|  |  |
| --- | --- |
| 3GPP TR 22.856 V0.1.0 (2022-05) | |
| Technical Report | |
| 3rd Generation Partnership Project;  Technical Specification Group TSG SA;  Feasibility Study on Localized 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.  © 2022, 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 3

Introduction 4

1 Scope 5

2 References 5

3 Definitions of terms, symbols and abbreviations 5

3.1 Terms 5

3.2 Symbols 5

3.3 Abbreviations 5

4 Overview 6

5 Use Cases 6

5.A Use case of A 6

5.A.1 Description 6

5.A.2 Pre-conditions 6

5.A.3 Service Flows 6

5.A.4 Post-conditions 6

5.A.5 Existing feature partly or fully covering use case functionality 6

5.A.6 Potential New Requirements needed to support the use case 6

6 Relation to other standards activities 6

7 Considerations 6

8 Consolidated potential requirements and KPIs 6

8.1 Consolidated potential requirements 6

8.2 Consolidated potential KPIs 6

9 Conclusion and recommendations 6

Annex <A> (informative): <Informative annex for a Technical Specification> 7

Annex <Y> (informative): Bibliography 8

Annex <Z> (informative): Change history 9

# Foreword

This Technical Report 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.

# Introduction

Editor’s Note: Add an Introduction will be added as the second unnumbered clause.

# 1 Scope

The present document investigates specific use cases and service requirements for 5GS support of enhanced XR-based services, (as XR-based services are an essential part of "Metaverse" services considered in this study,) as well as potentially other functionality, to offer shared and interactive user experience of local content and services, accessed either by users in the proximity or remotely. In particular, the following areas are studied:

- Support of interactive XR media shared among multiple users in a single location, including:

- performance (KPI) aspects; e.g. latency, throughput, connection density

- efficiency and scalability aspects, for large numbers of users in a single location.

- the combination of haptics type of XR media and other non-haptics types of XR media.

- Identification of users and other digital representations of entities interacting within the Metaverse service.

- Acquisition, use and exposure of local (physical and digital) information to enable Metaverse services, including:

- acquiring local spatial/environmental information and user/UE(s) information (including viewing angle, position and direction);

- Exposing local acquired spatial, environmental and user/UE information to 3rd parties to enable Metaverse services.

- Other aspects, such as privacy, charging, public safety and security requirements.

The study also investigates gaps between the identified new potential requirements and the requirements already specified for the 5G system.

It is acknowledged that there are activities related to the topic Metaverse outside of 3GPP, such as the W3C Open Metaverse Interoperability Group (OMI). These activities may be considered in the form of use cases and related contributions to this study, but there is no specific objective for this study to consider or align with external standardization activities.

A difference between this study and study on Study on supporting tactile and multi-modality communication services in TR 22.847 is that Metaverse services would involve coordination of input data from different devices/sensors from different users and coordination of output data to different devices at different destinations to support the same task or application.

# 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] 3GPP TS 22.173: "IP Multimedia Core Network Subsystem (IMS) Multimedia Telephony Service and supplementary services".

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

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

[6] Talaba D, Horvath I, Lee K H, "Special issue of Computer-Aided Design on virtual and augmented reality technologies in product design," Computer-Aided Design 42:361 - 363.

[7] Wang X, Tsai J J-H, "Collaborative Design in Virtual Environments," In: ISCA 48:1-2.

[8] Benford S, Greenhalgh C, Rodden T, "Collaborative virtual environments." Communications of the ACM 44: 79–85.

[9] M. Eid, J. Cha, and A. El Saddik, "Admux: An adaptive multiplexer for haptic-audio-visual data communication", IEEE Tran. Instrument. and Measurement, vol. 60, pp. 21–31, Jan 2011.

[10] K. Iwata, Y. Ishibashi, N. Fukushima, and S. Sugawara, "QoE assessment in haptic media, sound, and video transmission: Effect of playout buffering control", Comput. Entertain., vol. 8, pp. 12:1–12:14, Dec 2010.

[11] N. Suzuki and S. Katsura, "Evaluation of QoS in haptic communication based on bilateral control", in IEEE Int. Conf. on Mechatronics (ICM), Feb 2013, pp. 886–891.

[12] E. Isomura, S. Tasaka, and T. Nunome, "A multidimensional QoE monitoring system for audiovisual and haptic interactive IP communications", in IEEE Consumer Communications and Networking Conference (CCNC), Jan 2013, pp. 196–202.

[13] A. Hamam and A. El Saddik, "Toward a mathematical model for quality of experience evaluation of haptic applications", IEEE Tran. Instrument. and Measurement, vol. 62, pp. 3315–3322, Dec 2013.

[14] 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.

…

[x] <doctype> <#>[ ([up to and including]{yyyy[-mm]|V<a[.b[.c]]>}[onwards])]: "<Title>".

# 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].

**example:** text used to clarify abstract rules by applying them literally.

## 3.2 Symbols

For the purposes of the present document, the following symbols apply:

<symbol> <Explanation>

## 3.3 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].

<ABBREVIATION> <Expansion>

# 4 Overview

# 5 Use Cases

## 5.1 Localized Mobile Metaverse Service Use Case

### 5.1.1 Description

This use case will consider the potential service opportunities that arise when advanced location information is available to trigger AR based services.

A precursor to this use case is briefly considered: i-mode service, introduced by NTT DOCOMO in 1999. The discussion of i-mode serves as an inspiration. This service was extremely successful, was one of the early mobile services beyond messaging and voice, and has many potential similarities with metaverse services. This service in many ways preceded and foresaw mobile internet services that would arise 10 years later. Users could access data on-line concerning weather, traffic, etc. While there were many revolutionary aspects to i-mode, three are particularly relevant for this use case:

- **i-area** – a location information service that enabled the user to identify *locally relevant information* – concerning traffic, maps and retail store information for business in the user’s proximity.

- **i-channel** – *a distribution service of ‘latest information’* that the user could further investigate (through interaction) and *whose display was user configurable*.

- **a fully decentralized content and service creation framework** allowing third parties to easily provide content, especially relevant: *location specific content*. This made it possible for small businesses to provide information to potential customers in the proximity such as opening hours, special offers, etc. It was even possible for those in the same location to meet and join 'virtually,' e.g. to play a computer game with other passengers in the same train car or bus.

In this use case, analogously, services that are locally relevant can be accessed, with relevant information retried. The user will have the ability to selectively control which of this information to display. The content that is obtained comes from decentralized sources - in this case different individual merchants will provide this information.

The use case described in this clause does not rely on or recreate the i-mode service. Rather, some of the ideas in i-mode service are carried forward given the new opportunities enabled by localized 3D interactive media. Localized mobile metaverse services creates exciting new opportunities to receive locally relevant content and interact with services.

These capabilities taken to a much greater level will form the essence of the coming mobile metaverse user experience. What distinguishes this service most is that it will provide the user with information services *integrated into their ordinary experience.* Consider an example of a commuter navigating an underground passage. Diverse relevant information is integrated into the user's field of view, as shown in Figure 5.1.1-1 below.



Figure 5.1.1-1: Localized Mobile Metaverse Services offering relevant information

Here, the AR annotation provides much more than an augmented map. The user is going to catch a train. (a) The path to the platform is shown without obstructing the user's perception of their proximity, where the contrast is good and no distractions appear. The (b) store on the right can provide content that may be relevant to the potential shopper, here the store's opening hours. Further along, (c) a restaurant provides a personalized message, reminding the user that they ate there in the past and ordered soup. These services are linked to the space that the user is in. See Figure 5.1.1-2, below.



Figure 5.1.1-2: Services offering relevant information are anchored in space

The three information augmentations that are displayed are the result of different source of information. The path information (a) can be presented anywhere that it fits into the scene, where the information for (b) and (c) are anchored in space. The information depends on interaction between the user (or the user's preferences, application context, etc.): the content shown depends on the user's interest: (a) they are travelling, and the navigation app knows what information the user needs to see. (b) The user has sought Blue Lotion and it is available here - the user's 'persistent search' shows local results. (c) At this time of day, the user often eats, so the restaurant on the left's reminder is welcome.

The 5G system enables this 'access to local services' in a number of ways.

### 5.1.2 Pre-conditions

Ulysses uses his AR capable glasses while travelling. These are tethered to a UE that he carries. This UE receives 5G service from the mobile network operator he has a subscription with, M.

Ulysses is interested in using his AR capable glasses to receive relevant information, so he has selected *which applications* are 'relevant.' These applications are therefore configured with the operator M to be 'active.' The purpose of this configuraiton will be described below.

Local services, that is services that are localized, have anchors that can be relevant to applications. A pre-condition of this use case is that such services exist and have *spatially defined* access. This is considered in this use case as an 'anchor.' The local service provider associates a service with an anchor, potentially as well as metadata concerning the service (e.g. 'is a restaurant'.) This information is available to M, either because the local services have been registered with M directly, or are available in maps, registries, or other information sources that M has access to.

Editor's Note: The nature of this anchor and how it is defined and managed is not the focus of this use case. This will be further developed in another use case.

### 5.1.3 Service Flows

Ulysses transfers at the Osaka train station. He has some time before he catches his connecting train. As he is hungry, he activates a 'persistent search' on his mobile device to inform him of opportunities to eat as he traverses the station. He also has a shopping list (a set of items of interest) from retail stores. He makes use of a navigation facility so that he will neither lose his way nor lose track of time.

This has the following consequences:

a) As a result of the applications activated by the user, and the user's preferences, the UE requests the 'localized mobile metaverse' service enabler offered by M, providing a list of 'persistent search' information.

b) M receiving the persistent search, engages localized service activation. The location of the UE is compared against the search criteria and information available to M of spatially defined access points.

Editor's Note: It is for further study how the UE can obtain information regarding spatial positions that (in step b) are 'locations' that can be compared to 'spatially defined access points.'

Editor's Note: It is for futher study if this 'acquisition' and use (in this use case for 'search' and 'matching' with spatially defined service access points) is in scope of 3GPP.

c) M identifies a match - essentially a 'JOIN' of user preferences / application persistent search criteria "restaurants" AND location (in the user's field of view) AND local spatially defined access exists. ]

d) This match is provided to the UE, for further processing by the application.

Example 1: The application [service] associated with 'restaurants' has stored information the user has eaten there and indicated it was 'good.' Thus the annotation 'Soup you liked last time' is displayed.

Example 2: The application [service] associated with 'shopping' has stored a shopping list. When a local 'shopping service access point' has been identified, the shopping application queries the service provider directly. If there is success, the result is displayed. Here: "Blue Lotion you want ¥2000."

### 5.1.4 Post-conditions

Spatial information has successfully been employed to allow a user to identify services.

The user's location has been applied.

The information output of different applications are integrated into the media that the user sees through the AR display device. The information associated with the services is displayed in the proper location in the user's field of view.

Ulysses may choose to eat soup or buy Blue Lotion.

M may charge Ulysses for this service, e.g. for the use of a persistent search and for each successful result provided.

### 5.1.5 Existing feature partially or fully covering use case functionality

Location based services exist, to identify the position of the UE.

Editor's Note: The gap compared with the existing features to support the described use case are for further study.

### 5.1.6 Potential New Requirements

[P.R.-5.1.6-1] The 5GS shall enable third parties to make known the availability of application services (i.e. provided by Application Servers) associated with a precise location.

[P.R.-5.1.6-2] The 5GS shall provide suitable exposure mechanisms for application services (i.e. provided by Application Servers) associated with a precise location available in the user's proximity (e.g. within line of sight), such that the application services can conform tospecific service constraints.

Editor's Note: As EdgeApp (SA6) already supports the EAS registration with location information it is FFS to demonstrate that this requirement is not an already supported feature.

[P.R.-5.1.6-3] The 5GS shall provide suitable means to discover application services (i.e. provided by Application Servers) associated with a precise location available in the user's proximity.

Editor's Note: It is for further study whether P.R.-5.1.6-3 is in 3GPP scope.

Editor's Note: The term 'Application Service' in the above potential requirements is for futher study, as it needs to be further explained.

## 5.2 Mobile Metaverse for 5G-enabled Traffic Flow Simulation and Situational Awareness

### 5.2.1 Description

Smart transport is a very important area for 5G system as well as metaverse. To reduce traffic jam and minimize traffic accident, 5G, including cellular based V2X technologies become more and more essential to enable, 5G system can be utilized to support real-time information & data delivery for the traffic participants including pedestrians, bicycle riders, vehicle with or without autonomous driving mode. As shown in Figure 5.2.1-1, In general, the physical objects including road infrastructure and vehicles including cars and trucks in each lanes, will have a corresponding digital twin in the virtual world, and then the virtual and physical objects form the mobile metaverse. Digital Twin means a digital representation of the physical object or called virtual object in metaverse. In this use case, there are virtual objects which actually representing the physical objects including vehicle, road and also pedestrians. This is the basis to enable smart transport applications like traffic flow simulation and situational awareness.

公路上的汽车

描述已自动生成

Figure 5.2.1-1 Example of Smart Transport Metaverse

With the support of 5GS, real-time information and data about the real objects can be delivered the virtual objects of the road infrastructure and traffic participants including vulnerable road users can form a smart transport metaverse as shown in Figure 5.2.1-2 Then real-time processing& computing can be conducted to support traffic simulation and also situational awareness and real time path guidance and real-time safety or security alerts can be generated for ICVs as well as the driver and passengers.

日程表

中度可信度描述已自动生成

Figure 5.2.1-2 Scenario of 5G-enabled Traffic Flow Simulation and Situational Awareness

In order to support traffic flow simulation and situational awareness service, the 5G network need to provide low latency, high data rate and high reliability transmission, and in addition, the 5G network may also need to be further enhanced to meet the service requirements for 5G-enabled traffic flow simulation and situation awareness. Meanwhile, in addition to the real objects which may host the UE for cellular system, their corresponding virtual objects is also capable of interacting with each other and also interact with physical objects via 5GS.

### 5.2.2 Pre-conditions

1. Traffic participants may or may not be equipped with 5G-enabled terminal which can send and receive data information via 5G network.

2. Computing and storage resources are provided for the mobile metaverse servers deployed locally or over the central cloud to allow real-time processing of huge data produced by human, vehicle, camera, radar etc.

3. Wired or wireless communication resources are configured among the road infrastructure and mobile metaverse server so that the server has real time information of traffic participants perceived by these sensors. The road infrastructure include camera, radar/lidar and also other devices such as traffic control and guidance etc. The infrastructure equipment may have wired connection e.g. fiber or ethernet or any other wired connection if available. However, cellular wireless network provides a more flexible way when wired connection is not available.

### 5.2.3 Service Flows

1. Wired or wireless communication path are configured among the road infrastructure and mobile metaverse server so that the server has real time information of traffic participants perceived by these sensors.

2. Sensors deployed in roadside are initialized and enter normal working mode which means it can start to capture the traffic participants.

3. Data connections between the vehicle driver’s UE and server are established with 5G UE module being registered with the server. Vehicles without driver are equipped a 5G UE module which is coined T-Box.

4. The traffic participants including pedestrians, bicycle riders, vehicles, via 5G UE, send their real-time information towards the server. The real-time status information includes position, speed, heading, brake status etc. These traffic participant acts as real objects in physical world, actually sends their property and status information to digital-twins objects in a virtual world. These real-time information may include structured and unstructured data. Structured data normally means data which has been processed and thus formatted in a certain way, can be easily stored in database and when transmitted via 5G wireless network, normally less transmission resource (e.g. lower data rate) is needed.  Unstructured data are not formatted in a certain and pre-defined way and is not easy to store in database and when transmitting over 5G network, more transmission resource would be needed.

5. Within the mobile metaverse for 5G-enabled traffic flow simulation and situation awareness, real-time information of the physical objects including the road infrastructure, the traffic participants as well as other information from traffic light signal, camera, radar, etc, are synchronized to the virtual objects and real-time simulation are conducted. The virtual objects can also interact with each other within virtual world and interact with physical/real objects via 5GS. It is noted that physical objects may include static and dynamic ones. Some static objects deployed along road side like light poles may not move but their properties can be synchronized with their virtual objects if such properties may impact traffic simulation and situational awareness and also visualization processing of the physical world.

6. Traffic flow simulation are conducted which can predict whether there will be traffic jam and which path is optimal for a certain vehicle and generate traffic assistance or guidance in a real-time manner towards the traffic participants.

7. The mobile metaverse server send the traffic assistance or guidance information to the UE which serves the pedestrian, bicycle rider, vehicle driver or autonomous vehicle.

### 5.2.4 Post-conditions

1. The mobile metaverse server conducts big data analysis to further refine the accuracy of traffic flow simulation and situation awareness.

2. Both UEs serving the real objects and the digital twins objects of the UEs can be identified by the mobile metaverse.

### 5.2.5 Existing features partly or fully covering the use case functionality

QoS framework of 5G system has been defined to support low latency, high reliability or high data rate transmission of application layer data traffic.

### 5.2.6 Potential New Requirements needed to support the use case

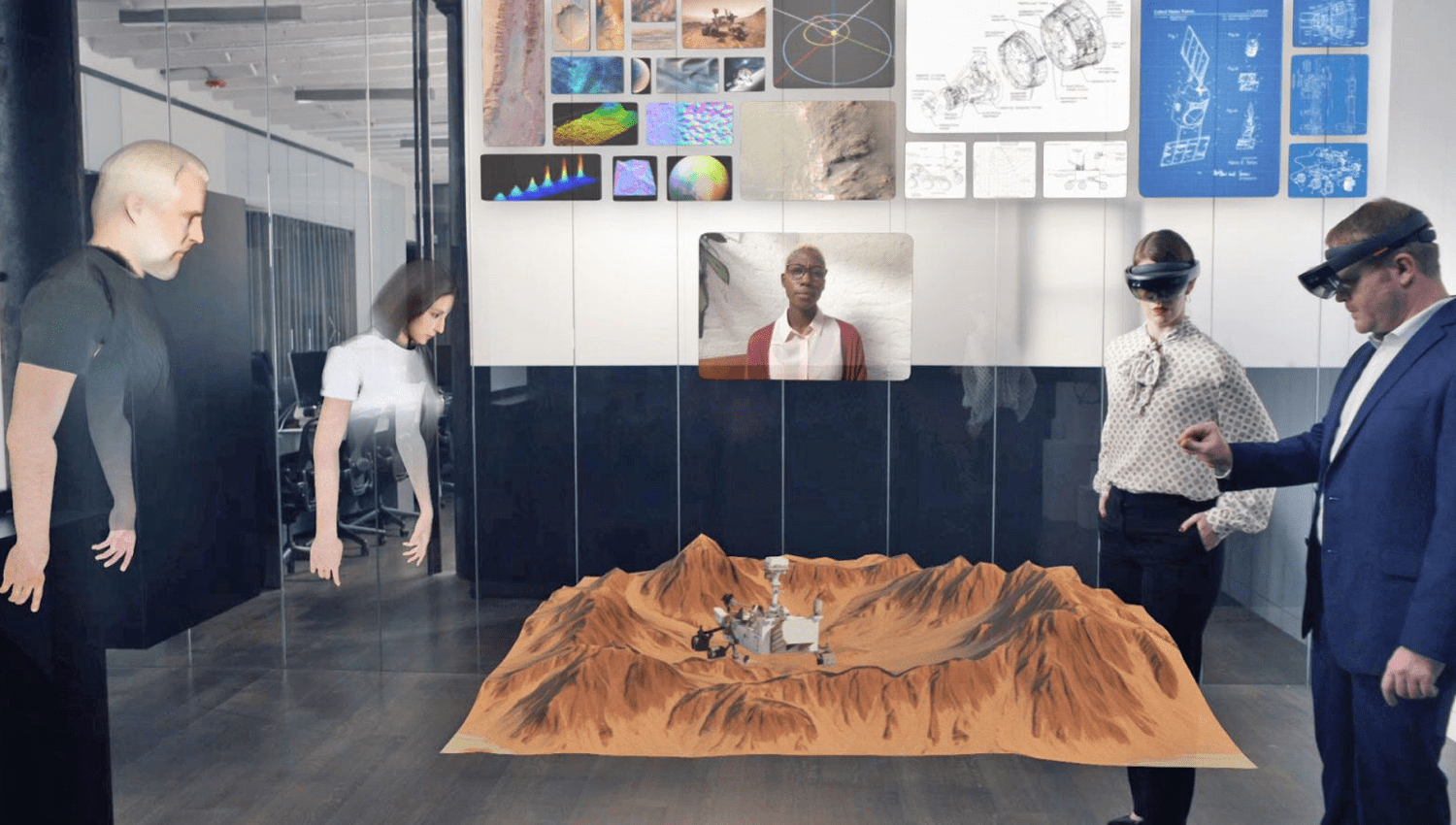
[PR 2.1.6-1] The 5G system shall provide low latency, high reliability and high data rate transmission for traffic between a large number of UEs and application server (e.g. mobile metaverse server).

Editor Note: This requirement needs to be revisited when concrete KPI values are provided.

5.3 Collaborative and concurrent engineering in product design using metaverse services

5.3.1 Description

Since the industrial age, engineering design has become an extremely demanding activity. Collaborative and concurrent engineering occur as a concept and methodology at the end of the last century and was defined as a systematic approach to integrated and co-design of products and their related processes. The diversity and complexity of actual products, requires collaboration of engineers from different geographic locations to share the ideas and solutions with customer and to evaluate products development. VR and AR technologies have found their ways into critical applications in industrial sectors such as aerospace engineering, automotive engineering, medical engineering, and also in the fields of education and entertainment. The range of technologies include Cave Automatic Virtual Environment (better known by the recursive acronym CAVE) environments, reality theatres, power walls, holographic workbenches, individual immersive systems, head mounted displays, tactile sensing interfaces, haptic feedback devices, multi-sensational devices, speech interfaces, and mixed reality systems [6].



**Figure 5.3.1-1: XR enabled collaborative and concurrent engineering in product design   
(Source: https://vrtech.wiki/)**

One of the key challenges is to how to enable a distributed virtual environment (DVE) allowing multiple users from different geographical locations (some of them are present at the same location) to interact over a network. DVEs are defined as multi-user virtual realities that actively support communication, collaboration, and coordination [7], 3D place-like environment in which participants are provided with graphical embodiments called avatars that convey their identity, presence, location, and activities to others [8]. DVE are the simultaneous existence of multiple users in the same virtual space represented as avatars, their communication, the shared exploration of 3D visualizations, and the collaborative construction of new content. This avatar representation is essential for every user knows about the actual perceptions of other users. The users can communicate with each other. They can interact with other users and with the virtual environment.

To support DVEs for the collaborative and concurrent engineering, 5G system need to fulfil the basic KPIs, such as latency, data rate, reliability. Moreover 5G system (with metaverse services) are expected to support the fundamental features including:

- multimedia communications with XR support among multiple users;

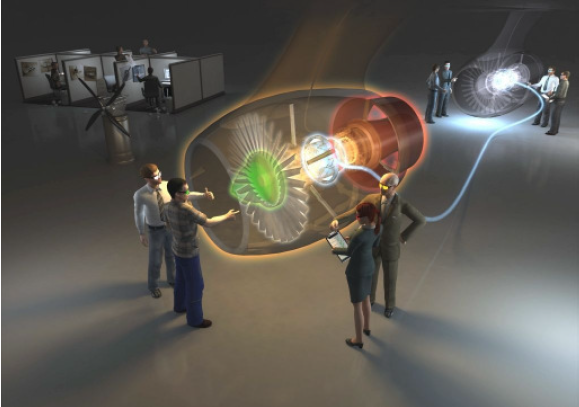
- user identity management;

- data security.

5.3.2 Pre-conditions

Novitas, an innovative start-up company, has set up a distributed virtual environment (with the corresponding 5G communication subscriptions) for collaborative and concurrent engineering in their product design with engineers participating locally and remotely. They have been granted a contract to work together with several partner companies to design and produce a new model of aeroplane engine.

In the current phase, Novitas need to collaborate closely with Nyhet, a partner company, to design the key parts of the engine. The service flows below illustrate how engineers interact with each other using services provided by 5G system.



**Figure 5.3.2-1: Illustration of Collaborative Workspace (Source: ESI-Icido GmbH)**

5.3.3 Service Flows

1. Having completed the authentication of the participants, the multimedia communication session/sessions are set up among multiple users as well as the associated devices in the mixed reality systems (e.g. head mounted displays, tactile sensing interfaces, haptic feedback devices, multi-sensational devices). This can be done by means of the IMS (including IMS CN with Data Channel capability) or via OTT applications.
2. When a session starts, multiple streams are established over the 5G network between the corresponding devices that carry multiple modalities data. Table 5.3.3-1 depicts the typical QoS requirements that have to be fulfilled in order for the users’ QoE to be satisfactory.

Table 5.3.3-1 Typical QoS requirements for multi-modal streams [9] [10] [11] [12] [13]

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Haptics** | **Video** | **Audio** |
| Jitter (ms) | ≤ 2 | ≤ 30 | ≤ 30 |
| Delay (ms) | ≤ 50 | ≤ 400 | ≤ 150 |
| Packet loss (%) | ≤ 10 | ≤ 1 | ≤ 1 |
| Update rate (Hz) | ≥ 1000 | ≥ 30 | ≥ 50 |
| Packet size (bytes) | 64-128 | ≤ MTU | 160-320 |
| Throughput (kbit/s) | 512-1024 | 2500 - 40000 | 64-128 |

1. The haptic information, video and voice are generated at one party and distributed to all other parties continuously.

5.3.4 Post-conditions

The 5G system enables efficient communication, with enhanced security and identity management, in support of DVEs for the collaborative and concurrent engineering.

5.3.5 Existing features partly or fully covering the use case functionality

The service requirements on the support of multimedia communication among multiple users have been captured in TS 22.228 [2] with the following key definitions:

***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.*

***Telepresence:*** *A conference with interactive audio-visual communications experience between remote locations, where the users enjoy a strong sense of realism and presence between all participants by optimizing a variety of attributes such as audio and video quality, eye contact, body language, spatial audio, coordinated environments and natural image size.*

***Telepresence System:*** *A set of functions, devices and network elements which are able to capture, deliver, manage and render multiple high quality interactive audio and video signals in a Telepresence conference. An appropriate number of devices (e.g. cameras, screens, loudspeakers, microphones, codecs) and environmental characteristics are used to establish Telepresence.*

***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).*

Support of Multi-device and Multi-Identity in IMS MMTEL service is captured in TS 22.173 clause 4.6 [3]:

*The support of multiple devices is inherent in IMS. In addition, a service provider may allow a user to use any public user identities for its outgoing and incoming calls. The added identities can but do not have to belong to the served user. Identities may be part of different subscriptions and different operators.*

In addition, TS 22.101 [4] has specified in clause 26a a set of service requirements on User Identity:

*Identifying distinguished user identities of the user (provided by some external party or by the operator) in the operator network enables an operator to provide an enhanced user experience and optimized performance as well as to offer services to devices that are not part of a 3GPP network. The user to be identified could be an individual human user, using a UE with a certain subscription, or an application running on or connecting via a UE, or a device (“thing”) behind a gateway UE.*

*Network settings can be adapted and services offered to users according to their needs, independent of the subscription that is used to establish the connection. By acting as an identity provider, the operator can take additional information from the network into account to provide a higher level of security for the authentication of a user.*

*The 3GPP System shall support to authenticate a User Identity to a service with a User Identifier.*

The functional requirement and performance KPIs in support of XR applications are mainly captured in TS 22.261:

- clause 7.6.1 AR/VR;

- clause 6.43 Tactile and multi-modal communication service

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

Clause 8 of TS 22.261 specifies the security related requirements covering aspects such as authentication and authorization, identity management, and data security and privacy.

Additional consideration need to be given to allow multiple users from different geographical locations to interact using XR techniques.

5.3.6 Potential New Requirements needed to support the use case

5.3.6.1 KPIs for the collaborative and concurrent engineering in product design

[PR 5.3.6.1-1] The 5G System shall provide the appropriate connectivity KPIs for the use case of collaborative and concurrent engineering in product design, see table 5.3.6.1-1.

**Table 5.3.6.1-1 – Potential key performance requirements for collaborative and concurrent engineering in product design**

| **Use Cases** | **Characteristic parameter (KPI)** | | | | **Influence quantity** | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Max allowed end-to-end latency** | **Service bit rate: user-experienced data rate** | **Reliability** | **Area Traffic capacity** | **Message size (byte)** | **UE Speed** | **Service Area** |
| Collaborative and concurrent engineering | [10] ms(note 1) | [1-100] Mbit/s  ([14]) | [> 99.9%] ([14]) | [3.804] Tbit/s/km2  (note 2) | Typical haptic data:  1 DoF: 2-8  3 DoFs: 6-24  6 DoFs: 12-48  Video: 1500  Audio: 100  ([14]) | Stationary or Pedestrian | typically  < 100 km2  (note 3) |
| NOTE 1: 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.  NOTE 2: 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 3: 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. | | | | | | | |

Editor’s Note: The KPIs in the above table need to be revisited.

5.3.6.2 Service requirements for collaborative and concurrent engineering in product design

[PR 5.3.6.2-1] The 5G system shall enhance the interaction between IMS CN and 5G CN to allow 5G CN to provide the IMS CN with real-time feedback in support of multiple media modes to multiple users simultaneously.

Editor’s Note: The above PR needs to be revisited.

## 5.A Use case of A

### 5.A.1 Description

### 5.A.2 Pre-conditions

### 5.A.3 Service Flows

### 5.A.4 Post-conditions

### 5.A.5 Existing feature partly or fully covering use case functionality

### 5.A.6 Potential New Requirements needed to support the use case

# 6 Relation to other standards activities

Editor's Note: This clause may capture relations to other standards activities, especially other studies and work items pursued in SA1.

# 7 Considerations

Editor's Note: This clause may capture other considerations, such as privacy, charging, public safety and security requirements.

# 8 Consolidated potential requirements and KPIs

## 8.1 Consolidated potential requirements

## 8.2 Consolidated potential KPIs

# 9 Conclusion and recommendations

# Annex A (informative): Avatar Service Considerations

The term Avatar originated in writings associated with Hindu religion, referring to an incarnation of a divine being on Earth, significantly Vishnu.

In computing an avatar is a graphical representation of a user or user’s character or persona. [Wikipedia-Avatar] The term was used to describe the player’s character in a number of games in the late 1970s into the late 1980s. In 1992, Neal Stephenson used the term to describe virtual simulation of the human form in his novel Snow Crash, in which he also coined the term metaverse. [2]

Avatars are used in a number of ways today, besides as digital representations of characters in video games. The representation is often thought to be one to one (one person is represented by one digital representation), but this cannot be generalized. Some people are represented in multiple ways (especially over time), some groups use an avatar to represent them, sometimes programs or automated services are represented with an avatar (and these aren't human users at all.) In most applications, people can choose their own avatars and they may change these frequently, even adopting the avatars of other users if there is no policy to prevent this.

Avatars may serve as a digital representation of a user in Internet forums. These are often a kind of cartoon version of a person’s face or an image representing them, often. For example, this is an avatar on Boardgamearena.com, for a community member known as “tree mile.”



Figure A-1: Avatar as Iconic User Representation

This digital representation is static (that is, it is generally not animated,) and serves to provide a user with a memorable and unique personality in the on-line forum, but without divulging my actual appearance. This is a common use on social media platforms.

A social forum, in which avatars are remote controlled, animated. An early example of this was SecondLife. [Linden Lab] This is an example group of avatars in discussion.



Figure A-2: Avatar as Animated User Representation

This platform does not feature a ‘game.’ Rather players interact, build things, share information, purchase virtual accoutrements. Some institutions built an on-line virtual presence, such as universities, private corporations even political parties to enable interaction between users represented as avatars.

Avatars have been used as a way to improve interaction between people using software or accessing on-line services and software. An example is “Clippy” a paperclip ‘help feature’ in Microsoft Office 97.

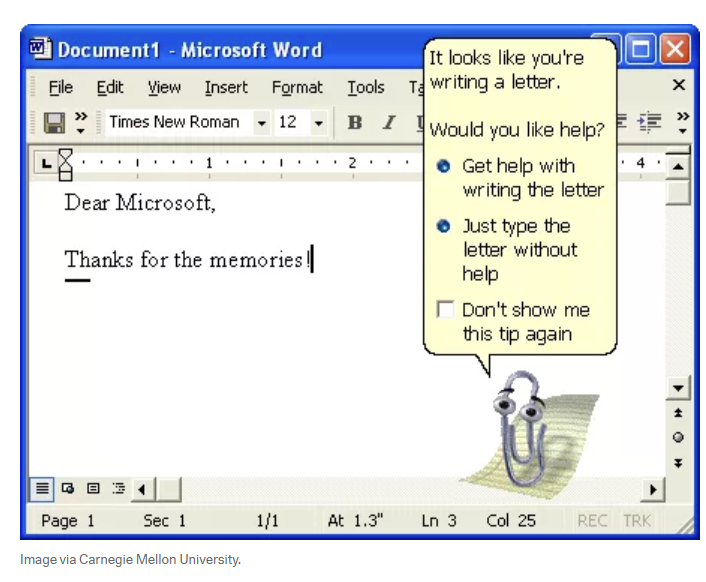


Figure A-3: Avatar as Animated Interactive Automaton

There are many other such digital representations that are used, e.g. for on-line chat services for service desks, etc.

Motion capture / animated avatars are used to stand in for a person. They model and reproduce or mimic the user’s movements, facial expressions and often represent specific facial animation for ‘talking’ in a way reminiscent of cartoons. One area where this has developed is a kind of content production by ‘vtuber’ contributors. Tools to create avatars (vtube animation software) can be coupled with motion capture software to allow contributors to generate video content in the form of animation. The creator is represented by media generated by means of a model and cameras. Sound can be added or recorded along with the video input.

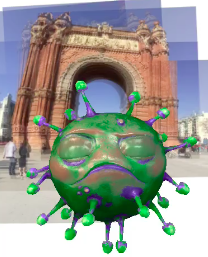


Figure A-4: Avatar Live Animation Generated from Camera and Microphone Input

A ‘live’ social media application can be designed around the techniques of animation and visual capture (as in the previous bullet) can provide an opportunity for users to communicate as cartoon digital representations of themselves with encoding and presentation in real-time. The communicating partner may be a human user or a ‘bot.’ Generally ‘chatbot’ services do not include such an animated figure – an icon or static image is used to represent the AI.  
  
A sophisticated ‘video capture,’ then transformation into a cartoon form with audio, and rendering this into media, is a very computationally intense task. There are many tools to create avatars and vtube video clips, however these are generally not ‘live.’ Figure A-5 presents ‘Kizuna AI’ a pioneering successful Vtuber personality. The media featuring these avatars is generated through tools that often involve animation editing and audio visual production operations. Pure animation techniques can be enhanced with motion capture and facial expression capture.



Figure A-5: Kizuna AI – an avatar celebrity

References

[Wikipedia-Avatar] <https://en.wikipedia.org/wiki/Avatar_(computing)>

[Stephenson] Stephenson, Neal “Snow Crash,” Bantam Books, New York, 1992.

[Linden Lab] https://secondlife.com/

Annex <Y> (informative):  
Bibliography

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

<Publication>: "<Title>".

Editor’s Note: It is likely that informative references may be provided in this study.

Annex <Z> (informative):  
Change history

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Change history** | | | | | | | |
| **Date** | **Meeting** | **TDoc** | **CR** | **Rev** | **Cat** | **Subject/Comment** | **New version** |
| 5.2022 | SA1#98e | S1-221264 | - | - | - | Initial Skeleton | 0.0.0 |
| 5.2022 | SA1#98e | S1-221265  S1-221266  S1-221267  S1-221268  S1-221269 | - | - | - | Output of approved pCRs from SA1 #98e. | 0.1.0 |