# MediaVerse

A universe of media assets
and co-creation opportunities

# D5.2

# Immersive Storytelling Authoring Tools v1

| | |
|---|---|
| **Project Title** | MediaVerse |
| **Contract No.** | 957252 |
| **Instrument** | Innovation Action |
| **Thematic Priority** | ICT-44-2020 Next Generation Media |
| **Start of Project** | 1 October 2020 |
| **Duration** | 36 months |

| | |
|---|---|
| **Deliverable title** | Immersive Storytelling Authoring Tools v1 |
| **Deliverable number** | D5.2 |
| **Deliverable version** | V1.0 |
| **Previous version(s)** | N/A |
| **Contractual Date of delivery** | 30.09.2021 |
| **Actual Date of delivery** | 30.09.2021 |
| **Nature of deliverable** | Other |
| **Dissemination level** | Public |
| **Partner Responsible** | VRAG |
| **Author(s)** | Stephan Gensch (VRAG), Dimitrios Ververidis (CERTH), Pilar Orero, Victor García, Anna Matamala, Susannah Fraser (UAB) |
| **Reviewer(s)** | Violeta Vasileva (AS), Francesco D'Andria (ATOS) |
| **EC Project Officer** | Alberto Rabbachin |

| | |
|---|---|
| **Abstract** | This reports on the initial iteration after the requirements elicitation, extending background technologies and integrating first research and development results in the authoring tools in MediaVerse. It demonstrates the initial capabilities of a) video authoring, b) stories authoring, and c) enhancing with 3D models and includes low- to mid-fidelity prototypes, such as mock-ups and clickable dummies, as well as a deployed prototype v1 platform. |
| **Keywords** | Accessibility, Authoring Tools, Immersive Storytelling, VR Sandbox |




MediaVerse is an H2020 Innovation Project co-financed by the EC under Grant Agreement ID: 957252.
The content of this document is © the author(s). For further information, visit mediaverse-project.eu.

## Revision History

| Version | Date | Modified By | Comments |
|---|---|---|---|
| V0.1 | 30/06/2021 | Stephan Gensch | First Draft Table of Content |
| V0.2 | 05/07/2021 | Stephan Gensch | Final Table of Content |
| V0.3 | 20/07/2021 | Dimitrios Ververidis | Adding developments for T5.5 |
| V0.4 | 24/08/2021 | Dimitrios Ververidis | Moving content from Section 4 to Section 2 |
| V0.5 | 31/08/2021 | Stephan Gensch | Adding development for T5.3 in Section 3 |
| V0.7 | 14/09/2021 | Stephan Gensch | Final Second Draft |
| V0.8 | 26/09/2021 | Violeta Vasileva | Review |
| V0.9 | 27/09/2021 | Francesco D'Andria | Review |
| V1.0 | 30/09/2021 | Stephan Gensch | Final Document |

## Glossary

| Abbreviation | Meaning |
|---|---|
| GDPR | General Data Protection Regulation |
| HMD | Head-mounted Display |
| WP | Work Package |
| DAM | Digital Asset Management |
| XR | eXtended Reality |
| VR | Virtual Reality |

# Table of Contents

# Index of Figures

# Index of Tables

# Executive Summary

Media and entertainment industries do not only require intelligent decentralized media services on individual content items, but also want to use the provided advanced services and media assets for the creation of more complex content. As such, the MediaVerse project provides a set of authoring tools that give access to media on the distributed nodes and that allow to create XR content.

This deliverable contains an elaborate section on requirements engineering, target groups, existing solutions and how the planned authoring tools integrate into the MediaVerse landscape. The general assumption for MediaVerse is to target users that have not acquired deep knowledge on XR development. The authoring tools should enable them to easily create content based on three main user scenarios. These scenarios include a) mobile 360-degree authoring with a fast turnaround using a template-based approach, b) co-creation scenarios to enable multiple users to collaboratively create 360-degree content, and c) an XR-sandbox tool to collaboratively enrich, experience and share a real 3D virtual environment.

Two baseline technologies are described and used within this work package. One of them is a SaaS-Solution for interactive 360-degree storytelling, called Fader, provided by partner VRAG. The other one is called VRodos, provided by partner CERTH. This deliverable provides insights on the specific use cases and requirements for either of the two and introduces an implementation plan to achieve the proposed requirements.

# 1 Introduction

One of the new trends is the emergence of XR (eXtended Reality) content in many media applications. XR content is a term introduced to cover Virtual Reality and Augmented Reality and, in most cases, requires 3D models to be constructed. Social media are continuously embedding XR content, for example, TikTok AR filters and the Facebook Horizon platform. It is necessary to find new authoring systems that also allow non-experts (non-programmers or non-designers) to be involved in the process and thus tackle the transformation efficiently without excluding sensitive groups from the production. Journalism, Art, and Cultural Heritage (CH) in general need also to conform with this XR norm so as not to risk losing part of their audience. Broadcasters, museums, galleries, music event organizers, artists, actors, and other stakeholders need tools that will allow their content to be consumed through virtual experiences, to attract their audiences. What should be done more to boost the XR content creation by non-experts is a crucial question that we have investigated throughout this deliverable.

The music industry in Europe is constantly evolving and the digital music revenues in the United Kingdom are estimated at over a billion euros, 0.9 billion euros in Germany, and almost half a billion euros in France (Statista, 2019). As regards to Use Case 3 on cultural spaces, artistic, and cultural productions are good options for entertainment without leaving home. In today's digital age, streaming services, including the upcoming 5G network, allow online content transmission technology. However, the music industry must evolve with technology and enable new ways of offering benefits not only to its consumers but as well to its professionals. Real-time music remote interaction through the internet is not yet a reality although it is technically possible. For example, musicians in different locations in the world being able to rehearse, record and perform online is not yet a standard. There are some technical options[1] but they can provide high quality services in the 4G network. Therefore, we have seen in MediaVerse as a good opportunity to collaborate with another project through partner ArtShare (see Section 4) that promotes the 5G technology that offers a new digital path for the music industry in Portugal and the rest of the world.

In system engineering and development, the most crucial part is the definition of the requirements. Therefore, we have executed a requirement engineering process for driving the design of new systems that will allow for enhancing 3D media productivity, for lowering the entry barrier in 3D media creation, and for innovative media forms across many media types (cross-media). This work has been carried out with the perspective of enhancing recovery and transformation as the pandemic has driven many professionals in culture to zero income. Toward this goal, we perform a requirements engineering process based on the standard for requirements specification[2]. It allows us to elucidate system requirements through existing (AS-IS) and envisioned (TO-BE) scenarios affected by the latest trends on design methodologies and content promotion in social media. A total of 34 tools for content creation, promotion, and monetization are reviewed, and 10 TO-BE scenarios were engineered and validated. The validation was performed through a survey of 24 statements on a 5 Likert scale by 47 individuals from the domains of Media, Fine arts, Architecture, and Informatics. Useful evaluation results and comments have been collected. The analysis is also published recently (June 2021) in a scientific journal related to sustainability (Mouzakis et al. 2021).

[1] Jamulus, Networking audio streaming services, URL: https://jamulus.io/

[2] IEEE Standard 830-1998. *Recommended Practice for Software Requirements Specifications*; IEEE Press, 1998.

## 1.1 Purpose of this Deliverable

The purpose of this deliverable is to describe the 1st year developments on the tools that provide 360 VR, namely Fader by Vragments, and native VR, namely VRodos by CERTH, provided in Sections 3 and 4, respectively. Background knowledge is provided on tools and technologies that existed before the launch of MediaVerse, and we have first addressed the scope, namely how they can be transformed to fulfil the contemporary requirements for XR applications (see Section 2) and MediaVerse needs (see MediaVerse Deliverable D2.1). Then we have listed the developments that have been done so far in Year 1 on these tools. In addition, we describe how the use cases that have been set up for MediaVerse make use of the authoring tools. Also, apart from the technical description of the proposed tools, our related work review and requirements study can be useful to research organizations, companies, and policymakers that seek to develop new platforms related to 3D media content.

## 1.2 Relation with Other Activities in MediaVerse

The immersive storytelling tools and the VR sandbox are high-level applications that make use of a variety of MediaVerse services, such as media processing and annotation, digital asset management, or accessibility tools, to name a few. These services are described in the deliverables D3.1, D4.1, and D5.1. As such, there are many activities in other WPs that directly influence the work on the authoring tools. D2.2 gives a good indication on the envisioned architecture that serves as a foundation to integrate the authoring tools that are part of this deliverable.

**WP2**

The use cases and requirements engineering in WP2 are a crucial part to determine what users expect from the authoring tools. Section 2 of this deliverable deals directly with the requirements engineering process and the outcome. Each tool has its own set of requirements presented in detail in the respective section.

Also, within WP2, a general architecture design process was made that determines, how the authoring tools are integrated into the MediaVerse architecture. The authoring tools presented in this deliverable are of a third-party nature and as such, many challenges need to be solved, from authentication to media management. A specific challenge is, how the authoring tools reference media files, deal with (temporary) local copies for processing, and much more.

**WP3**

Several tasks in WP3 deal with media processing. In T3.3, for example, an improved 360-degree video transcoding microservice will be developed, based on the OMAF file creation tools[3] by Fraunhofer HHI.

**WP4**

To make media from MediaVerse nodes available to the authoring tools, the digital asset management (DAM) must ensure the correct usage of media assets by internal and third-party tools. This includes proper adherence to copy- und usage rights, compensation, and attribution.

[3] HTML5 MSE Playback of MPEG 360 VR Tiled Streaming, URL: https://github.com/fraunhoferhhi/omaf.js/tree/master/omaf-file-creation (accessed 2021-09-13)

**WP6**

This WP delivers the foundation of the MediaVerse core software components that the authoring tools need to make use of. Authoring tools developed against a MediaVerse node API must adhere to the specified interfaces.

**WP7**

As the authoring tools are being used by real users, a user centric development and evaluation process is key to building good products. Over the course of the project milestones, there will be several opportunities to gather valuable feedback by users to further improve the tools being built.

# 2 Existing Tools, Requirements Engineering, and Overview on the Use of Immersive Authoring Tools within MediaVerse

Below we present a certain requirements analysis especially focusing on 3D media authoring systems. The conclusions inferred from the analysis allow us to develop many side-products during the developments of T5.3 and T5.5. The side-products enhance the probability of self-sustainability - as a mitigation plan - after the end of MediaVerse. We are investigating crucial features which new platforms should have to be attractive to experts, but also to enable non-expert citizens to participate in the creation process. The skills elevation of non-experts in programming and designing is considered by Gartner as a crucial factor for economic growth (Costello & Rimol, 2019). According to European Commission reports, the culture and creative ecosystem have been deeply affected by the pandemic (EC784/2020)[4]. European media SMEs face severe issues, while unemployment has increased, and many media professionals in culture—particularly those who are subject to precarious employment conditions or are freelancers—have found themselves with no income. Cinemas suffered a collapse in revenues (with losses estimated at 100,000 euros per screen per month during lockdown), whilst the shooting of new films, programs, and TV series has been in many cases halted. In parallel, the crisis has accelerated major trends in digital technology. Online platforms have strengthened their market position, launched new services, and attracted new audiences during the lockdowns. New online social media platforms—largely based on audio-visual content—have also hit records in downloads. The audio-visual industries are facing many changes.

## 2.1 Review of Existing Tools in XR Media Creation

This section provides an overview of the existing tools for the creation and promotion of 3D content. A list of these tools can be found in Table 2.1 and are explained in the following lines.

*Table 1: Authoring tools for VR experiences in Arts and Culture*

| # | Name | Target Group | Type | Medium | Scope | Features |
|---|---|---|---|---|---|---|
| 1 | 3D Vista Pro https://www.3dvista.com/ | Amateur VR experiences designers | App | Desktop | VR experience design. Author experiences based on 360 videos | Allows one to place tags onto video objects while they are moving. |
| 2 | Adobe Illustrator/XD https://adobe.com/products/illustrator.html | Graphic designers | App | Desktop | Authoring digital visual arts designs based on 2D and 3D geometries | 3D and 2D editor; Dynamic brushes; Painting; Export interface designs. |
| 3 | Adobe Medium | Graphics designers | Application | VR | Create 3D geometries | Sculpting and other tools that |

[4] European Commission. *Communication from the Commission to the European Parliament, Europe's Media in the Digital Decade: An Action Plan to Support Recovery and Transformation*; COM/2020/784 final; European Commission: Brussels, Belgium, 3/12/2020.

| | | | | | | |
|---|---|---|---|---|---|---|
| | https://www.adobe.com/products/medium.html | | | | and textures inside VR space | simulate real environments. Maximum level of detail for professional use. |
| 4 | Adobe Premiere https://www.adobe.com/products/premiere.html | Professional video makers | Application | Desktop | Video and audio editor; Author videos for stereoscopic VR glasses. | Allows one to place tags onto video objects; Media uploader; Allow 3rd party plugins for VR effects, e.g., Torusmedialabs Canvas 360. |
| 5 | A-frame https://aframe.io | Professional VR designers | Code | Code Editor | Framework to program fast VR applications | Allows with html tags to setup a VR environment. Abstraction of Three.js framework. |
| 6 | Amazon Sumerian https://aws.amazon.com/sumerian/ | Professional VR designers | Web | Desktop | Interface to author general purpose VR environments with avatars | 3D editor targeting programmers; Amazon speech synthesis and recognition addons. |
| 7 | Cospaces https://cospaces.io/edu | Educators for art and design in schools | Web | Desktop | Allows children to author and program 3D cultural experiences | 3D editor; visual programming interface; class management; VR support. |
| 8 | Dataverse.xyz https://dataverse.xyz/ | Amateur VR designers | Web | Google sheets | Allows artists to present a story in 3D space through table sheets as input. | Editor based on google sheets; Multiple layouts of 360 VR scenes based on 5 templates; Open source. |
| 9 | Desktop 3D design tools Blender, Maya, Cinema 4D, 3DS Max | Graphics designers | Application | Desktop | Author digital visual arts designs and sculptures based on 3D geometries. | 3D editor; dynamic brushes; 3D painting; poly integration; animation control. |
| 10 | Experizer https://experizer.com/ | Amateur VR designers | Web | Desktop | Allows the presentation of | 3D Editor; several templates; |

| | | | | | | |
|---|---|---|---|---|---|---|
| | | | | | information in 3D. | quizzes support; Storyline support. |
| 11 | Fader https://getfader.com/ | Amateur VR designers | Web | Desktop | Author stories and tours through 360 images and videos | Tags placement; media uploader; tags insertion; scenes interconnection. |
| 12 | Filmmaker Live https://filmmakerlive.com | Amateur VR Designers | Application | VR | Author a story with audiovisual content. | Video editor; camera editor; media uploader; multiplaying. |
| 13 | Google Blocks https://vr.google.com/blocks/ | Amateur VR Designers | Application | VR | Author digital designs and sculptures based on 3D geometries. | 3D editor; VR support; dynamic brushes; 3D painting; Google Poly repository integration. |
| 14 | Google Tilt Brush https://www.tiltbrush.com/ | Professional VR designers | Application | VR | Author digital designs and sculptures based on 3D geometries. | Dynamic brushes; 3D painting; intuitive interface; Google Poly repository integration. |
| 15 | Google Tour Creator https://arvr.google.com/tourcreator/ | Amateur VR Experiences designers | Application | Desktop | Basic VR experiences using 360 media | Allows one to place tags onto video objects while they are moving. |
| 16 | Isadora https://troikatronix.com/ | Novice programmers | Code | Code editor | Create audiovisual effects that can be used inside immersive experiences. | Show control; lighting editor; multimedia editor; visual programming interface. |
| 17 | MaxMSP https://cycling74.com/ | Multimedia Creators | Code | Code editor | Author audio content for VR applications. | Visual programming interface; multimedia editor; script editor; open source |

| | | | | | | |
|---|---|---|---|---|---|---|
| 18 | MediaPipe https://google.github.io/mediapipe/ | Professional media designers | Code | Code editor | Apply Video processing AI algorithms | Apply AI models on streams, e.g. selfie segmentation or pose recognition |
| 19 | Mozilla Hubs https://hubs.mozilla.com/ | Amateur VR Designers | Web | Desktop | Author 3D spaces for social interaction | Multiplaying; Google Poly and Sketchfab asset fetching; open source. |
| 20 | Network-aframe https://www.npmjs.com/package/networked-aframe | Professional VR designers | Code | Code editor | Extend aframe with multiplaying capability | Use WebRTC server to provide peer2peer multiplaying capability |
| 21 | Nuke https://www.foundry.com/products/nuke | Professional VR Designers | Application | Desktop | Author digital visual arts designs based on 360 media | Scripts editor; 2D and 3D compositing and visual effects; visual programming interface. |
| 22 | OpenFrameworks https://openframeworks.cc/ | Multimedia programmers | Code | Code editor | Author visual arts experiences that can be used inside VR. | Script editor; camera editor; open source. |
| 23 | PlayCanvas https://playcanvas.com | Professional VR Designers | Web | Desktop | Author a VR game without programming skills. | Asset uploader; animation control; scripts editor; multiplaying. |
| 24 | Processing.org https://processing.org/ | Novice Programmers | Code | Code Editor | Create artistic content for VR environments. | Script editor; movie editor; open source; visual editor. |
| 25 | Runtime Graphics Engines Unity, Unreal, Unreal | Storyboard artists and VR professionals | Application | Desktop | Author general purpose VR environments with avatars for storytelling | 3D editor; asset uploader; animation control; scripts editor; visual programming interface; |
| 26 | Sketchbox https://www.sketchbox3d.com/ | Amateur VR experiences designers | Application | Desktop | Educational tool for VR training. | Multiplaying; VR support; asset uploader; 3D editor. |

| | | | | | | |
|---|---|---|---|---|---|---|
| 27 | SynthEyes https://www.ssontech.com/ | Professional VR Designers | Application | Desktop | 3D camera tracking and realistic exposition of objects in VR based on 360 media. | 3D camera tracking; camera editor; Image preprocessor; 360VR stabilization. |
| 28 | Thinglink https://www.thinglink.com/ | Amateur VR experience designers | Web | Desktop | Author basic experiences based on 360 videos. | Allows one to place tags onto video objects while they are moving. |
| 29 | Three.js https://threejs.org/ | Professional VR designers | Code | Code editor | Authoring advanced 3D visualizations accessible through a web browser. | Allows one to visualize 3D models and 360 videos, VR support, wide community support. |
| 30 | VFX software Adobe After Effects, MOCHA, Canvas 360 | Professional Experiences designers | Application | Desktop | Author digital visual arts designs based on 360 media. | Apply a posteriori video effects on video. Used in creative industries for various movies effects. |
| 31 | VVVV https://vvvv.org/ | Novice programmers | Code | Code editor | Author interactive visual creations and creative art that can be used in VR. | Script editor; visual programming interface; animation control; open source. |
| 32 | VeeREditor https://veer.tv/veer-editor | Amateur VR Designers | Application | Mobile | Author VR videos and photos | Mobile editor with video effects, stickers, filters and background music. |
| 33 | WondaVR https://www.wondavr.com/ | Professional experiences designers | Application | Desktop | Author social VR environments for storytelling. | 3D editor; asset uploader; animation control; scripts editor; multiplaying. |
| 34 | Youtube VR https://vr.youtube.com | Amateur VR experience designers | Web | Desktop | Author basic virtual tour experiences based on 360 or 180 media. | Media uploader; various VR headsets support. |

The tools presented in Table 1 can be categorized according to their use into the following categories, shown in Table 2.

*Table 2: Categorization of authoring tools*

| CATEGORY | DESCRIPTION |
|---|---|
| **Graphics designing tools** | Tools that allow one to create 3D models with colour, texture, and animation, for example, Google Tilt and Blocks, Adobe Medium, Blender, Maya, 3DS MAX, and Cinema4D. |
| **Professional VR Experience design tools** | Tools that require extensive knowledge and skills in editing or programming, for example, tools like Unreal, Unity, 3D Vista pro, Amazon Sumerian, PlayCanvas, and WondaVR. |
| **Amateur VR Experience design tools** | Web-based tools that are easily accessible to anyone, for example, Google Tour Creator, Youtube VR, Dataverse.xyz, Experizer, Fader, Mozilla Hubs, and Thinklink. |
| **Education on tools for VR** | Tools that allow children to easily learn about synthesizing VR experiences, for example, Cospaces. |
| **Storyboard VR and VFX design tools** | Graphics Engines such as Unity or Unreal that allow one to make a VR movie, or tools that allow to post-edit a movie such as Adobe After Effects, Nuke, SynthEyes, and MOCHA. |
| **Open-source coding libraries** | Free libraries such as A-frame, Babylon.js, Isadora, MaxMSP, MediaPipe, Networked-Aframe, OpenFrameworks, Three.js, and VVVV allow more freedom into synthesizing impressive visualizations. |

**Graphics design tools:** Google Tilt Brush and Adobe Medium are useful tools for artists and designers to create visual arts in a virtual world and author virtual tours and galleries. Google Tilt Brush supports features like dynamic brushes, and it has intuitive interfaces mostly suitable for experts on artistic design. It is compatible with Vive, Oculus, Oculus Quest, Windows Mixed Reality, Valve Index, and Playstation VR headsets. It requires the downloading of the software from Steam, Humble Store, Vive, Oculus, or PlayStation Store. Recently, Google Blocks has been introduced which is similar to Tilt but much more simplified targeting for novice designers. Adobe Medium is a similar application available only for Oculus Rift and Oculus Quest headsets. It supports features like 3D editing and multi-playing. Other design tools available in the market are the "traditional" ones, that is, those based on a desktop screen interface such as Blender, Maya, 3DS Max, and Cinema4D.

**Professional VR experiences design tools:** Many professional artists rely on programming interfaces and tools to generate content. In this category, Unreal, Unity, and Godot Graphics Engines can be found which are often used to generate VR experiences. Amazon Sumerian is another tool that is based on web technologies for VR applications. It is compatible across VR headsets such as HTC Vive, Oculus Go, Google Daydream, and HoloLens. It can be used to create VR tours like virtual museums and other subjects of cultural interest, for example, a 360-video presentation that provides an immersive real-world experience to help travellers select their destination, accommodation, or adventure. It has a 3D editor, Amazon Web Services speech synthesis and recognition, ready to use templates and assets. PlayCanvas is a 3D editor for VR applications targeting novice programmers. The application has a cloud-hosted creation platform that allows multiple users to interact within the project. It supports features like asset uploader, animation control, scripts editor, multi-playing, etc. Wonda VR is an application for novice programmers based both on 3D geometries and 360 media to generate VR experiences. The scope of the application is to author social VR environments for immersive storytelling and interactive branching narratives. It supports features like 3D editor, asset uploader, animation control, scripts editor, and multi-playing.

**Amateur VR experiences design tools:** Artists, journalists, and other amateur content creators often resort to the solutions of this category where VR experiences are easily authored through 360 media. Several tools can be found in this category:

(1) Google Tour Creator is an authoring tool for anyone to create a VR experience through desktop, Android, or iOS devices. The user imports photos and does basic editing like adding audio or points of interest. It can be used for virtual art galleries and other subjects of cultural interest with a low need for interactions.

(2) Thinglink is a similar application. Thinglink is an exposition tool for virtual galleries and other subjects of cultural interest. Supports interactive tags, 360 media editor and is compatible across VR headsets such as Oculus GO, Google Daydream, Hololens, etc.

(3) Experizer is another such type of authoring tool accessible through a web browser. The scope of the application is to author immersive interactive virtual tours and presentations. It also supports quizzes, a storyline, and can analyse the activity of users and tracks scores. It can create immersive 360 experiences for a quest.

(4) Fader is another similar authoring tool. It also offers the possibility to embed the projects directly to the developers' website and supports live editing on them even after the publication.

(5) Mozilla Hubs is an open-source VR authoring tool for web browser users. The scope of the application is to author 3D spaces for social interaction by making virtual rooms. Hubs supports features like multi-playing, importing 3D models, and allowing the positioning of 3D emoticons in the environment. It has a 3D editor named Spoke that can be used to create new levels. A novel feature of Spoke is that it performs automatic estimation of the navigation mesh on the whole scene in order to make a compact form of colliders and preserve resources, which is very useful for web applications. It can also change lights positions and provide the ability to manage animations. It can be used for authoring virtual galleries, museums, and other types of venues. It can support more than 100 concurrent users through Amazon hosting services.

(6) YouTube VR is compatible with all VR headsets. YouTube VR, allows the users to watch any video on YouTube and provide an immersive VR experience whether it is on 360-degree video or a standard video shown in a cinema-style mode. For editors, supports both 360 and 180 VR format.

(7) 3D Vista Pro allows users to place tags onto video objects while they are moving. It also allows the user to combine multiple 360-degree photos into one to create a time-lapse effect. Other features include the ability to share to social media and providing the guest immediate access to maps, floor plans, and panoramic images of the tour.

(8) Dataverse.xyz is an open-source VR authoring tool for 360 media, which exploits google sheets as an editing environment. It allows artists or journalists to present a story in VR through eight visualization templates (Geo Viz, Map, isotypes, 360 photo, 360 video, PhotoGrid, TimeLine, and TreeMap).

**Free coding tools for Art:** Many artists prefer to create designs by coding in free programming tools. These tools have gained particular interest as they can be used to make audio-visual content for VR environments. Prominent examples are Isadora, MaxMSP, OpenFrameworks, Three.js, Processing, and VVVV.

(1) A-frame is a popular JavaScript library that abstracts further Three.js library for Web VR applications. A-frame success is related to the fact that totally hides JavaScript layer, but programming is achieved by only using html tags. In this manner, the code needed is significantly less than Three.js.

(2) Babylon.js[5] is a JavaScript library for beginners and advanced developers to create games or web apps and is comparable to Three.js. It offers WebXR[6] features that allow building VR and AR experiences for the web.

(3) Isadora is an authoring tool based on programming that allows one to create immersive experiences and visual arts designs. Supports a variety of uses during live performances, for example, from simple video playback up to complex interactive scene control. It has a visual programming interface to help artists create immersive interactions and editing interfaces.

(4) MaxMSP is a visual programming tool for multimedia developers to author interactive software. As with most of the tools mentioned above, it has a visual programming interface. Other features are a JavaScript editor and an audio editor to create interactive channels.

(5) MediaPipe is a recently published library that includes several AI algorithms for processing video streams on nearly all APIs, namely JavaScript, C++, python, Android and others. It offers about 10 solutions such as Selfie-Segmentation, Pose recognition, SLAM and others.

(6) Networked-aframe is a library build on top of A-frame that provides the multiplaying features for VR applications. It is using the WebRTC protocol with Janus or EasyRTC adapters to achieve the replication of player movement, video, audio, as well as for streaming custom data, e.g., hands position from VR controllers (Oculus Quest 2).

(7) OpenFrameworks is an open-source framework for professional programmers to author visual arts experiences. It requires knowledge of C++. It provides a simplified interface for hardware control, media handling, and network communication.

(8) VVVV is a coding tool for novice programmers to author interactive visual installations and creative art. It requires basic knowledge of C#. It has a visual programming interface, a script editor, animation control, and a video editor. The VVVV.OpenVR pack contains nodes to get pose data and render a 3D scene into any VR headset supported by SteamVR. VVVV is free only for non-commercial use.

(9) Processing is an open-source graphical tool for novice programmers in order to create digital visual arts. It started as an introduction tool for the basics of computer programming in a visual context and evolved into a professional programming environment. It supports features like editing in JavaScript and a movie editor.

(10) Three.js is a popular open-source framework that builds on top of WebGL to offer 3D visualization, animation, and sound. It requires basic knowledge of HTML and JavaScript. It has been evolved through its related A-frame framework into a tool for creating VR environments with few lines of code.

**Education on VR design:** This category contains useful tools for educational purposes that help students or employees to easily learn subjects of interest.

(1) CoSpaces is an authoring tool for children that allows them to generate artistic creations with primitive 3D objects, avatars, and 360 media. It is a simplified version of more complex tools such as Blender or Unity. CoSpaces has a 3D editor, a visual programming interface, a class management interface, a physics engine (e.g., for object collisions), and VR capability. It is compatible across VR headsets such as Mobile VR, Oculus Go, Google

[5] https://www.babylonjs.com/

[6] W3C, World Wide Web Consortium. WebXR Standard Specification, WebXR. Available online: https://www.w3.org/TR/2020/WD-webxr-20200724/ (accessed on 20/7/2021).

Daydream, Samsung VR, and Class VR. With CoSpaces, kids can design a virtual exhibition, build an immersive virtual tour based on 360 media, program their own game, and create an interactive story in a virtual world.

(2) SketchBox is an application for learning in VR. Collaborative with 3D design applications such as Tinkercad, it is an easy-to-use tool for simulation. It supports features like a 3D editor, an asset uploader, and multiplaying. It is also a tool to sketch ideas and explain them to others.

**Storyboard and VFX design tools:** This category contains tools that allow experts in design to generate storyboards, virtual productions, real-time XR productions, and apply special effects to movies. Graphics Engines such as Unity, Unreal, Godot, and Blender can be found in this category. Other editing tools are movie editing tools such as Adobe After Effects and TorusMediaLabs 360 Canvas for Adobe Premiere that allow to easily change background font or foreground characters with 3D graphics.

## 2.2 Requirements Engineering Procedure

A requirement is defined by IEEE (IEEE830; MSEE, 2012; Marcelino et al, 2014) as: (1) a condition or capability needed by a user or a system to solve a problem or achieve a goal; (2) a condition or capability, which has to be provided by a system to fulfil a contract, a standard, a specification or any other formal documentation; and (3) a documented representation of a condition or capability. In our work, particular gravity was given to the third definition as we wanted to find a well-justified and documented representation of a capability for future developments during a Requirements Engineering (RE) process. Overall, the RE process is depicted in Figure 1 and it consists of the following steps: (1) Preparation, (2) Elicitation, (3) Analysis, (4) Specification, and (5) Validation.

The Preparation Phase consists of three steps. The first is related to scenario modelling. In this step, a “Template for collection of scenarios” is defined. The aim is to distribute this form among artists and journalists so they can perform their contributions to the art industry scenarios identification. Based on this input information, modelling can be achieved. This step intends to provide the foundations and guidelines for the representation of various types of artists, so that the current process may be analysed and improved. Modelling or representing the current (AS-IS) situation is the basis for identifying shortcomings and potential improvements and forms the basis for the design of adequate models (TO-BE). The results of the template, namely the AS-IS and TO-BE models for art are summarized in 10 cases in Table 3.

Both the AS-IS model, which represents the current situation as it is, and the TO-BE model, resulting from incorporating the desired improvements, are equally important. Catalysts are the new trends that appear in society and particularly in electronically generated visual arts. These will allow the proposed system to find a track in the market of creative industries.

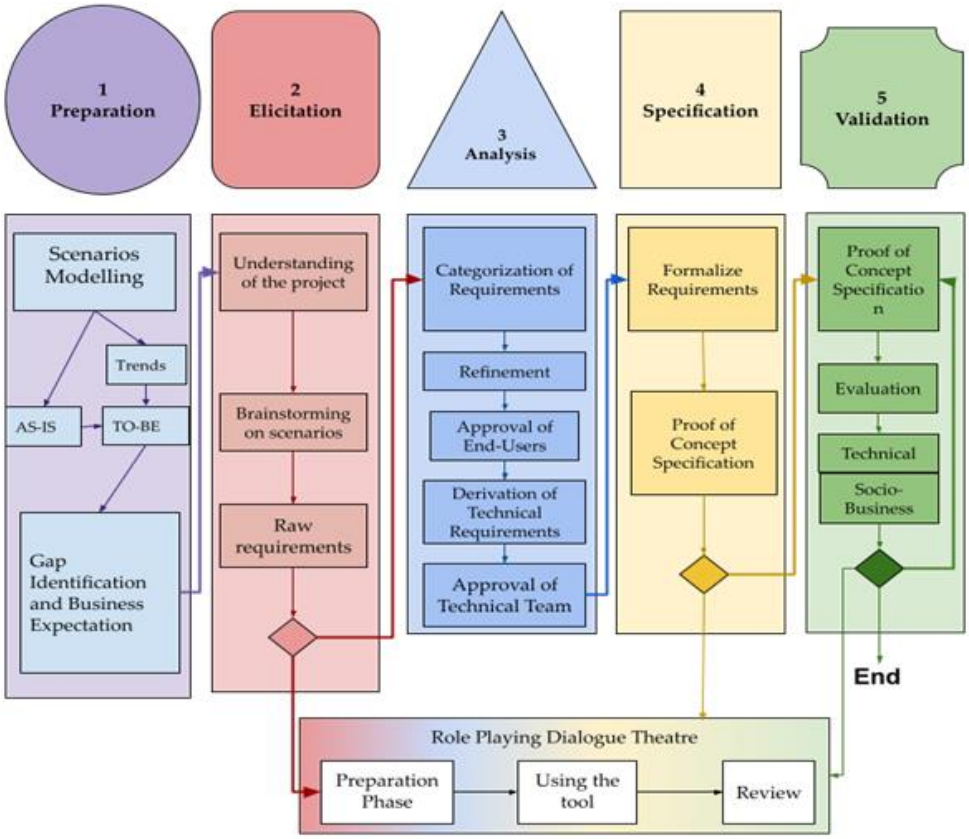


*Figure 1: Requirement Engineering phases*

*Table 3: Target user group activities and requirements engineering elicitation phase*

| # | Artist | Activity | Current Tools (As-Is-Scenarios) | Trends (Opportunities) | Brainstorming (To-Be-Scenarios) |
|---|---|---|---|---|---|
| 1 | Web UX artist | Make front-end web visualizations | Adobe Illustrator and XD, Figma, WordPress, Three.js | Neomorphism, 3D, Glassmorphism, VR, AR | A flexible 3D viewer component for web pages does not exist |
| 2 | Multimedia -Graphics Designer | Developing audiovisual storyboards | Adobe After Effects, Premier, Photoshop | Content and ads creators in YouTube, Facebook, Instagram | Adobe suite is too expensive. A commercial gap exists. |
| 3 | VFX artist | Create post-production effects | Adobe After Effects and Premier, Nuke, Maya, MOCHA | Social media content creators on TikTok, YouTube, etc. Animators code in Python AI tools. | Setting rotoscopy object borders takes too much time; no high resolution is needed for social media. |
| 4 | 3D models artists | Design the geometry of 3D models | Maya, Blender, Cinema4D, Pix4D, Agisoft | Design within VR with Google Tilt or Adobe Medium; Capture with Photogrammetry. | Tilt and Medium are difficult to operate; Google Blocks is easier but it has fewer capabilities. Photogrammetry retrieves too noisy meshes. |
| 5 | Digital 3D Animation artist | Animate a 3D model | Blender, Maya, Cinema4D, Zbrush, 3DS Max, iClone, Daz3D, Rokoko | Prefixed, surreal, and retro 3D filters in TikTok applied with localization | Pre-fixed animations as in mobile apps become boring. Keyframing is a tedious procedure. |

| | | | | | |
|---|---|---|---|---|---|
| | | | | on top of videos. Bodysuits. | Rotoscoping needs automation to find the borders of objects automatically. |
| 6 | 3D Texture artist | Paint 3D models material layers. | GIMP, Photoshop, Blender, Maya, Mari, Substance Painter | Represent materials with high realism exploiting the latest GPU technologies. Problems in Photogrammetry. | Materials are not compatible across graphics engines, for example, Maya with Unity. Photogrammetry assesses all layers in one material map which should be decomposed into layers. |
| 7 | Scientific illustrations artist | Draw 3D models of scientific elements | Adobe Illustrator, Blender, Maya, Gimp, Photoshop | Many artists have WordPress personal blogs for their work promotion. COVID-19 has attracted the attention of the public. | WordPress does not support 3D models. The COVID biological processes are too complex for the audience to learn and remember. 3D models need proper sound coverage to be presented in an immersive way. |
| 8 | Story-telling artist | Make movies or cutscenes | Unity, Unreal, Godot, Blender Graphics Engines | Real-time productions: Visualize the imaginary in real-time. | COVID has been an impediment to real productions. There is a need to allow actors to collaborate remotely. |
| 9 | School STEAM teacher | Educate children about 3D technologies | Cospaces, Tinkercad | Visual programming is gaining ground for setting the behaviour of objects. | WebGPU standard will change the landscape of web browsers graphics with more realistic graphics. Cospaces and Tinkercad have low resolution. |
| 10 | Researcher Artist | Generate and visualize 3D content | Three.js, VVVV, MaxMSP, Isadora, Processing, OpenFrameworks | Artists become programmers and explore code and node-based editing | Connecting music to textures to 3D is something it is missing from the market, for example, connect MaxMSP to Three.js |

**Elicitation:** The requirements elicitation phase represents all the actions performed to acquire raw requirements related to what is intended to develop in the project. The purpose of our project is to construct a platform that can help artists such as those in the first column of Table 2 to do their job more efficiently, quicker, with lower cost, or publish and disseminate it better. In principle, the requirements are defined according to the traditional internal expectations such as an increase in profits, production cost savings, streamlining of processes, reduction of creation times, shortening of processing times, receive up-to-date information, achieve better communication between production units, or minimizing of idle times. On the other hand, some of the common external customer and/or market-oriented socio-business expectations are higher process quality and resulting product quality, closer proximity to customers and better customer commitment, and faster communication with market partners (MSEE, 2012). These socio-business indicators are clearly defined in order to be later used for progress measurement.

A formalization of the requirements according to the IEEE 830 standard characteristics is presented in Table 4. These characteristics are Unambiguity, Completeness, Consistency, Verifiability, Relevancy, and Feasibility. The most important is Unambiquity which is a perceivable definition of the requirement. Many Requirements Specification Languages (RSLs) exist from the early 80s. However, we preferred to define the requirement in Attempto Control English (ACE), which is a natural language (Fuchs et al, 1998). Although not tested with an ACE parser, the premature descriptions allow one to confine the definition of the component and to let it be a seed for future official tests. With respect to Completeness, a short concise description of the requirement is provided. As regards Consistency, it involves conflicts in the implementation of the proposed requirement. For example, some requirements have inconsistencies out of which the most important is security. Although WebGL standard is designed according to security standards, a poor implementation might be open to penetration attacks. As regards the use of web cameras for applications targeted to children, they will have to grant access to a web page or a native program to access their camera. This may be an impedance to application adoption. As regards Verifiability, several requirements do not have a quantitative result that can be measured directly, for example, how good are visual effects on a video or how well are the borders of objects are defined on a video stream. In such cases, user evaluation quality tests are required. Relevancy is correlated to the scope, the budget, and the contract terms of the project, namely, to develop a 3D-media-related system for artists. As regards feasibility, budget and current state of the art were taken into consideration.

*Table 4: Understanding the requirements*

| # | Requirement | Unambiguity | Completeness | Consistency | Verifiability | Relevance | Feasibility |
|---|---|---|---|---|---|---|---|
| 1 | Enhance Web page user experience with 3D models | A web widget to view 3D models in a web browser | Major WebCMSes support; 3D commerce standard support; animation and sound support; CSS positioning support; no iframes and rectangular boxes; | - | Yes | Yes | Yes |
| 2 | Apply visual effects on short term videos | Web 3D Effects composer | Seamless 3D effects on video; web interface; no coding interfaces; access through major social media. | - | Subjective verification | Yes | Yes |
| 3 | Separate persons in video streams | Auto-rotoscope separation algorithm | No human segmentation process; web interface; no more than five clicks to finish the process. | - | Subjective verification | No | Maybe |
| 4 | Photogrammetry geometries correction | Fix Photogrammetry geometries | Auto-filling gaps in 3D models; auto-remove vertices when they are too dense; symmetrize objects. | No information for gaps | Subjective verification | No | Maybe |
| 5 | Rotoscope video for applying movement to 3D animation | Photo to 3D animation algorithm | Estimate human skeleton automatically from RGB camera ; apply movement to 3D models rigs. Web interface accessible to non-programmers. | - | Yes | No | Yes |
| 6 | Separate Photogrammetry generated texture layer to n-layers | Materials Demux algorithm | Provide one image, receive many images of the same size but for different layers (albedo, roughness, metallic, | Multiple solutions problem | Subjective verification | No | Maybe |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | | normal); web interface to test on a 3D model live. | | | | |
| 7 | Promote 3D work of scientific illustrator artists | Personal WordPress blog with 3D content | Allows one to upload, view, and interact with a 3D model. Support main formats: FBX, OBJ, GLB, 3DM, and PDB for biology. Walk inside virtual spaces. Security of models. | WebGL is not 100% secure. | Yes | Yes | Yes |
| 8 | Make an environment for real-time productions from home | Distant real-time production | Place two actors in the same context even while they act remotely. Support of Skeleton estimation through web RGB cameras (MediaPipe; Cao et al, 2019). Apply animation to 3D models that act as puppets of actors. | - | Yes | No | Yes |
| 9 | Allow children to provide realistic animations to avatars in 3D spaces | Animate 3D models with camera input | A child can apply 3D animation to a 3D avatar using only an RGB camera (MediaPipe). Security issues of cameras should be dealt. Create cameras that embed safety mechanisms on hardware. | Camera security issues | Yes | No | No |
| 10 | Allow music composers to visualize photorythmics as textures of 3D models | Audio to texture to 3D visualization component | Generate immersive 3D spaces with photorythmics and sound. Web interface. Blogspot generation capability. Node editor for real-time changes. 3D editor for objects positioning. Uploading of new 3D models. Position primitive 3D models, for example, cubes. | - | Yes | Yes | Yes |

In the second step, brainstorming based on the aforementioned scenarios was carried out to discuss and present ideas on the tools/solutions necessary to develop/implement to accomplish the TO-BE scenarios among stakeholders, namely an experienced artist, an experienced journalist, and the project technical manager. Requirements' elicitation was an iterative activity, where brainstorming and interviews were used. A specific template form for the requirements definition process was defined as an extension of Table 4 with more project-related details that are out of the scope of the paper to present. Next, in requirements analysis, user requirements were clarified, categorized, and documented to generate the corresponding specifications. A crucial step is the "Approval of end-users". Therefore, we conducted a survey in order to mine the opinion of the end-users, namely artists, journalists, architects, and informatics scientists, regarding the posed requirements.

## 2.3 Requirements Validation

We have allowed media content creators to evaluate and validate the proposed TO-BE scenarios and the requirements that were found during the Preparation and Elicitation phase. A total of 47 individuals participated in our survey who were reached from social media such as Facebook, Twitter, and LinkedIn. The survey consisted of four parts, Part I that consists of questions regarding the demographics of the participants. Part II that contains the previous experience of the participants (the AS-IS scenarios). Part III that contains the main results, namely the validation results of the proposed scenarios (TO-BE scenarios). Part IV that is an overview of the comments that participants have provided.

**Part I – Demographics:** The nationality distribution of the participants can be seen in Figure 2. Most of the participants were located in Greece (75%) whereas the rest of the participants were located in Belgium, Bulgaria, Germany, Hungary, Italy, Malta, Spain, Switzerland, and the United Kingdom. As can be seen in Figure 3, 18.39% of the participants are from Journalism and Mass Media Communication (MMC). Significant participation is from Fine Arts (8.17%), Informatics (8.17%), and Architecture (6.13%). Other participants are from Advertising, Design, Education, Gaming, Humanities, Marketing, and Social Sciences with 1.2%. The age group distribution of the participants can be seen in Figure 4. Most of the participants are from the 18–25 yo group (49%), followed by the 25–35 yo group (28%), 35–45 yo group (15%), and the 45–55 yo group (8%). As can be seen in Figure 5, most of the participants are from universities (55%), Private Companies (24%), and Research Centres (13%). Other types of employment are Freelancers (4%), Media Organizations (2%), and NGOs (2%). As regards individuals from universities, their age distribution can be seen in Figure 6. Mainly Bachelor and M.Sc. students participated with a percentage of 80%.

From the analysis of demographics, it is seen in Figure 2 that all participants are Europeans, with a 75% stemming from Greece which makes the research mostly focused on the Greek media status quo. As regards the educational background of the participants, it is seen in Figure 3 that Journalism and Mass Media Communication, Fine Arts, Informatics, and Architecture dominate. These disciplines have flexibility regarding the use of tools both for 2D and 3D design. As regards the age group of the participants, high participation of youth (18–25yo) was observed at about 50%, and there is a gradual decrease to half when observing older groups, until the 45–55yo group with 8% (Figure 4). A remark is that no participants were above 55yo. This can be interpreted to meant that youth is keener on using electronic tools for artistic content creation. As regards the organizations that participants stem from, it can be seen in Figure 5 that most of the participants were from academia. Private companies also have a strong presence with 24%. From Figure 6, it can be observed that participants from universities are mostly students with a percentage of 80.

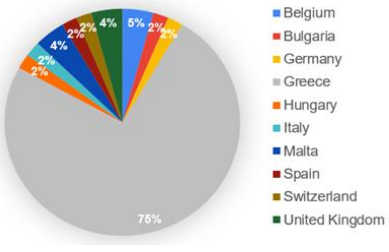


*Figure 2: Nationality of the participants*

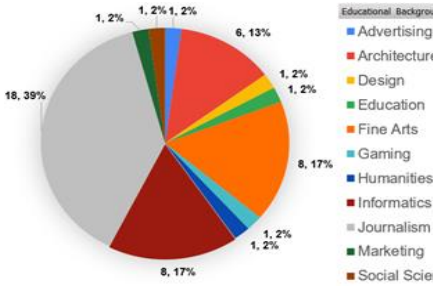


*Figure 3: Educational Background*

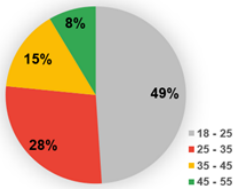


*Figure 4: Age group*

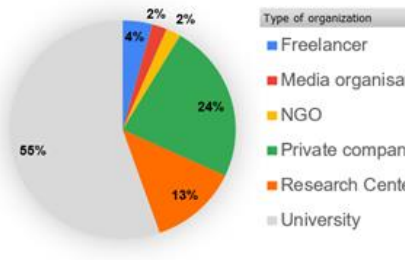


*Figure 5: Organization of participants*

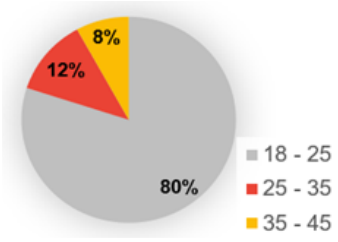


*Figure 6: Age group from universities*

**Part II - Previous Experience:** In this question, participants have provided information about the experience with existing software for creation activities. The results are shown in Figure 7. Most of the participants use Adobe Illustrator (46.8%) for graphics design, Adobe After Effects (40.4%) for applying effects on videos, and Adobe Premier for editing videos. Next, Unity3D graphics engine is used for 3D experience production (27.7%). Rhino3D is popular with 17% and Blender is popular with 14.9% which are both 3D design programs. Adobe XD for mobile and web experience design, Google Blocks for 3D design in VR, and Unreal graphics engine for 3D experience creation are also popular with 12.8%. Cinema4D, Maya-3DSMax design tools and Python programming language show some indication for increased usage.

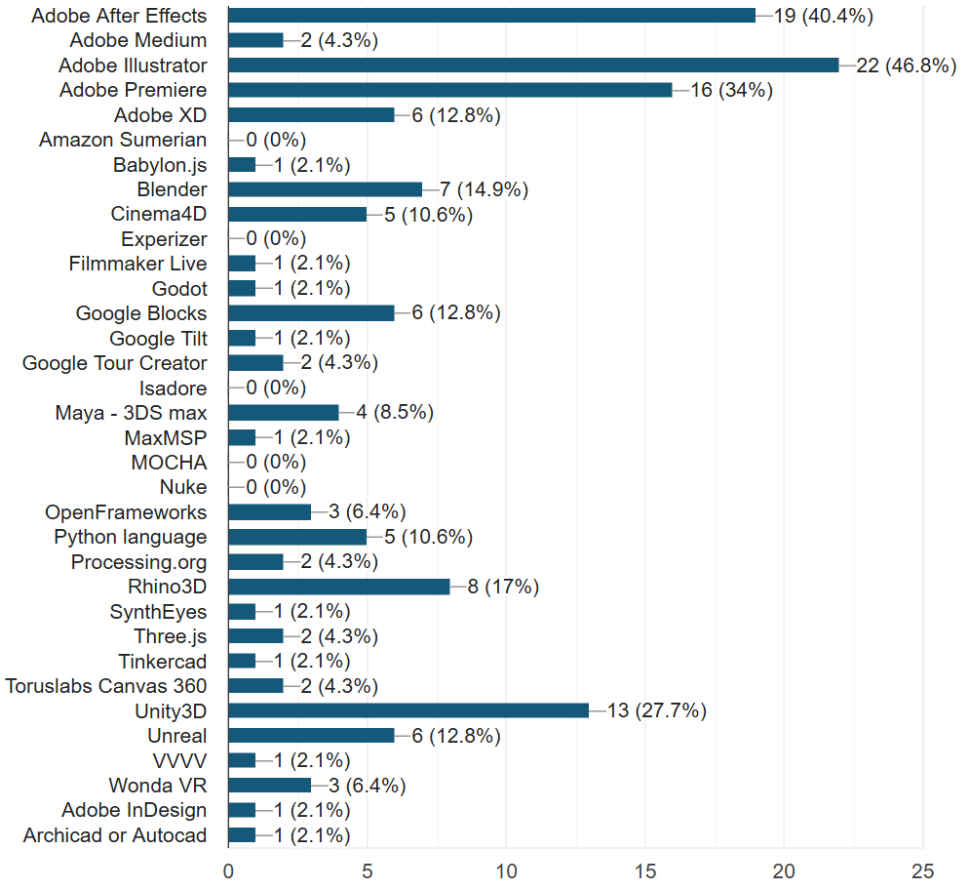


*Figure 7: Most used software tools for creativity*

The software tools popularity not only shows the popularity of tools but also the popularity of each type of content. 2D content creation tools of Adobe such as Illustrator, After Effects, Premier, and XD are very popular. This can be explained by the fact that most of the participants stem from media and fine arts. Unity3D and Unreal graphics engines show increased popularity as well, which indicates that game creation and storytelling are popular. However, they have lower rates than 2D creation tools which can be explained by the fact that they need programming skills and require more training to produce the final result. Rhino, Blender, Maya-3DS Max, and Cinema4D are also popular tools for 3D designing content. The increased rates for Rhino can be explained by the participation of Architects in the research as it is a tailored solution for them. Google Blocks is a surprise as it was not expected to receive more votes than Adobe Medium and Google Tilt. These are all 3D design tools that are used through VR glasses, but Google Blocks is only for low polygon models whereas the other two can achieve more realistic models. It seems that low polygon design is more attractive to be done inside VR environments. Python language is also popular, but it seems to be affected by the participation of the informatics experts.

Participants voted on which software tools they use for promoting and monetizing their work. The distribution of the most voted tools is shown in Figure 8. It can be seen that social media such as Instagram is prevalent with 78.7%, Facebook follows with 72.3%, next is YouTube with 55.3%. LinkedIn is also popular with 42.6%, WordPress personal blog received 27.7%, whereas TikTok and Blogspot received 25.5% each. Lesser used software are Twitch gaming social media and Wix web page creator with 10.6%. CGtrader and Turbosquid repositories for selling 3D models have low percentages with 6.4%. Mozilla Hubs and Playcanvas as 3D space creation tools for multiplying activities received 4.3%.

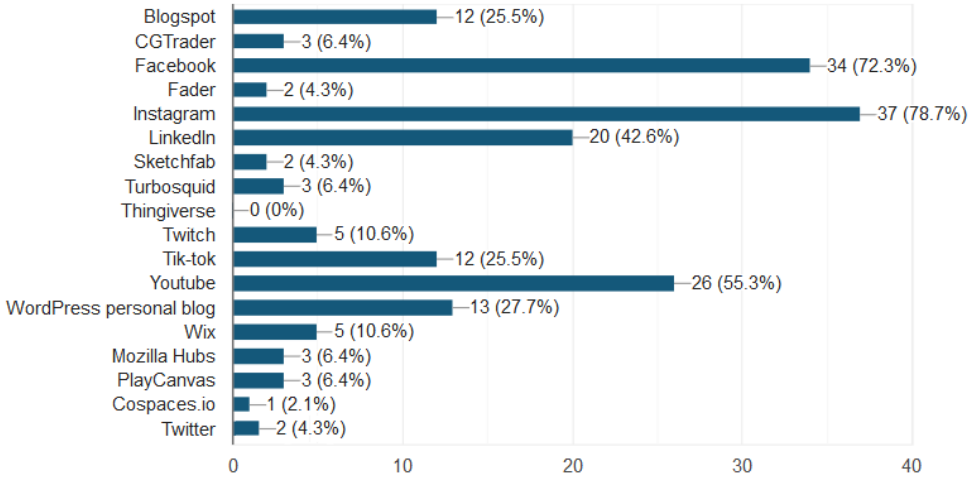


*Figure 8: Most used tools for promotion and monetization*

The results of the question regarding promotion and monetization software tools and platforms indicate that major social media platforms such as Instagram, Facebook, YouTube, and LinkedIn are widely adopted by media creators to promote and monetize their work. “Facebook Creator Studio” and “YouTube Studio” allow monetization through advertisement, whereas LinkedIn is for promoting career opportunities and portfolios. Instagram is a smartphone-centred application without a desktop front-end. The latest addons of Instagram such as Live Shopping allows one to monetize non-electronic art such as handmade paintings. Another pathway for promotion and monetization is through personal websites achieved by WordPress, Blogspot, and Wix. All these three platforms provide a simple way to make personal websites through templates without requiring programming.

**Part III – Questionnaire results for Scenarios Validation:** This is the main part of the survey. Six TO-BE scenarios were selected from Table 2, which are relevant to our project. Then, several statements per scenario were composed according to Table 3. Each statement can be rated on a 5-grade Likert scale where 1 corresponds to a disagreement and 5 stands for an agreement with the statement. The results for the evaluation of each of the six scenarios are presented in Figure 9, Figure 10, Figure 11, Figure 12, Figure 13, and Figure 14, respectively.

Scenario 1 refers to the visualization of 3D models in web pages. The participants disagreed with "Current web pages are adequate, there is no need to visualize 3D models" as indicated by the average score of 2.15 (Figure 9a). This reveals that participants consider the presence of 3D models in web pages as a great need. According to the results in Figure 9b, participants believe that "The 3D models will significantly increase downloading time and will require high-end client devices". The need for 3D models in web pages stems from the fact that the current web page design software does not support the insertion of 3D models in personal websites pages as can be inferred from the results Figure 9c.

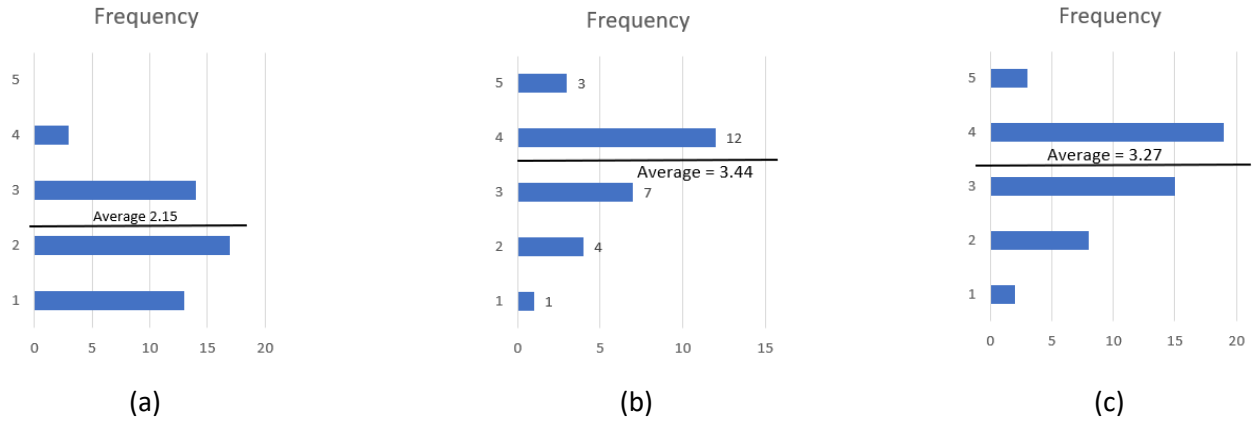


*Figure 9: Scenario 1 validation results*

As regards Scenario 2 about the question "Do 3D special effects in web or mobile?", the results are as follows. For the statement regarding the belief that mobile and web software will never have the potential of desktop devices, most participants disagreed with a score of 2.27 (Figure 10a), indicating that users have a strong belief that mobile and web technologies are capable of being incorporated in the media production chain. As regards the choice among web or mobile technologies, most participants answered with web technologies with a score of 3.19 (Figure 10b) vs. 3 for mobiles (Figure 10c). Further analysing the results, 17 participants voted for web technologies against 12 for mobile technologies.

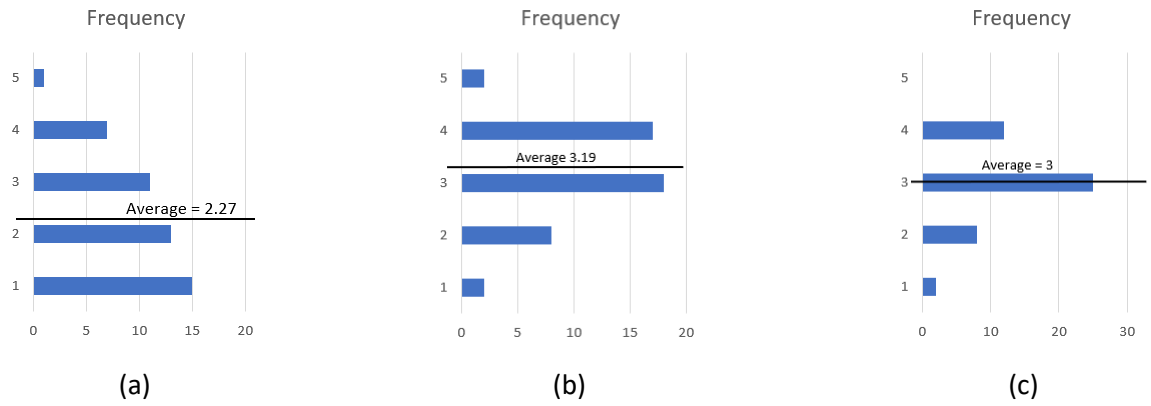


*Figure 10: Scenario 2 validation results*

As regards Scenario 3 about "Making edits on video streams inside VR worlds", the participants agreed that it is difficult to do video edits inside a 3D environment with an average score of 3.25 (Figure 11). In general, the participants were sceptical about this scenario.

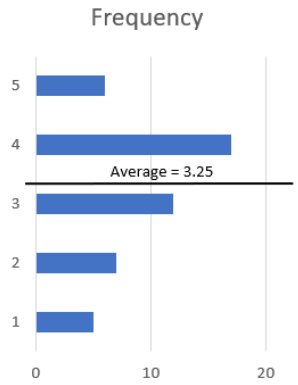


*Figure 11: Scenario 3 validation results*

As regards Scenario 4 about "Promoting artistic media inside VR worlds", the statement about the commercial saturation of the media due to high exposure has received strong disagreement with a score of 2.34 (Figure 12a). This reveals that VR worlds are suitable for promoting artistic and media work. Almost the same was the response about the risk of the 3D models being stolen if exposed in VR environments with a score of 2.19 (Figure 12b) indicating that participants consider VR spaces as a safe place for 3D content. As regards the introduction of a fee to enter VR spaces, the users show a disagreement with a score of 2.70 (Figure 12c). As regards the exposure to and accessibility from social media, the participants have shown great agreement with a score of 3.55 (Figure 12d), indicating that collaboration with social media is an important factor for such a scenario. Even stronger was the opinion that the existing repositories do not promote and visualize content adequately with a score of 3.59 (Figure 12e).

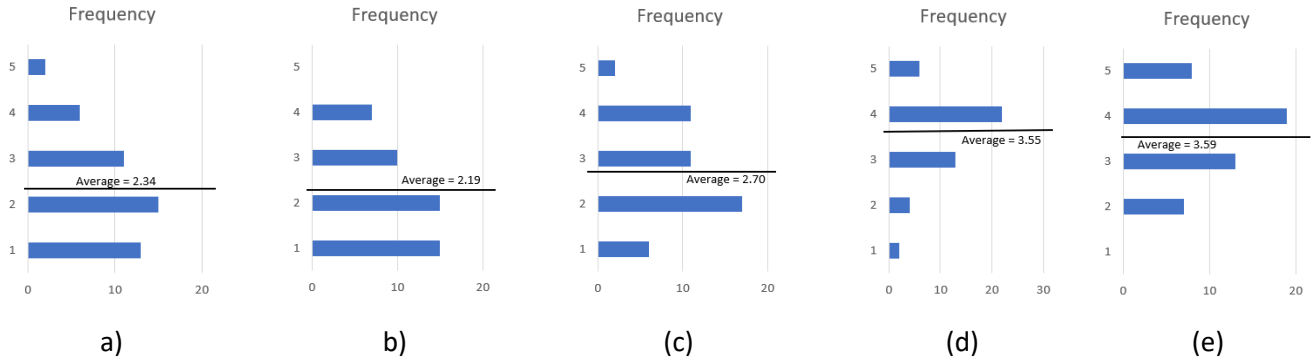


*Figure 12: Scenario 4 validation results*

As regards Scenario 5 about the "Theme dedicated VR experiences for personal or joint blogs", the participants disagreed with the statement "It is a bad idea since next generation social VR environments are more attractive" with an average score of 2.59 (Figure 13a). This indicates that there is innovation potential in such a scenario. In general, they found the idea interesting because the features offered can be tailored better to personal interests (average score 3.82, Figure 13b). Also, about the idea of "COVID19 VR nano-worlds", they found it interesting with the highest score observed among all questions, that is, 3.97 (Figure 13c).

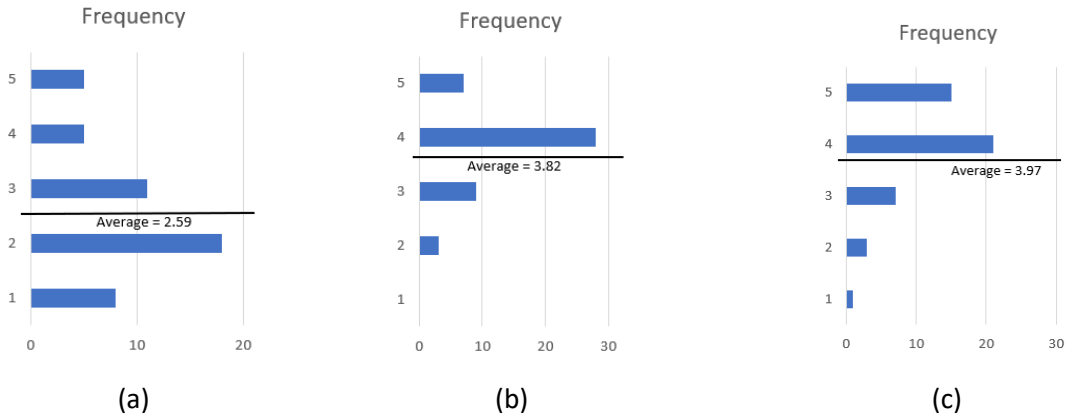


*Figure 13: Scenario 5 validation results*

As regards Scenario 6 about the capability of designing 3D models inside VR spaces, the participants disagreed with the statement "The idea is bad because traditional software allows easier to design with mouse and keyboard" with a score of 2.23 (Figure 14a). They also disagreed that such a design method will be for amateur artists with a score of 2.59 (Figure 14b). Obviously, the idea of designing 3D models inside VR spaces is good for both amateur and professional creators, however, as it is inferred also from the comments received (see Part IV below), the interfaces of existing approaches are not very intuitive.

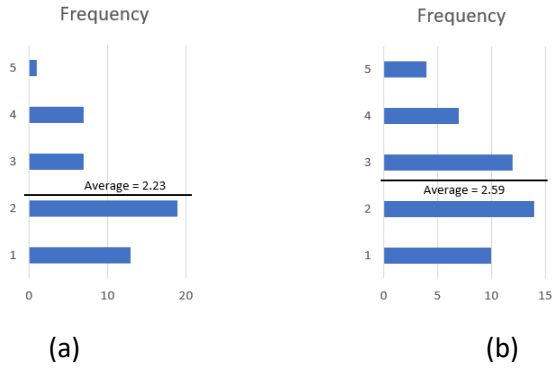


*Figure 14: Scenario 6 validation results*

**Part IV – General Comments:** The overall comments received by participants were divided into four categories, namely Interfaces, Barriers, Benefits, and Personalization. The comments regarding Interfaces are:

1. Easy to use interfaces.
2. Easy ways of editing in a 3D environment.
3. Accessible, cloud-based co-design software with a solid source-control system for digital assets of all kinds.
4. VR can be a multimedia converged experience, and users should have the freedom to create and project their contents inside and outside the VR space.
5. An online platform for VR content is interesting for idea exchange/moving the technology beyond the status quo. The same way that 3D models are available for sketch-up, perhaps there could be a repository for VR environments, for drag and drop solutions for scenes, educational subjects, etc.
6. VR environment to create 3D models.
7. VR/AR engineering and training application including detailed human–computer interaction via fingers haptic devices.

Comments 1 and 2 refer to the easiness of the interfaces and particularly the editing capability inside the 3D environment. Comment 3 refers to the accessibility of the solution mostly achieved through the cloud and to the coverage of all types of media. Comment 4 refers to media convergence, that is, to allow the combination of several types of media to create a new solid type of media. Freedom to create, project personal content, and expose media outside VR environments should be also possible. Comment 5 is on an online platform for VR content, namely, to make a repository for whole scenes as pre-built solutions for certain scenarios. Comment 6 is on the idea to create 3D models inside the VR environment. Comment 7 suggests that VR applications for training are also interesting if they also exploit peripheral devices such as haptic devices.

It was observed that most individual comments regarded the interfaces of VR environments. They are related to easiness, accessibility, and haptic devices interconnection. It is true that current VR systems require a process to download, install, and learn the interfaces of the design tool. As regards accessibility, many devices still do not allow the use of VR headsets with prescription glasses. Another issue is the cost, as they are still expensive for amateur creators with a cost of around 400 euros. All these seem to be the most important barriers for content creators. Another evident issue is the cloud availability and the co-design capabilities. Many creators would like a cloud-oriented application where assets, scenes, and projects can be easily created, shared, and co-edited in joint repositories. For the time being, most design applications limit the design capabilities to local use or single-user cloud repositories. The low penetration of VR design tools due to the high cost of devices might be a reason that such repositories did not gain acceptance from users.

The comments regarding Barriers are as follows:

1. I'll stick to 2D, 3D is overpriced.
2. Basic training material.
3. Online VR/AR applications that utilize 5G to render 3D models in real time.
4. I believe that collaboration tools would be great to have but sharing spaces to create 3D seems one or two steps too far or too fast.
5. A better user experience while drawing/designing and a sandbox toolset for publishing my work in a device-independent environment.
6. The key is seamless interfaces where creators do not wow at the tool itself, but rather it becomes transparent, and they can just focus on their creative processes.

Comment 1 refers to the high price of 3D models with respect to the 2D media. As regards Comment 2, it highlights the need for training materials for using the authoring tools. Comment 3 regards the internet speed that allows one to download 3D models on VR/AR applications on the fly improving thus the user experience. Comment 4 refers to the collaboration of many users for VR environments or 3D model creation but the sharing of spaces for collaboration with other users is a far-fetched goal. Comment 5 focuses on the device-independent environment for publishing the results, whereas Comment 6 is highlighting the need for reducing cognitive burden so that the user can focus on the creative process. It seems that users have many concerns regarding 3D content creation. Most of the Barriers refer to the design inside VR environments and only one refers to the expensiveness of 3D content as for the time being too expensive to be used as an artistic improvement for various applications. As regards VR environments, it seems that users are missing enough training material. There are limited resources for viewing what other designers are designing in VR as the approach to create 3D models in VR space differs from the traditional way of designing 3D models. Shared spaces for multiple users to design seems to be a far-fetched target for the time being as many issues are not solved for a single user. A comment also refers to the 5G connectivity that will enhance downloading times.

As regards the category Benefits, the comments received are as follows:

1. Accessible and mutual professional benefits.
2. Variety of subjects and costs; interactions between the designers.
3. New mediums that will expand my creativity.

It is inferred from Comment 1 that creators should have some benefits after creating the content. Comment 2 refers to the variety of media so the space can be commercially exploitable, as well as to the interaction among the designers in order to improve the result. Comment 3 refers to the combination of media that will allow one to enhance creativity.

As regards the Personalization category, two comments were received, namely:

1. VR personal spaces rather than ad hoc 2D web pages.
2. VR experience can be both realistic and illusionary, and visualization styles should not limit the way users experience the virtual reality space.

Comment 1 refers to the use of VR spaces as a replacement for the traditional 2D web pages. Comment 2 is addressing the fact that VR experiences can be expanded to illusionary spaces rather than limited to the representation of real spaces. In general, the comments refer to monetization, namely the mutual professional benefits and the existence of a variety of subjects and costs. This might be interpreted as an increased need for a monetization plan that allows the provision of mutual benefits to collaborators in the VR space. Another comment states that new mediums that might be interpreted as new media can actually merge all types of media.

## 2.4 Validation Results Summary

Several dimensions of our requirements engineering approach are topics for discussion. Firstly, the demographics of the survey suggest that the tools to be developed should be tailored to two extreme poles, namely a) fit well to youth, for example, using a low-price plan, or b) to companies with a high-budget plan. As regards the creation tools and the type of software, in general, we can infer that big companies dominate with tools tailored to experts in design and in programming. Adobe desktop tools for 2D content creation are widely used by designers, a trend that is difficult to change. As regards 3D content creation, tools are not limited to certain brands, apart from the Unity3D graphics engine that dominates in programming XR applications. Promotion and monetization are correlated to the exposure of content on Instagram, Facebook, YouTube, and LinkedIn. It seems that Instagram is more popular for 3D content creation promotion than any other social media platform. Another fact is that content creators tend to have their own website in WordPress, Blogspot, and Wix as these three platforms provide a simple way to make a personal website through templates.

As regards the evaluation of scenarios, Scenario 1, namely the exposure to 3D models on websites is ranked highly according to the responses of the users. Content creators would like to publish their 3D models on web pages as long as the downloading time and the end-user device requirements are kept low. It is also inferred that creators do not have or do not know a certain methodology to expose 3D content. If this is combined with the conclusions of the previous paragraph, it seems that the content creators do not know how to embed 3D models in WordPress, Blogspot, and Wix. Exposing 3D content (or images of it) on Instagram, Facebook, YouTube and LinkedIn is very important for dissemination purposes. Scenario 2 about achieving visual effects on web or mobile is bending towards web applications for both amateur and professional users. This might be interpreted to mean

that the latest technologies such as 5G internet, WebAssembly, and the upcoming standards (WebGPU[7]) have increased the creators' trust in web browsers. The conclusions from Scenario 3 are rather mild. Users are not inclined towards a direction indicating only that the editing of videos inside VR space is difficult or not interesting to them. Scenario 4 about the promotion of artistic media inside VR worlds has also raised the interest of participants. It is inferred that creators believe that their content is not promoted well on current repositories. The accessibility of VR spaces from social media and the exposure of content to social media is also important. As regards Scenario 5, the idea of a personal or joint blog in a VR space seems to be attractive to better express personal interests. The idea of a nano VR space with information regarding biology, and most specifically COVID-19 mechanisms, seems to be more attractive than illusionary spaces of other types. Towards this direction, one of the most unexploited resources in Art, but prominent in Biology, is the Protein Data Bank (PDB) (Berman et al, 2000). It provides rich data for visualization, such as the protein-related SARS-COVID-19 model as shown in Figure 15.

*Figure 15: Scientific 3D models repositories can be used as a source of 3D content (Berman et al, 2000)*

## 2.5 Use Case 1 Pilot: Citizen Journalism and Mobile Immersive Journalism

This use case comprises of two scenarios, a) *Citizen Journalism* and b) *Immersive Journalism*. While the citizen journalism scenario stresses fast mobile content creation, the immersive journalism scenario focuses on immersive content creation and consumption. The main idea is to allow a quick capture-and-provide process of immersive media towards MV and making this immediately available in a 360-story format by using templates. In a later process, the stories that get created this way remain editable and can be improved by editors in the newsrooms, subsequently.

**Target Groups**

Target group "Creation" (incl. experience level): Who will create content in your scenario?

- Citizen journalists, freelancers
- Professional journalists who supply materials online

Target group "Consumption": Who will consume the content in the end?

- General public

[7] https://www.w3.org/TR/webgpu/

### 2.5.1 Scenario Description for the Authoring Tools

This section highlights, how the authoring tools are being used in the context of the use case.

**Creation:** The demonstration has reached Trafalgar Square and people are gathering here, slowly filling the square. Henry has placed himself on top of the steps near the statue of King George IV to get a better view. So many things happen around him that he has a hard time deciding where to point the camera. This is a great opportunity to use his 360-degree camera: it points in all directions at once and the journalist can enhance the scenes later. The camera is put in a suitable place to take a 360-degree film which Henry uploads to his MediaVerse library (the storage for his media content).

Henry selects one of the handy templates, “Mass Event”, which provides some very useful best practice tips as to how to best enhance his basic image with relevant additional media and information. He follows these tips and adds 2D images, audio sequences, text banners, and plates to help his audience grasp the atmosphere of this demonstration. Within minutes he has produced quite a decent first version which he uploads for his editors to check and publish.

His editor, Mary, doesn’t want to waste time as the viewers demand fast information on this controversial demonstration. She publishes a link to the 360-degree film to the main website where the demonstration is featured as a top article. An event like this is both hot news and an important event for so many Londoners that Mary starts preparing a high-end version in parallel. Based on Henry’s first “quick and dirty” version, Mary and her team add further footage and images of demonstrators, police, and bystanders, interviews that were taken after Henry’s first upload, and even posts on social media. As these additions need to be treated with care concerning personal rights and intellectual property, Mary and her team add these one by one and only after they have cleared the rights.

**Consumption:** Graham and Laura have been following the news in their home in Muswell Hill. Now they feel the need to join the demonstrators, and they keep scanning the media on their way into town to make sure they don’t get caught in any violence. They are relieved when Laura receives a chat message from her friend Carol that contains a link to a 360-degree video. “Look at this! Looks fairly safe, I think!” It shows that the demonstration on Trafalgar Square, while there are lots of people, is quite peaceful. Holding the smartphone at arm’s length and turning all ways Graham and Laura almost feel like they are already amid the action. Clicking on the hotspots to hear interviews with police officers on site or listening to the relative silence on the square reassures them further.

Meanwhile they have arrived at Charing Cross and they can hardly push through the massive crowd. Graham takes a short video of the crowd of what are clearly peaceful demonstrators, though many of them without masks. Laura encourages Harry to upload the video to the website they saw. She had seen a red banner there asking users to share their recordings. Graham thinks this may be important information for other people like them and he uploads the clip, adding his name, and a short caption (“peaceful protesters at Charing Cross”) and clicks the “free to use” checkbox.

Back at the studio, Mary receives an “incoming content” alert, checks the video, likes it, contacts the contributors in a short message and adds their content as soon as all rights are cleared.

### 2.5.2 Description of Technology, Planned Development and Deployment

As a baseline technology, the use case will make use of the *FADER* SaaS (https://app.getfader.com), initially employing it via its native user and data management. In a later development stage, the platform is integrated as a 3rd party tool using MediaVerse node(s) for registration and authentication as well as accessing media from the respective node(s).

Content created within this use case shall be available via streaming/download from MediaVerse nodes, SocialMedia (embedded), Website (embedded).

**Templating:** To provide an entry point for the technical implementation of the templating functionality partner DW is setting up a template *FADER* project. Using the standard web-based editor, the template will express the content items and positioning, scene defaults (e.g., intro scene, menu scene) and order and scene that should be available. This will allow the creation of a very low effort mobile application to quickly produce 360-degree contents by simply providing the media files and assigning them to the templates.

## 2.6 Use Case 2 Pilot: Co-creation of New Media Formats

The use case was defined to test the availability of tools towards co-creation of accessible immersive storytelling content. The objective was twofold: to understand the functionality of the existing tools towards generation of accessible immersive media content, and to understand the potential capabilities by non-professional consumers towards becoming producers of content.

### 2.6.1 Scenario Description for the Authoring Tools

Two scenarios were created for the two objectives.

**Scenario 1 Accessibility in Immersive Media Storytelling**

UAB team set up to generate an accessible immersive story using *FADER* SaaS (https://app.getfader.com). The three services: subtitling, audio description and sign language were chosen. The reason is the three different formats required by each service, and the IT development that will be needed for each case: subtitles/text, audio description/sound, and sign language/video.

Generating content (subtitles, audio description, and sign language) by hand was time consuming, and the results are neither accessible nor aesthetically pleasing, still it has been an important exercise towards MV partners understanding the accessibility services/formats challenges.

**Scenario 2 Co-creation of new media formats**

Two activities took place:

1. Fader was presented at Som Fundació.

UAB started to collaborate in November 2020 with Som Fundació (an association of persons with cognitive disabilities). The objective of this collaboration was to identify possible areas of co-creation, where MediaVerse technologies could be relevant to Som Fundació's daily activities and user interaction. As a result, three areas were selected: a) to raise awareness of the persons with cognitive disabilities and their rights, b) to work against isolation and to increase quality of life, and c) to support decision making. A first pre-pilot scenario was agreed

with Som Fundació carers. Jointly with UAB an accessible video was produced to assess the challenges linked to accessibility storytelling in 360.

2. Fader was presented to a group of young migrants.

The UAB was chosen to present MEDIAVERSE research in an Arts and Technology residence organised by FABERLLULL in Olot (Catalonia). One of the residence main objectives was to have a local impact in the city of Olot. To this aim UAB contacted two local NGOs supporting young migrant integration: Cepaim and the Xarxa Autònona d'Acollida de la Garrotxa (XAAG). UAB team aimed at identifying participants' areas of interest where MEDIAVERSE technology can be piloted. In order to collect primary research data, UAB carried out focus groups with the same participants that had been recording in 360. Focus groups aim at learning about users' attitudes, beliefs, and reactions to concepts (Bryman 2016). Two focus groups were run, one with three young migrants (two men and one woman) and the second with social workers from the NGOs (three women and one man).

Young migrants recorded 360 videos in local ateliers of La Iera Association. Young migrants followed the suggested script prepared by UAB, but there were decisions to take such as deciding the location of cameras or preparing the questions for an interview with La Iera artisans. Two groups of 4 males and 4 females accompanied by facilitators and members of the UAB and NGOs social workers recorded the 360 videos and photographs. The main objectives for the videos audience were to 1) Empathise with the young migrants, 2) See the relationship between the young migrants and the city of Olot, 3) See the relationship of the young migrants with someone from Olot, and 4) See the feedback from young migrants at the end of the activity.

**Resources**

Group A: https://app.getfader.com/projects/3308e1ce-f334-4208-b2fa-9ea44197c0f2/publish

Group B: https://app.getfader.com/projects/80420066-cbe2-418f-ae5d-ac9d8e70625a/publish

**Discussion**

- Participants agreed that the activity has been innovative (they have never done anything like it before) and they enjoyed it.
- They highlighted three items of experimenting with new technologies (360 cameras), working in groups, and meeting new people from outside their environment.
- Confidence stands out as a key term for all participants. The activity helped them to build confidence in themselves and within a group.
- A topic of interest was finding activities and environments where they could share life experiences.
- They ranked the activity as positive; they liked the fact that had not been treated as "illegal immigrants" and that they were given room to be themselves and explain their life events.
- As for the future, all participants were interested in expanding their technological and audio-visual knowledge and skills for both leisure and work purposes.

**Conclusion**

- 360 technology is useful to break stereotypes through awareness raising campaigns.
- Creating 360 technology training as a mean for improving young migrants labour inclusion.

**Outcome by UAB team**

- Participants were not ready to use immersive technologies without training or professional guidance.

- There is a clear impact of the transformative personal and social nature of storytelling, and the immersive media offers the wow factor to interest young migrants.
- Accessibility in immersive story telling in MV must be stalled until technology is ready for the production and visualisation of accessibility services.
- Accessibility research and innovation activity is now moved to another of the MV descriptor: copyright management which will be studied and analysed in the *RACU* workflow (Use Case1) (see D5.1 User Case 2 and D7.1)

### 2.6.2 Description of Technology, Planned Development and Deployment

As a baseline technology, the use case will make use of the *FADER* SaaS (https://app.getfader.com), initially employing it via its native user and data management. In a later development stage, the platform is integrated as a 3rd party tool using MediaVerse node(s) for registration and authentication as well as accessing media from the respective node(s).

Two issues were identified and reported as WP2 requirements:

a) the need for a 360 editor for accessibility services: subtitling, audio description and sign language.

b) the need for a 360 accessible media player to allow for accessibility services

These requirements have been noted and discussed within MV consortium and their availability will be considered at later stage of the MV project development.

Still the opportunity to experiment with an immersive video storytelling tool allowed us to generate a first prototype of what may be and how it may look and feel accessibility services in such environments, pointing also towards clear challenges.

To this aim UAB took advantage of *FADER* "hot spots" to generate 2D accessibility services. The result is a movie created for the GAAD2021.

The experience aims to transport the viewer to a series of 360-degree cultural heritage locations in nature near the city of Banyoles (Spain). In each scene, closed captions, audio description and sign language interpretation have been integrated to offer content that is accessible to all citizens.

## 2.7 Use Case 3 Pilot - Hybrid Intelligence Experimental Artwork Series

Use case 3 pilot is about cultural spaces. It will be demonstrated aside another project which is related to the installation of a 5G network in a small scale. The target of the 5G project is to make a music concert where each member of the music band is situated at a different location. However, through 5G network they can play all together via electronic means, namely through an iPad tablet and a Unity3D application. Also, the people that are connected within the 5G network will be able to listen to the concert as one band. For MediaVerse, the following scenario is foreseen. As this event has only audio collaboration, we consider it as a challenge in MediaVerse to have also visual collaboration, i.e., to try to make a low-cost virtual production by putting all the members of the band into a virtual stage.

### 2.7.1 Scenario Description for the Authoring Tools

The collaboration of many musicians or actors into one scene is a challenging task in media production. It involves the use of expensive lights, chroma key backgrounds, and one stage where all actors are located. With the advent of new video technologies that are described below we can devise a very low-cost virtual production with a web browser as an end-user client software. The role of the authoring tools is to setup the stage where all actors perform and derive several instances of the stage for generalization and upscaling.

### 2.7.2 Description of Technology, Planned Development and Deployment

The overall concept for T5.5 aiming at use case 3 is depicted in Figure 16. Starting from the left, in the first step the actors (musicians) open a URL link using one of the most common devices, such as Desktop, Laptop, Tablet, or Smartphone - given that it has a web camera (preferably the selfie camera). The URL opens a website that asks how users want to be represented, either as camera stream or 3D model or both. In the case of camera stream, the system isolates the actor from the background using MediaPipe AI[8]. This library is using the latest state of the art algorithms that are based on artificial intelligence and deep-learning for a fast and accurate - as much as possible with the low budget lights and background conditions - implementation of selfie segmentation. The actor video stream is attached to a 3D surface, i.e., a plane, that hovers onto a 3D scene, thus the user can walk as a video stream anywhere in the 3D scene. Multiple players can be immersed simultaneously in the scene with the help of a WebRTC server (Networked-Aframe[9], see Section 4). The actors can disseminate the link to other actors or audiences to join. Thus, a multi-playing VR environment has been constructed. 3D audio is also supported, and the audience can socialize privately without been heard from distant avatars.


*Figure 16: General concept of use case 3*

In a second step, the experience will be transformed into a template and will be inserted in an authoring tool named VRodos. The authoring tool, a plugin for WordPress, can modify the 3D models of the scene and generate new links either for actors or audiences. MediaVerse nodes can be used to fetch 3D models, sounds, videos, and images to be inserted in the VR environment.

[8] https://google.github.io/mediapipe/
[9] https://www.npmjs.com/package/networked-aframe

# 3 Immersive Storytelling Toolset

## 3.1 Context and Baseline Description

Within T5.3, an immersive authoring toolset will be developed for MediaVerse by integration of innovative authoring tools that abstract all the coding and programming skills needed for authoring immersive media experiences, allowing content creators to experiment with interactive 360-degree storytelling in an easy and fast way, diminishing technical entry barriers. The authoring environment includes asset provision, behavioural templates for application logic, and deployment procedures to allow platform-specific experience development for non-expert users. The specific objectives for T5.3 are:

- Provision of tools for enhancing 360-video with interactive capabilities
- Integration of capabilities for advanced 360 photo and video image processing to allow quality improvement, selective removal, object location and tracking. This is deeply tied to the outcome of the processing services developed within WP3.

## 3.2 Baseline Description and Previous Work

The baseline technology for the immersive storytelling toolset is the *FADER* SaaS (https://app.getfader.com). It is developed, maintained, and operated by Vragments. Fader allows content creators a low-level entry into authoring interactive 360-degree stories.

### 3.2.1 Backend and Backend Applications

The backend provides user, project, and media management functionality, as well as a series of processing services. When a user logs into Fader, several static page structures are rendered to allow management of projects, media assets and user data. It also serves the *Discover* page, a listing of all available public projects on Fader. Viewers can search for content or browse for projects by users.

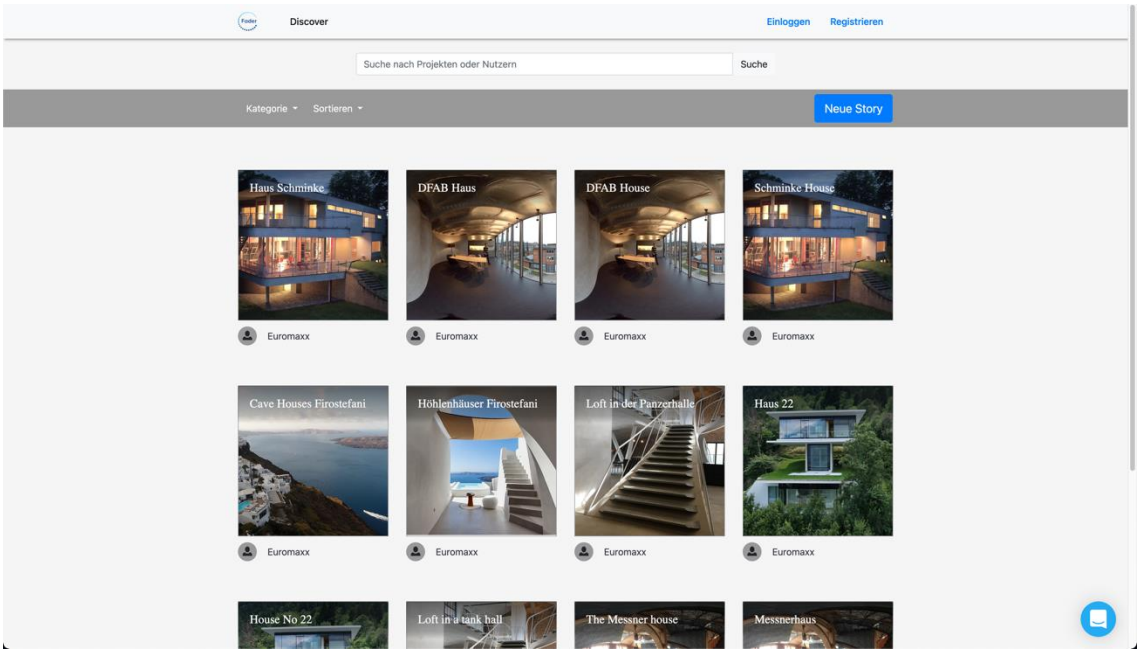


*Figure 17: Discover static page of FADER at https://app.getfader.com*

Logged in users have access to three additional page sections, *Media Library*, *Projects*, and *Analytics*.

Projects is the default view and empty at the beginning, prompting a user to get started on the first project. A user can start a new story, or edit, copy, and delete and existing story, or view its analytics. Starting a new story or editing an existing one, takes a user to the editor, the real frontend application of Fader, which is described in more detail in its own section below.

The *Media Library* consist of all media; a user has uploaded. Just like in *Discover* or *Projects*, the elements can be searched, filtered, and sorted. Clicking on a tile reveals the asset detail information.

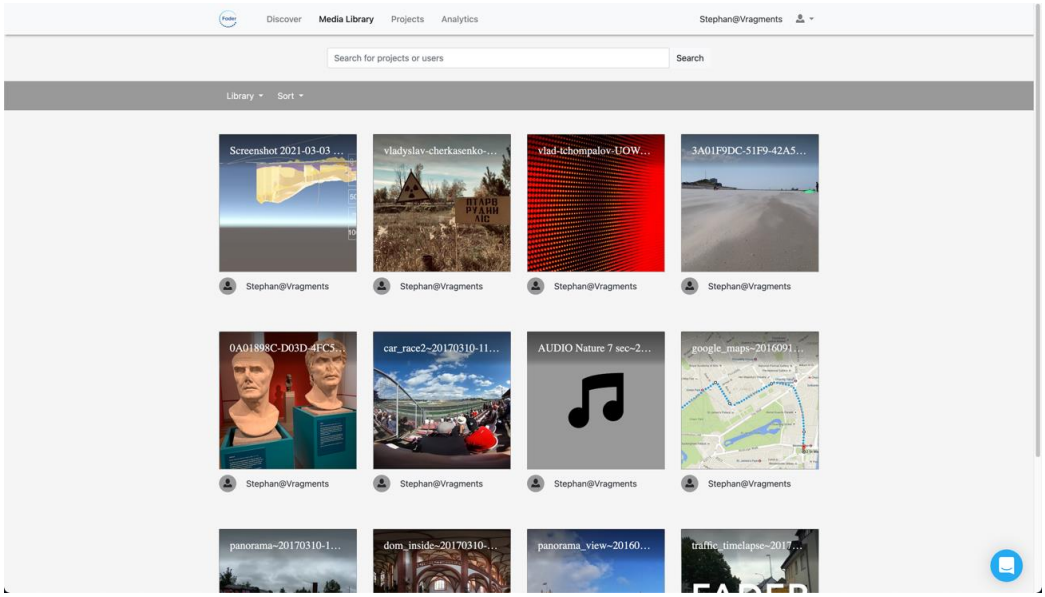


*Figure 18: Media Library overview*

The detail information of an asset includes a larger preview, some high-level metadata, and the license information. A *private* license marks an asset as only available for the user. A *public* license allows the asset to be used by anyone on the platform. Currently, these are the only license types, but a more fine-grained license model can be applied, if the requirements in the project demand such adaptation.

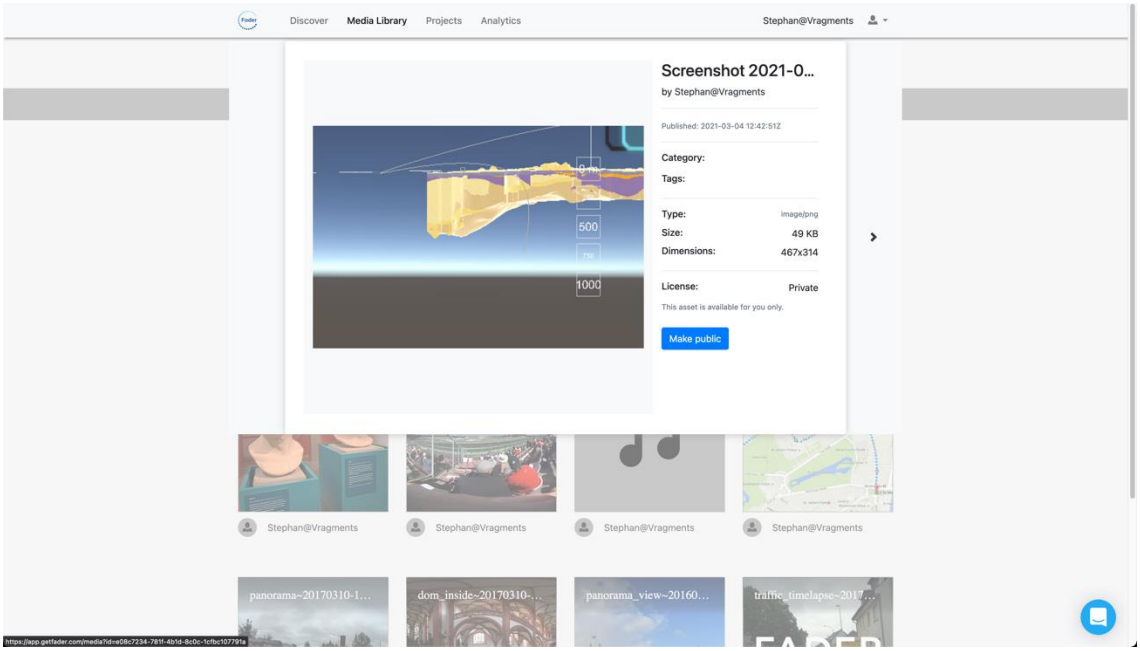


*Figure 19: Media Library asset detail*

The *Analytics* page provides a high-level overview on some key metrics for a users' stories and assets. Currently, Google analytics is used as the analytics provider. An analytics page is also available per story from the *Projects* page.

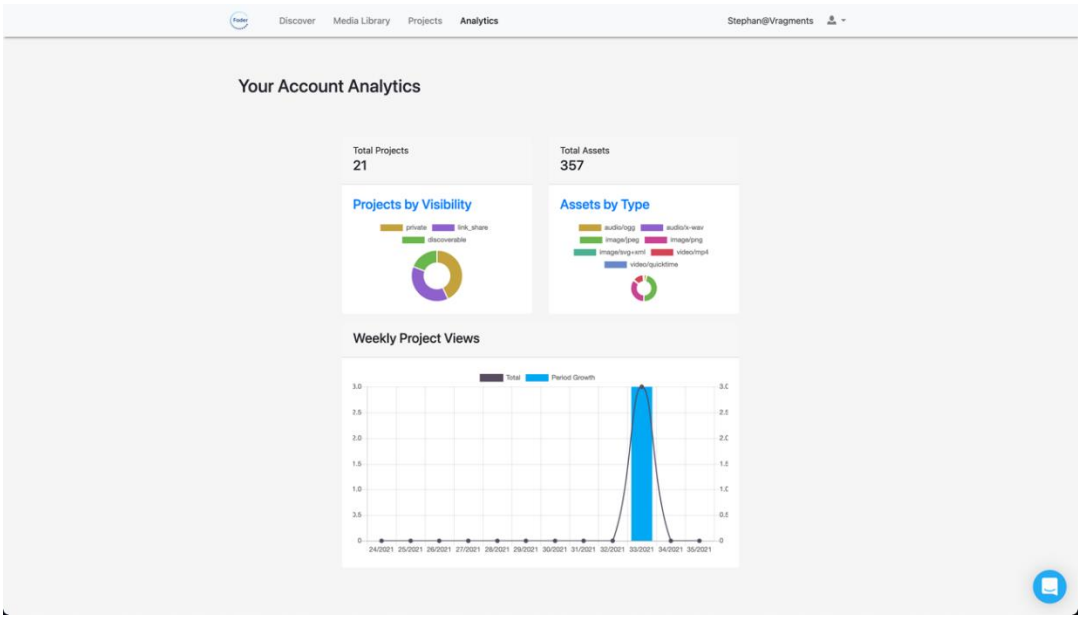


*Figure 20: Account analytics overview*

The backend also provides an administrative interface to view and manage assets, asset leases, projects, users, and to view statistics.

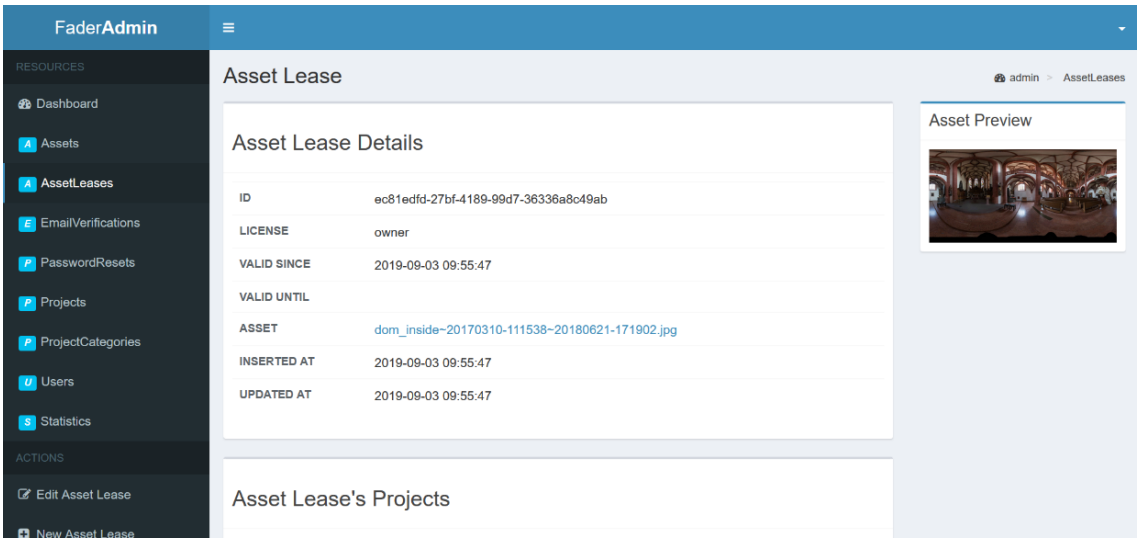


*Figure 21: Admin dashboard of FADER*

**User:** The user model is the central element in the FADER data structure, that has associations to most other models. Each user has an account plan and user data association, that determines feature and access levels.

**Account Plan:** Each user has an account plan that determines the available feature range. There are two major types of capabilities to distinguish users. For once, the nr_*_projects settings define, how many stories can be created and/or published. Other settings allow the customization of the appearance of the player. As such, customers of a higher account plan tier can choose to make their account appear tailored to their companies' corporate identity.

**User meta data:** The user data holds specific settings and optional parameters for a user. It comprises of the account plan data and specific settings. If a user does not have access to custom configurations, then default platform values are used.

**Asset:** An asset is the term for any media item that users can upload to Fader. They can be videos, images, or sound files. Visual media can be simple 2D or equirectangular 360-degree. A more detailed description of the assets is given in the section on media transcoding.

**Asset Lease:** An asset lease is a way to add a license to an asset and potentially share it among users and projects. Per default, when a user uploads media to fader, an asset lease is created with an *owner* license set. An asset lease can have associations to 0, 1, or more projects – depending on where it is used.

**Project:** The project model is a central part in Fader. It is a model that contains all relevant info to represent story meta-data, scenes, and content. Its meta data contains, among others, information on the owner, a preview image, and sharing status.

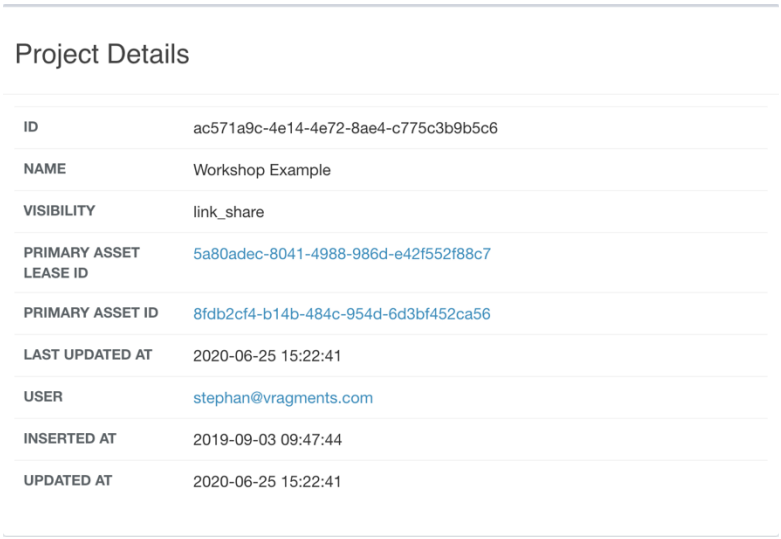


*Figure 22: Project metadata*

### 3.2.2 Media Transcoding

When media assets get uploaded to Fader, they are not only stored, but depending on the media type, a set of processes is started, from media analysis to creating media variants that allow adaptive presentation. For video files, we use Dynamic Adaptive Streaming over HTTP (DASH or MPEG-DASH). Image files get transformed into low, medium, and high-quality variants to support loading of lower quality images in case of lower bandwidth or performance of the device.

**Video transcoding:** The current video transcoding process is built on DASH-MPEG and delivers 9 quality steps: 720p, 1080p and 1440p with Bitrates ranging from 1000 to 9000K. The toolchain used is built on ffmpeg with an output manifest in m3u8 format using H264 codec and yuv420p as pixel format. This had the effect, that on iOS, the player had to be adapted for some versions, since there has been a switch in colour profiles.

The current video transcoding process will be greatly enhanced by the outcome of T3.3, which will allow OMAF compliant media to be displayed at much higher resolutions (up to 6k)

**Image transcoding:** An image that is being used in Fader can either be a 360-degree equirectangular image, or a traditional 2D image. This determines, how the image is being processed internally. In all cases, we render images in 3 quality steps (4096x4096, 2048x2048, and 1024x1024 pixels) regardless of source image. This has to do with the requirement to use textures in 3D that are power of 2 in dimensions. For 2D images, we also store the original dimensions and aspect ratio to correct the distortion by modifying the plane geometry that the 2D textures are rendered on.

Lower resolutions are also chosen, if bandwidth is low, or the overall performance of the viewing device is low. To support a higher framerate, the player will switch to downloading and rendering low resolution images.

Accepted image types are *image/jpeg*, *image/png*, and *image/svg+xml*.

**Audio transcoding:**

The output of audio transcoding is in Opus format encapsulated in Ogg containers (*audio/ogg; codecs=opus*). All original metadata is stored.

Accepted audio types are *audio/ogg*, *audio/mp4*, *audio/opus*, *audio/vorbis*, *audio/mpeg*, *audio/wav*, and *audio/mp3*.

**Media Library:** Each user has their own media library that all projects can source from. They are available both in the management view and the editor to pick media from. In the management view, assets can be set public or private.

**API:** The Fader backend exposes all functionality via API which allows different clients to be developed. Where applicable this API uses the `application/json` media-type to represent objects and response content.

### 3.2.3 Editor

The Fader editor is the main frontend application that allows users to create interactive 360-degree stories. The top bar contains all story specific settings and options, like setting title and author, adding media (from the library or via upload) and viewing or sharing the story. The bottom bar shows all the story scenes in order, allows to preview play the current scene, give it a name and register it in the scene panel for the player frontend. The right bar contains all media tools for the currently selected scene to add, edit, and remove media assets from the scene.


*Figure 23: Editor view of FADER*

The following image shows the editor window with the 360-tool open. It shows the currently used asset and allows to change orientation, vertical position and reveals some meta data. The asset can be removed or replaced in this UI element.


*Figure 24: Editor asset detail view of FADER*

The Add Media button opens a modal that shows the media library on the left side, and the used media elements in the story on the right side. Media can simply be dragged and dropped from the own or public library into the story. Users can also upload their media within this screen. A media upload triggers the media processing toolchain described in the backend section. UI feedback is given on the processing steps, success, or failure.

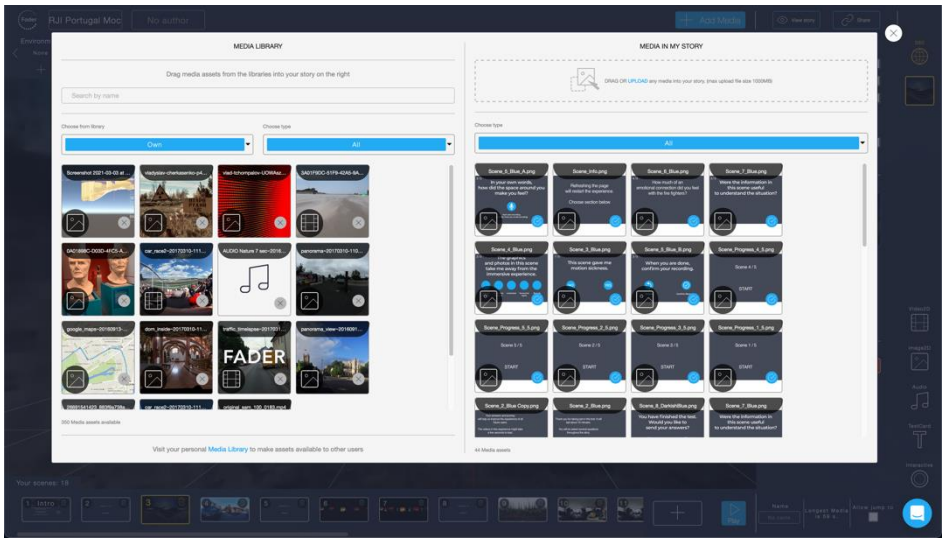


*Figure 25: Editor add media view of FADER*

### 3.2.4 Player

The player uses the same underlying rendering technology (A-frame/Three.js) to load and display the story. It has some general and some optional components. Before starting, it shows a hero image of the story, a load counter (or start button when finished), logo and title of the story, and sharing buttons.

The player has some controls, depending on whether in VR or magic window mode. There is a button to mute / unmute sounds and activate full screen mode. When scenes in a story are checked with the allow jump to-attribute, a scene bar including these scenes can be shown in the bottom to quickly access scenes directly.


*Figure 26: Player view of FADER while loading a story*

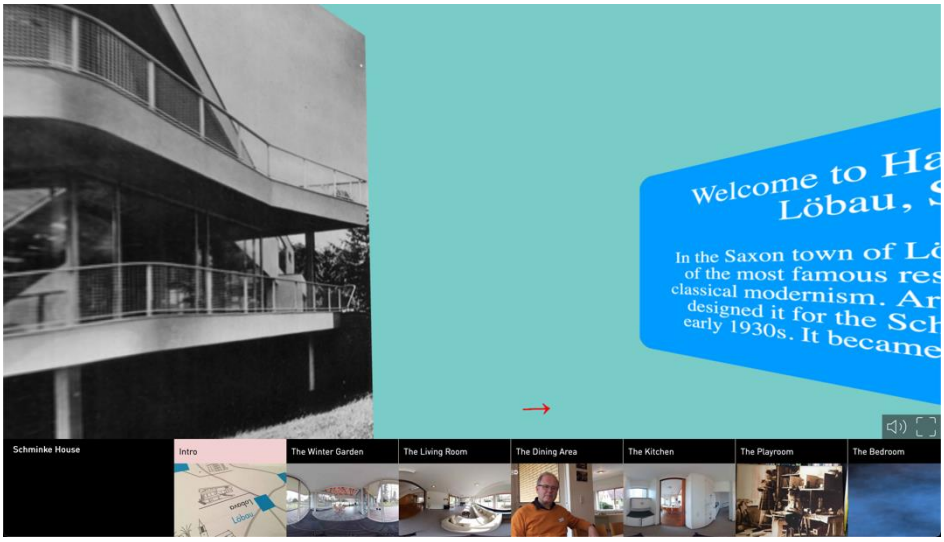


*Figure 27: Player view of FADER of a started story*

The player is also a part of the editor and as such allows WYSIWYG-style editing of stories.

### 3.2.5 Description of Requirements

The requirements for MediaVerse in general have been elicited in WP2. A detailed list is given in Table 13.

The requirements categories are user experience, collaboration, player technology, rights broker, authoring tools, eXtended Reality, rights management, and monitoring.

## 3.3 Implementation Plan

To support the requirements and use cases of MediaVerse, there are activities to take place in three main application components of the existing Fader platform. Additionally, a new mobile frontend application for on-site reporting and templated immersive storytelling will be implemented over the course of the project.

*Table 5: Developments to support the requirements across the technical components.*

| KEY / DESCRIPTION | COMPONENT(S) | | 2022 | | | | 2023 | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Q4 | Q1 | Q2 | Q3 | Q4 | Q1 | Q2 | Q3 |
| UX-02 / **Content Information will be easily visible.** | Fader Media Library | | ■ | | | | | | |
| COLL-02 / **The dashboard provides a concise overview of various types of information.** | Fader Projects | ■ | ■ | | | | | | |
| COLL-04 / **Project participants communicate via simple chat.** | MediaVerse Node | | | | ■ | ■ | | | |
| COLL-05 / **Project participants can add notes in the project.** | Fader Backend and Editor | | | | ■ | ■ | | | |
| PLAY-01 / **The player enables display/overlay of subtitles from .srt files.** | Fader Backend and Player | ■ | ■ | | | | | | |
| PLAY-02 / **Accessibility extras can be switched on and off.** | Fader Player | ■ | ■ | | | | | | |
| PLAY-04 / **MediaVerse content can be consumed/displayed in different devices.** | Fader Player | ■ | ■ | ■ | ■ | ■ | ■ | | |
| BRKR-04 / **Incorporating segments will have implications on the usage rights of the project.** | Fader Media Library and Editor / Mobile Application | | | | | ■ | ■ | | |
| AUTH-01 / **Mobile and desktop interfaces.** | Fader Editor / Mobile Application | ■ | ■ | ■ | ■ | ■ | ■ | | |
| AUTH-02 / **MediaVerse enables content creators to do basic editing on audio, video, 360 Video.** | Fader Editor / Mobile Application | ■ | ■ | ■ | ■ | ■ | ■ | | |
| AUTH-03 / **External authoring tools can be connected with MediaVerse for easy updates.** | Fader Editor / Mobile Application | ■ | ■ | ■ | ■ | ■ | ■ | | |
| AUTH-04 / **The authoring UIs need to support various languages.** | Fader Editor / Mobile Application | ■ | ■ | | | | | | |
| AUTH-05 / **Creators can add and edit various accessibility enhancements to content.** | Fader Editor / Mobile Application | ■ | ■ | | | | | | |
| XR-02 / **A 3D Scene can be exported as a video clip.** | Fader Editor / Mobile Application | | | | | ■ | ■ | | |
| XR-05 / **360-video can be enhanced by adding objects/assets.** | Fader Editor / Mobile Application and Player | ■ | ■ | ■ | ■ | ■ | ■ | | |
| XR-06 / **360-video can be enhanced by hotspots with different functions.** | Fader Editor / Mobile Application and Player | ■ | ■ | ■ | ■ | ■ | ■ | | |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| RGHT-03 / **The system must be able to trace content sources (inside MV).** | Fader Backend | | | | | | | | |
| MNTR-01 / **Monitor the consumption and re-use of content.** | Fader Backend | | | | | | | | |

### 3.3.1 Fader Backend

The Fader backend requires substantial changes in user and asset management to integrate with MediaVerse nodes. A key category will be access to and delivery of MediaVerse data to immersive storytellers.

To make use of SRT files, provided by STXT and their accessibility toolkit as mentioned in D5.1, the meta-data for video files as it currently exists needs to be extended to allow retrieving them from a MediaVerse node to be included in a Fader story.

An additional feature will be to create templates from either existing Fader stories or scenes. This is possible since the internal representation of objects in a scene does not strictly couple placement information and content. As such, one can create placeholders that can be stored as blueprints and later be assigned with other media items.

The requirements to satisfy in this component are:

UX-02, COLL-02, COLL-05, PLAY-01, BRKR-04, RGHT-03, MNTR-01

### 3.3.2 Fader Editor

The current desktop-based Fader editor will be integrated and usable for MediaVerse nodes as well. The main parts are usage of the improved 360-degree video transcoding, access to MV nodes media.

Also, the improved story and scene templates need to be made accessible in the editor – both web and mobile version. Users need to access the templates available and have a way to replace content without having to use the sandbox functionality of the current web editor.

The requirements to satisfy in this component are:

COLL-05, BRKR-04, AUTH-01, AUTH-02, AUTH-03, AUTH-04, AUTH-05, XR-02, XR-05, XR-06

### 3.3.3 Fader Player

A major part of the implementation for the Fader player will be the integration of subtitles and other features that improve the accessibility of a Fader story. Also, additional interactive elements need to be integrated and played out.

Taking into consideration the required changes in the Fader backend to use SRT files for subtitles, the display features in Fader need to be adapted to also play timed subtitles in a story. This should be the case for any video or audio piece, regardless of a 2D or 360-degree nature.

The requirements to satisfy in this component are:

PLAY-01, PLAY-02, PLAY-04, XR-05, XR-06

### 3.3.4 Mobile Editor Application

The main driver for an updated, easy to learn and use mobile application is the citizen journalism/ immersive journalism use case. While currently, editing stories in Fader is much more convenient on a desktop system, breaking news contexts with an emphasis on mobile journalism cannot be catered to, yet.

The requirements to satisfy in this component are:

BRKR-04, AUTH-01, AUTH-02, AUTH-03, AUTH-04, AUTH-05, XR-02, XR-05, XR-06

# 4 VR Multi-user Collaboration Sandbox (CERTH)

## 4.1 Context and Baseline Description

The VR multi-user collaboration sandbox is the outcome of MediaVerse T5.5 which aims in developing VR spaces that allow users to consume, comment, discuss, and create media with other users seeing their reactions and discussing with them as in real-life. Analysing the requirements of T5.5, we can divide them into two subtasks, namely: a) *Personal room authoring* where *t*his subtask allows the author to make his own room that allow to act as a personal space for social interaction, and b) *Embedding media in VR and enhancing them with 3D content, i*.e., enable the avatars (content makers) to do post-like interaction in VR such as inserting videos and images on their personal rooms with gestures. Avatars will be able to add 3D content on top of videos and make similar edits working collaboratively in 3D VR, rather than editing/viewing 360 VR solely, that offers the capability to increase the quality of the result. The media will adapt to user context such as Desktop VR or Mobile VR and play out within performance and bandwidth levels, as well as UX schemes, of the consumer platform.

We have isolated the outcome of this task to be aimed into Use Case 3 about cultural multi-presence spaces, without ignoring, however, the other use cases by allowing the 3D environment to be customized with an authoring tool. To raise interest towards multi-presence VR environments authoring we have gathered all of our developments into a WordPress plugin named as VRodos and setup a WordPress website with it (https://vrodos.iti.gr). This allows for a central service for VR spaces authoring. VRodos can be used for the generation of VR rooms using a template guided procedure avoiding thus the overstressing of the non-experienced users. Several of the authoring capabilities of VRodos are background knowledge of CERTH that was obtained from previous work (H2020 projects DigiArt, Envisage, and Helios, See MediaVerse proposal).

## 4.2 Implementation Roadmap

More specifically the actions done in the first year are depicted in Figure 28 and are as follows:

**A) Define target use:** Focus on cultural events and how can be done remotely with the collaboration of artists or actors in a virtual space, as well as have the presence of audience as avatars.

**B) Setup an access point – website:** Setup a site where users can easily find the service. Define side products to increase developments exploitation, e.g., 3D models repository, 3D models visualization service as iframes in webpages and others (See MediaVerse proposal and Deliverable D2.2 for more information about VRodos).

**C) Define functionalities of the template of cultural spaces:** Select technologies that implement these functionalities (e.g., multi-presence, video-stream, selfie-segmentation from background), make unit tests and integrate these technologies.

**D) Generalize and upscale withe VR authoring tool:** Give the opportunity for organizers to make their own setup using a web-based scene 3D editor and integrate the template of the functionalities. Then make review iterations and improve it.

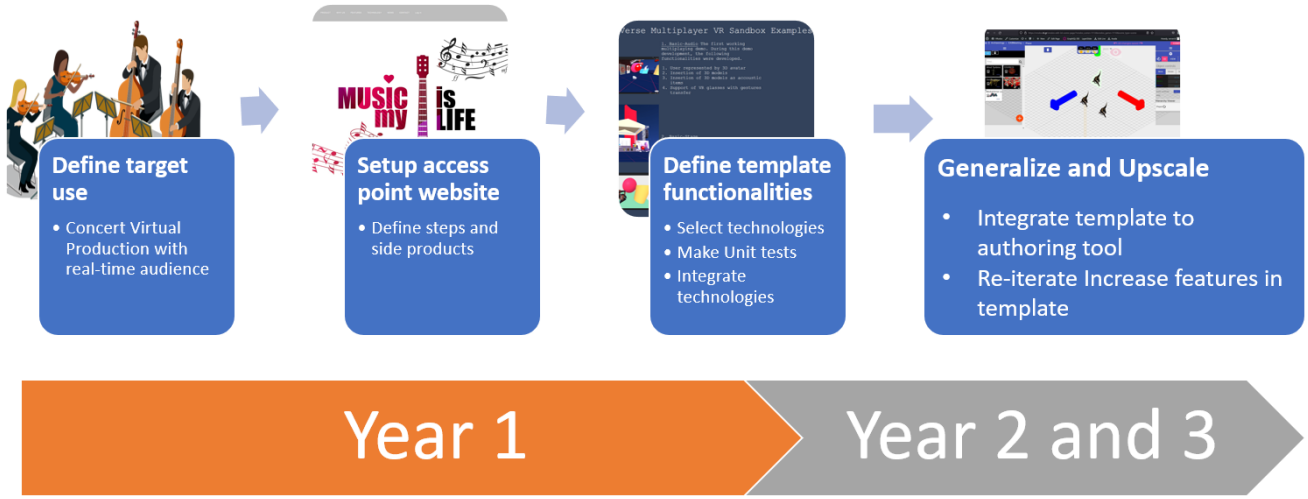


*Figure 28: Achievements for T5.5 in Year 1 and plans for Years 2 and 3*

## 4.3 Use Case Scenario

The use case scenario that VR social collaboration spaces will be applied is use case 3, which is about cultural spaces. Use case 3 will be demonstrated aside another project which is related to the installation of 5G network. In the city of Aveiro (Portugal), a big dissemination event will take place in October 2021 for the promotion of 5G technology. The purpose of the event is demonstrating the 5G network capabilities in a small-scale experiment. Only, a certain area of the town is covered with 5G under a local network (Figure 29). The event to take place is a music concert event where each member of the band is located at a different location. However, through 5G network they can play all together via electronic means and experienced simultaneously through an iPad tablet. Also, the people that are connected in this 5G network will be able to listen to the artists merged electronically as one band.

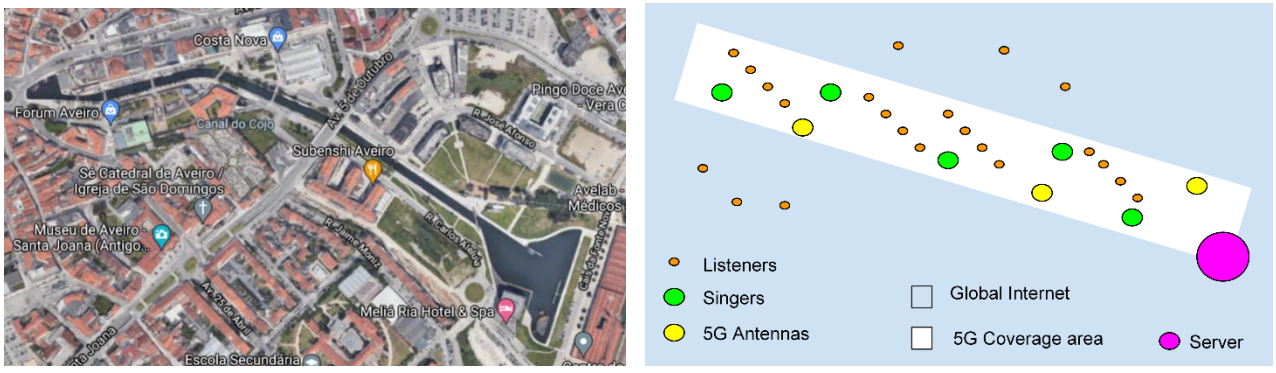


*Figure 29: Topology of singers and listeners in Use case 3 scenario. Left: geographic map. Right: Coverage map*

The 5G project consists of the use of an Augmented and Virtual Reality App, using 5G technology, to create and record musical content remotely. From a technological point of view, it explores the capacity of 5G to allow musical interaction in real-time since the musical content needs to be synchronized. Taking this into account, the Internet of Things plays a very important role in materializing innovative concepts in an artistic way. It allows

connection to the terrain and surrounding environments through sensing and actuation while 5G allows collaborations in real time. The goal is to promote the economic growth of Aveiro, with emphasis on the development of musical products with the use of emerging technologies thus stimulating digitization of the music industry, creating new content and artistic and creative experiments. Presently, the geographic area with 5G connectivity compromises the Aveiro's city centre, which includes Ria's Central Channel, between Aveiro Congress Centre and General Humberto Delgado Square. Use case partner Artshare is developing a Unity3D application for iPads that will allow for the singers/musicians to stream content and for the listeners to listen to it in the context of the 5G project. Our aim in MediaVerse is to hook on this scenario to also provide visual representation in a virtual production. Not only the listeners will be able to hear the singers, but the singers will be seen in a tablet, PC, smartphone, or VR glasses as they were together in the same scene, and the listeners will be represented as 3D avatars and discuss together as in a normal concert. The technologies to implement this scenario are web technologies that are outside the Unity application, and they can be used through a web browser.

## 4.4 Implementation Plan

The scenario describe in Section 4.3 is a virtual production scenario with remote control capabilities. The pathway to achieve this scenario is challenging as virtual production always involved the use of special equipment such as chroma key backgrounds, advanced graphic engines such as Unity or Unreal, and dedicated cable network connections. We aim now towards having a cheaper solution addressing low budget projects and low needs in equipment, with everyday low-cost devices. The quality of the result might not be for professional movies, but it will allow anyone to enter to the scene even with a mobile device. The development of the VR multi-user virtual production involves two steps, namely:

**STEP 1 Template development:** Development of the functionalities necessary to support multiuser and avatar representation either as 3D element or as video stream.

**STEP 2 Generalization:** Incorporation of the template into the VRodos authoring tool.

In the first year, we have focused mainly on the STEP 1, i.e., on the development of the template functionalities whereas minimum effort has been given in the insertion of the template within the existing authoring tool, where the latter is the plugin for WordPress. Apart from the integration of the template with the authoring tool, on years 2 and 3, we will also integrate the authoring tool with MediaVerse nodes that will allow the fetching of more professional 3D models and multimedia in general content for setting up the 3D scenes.

### 4.4.1 Proposed Architecture to Achieve Low-Cost Virtual Productions with Web Technologies

The overall concept of the proposed multiuser tools is presented in Figure 30. Starting from bottom-up, the users are provided with a URL link to enter the web application from whatever device they have using a web browser. Then their web camera stream or 3D avatar is represented in a virtual multiuser stage. If the user is represented by a camera stream, then his or her selfie is segmented from the background. Then the audience is provided with another link to enter the VR experience from any device. The audience can be represented in VR space either through a camera stream or a 3D avatar.

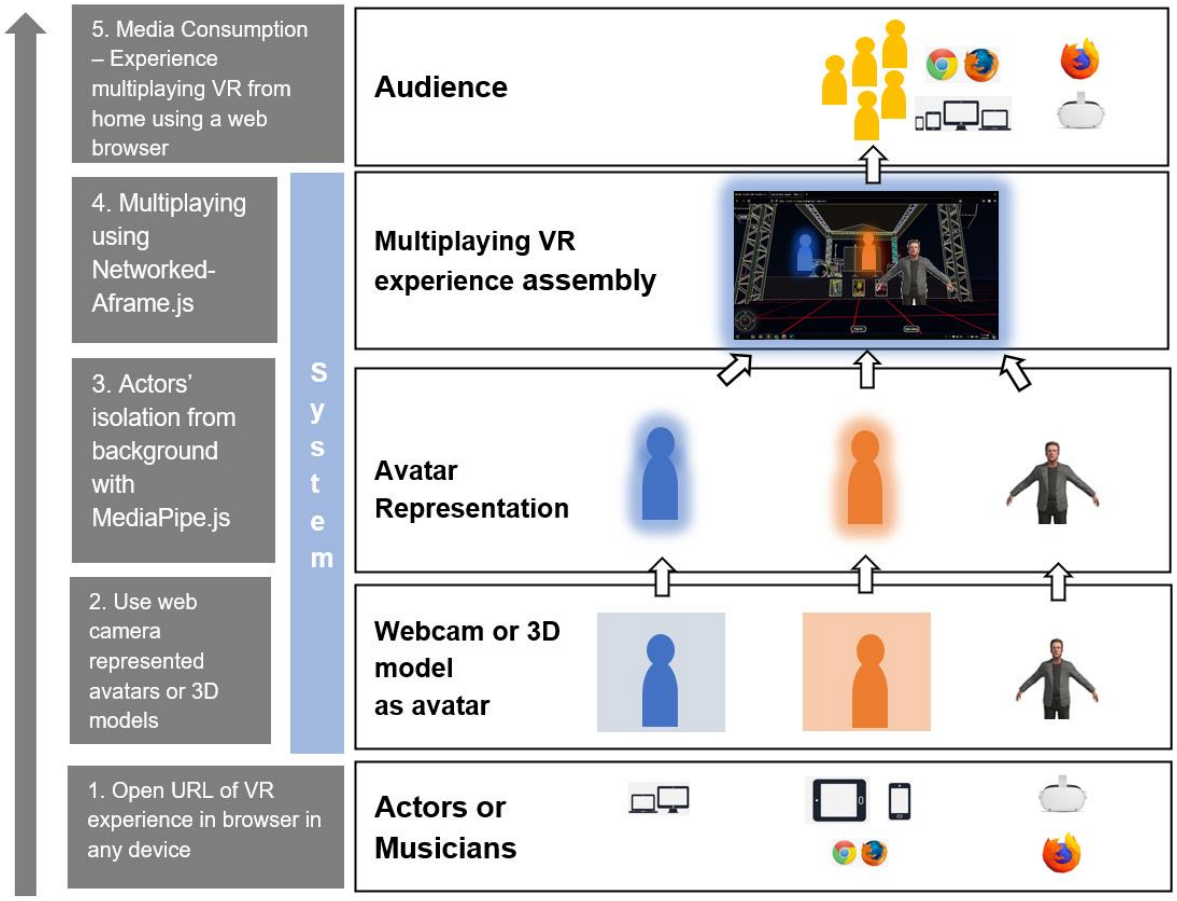


*Figure 30: Overall VRodos multiuser environment for covering Use case 3*

### 4.4.2 First Year Developments

As regards STEP 1, the template development with multiuser features, it is comprised of several sub-steps to achieve all the necessary functionalities and support many types of inputs, as follows.

**STEP 1.1 - Setup a 3D environment**

In this step, we have selected a framework that it is able to achieve all the necessary VR functionalities for desktops, laptops, mobile devices, or headsets. The most suitable framework for this purpose is A-frame[10] which is based on the popular Three.js library for 3D graphics on the web. A-frame is suitable for MediaVerse due to its web nature that allows for easily accessible 3D environments with one URL through web browser, without having to download binaries and install them. A-frame allows to build a 3D environment with just few lines of html code because it has components that are abstracting the low-level commands of Three.js. An example is shown in Figure 31 that consists of 15 lines of code that allow to make a VR environment with a box, a sphere, a cylinder, and a plane without having to write

[10] https://aframe.io/

JavaScript code. This compact of way of setting up scenes is also suitable for converting the code into a template for the authoring tool of STEP 2.

```
<!DOCTYPE html>
<html>
  <head>
    <script src="https://aframe.io/releases/1.2.0/aframe.min.js"></script>
  </head>
  <body>
    <a-scene>
      <a-box position="-1 0.5 -3" rotation="0 45 0" color="#4CC3D9"></a-box>
      <a-sphere position="0 1.25 -5" radius="1.25" color="#EF2D5E"></a-sphere>
      <a-cylinder position="1 0.75 -3" radius="0.5" height="1.5" color="#FFC65D"></a-cylinder>
      <a-plane position="0 0 -4" rotation="-90 0 0" width="4" height="4" color="#7BC8A4"></a-plane>
      <a-sky color="#ECECEC"></a-sky>
    </a-scene>
  </body>
</html>
```

*Figure 31: A-frame provides a compact way of storing and publishing VR applications*

Our developments involved the setting up of an environment as shown in Figure 32, which demonstrates a 3D of a music stage where all artists can virtually make a concert.

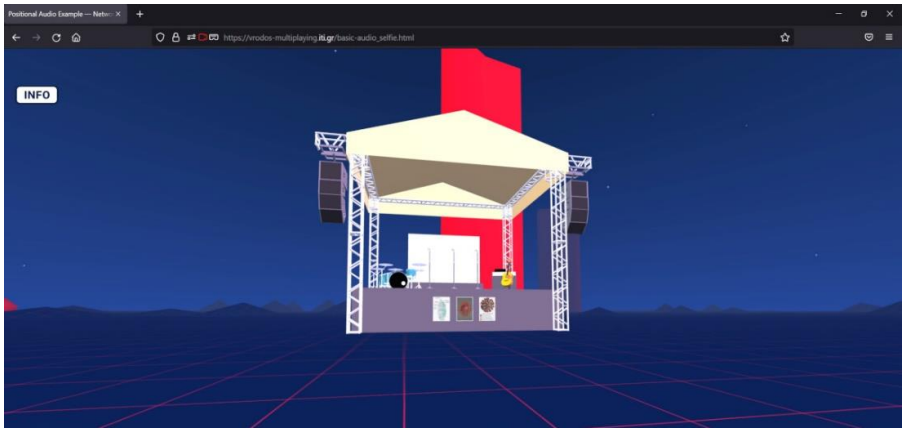


*Figure 32: VRodos multi-playing environment template with a 3D model of a stage*

**Step 1.2 - Development the interfaces for input and visualization for each device**

This step involves the developments towards the support of the most common devices as regards visualization and input. In Table 6, we summarize the support that the VR offers. The developments for PCs were done using Windows 10 tested for Chrome and Firefox in Desktops and Laptops. The tests for Tablets and Smartphones were done in Android 9 devices (Samsung and Xiaomi). The tests for VR Glasses were done using Oculus Quest 2 device.

*Table 6: Developments to support all types of devices.*

| Device | Requirements | | Interfaces | | |
|---|---|---|---|---|---|
| | Web Browser Support | OS & Devices | Input | Avatar as video Stream | Avatar features as 3D model |
| **Desktop** | Firefox, Chrome | PC or Mac | Keyboard / Mouse | Web Camera | Walk |
| **Laptop** | Firefox, Chrome | PC or Mac | Keyboard / Mouse / Touchscreen | Web Camera | Walk |
| **Tablet** | Firefox, Chrome mobile | Android or iOS | Touchscreen | Face Camera | Walk |
| **Smartphone** | Firefox, Chrome mobile | Android or iOS | Touchscreen | Face Camera | Walk |

For the input interfaces, several developments were done for the template page to run smoothly on all the devices. As regards touchscreen interfaces for mobile devices and some laptops, we have implemented two types of touch inputs, one by touching on the screen to click on objects or swipe on screen to rotate the player, and another one as virtual joystick on the screen for manipulating the forwards-backwards functionality and the strafe functionality (Figure 33). Another development was for the VR glasses, namely, to allow virtual walking, i.e., to walk with thumb-stick movement in the case that the player does not have enough space to move, which is almost valid for all immersions. In some cases, we have provided alternative methods to give the same type of input. The interface input methodologies are described in Table 7: Basic input methods.

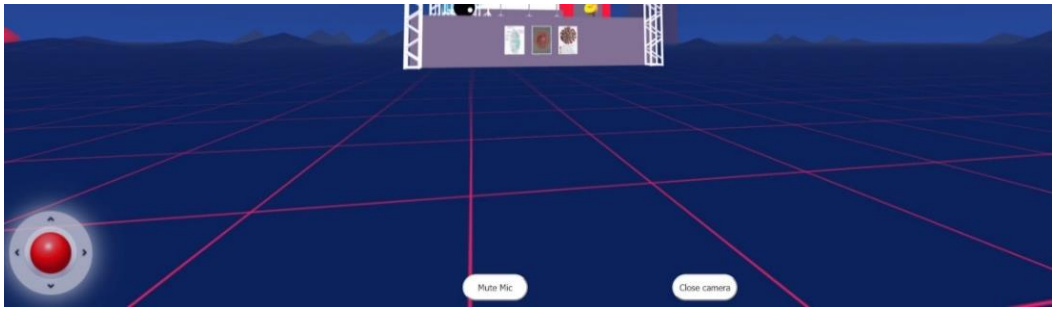


*Figure 33: Bottom Left, a virtual joystick can be used in touch devices for handling forwards-backwards movement*

*Table 7: Basic input methods.*

| Device | Basic VR Inputs | | | |
|---|---|---|---|---|
| | Forwards - Backwards | Rotate Player | Strafe | Clicking |
| **Desktop** | Method 1: W, S buttons<br>Method 2: Move the virtual joystick up / down | Drag Mouse | Method 1: A, D buttons<br>Method 2: Move the virtual joystick left / right | Left mouse button click |
| **Laptop** | Method 1: W, S buttons<br>Method 2: Move the virtual joystick | Drag Mouse or Touchpad | Method 1: A, D buttons<br>Method 2: Move the virtual joystick left / right | Left mouse / touchpad button click |
| **Tablet** | Move the virtual joystick up / down | Swipe on screen | Move the virtual joystick left / right | Touch on screen |
| **Smartphone** | Move the virtual joystick up / down | Swipe on screen | Move the virtual joystick left / right | Touch on screen |
| **VR Glasses** | Method 1: Walk<br>Method 2: Press left thumbstick towards up / down | Rotate head | Method 1: Strafe<br>Method 2: Press left thumbstick towards left / right | Use hand-controller as laser and click Button A on controller. |

**Step 1.3 - Multi-playing functionality**

Towards the development of multi-playing functionalities, we have exploited the Networked-Aframe library that builds on top of A-frame to offer the capabilities of many avatars into one environment (Networked-Aframe). Networked-Aframe exploits EasyRTC implementation of WebRTC protocol to achieve peer-2-peer connections. In these capabilities, the video and sound stream are also included. It is also network sensitive meaning that does not reserve bandwidth when nothing changes in the 3D experience. Networked-Aframe follows also the html

syntax as A-frame so it is easy to make a simple multi-playing experience without writing JavaScript code. An example is shown below. It generates an environment where each avatar is represented by a sphere.

```
<html>
  <head>
    <title>My Networked-Aframe Scene</title>
    <script src="https://aframe.io/releases/1.2.0/aframe.min.js"></script>
    <script src="https://cdnjs.cloudflare.com/ajax/libs/socket.io/2.3.0/socket.io.slim.js"></script>
    <script src="/easyrtc/easyrtc.js"></script>
    <script src="https://unpkg.com/networked-aframe/dist/networked-aframe.min.js"></script>
  </head>
  <body>
    <a-scene networked-scene>
      <a-assets>
        <template id="avatar-template">
           <a-sphere></a-sphere>
        </template>
      </a-assets>
      <a-entity id="player" networked="template:#avatar-template;attachTemplateToLocal:false;" camera wasd-controls look-controls>
      </a-entity>
    </a-scene>
  </body>
</html>
```

Several developments were done to support in all devices the transfer and replication for each player input data such as the selected name and colour for each avatar, their rotation, movement, and audio communication. Additionally, the video stream can be also transferred together with the network streams as shown in Figure 34. The challenge in the multi-playing environment was to also transfer the movements of the hand gestures when using VR glasses. This was achieved by creating networking entities and assigning hand 3D models for the entities. Position and rotation are supported for each hand.

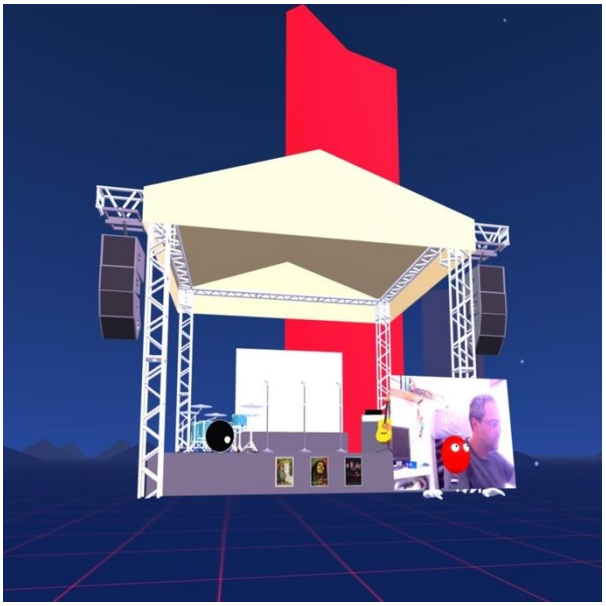

*Figure 34: Development of a multiplaying VR experience with video-streaming capability*

Another development was the incorporation of MediaPipe within A-frame for separating the user of a video stream from the background and inserting it as a streaming video texture over a 3D model. An example is shown in Figure 35 where the video stream is segmented from the background and inserted as a texture over all faces of a cube.

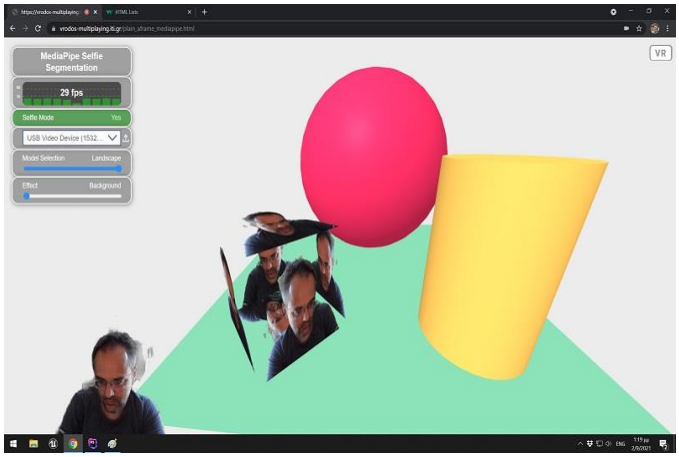


*Figure 35: A-frame and MediaPipe combined in order to isolate user from background*

Next steps involve the combination of Networked-Aframe with MediaPipe to allow multiple segmented streams within a VR space, and thus allow the template to support the necessary virtual production features for the use case.

In year 2 and 3, other activities are the interconnection with MediaVerse platform so that the 3D models and images can be retrieved from a content rich source and the embodiment of the template into the authoring tool to allow non-experts in programming to change the setup of the 3D environment.

# References

A-frame, VR framework for Web, URL: https://aframe.io/

Babylon.js, 3D web rendering library with WebXR capabilities, URL: https://www.babylonjs.com/

Berman, H.M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T.N.; Weissig, H.; Shindyalov, I.N.; Bourne, P.E. The Protein Data Bank. *Nucleic Acids Res.* **2000**, *28*, 235–242, doi:10.1093/nar/28.1.235.

Cao, Z.; Hidalgo Martinez, G.; Simon, T.; Wei, S.-E.; Sheikh, Y.A. OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields. *IEEE Trans. Pattern Anal. Mach. Intell.* **2019**, *43*, 172–186, doi:10.1109/tpami.2019.2929257

Costello, K.; Rimol, M. Gartner Identifies the Top 10 Strategic Technology Trends for 2020. In Gartner IT Symposium/Xpo 2019. Available online: https://www.gartner.com/en/newsroom/press-releases/2020-10-19-gartner-identifies-the-top-strategic-technology-trends-for-2021 (accessed on 20/7/2021).

EC784/20: European Commission. *Communication from the Commission to the European Parliament, Europe's Media in the Digital Decade: An Action Plan to Support Recovery and Transformation*; COM/2020/784 final; European Commission: Brussels, Belgium, 3/12/2020.

Fuchs, N.E.; Schwertel, U.; Schwitter, R. Attempto Controlled English—Not just another logic specification language. In *International Workshop on Logic Programming Synthesis and Transformation, Lecture Notes in Computer Science*; Springer: Berlin/Heidelberg, Germany, 1999; pp. 1–20, doi:10.1007/3-540-48958-4_1, ISBN 978-3-540-65765-1, 1998.

Google Poly 2021: Google Poly. Software. Available online: https://en.wikipedia.org/wiki/Poly_(website) (accessed on 20/7/2021).

Herx 2019: Herz, M.; Rauschnabel, P.A. Understanding the diffusion of virtual reality glasses: The role of media, fashion and technology. *Technol. Forecast. Soc. Chang.* **2019**, *138*, 228–242.

IEEE Standard 830-1998. *Recommended Practice for Software Requirements Specifications*; IEEE Press: New York, NY, USA, 1998.

Jamulus, Networking audio streaming services, URL: https://jamulus.io/

Marcelino-Jesus, E.; Sarraipa, J.; Agostinho, C.; Jardim-Goncalves, R. A Requirements Engineering Methodology for Technological Innovations Assessment. In *Moving Integrated Product Development to Service Clouds in the Global Economy*; IOS Press: Amsterdam, The Netherlands, 2014; pp. 577–586.

MediaPipe, Media Technologies AI framework, URL: https://google.github.io/mediapipe/

Mouzakis, C.; Ververidis, D.; Girao, L.M.; Patz, N.; Nikolopoulos, S.; Kompatsiaris, I. "Holistic Requirements Analysis For Specifying New Systems For 3D Media Production and Promotion." MDPI Sustainability, Special Issue "Cultural Heritage Storytelling, Engagement and Management in the Era of Big Data and the Semantic Web", Accepted for publication, July, 2021.

Mütterlein, J.; Kunz, R.E.; Baier, D.; Mütterlein, J.; Kunz, R.E.; Baier, D. Effects of lead-usership on the acceptance of media innovations: A mobile augmented reality case. *Technol. Forecast. Soc. Chang.* **2019**, *145*, 113–124, doi:10.1016/j.techfore.2019.04.019.

Manufacturing SErvices Ecosystem (MSEE) Project, Project ID 284860, Deliverable "Methodology for requirement engineering and evaluation" 51.1, M6, 30/04/2012, Available online: https://cordis.europa.eu/project/id/284860/reporting (accessed on 20/7/2021).

Nebeling, M.; Speicher, M. The trouble with augmented reality/virtual reality authoring tools. In Proceedings of the 2018 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), Munich, Germany, 16–20 October 2018.

Networked-Aframe, Multiplaying library for A-frame, URL: https://www.npmjs.com/package/networked-aframe

HTML5 MSE Playback of MPEG 360 VR Tiled Streaming, URL: https://github.com/fraunhoferhhi/omaf.js/tree/master/omaf-file-creation (accessed 2021-09-13)

Statista, Music Industry in Europe, 2019, URL: https://www.statista.com/topics/3903/music-industry-in-europe/

Ververidis, D.; Migkotzidis, P.; Nikolaidis, E.; Anastasovitis, E.; Papazoglou-Chalikias, A.; Nikolopoulos, S.; and Kompatsiaris, I. "An authoring tool for democratizing the creation of high-quality VR experiences," Springer Virtual Reality, June, 2021, DOI: https://doi.org/10.1007/s10055-021-00541-2

WebGPU. W3C Working Draft. 12 July 2021. Available online: https://www.w3.org/TR/webgpu/ (accessed on 20/7/2021).

W3C, World Wide Web Consortium. WebVR Standard Specification Legacy site. Available online: https://webvr.info/ (accessed on 20/7/2021).

W3C, World Wide Web Consortium. WebXR Standard Specification, WebXR. Available online: https://www.w3.org/TR/2020/WD-webxr-20200724/ (accessed on 20/7/2021).

XRAssociation20: XR Association. Augmented and Virtual Reality Survey Report—Industry Insights into the Future of Immersive Technology. 2020. Available online: https://www.perkinscoie.com/images/content/2/3/v4/231654/2020-AR-VR-Survey-v3.pdf (accessed on 20/7/2021).

# Annex I: Fader Data Model

*Table 8: User details*

| NAME | TYPE | DESCRIPTION |
|---|---|---|
| id | UUID | Unique identifier |
| firstname | String | First name of a user |
| surname | String | Family name of a user |
| display_name | String | Self-chosen name |
| username | String | Unique username |
| email | Email | Unique email |
| last_logged_in_at | Date | When user was last logged in |
| updated_at | Date | When user was last updated |
| inserted_at | Date | When user was created |

*Table 9: Account plan*

| NAME | TYPE | DESCRIPTION |
|---|---|---|
| show_intercom | Boolean | Show or hide the intercom chat for customer support. |
| nr_public_projects | Integer | # of public projects a user can create (-1 infinite) |
| nr_private_projects | Integer | # of private projects a user can create (-1 infinite) |
| custom_players_settings | Boolean | Setting, if a user can change visual elements of the player. |
| custom_logo | Boolean | Setting, if a user can change the logo of the player. |
| custom_icon_video | Boolean | Setting, if a user can change the scene video icon of the player. |
| custom_icon_image | Boolean | Setting, if a user can change the scene image icon of the player. |
| custom_icon_audio | Boolean | Setting, if a user can change the scene audio icon of the player. |
| custom_font | Boolean | Setting, if a user can change the font of the player. |
| custom_colorscheme | Boolean | Setting, if a user can change the color scheme of the player. |

*Table 10: User meta data*

| GROUP | NAME | TYPE | DESCRIPTION |
|---|---|---|---|
| features | See table /ref Table #: Account plan | | User settings, that allow to define specific user privileges based on an account plan and access rights. |
| custom_player_settings | | | Custom settings for player |
| | unmute | Boolean | Prefer to start stories with audio on (suppressed on mobile) |
| | show_author | Boolean | Allow setting a story author and show in story credits (for team accounts) |
| custom_logo | | | Replace the Fader logo |
| | url | String | URL for the custom logo |
| | path | String | Local path to the file |
| | name | String | Original file name |
| custom_icon_video | | | Replace the default video icon |
| | url | String | URL for the custom logo |
| | path | String | Local path to the file |

| | | | |
|---|---|---|---|
| | name | String | Original file name |
| custom_icon_image | | | Replace the default video image |
| | url | String | URL for the custom logo |
| | path | String | Local path to the file |
| | name | String | Original file name |
| custom_icon_audio | | | Replace the default video audio |
| | url | String | URL for the custom logo |
| | path | String | Local path to the file |
| | name | String | Original file name |
| custom_font | | | Replace the default font |
| | url | String | URL for the custom logo |
| | path | String | Local path to the file |
| | name | String | Original file name |
| custom_colorscheme | | | Replace the default video icon |
| | secondary | Hex | Custom secondary color |
| | primary | Hex | Custom primary color |
| | font | Hex | Custom font color |

*Table 11: Asset lease*

| NAME | TYPE | DESCRIPTION |
|---|---|---|
| id | UUID | Unique identifier |
| license | Enum [owner, public] | The license type for this asset lease. Public assets can be used by other users. |
| valid_since | Date | Date, when the asset lease became active |
| valid_until | Date | Date, when the asset lease was terminated |
| asset | URI | The actual asset that the asset lease is associated to |
| inserted_at | Date | When asset was created |
| updated_at | Date | When asset was last updated |

*Table 12: Project data*

| GROUP | NAME | TYPE | DESCRIPTION |
|---|---|---|---|
| Version | version | String | The project data structure version. Currently 1.4 is the latest. |
| UploadedAssetIds | uploadedAssetIds | List<String> | IDs of all assets that belong to the project |
| | unmute | Boolean | Prefer to start stories with audio on (suppressed on mobile) |
| | show_author | Boolean | Allow setting a story author and show in story credits (for team accounts) |
| custom_logo | | | Replace the Fader logo |
| | url | String | URL for the custom logo |
| | path | String | Local path to the file |
| | name | String | Original file name |
| custom_icon_video | | | Replace the default video icon |
| | url | String | URL for the custom logo |

| | | | |
|---|---|---|---|
| | path | String | Local path to the file |
| | name | String | Original file name |
| custom_icon_image | | | Replace the default video image |
| | url | String | URL for the custom logo |
| | path | String | Local path to the file |
| | name | String | Original file name |
| custom_icon_audio | | | Replace the default video audio |
| | url | String | URL for the custom logo |
| | path | String | Local path to the file |
| | name | String | Original file name |
| custom_font | | | Replace the default font |
| | url | String | URL for the custom logo |
| | path | String | Local path to the file |
| | name | String | Original file name |
| custom_colorscheme | | | Replace the default video icon |
| | secondary | Hex | Custom secondary color |
| | primary | Hex | Custom primary color |
| | font | Hex | Custom font color |

# Annex II: Immersive Storytelling Tools Requirements

*Table 13: Requirements for the Immersive Storytelling Toolkit*

| KEY | NAME | COMPONENT(S) | DESCRIPTION, RATIONALE AND IMPORTANCE |
|---|---|---|---|
| UX-02 | Display of content information | Fader Media Library | **Content Information will be easily visible.**<br><br>Once an item was selected, the interface will show some basic metadata, available extras, and copyright information<br><br>**Essential** |
| COLL-02 | Dashboard (Project Information) | Fader Projects | **The dashboard provides a concise overview of various types of information.**<br><br>A project owner needs an interface where s/he can edit and view all sorts of Information.<br><br>**Essential** |
| COLL-04 | Team communication | MediaVerse Node | **Project participants communicate via simple chat.**<br><br>Collaboration requires communication.<br><br>**Nice to have** |
| COLL-05 | Editing notes | Fader Backend and Editor | **Project participants can add notes in the project.**<br><br>Another means of communication, but closely related to the content the team is working on: Project participants can attach notes to segments of the project (e.g., scene in a video) as comments, mentions, or assignments.<br><br>**Important** |
| PLAY-01 | Subtitle format | Fader Backend and Player | **The player enables display/overlay of subtitles from .srt files.**<br><br>Inclusion: People in need of such extras (CC, SL, AD) should consume the same content as those who are not. |

| | | | |
|---|---|---|---|
| | | | **Essential** |
| PLAY-02 | Accessibility toggle | Fader Player | **Accessibility extras can be switched on and off.**<br><br>Inclusion: People in need of such extras (CC, SL, AD) should consume the same content as those who are not.<br><br>**Essential** |
| PLAY-04 | Output devices | Fader Player | **MediaVerse content can be consumed/displayed in different devices.**<br><br>Depending on the content, 1-4 device types will be able to show or use MediaVerse content: desktop (web interface), mobile (tablet/ smartphone), TV, VR.<br><br>**Essential** |
| BRKR-04 | Segment settings are inherited | Fader Media Library and Editor / Mobile Application | **Incorporating segments will have implications on the usage rights of the project.**<br><br>Whoever wants to use a fragment, must accept the rules, incl. its copyrights, moral rights, local restrictions, age restrictions, etc.<br><br>**Essential** |
| AUTH-01 | Devices | Fader Editor / Mobile Application | **Mobile and desktop interfaces.**<br><br>There are alternative interfaces for the search, retrieval, creation and editing of MediaVerse content:<br>1. web browser (All)<br>2. mobile interface (Editing, management, visualisation)<br>3. HMD (Editing, visualisation)<br>4. Connected TVs (visualisation)<br><br>**Essential** |
| AUTH-02 | A/V Authoring | Fader Editor / Mobile Application | **MediaVerse enables content creators to do basic editing on audio, video, 360 Video.**<br><br>These are the basics of “content co-creation”. |

| | | | |
|---|---|---|---|
| | | | **Essential** |
| AUTH-03 | External authoring | Fader Editor / Mobile Application | **External authoring tools can be connected with MediaVerse for easy updates.**<br><br>Content creators can use MV tools or external tools to edit content. If it is external, then it would be good to automatically synch MV folders/content with the content edited with the external tool. (plugin-style?).<br><br>**Essential** |
| AUTH-04 | Supported languages | Fader Editor / Mobile Application | **The authoring UIs need to support various languages.**<br><br>The tools will be used for demonstrators in (at least) 4 countries with these languages.<br><br>**Essential** |
| AUTH-05 | Authoring of accessibility enhancements | Fader Editor / Mobile Application | **Creators can add and edit various accessibility enhancements to content.**<br><br>These enhancements could be<br>● Add and edit subtitles (.srt) in different languages<br>● Add sign language videos (Overlay/hotspot)<br>● Add Audio Description (separate audio track)<br>● Add and edit translations (multi-language)<br><br>**Essential** |
| XR-02 | 3D model as video clip | Fader Editor / Mobile Application | **A 3D Scene can be exported as a video clip.**<br><br>3D experiences require special players to watch. Sharing them via other Social Media platforms will only be possible, if they can be exported as video files<br><br>**Important** |
| XR-05 | Enhancing 360 videos | Fader Editor / Mobile Application and Player | **360-video can be enhanced by adding objects/assets.**<br><br>A 360 scene can be adapted to the need of the situation, e.g., by adding 2D |

| | | | |
|---|---|---|---|
| | | | videos or images, 3D diagrams, text panels, etc., or hotspots that will open such related/attached media<br><br>**Essential** |
| XR-06 | Hotspots in 360° videos | Fader Editor / Mobile Application and Player | **360-video can be enhanced by hotspots with different functions.**<br><br>A hotspot is a button which triggers certain actions in the interactive scene.<br><br>**Essential** |
| RGHT-03 | Origins traceable | Fader Backend | **The system must be able to trace content sources (inside MV).**<br><br>In case a segment is being re-used, the new asset must carry information of the original source. This information should be based on small units (shots, scenes, objects) so that in case of a re-use of the new asset's segments, tracing these origins will still be possible.<br><br>**Essential** |
| MNTR-01 | Monitoring use of content | Fader Backend | **Monitor the consumption and re-use of content.**<br><br>Monitoring the use of the content is important and should provide information about where the content has been used in order to determine the reach, or engagement.<br><br>**Essential** |

MediaVerse is an H2020 Innovation Project co-financed by the EC under Grant Agreement ID: 957252.