR media.

Editor's Note: ------------------------- END 'Introduction' CONTRIBUTION -------------------------

Editor's Note: ------------------------- BEGIN 'Scope' CONTRIBUTION -------------------------

# 1 Scope

The present document provides Stage 1 normative service and performance requirements for diverse service enablers for applications that use XR. 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;

- Digital representation of users and avatar functionality;

- Digital asset management functionality;

- Operation efficiency, exposure, coordination of mobile metaverse services

Editor's Note: ------------------------- BEGIN 'Scope' CONTRIBUTION -------------------------

Editor's Note: ------------------------- BEGIN 'References' CONTRIBUTION -------------------------

# 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".

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

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

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

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

Editor's Note: ------------------------- END 'References' CONTRIBUTION -------------------------

Editor's Note: ------------------------- BEGIN 'Terminology & Abbreviations' CONTRIBUTION -------------------------

# 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 TS 22.228 [a].

**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 TS 22.228 [a].

**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 image (avatar), software licenses, gift certificates and files (e.g. music files) that have been purchased under a license that allows resale.

**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 [b].

**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 4: This definition was taken from TS 22.101 [c].

**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 5: This definition was taken from TS 22.101 [c].

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

NOTE 6: This definition was taken from TS 22.101 [c].

## 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 ClosedCircuit TeleVision

DoF Degrees of Freedom

DVE Distributed Virtual Environment

FACS Facial Action Coding System

FOV Field Of View

LiDAR Light Detection And Ranging

VRU Vulnerable Road User

Editor's Note: ------------------------- END 'Terminology & Abbreviations' CONTRIBUTION -------------------------

Editor's Note: ------------------------- BEGIN 'Overview' CONTRIBUTION -------------------------

# 4 Overview

The term metaverse has been used in various ways to refer to the broader implications of AR and VR. The present document uses this term to refer to a shared, perceived set of interactive perceived spaces that can be persistent. Metaverse in diverse sectors evokes a number of possible 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. The present document focusses on how to make these services function well, consistently and with divers support mechanisms over mobile telecommunications networks.

In addition to services that offer virtual or 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 in the locations 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 services, digital asset management and support for virtual entities.

Editor's Note: ------------------------- END 'Overview' CONTRIBUTION -------------------------

# 5. Functional requirements

Editor's Note: -------------------- BEGIN 'Description' + 'General requirements' CONTRIBUTION --------------------

## 5.1 Description

Editor's Note: describe the overall concept of a mobile metaverse service and how the subsections relate to each other. Do not imply that the subclauses covered here rule out further functions being defined in the future. This clause is intended as informative, containing no normative requirement.

Clause 5 of present document introduces functional requirements. These are grouped in subclauses of clause 5.2 for all general requirements.

The functional requirements for service enablers for mobile metaverse services that offer specific types of capabilities are specified in subclauses of clause 5.3.

The security, privacy and charging aspects are also considered general requirements, in clause 5.2, even when they pertain to specific functional requirements otherwise elaborated in clause 5.3.

## 5.2 General requirements

### 5.2.1 General

Editor's Note: It is FFS whether there are any general requirements or whether this clause should be removed.

### 5.2.2 Requirements

### 5.2.3 Security and privacy aspects of mobile metaverse service functionality

Subject to regulatory requirements, operator policies and user consent, the 5G system shall be able to support mechanisms to expose to a trusted third party (e.g. the conference focus) 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 3GPP scope.

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

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

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

### 5.2.4 Charging requirements for mobile metaverse service functionality

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 1: 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 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.

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

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

Editor's Note: -------------------- END 'Description' + 'General requirements' CONTRIBUTION --------------------

## 5.3 Specific functional areas

Editor's Note: ---------- BEGIN 'Localized mobile metaverse service functionality' CONTRIBUTION ----------

### 5.3.1 Localized mobile metaverse service functionality

#### 5.3.1.1 General

Localized mobile metaverse services are immersive, integrated into a user's ordinary experiences. These are location-related service experiences, AR, MR, etc., not location agnostic service experiences, as offered by VR 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 5.3.1.1-1: 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.

In order to determine which spatial anchors to discovery, the users's localization, that is their precise location and orientation, is important. 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.

#### 5.3.1.2 Requirements

Subject to operator policy, the 5G system shall provide a means to define and expose to a 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, 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.

Subject to operator policy, user consent and regulatory requirements, the 5GS shall provide a means for a UE to provide sensor data, (e.g. from UE sensors, cameras, etc.) to the UE's serving network in order to derive localization information, e.g. to produce or modify a spatial map or discover or find localized mobile metaverse services.

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

Editor's Note: ---------- BEGIN 'Localized mobile metaverse service functionality' CONTRIBUTION ----------

Editor's Note: ---------- BEGIN 'Digital representation of users and avatar functionality' CONTRIBUTION ----

### 5.3.2 Digital representation of users and avatar functionality

#### 5.3.2.1 General

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 precieve each others' 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 a digital representation of the user (e.g. 3D) that can be presented to other users. The system supports communication of this 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 expressions is sent from the UE to the network to be rendered as avatar encoded media.

Capabilities of UEs differ. For example: some UEs can render avatar encoding, 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), or culturally (e.g. due to spoken language diversity), conversational media can be transcoded. This is consistent with the objective for Total Conversation [d], clause 4.5. Avatar functionality provides new options for total conversation, as media can be transcoded to and from an avatar call where the user or the avatar media communicates by means of sign language.

#### 5.3.2.2 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 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 3: Avatar media can be transmitted on both uplink and downlink.

NOTE 4: Confidentiality of the data used to produce the avatar (e.g. from the UE cameras, etc.) is assumed.

Subject to 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.

The 5G system (including IMS) shall support transcoding between media such as text, GTT, video and avatar media in multimedia conversational communications.

NOTE 5: 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 6: 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 regulatory requirements, user consent and operator policy, 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.

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.

NOTE 6: The actual transmission and rendering of facial expression and movement and gestures of a person within a multimedia conversational communication is subject to that person’s consent.

Editor's Note: ---------- END 'Digital representation of users and avatar functionality' CONTRIBUTION ----

Editor's Note: ---------- BEGIN 'Operational efficiency, exposure, and coordination' CONTRIBUTION ------

### 5.3.3 Operational efficiency, exposure, and coordination of mobile metaverse functionality

#### 5.3.3.1 General

These capabilities whose service requirements are defined in clause 5.3.3.32 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 'physical 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 corresponding to means by which the 5G system provides 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 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 monitoried and exposed to third parties. 5.3.3.2 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 be able to provide a means to associate and coordinate service 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 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 take into account all available access network types, i.e. 3GPP and non-3GPP.

Subject to operator policy, the 5G system shall support a mechanism to enable one or more authorized third party(ies) to coordinate multiple service data flows 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).

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, subject to 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.

Editor's Note: ---------- END 'Operational efficiency, exposure, and coordination' CONTRIBUTION ------

Editor's Note: ---------- BEGIN 'Digital asset management functionality' CONTRIBUTION ------

### 5.3.4 Digital asset management functionality

#### 5.3.4.1 General

Mobile metaverse services can depend upon information that is associated with the user, e.g. IDs and personal data that are commonly required and have a known form. 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 may 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 information used by different services can be considered 'digital assets' in that the user needs or could benefit from having this information available when access mobile metaverse services.

A similar service is the EU digital wallet initiative. [EU- digital-wallet] 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.

#### 5.3.4.2 Requirements

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

Subject to user consent, the 5G system shall be able to allow a trusted 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.

Subject to regulatory requirements and operator policy, 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 the digital assets associated with a user.

Editor's Note: ---------- END 'Digital asset management functionality' CONTRIBUTION ------

Editor's Note: ------- BEGIN 'Performance requirements ' CONTRIBUTION ------

# 6 Performance requirements

## 6.1 Description

Editor's Note: Explain how the requirements refer to specific (non-normative) use cases, to give context and provide clarity to the KPIs in the next subclause. This clause is intended as informative, containing no normative requirement.

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.

## 6.2 Performance requirements

Editor's Note: Use the consolidated KPI table from TR 22.856 as the starting point. The title might be updated depending on how to organise the KPI requirements, which needs to take into account of the related content in 22.261.

Table 6.2-1: KPI 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]  [25]  (NOTE 6) | > 99.9% | [~39.6 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  [14]  (NOTE 7) | [1-100] Mbit/s  [14] | [> 99.9%]  [14] | [1.55] Tbit/s/km2  (NOTE 8) | Video: 1500  Audio: 100  [14] | - | - | Stationary or Pedestrian | typically  < 100 km2  (NOTE 9) | UL and DL audio/video |
| [5] ms UL  [1-50] ms DL  [14]  (NOTE 7) | [<1] Mbit/s  [14] | [> 99.9%] (without compression)  [> 99.999%] (with compression (NOTE 10))  [26] | [2.25] Tbit/s/km2  (NOTE 8) | 1 DoF: 2-8  3 DoFs: 6-24  6 DoFs: 12-48  [14] | 0.25-10 ms  [14] |  |  |  | UL and DL haptic feedback |
| Metaverse-based Tele-Operated Driving  (NOTE 16) | [100] ms [25] (NOTE 11) | [10~50 Mbit/s] [25] | 99% [25] | [~360 Mbit/s/km2 ]  (NOTE 14) | - | 20~100 ms [25]  (NOTE 12) | [10] cm [25] | [10-50] km/h (vehicle) [25]  Stationary/Pedestrian (user) | Up to 10km radius [25]  (NOTE 13) | UL real-time vehicle data (video streaming and/or sensor data) [25] |
| [20] ms [25] | [0.1~0.4 Mbit/s] [25] | 99,999% [25] | [~4 Mbit/s/km2 ]  (NOTE 14) | Up to 8Kb  [25] | 20 ms [25]  (NOTE 12) | [10] cm [25] | [10-50] km/h (vehicle) [25]  Stationary/Pedestrian (user) | Up to 10km radius [25]  (NOTE 13) | DL control traffic (commands from the remote driver) [25]. |
| 1-20ms  (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/Pedestrian (user) | Up to 10km radius [25]  (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. The end-to-end latency refers to the transmission delay between a UE and the mobile metaverse server.  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. The latency refers to the transmission delay between a UE and the application server. 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 [26].  NOTE 11: The end-to-end latency refers to the transmission delay between a UE and the mobile metaverse server or vice-versa, not including 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. [25]  NOTE 15: KPI comes from [5] 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 [25]. | | | | | | | | | | |

Editor's Note: ------- END 'Performance requirements ' CONTRIBUTION ------

Editor's Note: ------------------------- BEGIN 'Annex A' CONTRIBUTION -------------------------

# Annex <A> (informative): Avatars and avatar communication

Editor's Note: Provide context and explanation concerning avatars and avatar communication as used in service requirements and other clauses of the TR.

Editor's Note: ------------------------- BEGIN 'Annex A' CONTRIBUTION -------------------------

Annex <B> (informative):  
Bibliography

[EU-digital-wallet]: "COMMISSION RECOMMENDATION of 3.6.2021 on a common Union Toolbox for a coordinated approach towards a European Digital Identity Framework", https://ec.europa.eu/newsroom/dae/redirection/document/76610, accessed 6.7.23.

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** |
|  |  |  |  |  |  |  |  |