**3GPP SA4 RTC SWG AH *S4aR250179***

**Paris, FR, 3–5 Sept 2025**

**Agenda item:** 7

**Source:** Nokia

**Title: [AvCall-MED] On animation source streams**

**Document for** Discussion andAgreement

# Introduction

The present document discusses the source of animation data in an avatar call and how to ensure robustness during the session using voice-driven animation.

In a basic case a UE1 with an avatar will transmit an animation stream to the network or a UE2 where the animation stream is used to animate the base avatar of UE1 user. The animation stream is typically created using sensor input from the Party A, which may include camera, microphones (for voice-driven avatars) or other sensors.

When such sensory input is temporarily unavailable (e.g., the user moves out of view of the camera in cases of user video-based animation data generation, leaves the phone on the table while doing something else, etc.), voice can be used as fallback. It is possible for the sender (e.g., a UE or MF) of the animation data to switch to a new source data (e.g., to voice-driven animation) for creating the animation data. However, in some cases the remote UE2 may do the voice animation locally when, for example:

* UE1 and network cannot generate animation data from voice. If voice is transported UE-to-UE, the network may not have access to voice data for generating the animation data.
* There is congestion on the network. Sending voice and animation data based on voice creates redundancy if the remote UE is capable of generating animation data based on voice.

The required signalling for avatar calls can be divided into two categories, excluding the IMS data channel and media configurations that are inherited from MTSI calls:

* The initial SDP exchange required to communicate the capabilities of the UE and the network to establish the parties having an avatar within the call, the format of the avatars, the entities where these avatars will be animated and rendered and the animation streams.
* Management of the avatar scene via scene descriptions and other metadata exchanged during the call.

This contribution proposes required signalling for voice fallback during an avatar call.

# Identification of fallback animation source in DC negotiation

Additions on top of [S4aR250142\_signaling\_online\_IB\_EY\_Online.docx](https://www.3gpp.org/ftp/TSG_SA/WG4_CODEC/3GPP_SA4_AHOC_MTGs/SA4_RTC/Inbox/Drafts/S4aR250142_signaling_online_IB_EY_Online.docx) is highlighted below.

**Table 1 — REQUEST payload**

|  |  |  |  |
| --- | --- | --- | --- |
| **Name** | **Type** | **Cardinality** | **Description** |
| avatarProfile | string | 1..1 | Indicates the avatar profile as defined in clause 5.6.  [Allowed values: "2d" or "3d". The value shall match the ARF profile of the selected Avatar representation. ] |
| arfContainer | object | 1..1 | Reference to ARF container of the Avatar representation that is offered for the call. The container shall be retrievable from the BAR. |
| animationSourceData | array(URI) | 1..1 | Provides a list of supported animation data for the avatar.  The following values are allowed:   * A subset of the AnimationFrameworks as declared in the ARF container of the base avatar * Audio with the URN "urn:3gpp:avatar:v1:animation:audio"   Video with the URN "urn:3gpp:avatar:v1:animation:video"  At least one animation source data shall be included.  An audio stream can be used as a fallback for animating the avatar in the absence of an avatar animation stream. |
| renderingLocation | array(string) | 1..1 | Indicates the possible location(s) for animating and rendering of the offered avatar. The possible values are: “sender”, “receiver” or “mf”.  An MF may add “mf” in the offered list if it is capable of animating and rendering the offered avatar. |

**Table 2 — ACCEPT payload**

|  |  |  |  |
| --- | --- | --- | --- |
| **Name** | **Type** | **Cardinality** | **Description** |
| renderingLocation | string | 1..1 | Indicates the selected location for the rendering and animation of the offered avatar.  If the location is set to “mf”, it has to be confirmed by the transmission of an ACK message from the MF to the receiver. |
| animationSourceData | array(URI) | 1..1 | Includes the selected subset of the animation frameworks that the receiver prefers to receive for performing the avatar animation.  If an audio or a video source is used, the receiver or MF shall use one of the audio or video stream from the IMS session to animate the avatar.  An audio stream can be used as a fallback for animating the avatar in the absence of an avatar animation stream. |

When fallback is used, the SDP media description of the fallback audio stream shall include a MID value.

The MID value of the fallback stream shall be sent as part of the stopped/resumed message over ADC (defined in clause 3 below).

# Sending notification for animation data disruption

An AR-MTSI client that offers an avatar animation stream should notify the remote client with a message over the data channel with the URN **urn:3gpp:avatar:animation:stopped** when the animation stream becomes temporarily unavailable**.** An AR-MTSI client that receives the message shall switch to audio-driven animation if a source audio stream whose media description includes the SDP attribute avatar\_voice\_fallback was negotiated. Format of the stopped message is shown in Table 1.

Table 1: **Stopped** message for avatar animation

|  |  |  |  |
| --- | --- | --- | --- |
| **Name** | **Type** | **Cardinality** | **Description** |
| **id** | string | 1..1 | A unique identifier of the message in the scope of the data channel session. |
| **type** | string | 1..1 | A URN that identifies the message type: **urn:3gpp:avatar:animation:stopped** |
| **animationStream** | string | 1..1 | A URN that identifies the stopped animation stream |
| **reason** | string | 0..1 | An optional field indicating the cause of disruption in animation stream. Possible values include: "device error", "low-confidence for sensor data", "network issues", etc. |
| **startTime** | number | 0..1 | start time of the suspension of the animation data |
| **endTime** | number | 0..1 | end time of the suspension of the animation data if known |
| **mid** | number | 0..1 | MID value of the audio stream used as fallback when animation stream stopped. |

When the animation stream becomes available again, the AR-MTSI client that offers the avatar animation stream should notify the remote client with a message over the data channel with the URN **urn:3gpp:avatar:animation:resumed**. Format of the resumed message is shown in Table 2.

Table 2: **Resumed** message for avatar animation

|  |  |  |  |
| --- | --- | --- | --- |
| **Name** | **Type** | **Cardinality** | **Description** |
| **id** | string | 1..1 | A unique identifier of the message in the scope of the data channel session. |
| **type** | string | 1..1 | A URN that identifies the message type: **urn:3gpp:avatar:animation:resumed** |
| **animationStream** | string | 1..1 | A URN that identifies the resumed animation stream. |

# Proposal

Section 2 and 3 should be added to appropriate sections in the CR to TS 26.264.