TRANSMIXR

Ignite the Immersive Media Sector by Enabling New Narrative Visions

D4.1 Specifications of Alpha and Beta Versions of the Environment

LEAD AUTHOR: IMMERSION

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

Table of Contents

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

Table of Figures

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

Table of Tables

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

Document Versioning

VERSION	DATE	WRITERS	CONTENT
V0	*06/03/23*	*Juliette Vauchez (IMM), Thomas Piet (IMM)*	*Social XR state of the art*
V1	*29/03/23*	*Juliette Vauchez (IMM)*	*Integration of first specifications in regard of February workshop with professionals*
V2	*19/12/23*	*Juliette Vauchez (IMM), Joseph Fadel (IMM)*	*Second row of specifications after the December workshop with professionals*
V3	*12/01/24*	*Croci Simone (HSLU), Marie Hospital (AFP), Tupac Martir (Satore), Rasa Bočytė (NISV), Trevor O Clochartaigh (TG4), Juliette Vauchez (IMM)*	*Including the three other use cases. Integration of Croci Simone's feedback.*
V4	*19/02/24*	*Niall Murray (TUS), Pablo Cesar (CWI), Julien Castet (IMM), Juliette Vauchez (IMM)*	*Including Niall Murray's feedback and uniting use cases' vision after coordination workshop.* *Reduced Social XR state of the art.*
V5	*27/02/24*	*Conor Keighrey (TUS), Arno Verhofstadt (VUB), Croci Simone (HSLU), Adrian Brasoveanu (MODUL University), Juliette Vauchez (IMM)*	*Including workshops statistics from Arno Verhofstadt, and reviews from Croci Simone and Adrian Brasoveanu.*
V6	*11/03/24*	*Niall Murray (TUS), Conor Keighrey (TUS)*	*Formatting, Edits and updated figures.*

*In this document, singular 'they' is used instead of 'he' or 'she' to designate a user without referring to gender, to offer a more inclusive vision.

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

1 Executive Summary

The TRANSMIXR project aims to offer rich and meaningful immersive experiences to professionals and consumers in the domains of news media and broadcasting, performing art and cultural heritage. The unique combination of Artificial Intelligence (AI) and eXtended Reality (XR) core technologies will advance the state-of-the-art in media production, delivery, and consumption. The project aims to create a range of human-centric tools for remote content production and consumption.

The technical work packages focused on understanding (WP2), creation (WP3), and experience (WP4) are delivering the technical components to realize different four use case pilots. Each of the use case pilots requires a range of the various technical components that constitute their own pipelines. The use case pilots have their own specific needs, informed by significant ongoing efforts during the requirements and co-design efforts of WP1 and use case pilot work of WP5.

The use case pilots are the following:

- News media lead by AFP with support of SparkNews and RTVSLO
- News broadcast lead by TG4
- Cultural heritage lead by Sound & Vision
- Performing arts lead by Satore Studio

In this context, the purpose of this deliverable, D4.1, is to provide a starting point for further exploration and refinement in terms of the user experiences to be created within the TRANSMIXR project. It presents the modalities and scenarios for each of the four use case pilots to set the XR environment specifications. As mentioned, these specifications are building on the work already completed in T1.2 about user requirements, and professional and end user workshops T1.3 that were carried out. It also builds upon the ongoing work in T5.1. Here, the focus is on users' experience.

As all the TRANSMIXR experiences have a social dimension, this document also includes the current state of the art of "Social XR" and identifies areas of concern as an appendix. This was the initial version of this deliverable. Overall, the identified state of the art solutions do not fit to media needs, thus encouraging us to explore our own solution within TRANSMIXR project.

However, it is important to keep in mind that this does not strictly describe the final product: changes and revisions will be made according to the iterative feedback and insights gained from the next developments, tests, and workshops as per our human centric methodology. Overall, this document serves as a foundation for further

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

collaboration and refinement. It lays the groundwork for future discussions and iterations, also as part of piloting realized in WP5.

2 Foundations of TRANSMIXR Use Case Pilots

2.1 Experiences' Specificities

Initially, each use case pilot had pre-identified technical components they wanted to integrate. We then led a coordination action to ensure the proposed use case pilots were innovative, complementary, and exploring an entire scope of specificities. To do so, we set the following characteristics:

- **Broadcast:** Live with real-time content or replay with offline content.
- **Network:** Remote or local experiences. Remote might be web technology, where experiences are consumed via the public internet, whereas local refers to experiences within a local private network.
- **Platform:** TRANSMIXR is working towards making experiences all Immersive and can be deployed on multiple platform (2D screen, Head-Mounted Display (HMD) etc.). At this stage, we do not consider Virtual Reality (VR) nor Augmented Reality (AR) separately, but global XR experiences that are immersive.
- **Interactive:** All TRANSMIXR experiences proposed to be naturalistically intuitive and interactive.
- **Number of users:** Social or individual.

There are also further considerations, which are ongoing within the various use case pilots, in terms of degree of virtuality i.e. whether experiences included a physical object or were fully virtual as well as timeline i.e. whether there is a degree of persistence or non-persistence in the experiences.

The chosen game engines are Unity for News Media, News Broadcasting and Cultural Heritage, and Unreal for Performing Arts.

All the use case pilots include a social aspect with multiple users who might be able to interact. In addition, we will strive to support a range of consumption platforms per use case pilot to ensure accessibility and wide reach of the TRANSMIXR experiences. However, in the demonstrators we will mainly explore HMD based immersive platforms which are the more complex to implement.

The decision about what technologies to use was informed by an iterative technology demonstrations in Lisbon (Feb 2023) as an initial base, followed by 11 technology

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

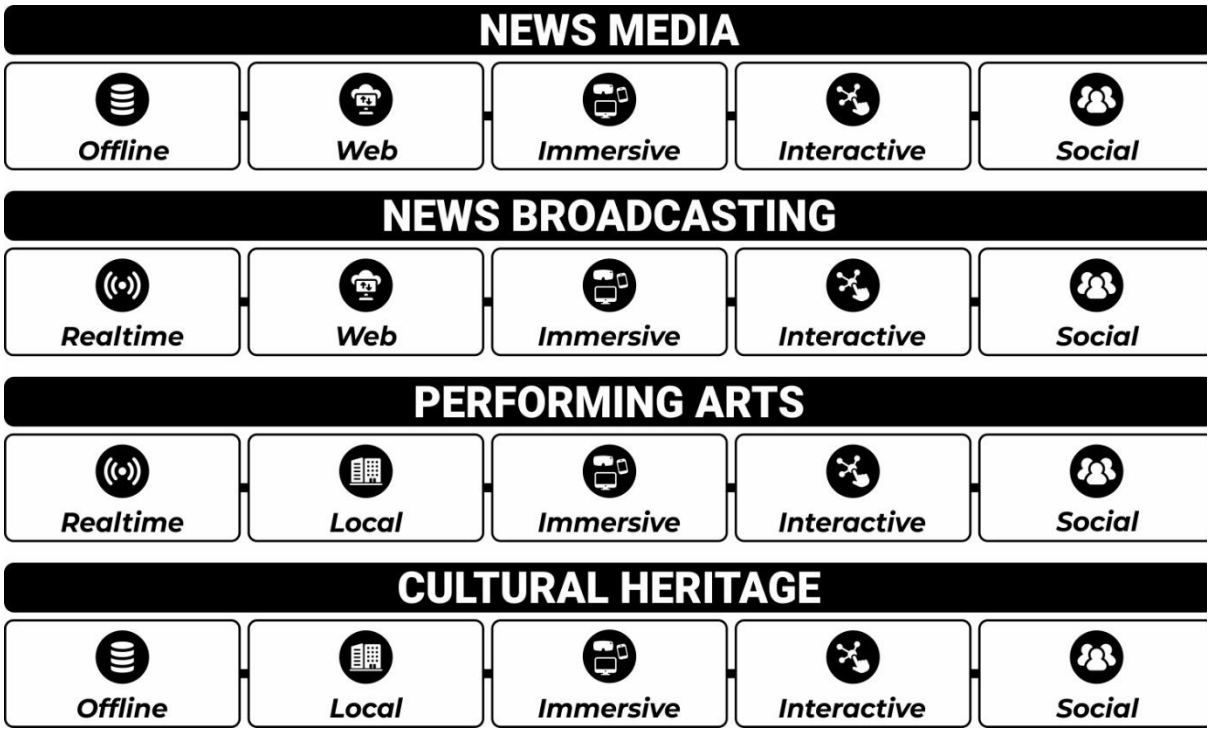


Figure 1: Use case pilot specificities

demonstrations at the General Assembly which was held on June 2023 at Amsterdam CWI headquarters.

We also distinguish two types of users within each use case pilot: professional users and end-users. Both will benefit the TRANSMIXR experience which are described in the following sections for each use case pilot. Nonetheless, they all should all ensure that general recommendations about usability are applied.

2.2 Common Specifications

Each use case pilot shall have an accessible and intuitive interface. Many Human-Computer Interfaces (HCI) usability guidelines already exist that might guide the User Interface (UI) choices and developments. For example, Bastien & Scapin (1993) provided ergonomic criteria for the evaluation of HCI designs. They should also be applied to XR applications design. In addition, several authors investigated User eXperience (UX) guidelines for designing XR app. For example, Vi et al. (2019) in their HMD design guidelines. Table 1 provides a high-level overview of some of these basic guidelines.

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

Table 1: Basic usability guidelines (Bastien & Scapin 1993, Vi et al. 2019).

Category	Specification
Guidance	Guide the beginners and facilitate interaction with layout thanks to prompting, grouping, feedback, legibility. Use cues.
Workload	Reduce perceptive or memory load thanks to brevity and relevant information density.
Explicit control	Allow the user to control the workflow thanks to explicit actions and cancel / return / continue actions.
Adaptability	Make the layout adaptable to every type of users thanks to interface personalization and differentiation of user experience.
Error management	Allow for trial and error. Prevent user from errors (avoid and reduce) and allow for errors to be fixed thanks to qualitative error messages.
Consistency	Global consistency (standardization) to avoid wasting time.
Code significance and denominations	Consider users' reality and existing environment (visual and semantic understanding).
Compatibility	Can be integrated into users' real life, according to their expectations and habits.
Spatial environment	Organize the spatial environment to maximize efficiency and to be flexible.
Comfort	Prioritize user's comfort and keep it simple, do not overwhelm the user.
Hardware	Design around hardware capabilities and limitations.
Multimodal	Create a compelling XR experience (senses).

Vairo et al. 2023 nonetheless highlights some relevant limitations of the integration of physical and virtual worlds in a convincing way. Within their "SUN project", they propose solutions mostly AI-based allowing:

- scalable and cost-effective XR app development,
- convincing mixing of physical and virtual world integrating not only visual alignment but also corresponding actions and effects,
- plausible human interaction with natural and realistic counterparts in the virtual world,
- diminished XR devices barriers dealing with computing, memory and networking resources.

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

2.3 Requirements Gathering Process of WP1

In order to establish the following scenarios and specifications, a range of user requirements, production workflow and design workshops were performed as part of the WP1 efforts. These efforts of WP1 were complemented by efforts in WP5.

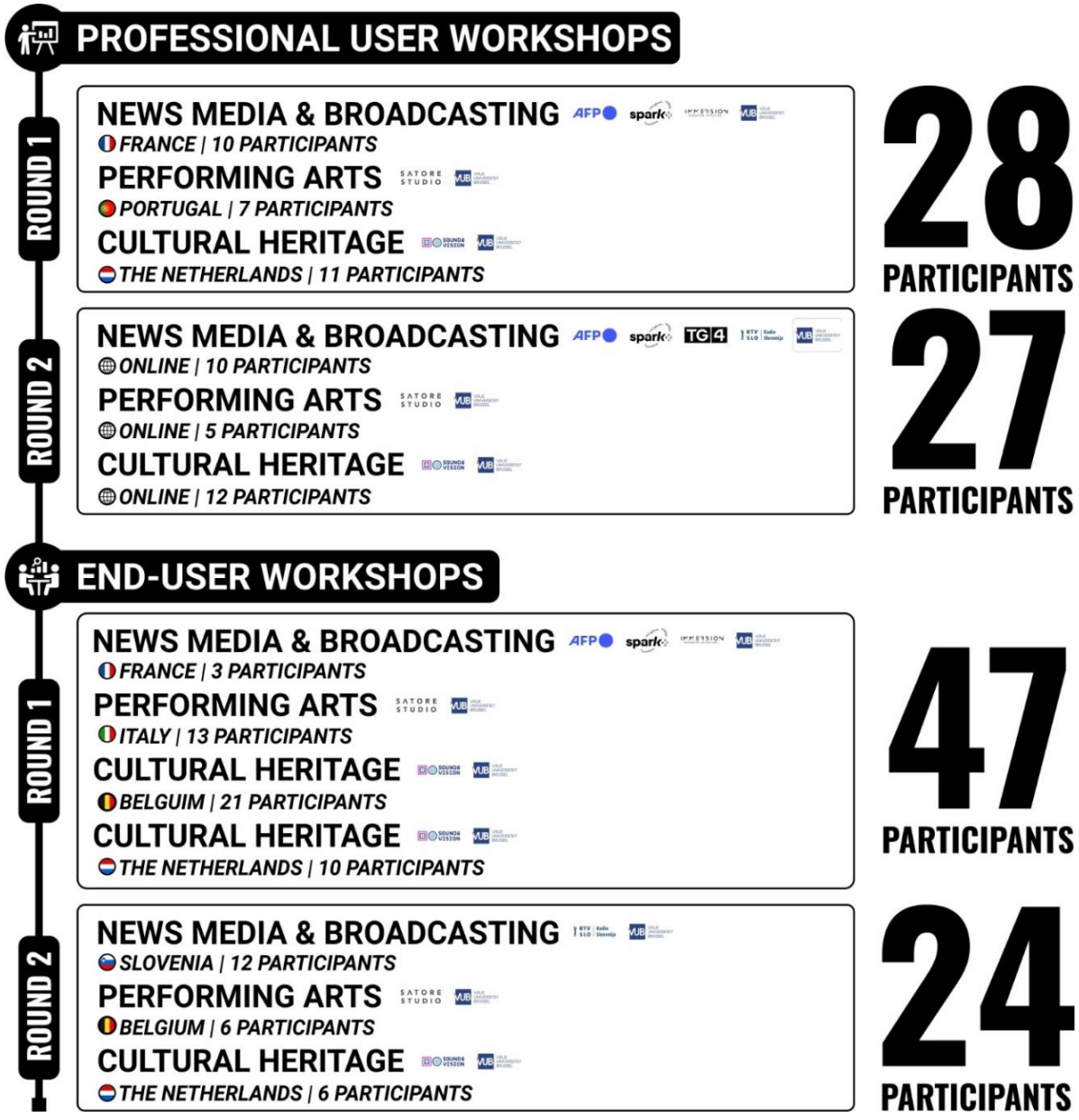


Figure 2: Participation over the 13 organized workshops presented during online GA Dublin of WP1 (6 Nov 2023).

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

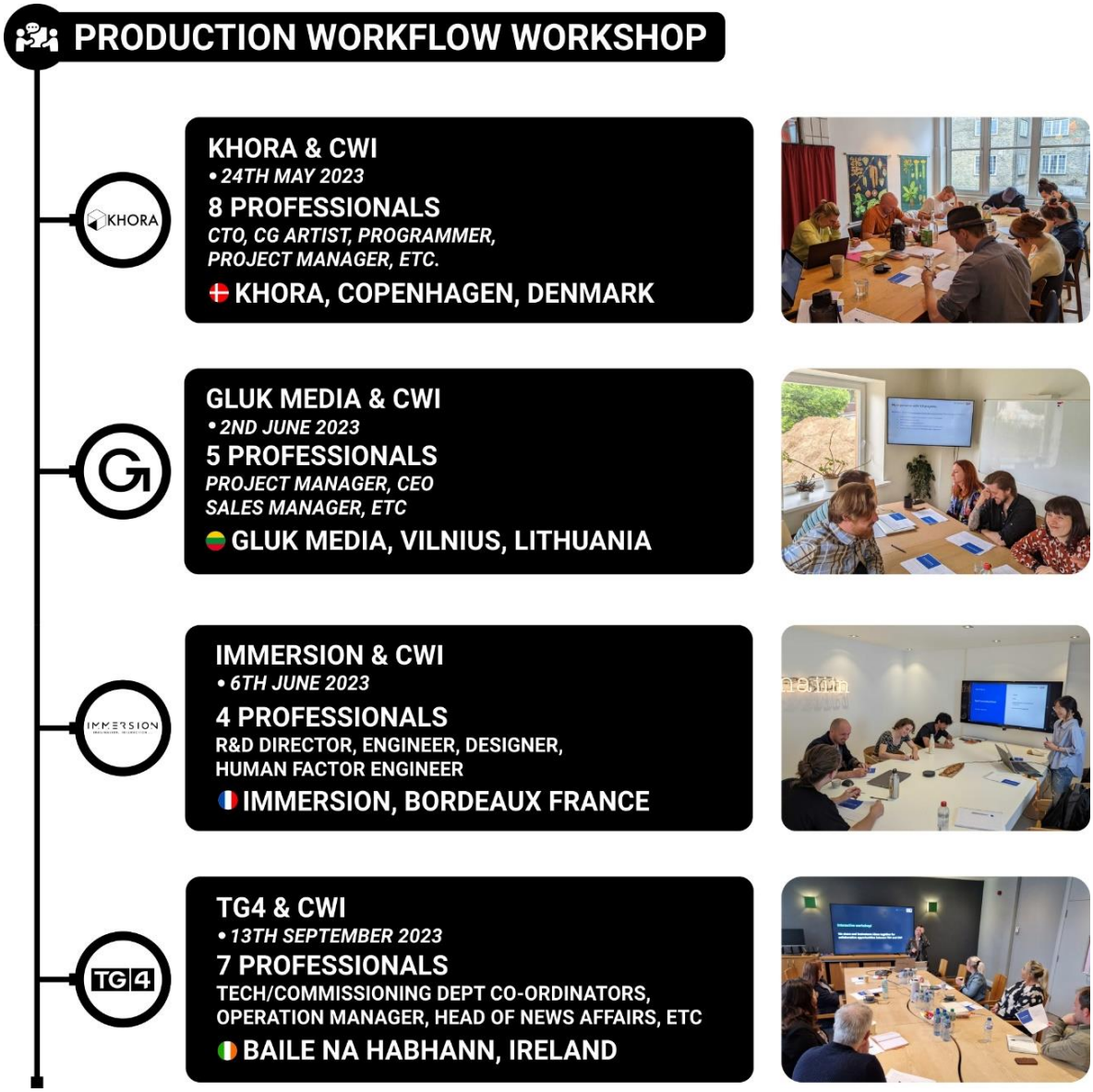


Figure 2: Participation over the 13 organized workshops presented during online GA Dublin of WP1 (6 Nov 2023).

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

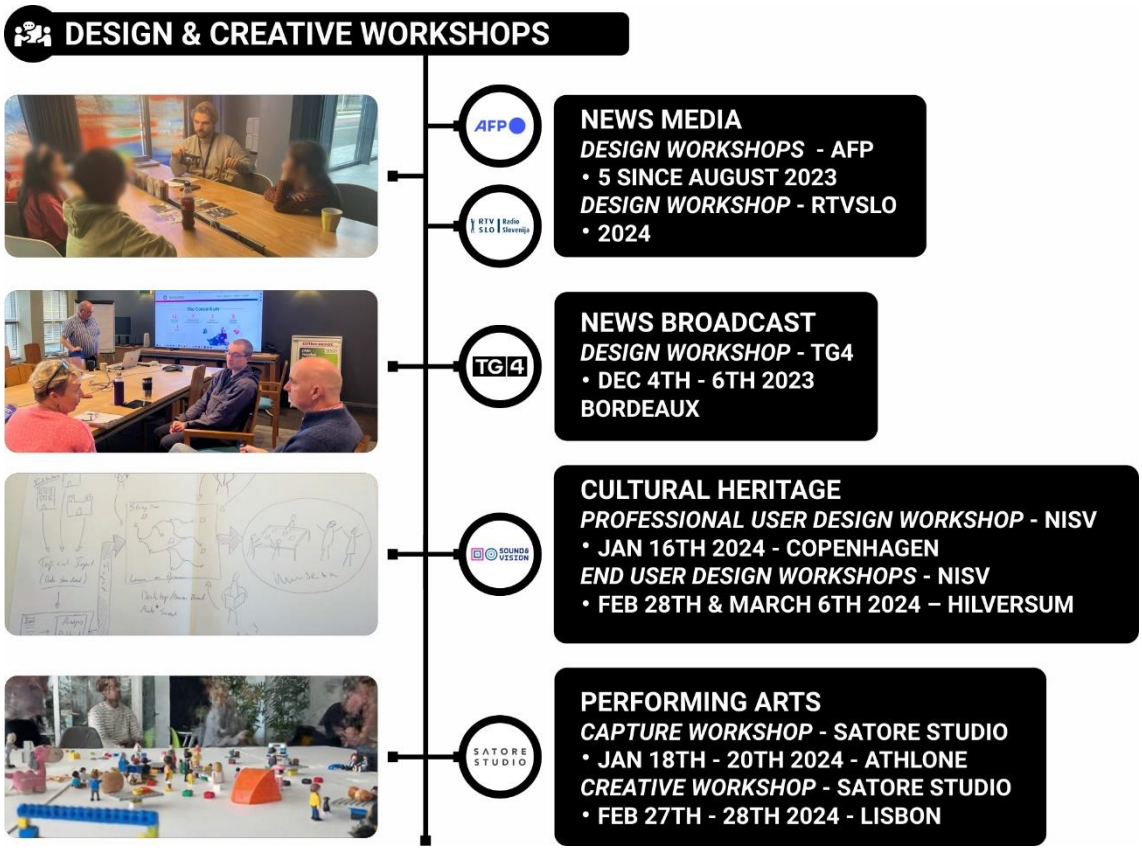


Figure 2: Participation over the 13 organized workshops presented during online GA Dublin of WP1 (6 Nov 2023).

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

3 AFP & RTVSLO – News Media Pilot

3.1 Scenario

This use case pilot (WP5) targets news media with AFP handling creation part (WP3) and RTVSLO handling experience (WP4). SparkNews and TG4 also provided inputs from their own perspectives.

Figure 3: News media characteristics.

It will impact the sector of news media with business opportunities by offering new content format to sell and show room for content. Furthermore, the proposed technologies can help on accelerate the creation process such as making infographics team working quicker to produce volumetric content (e.g. AFP use case pilot).

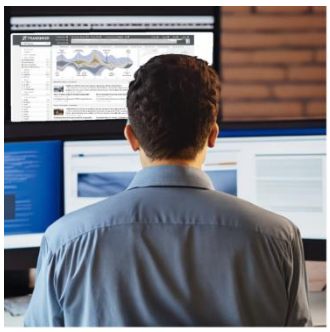
The professional user (any kind of journalist) needs a tool that can provide them with a competitive advantage and efficiencies in repetitive manual tasks, thus freeing up time for more added value tasks. This tool will include 1) a real time news monitoring tool and 2) accessible volumetric and formatting tools that integrate in the journalists' CMS.

The news-consumer user wants to be able to immerse themselves in an experience that takes them behind the scenes of an event that is out of their reach and gives them in-depth knowledge. To do so, the experience will be gamified, immersive, and interactive, also allowing for a personalized exploration. This will lead to a better understanding of a news event and topic.

The associated storyboards are presented in Table 2.

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

Table 2: News media storyboards.

Journalist Point of View	
The journalist "on duty" at the AFP Rennes office keeps an eye on the TRANSMIXR platform, to check for breaking news. The platform was fed with all the sources (local news websites + social network accounts) monitored by that newsroom. Before TRANSMIXR, the journalists had to keep checking the different websites + social network accounts manually.	
Another journalist preparing a story ahead of the Vendée Globe (the round of the world single-handed, non-stop, non-assisted round-the-world sailing race). This event takes place every 4 years. The next one is in November 2024. The journalist uses the TRANSMIXR platform to research on the subject and sets up a report with one of the sailing teams taking part in the event.	
On the day of the report, a video journalist or photographer shoots drone images of the sailing boat and mobile phone background images specifically for the infographics team to generate 3D volumetric images of the boat.	

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

The video journalist or photographer then sends the drone footage to the infographic department that use a tool similar to LumaLabs to produce a 3D volumetric image of the boat that will then be used in a videography to describe and explain the “giants of the seas” from outside and inside.	
Once the report is ready, the video desk adapts it according to the different channels where the report will be shown. To format the report according to the channel (length, vertical-horizontal format), the video editor uses the CERTH summarization tool.	
Viewer Point of View – VR Headset at Home	
User uses VR headset to enter the place (before or after the event). This gives him an opportunity to learn more about the setting, the sport, the history of the event and the Nordic disciplines.	
User interacts with a reporter and starts learning.	

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

User explores various points of the venue, augmented with infographics (the jumping device, athletes' room, tunnel, referee tower). Reporter is still present, but this time only with voice (spatial sound experience).	
User experiences the jump and immerses in other experiences like giving an interview in the mixed zone.	
User learns about the rules, physics, meteorology, history of ski jumps. Narration of the reporter (explainer)	
User "flies" over the venue and discovers the geography of the valley using gestures. User can learn about the geological history of the glacier valley and how it evolved.	
User gets access to other material and is directed to further sources or restarts the experience.	
User can also experience the platform within their mobile phone with headphones to explore other facilities during an interval. He is at the venue or at home. User interacts with a reporter, they explore the venue (AR), experience the jump, learn about the rules, physics, meteorology, and history of ski jumps. Finally, they get access to other sources.	

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

In the example, for the end user as a demonstrator, we have chosen to create an experience of the 45th FIS Ski Jumping and Ski Flying World Cup, Planica, hosting every ski championship in Slovenia. The set of associated features is described on the following "Needs and Specifications" section. Again, the specifications below are a work in progress informed by the ongoing requirements and design workshops of WP1.

3.2 Needs and Specifications

Table 3: News media specifications.

Need	Associated specifications	Envisioned Technical Solutions
All-in-one platform to work faster and be more efficient.	Ability to keep an eye on everything. Centralizing all sources, including social network.	WLT Dashboard. Alerting for breaking news. Monitoring.
	Centralizing different content production tools.	Such as WLT, CERTH and 3D volumetric tools.
	Ability to do research and organize the findings.	Search tool with selected sources (text, video, photo, VR, graphics, databases)
Be easily integrated in the existing workflow	Adaptability of the format to different channels. Edit the results if needed.	Format converter tool. CERTH summarization tool.
	Trustable. Be able to control and select the sources.	To be integrated in their CMS. WLT Dashboard.
Create immersive informative content	Put infographic department in the loop and create new format.	3D volumetric capture with drone and mobile phone // content enhancement with tools like Luma Labs.
	Ability to explore various points of the venue.	Interactive experience.
User can learn things	Informative menu with the opportunity to learn about the setting, sport and history of an event.	The user can make statistics / information appear during experience.
	Engaging and interactive experience.	Pre-recorded / not real time interaction with the reporter.

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

	Ability to access additional sources.	Possibility to access archives images of the event (captured by RTV SLO and accessible via WLT).
Appealing experience for the user	Personalized exploration.	User profile with localisation, age, interests…
	Innovative experience.	Immersive experience as described above.

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

4 Satore Studio – Enriching Performing Art Experiences

4.1 Scenario

This use case pilot (WP5) targets performing art, including dance, and acting, through an immersive theatre concept. Even though the main concept is real-time and local, we will create a "CCTV view" of the experience to enable remote experience across multiple platforms for accessibility. The piece hence is then also available across different countries.

Figure 4: Enriching performing art characteristics.

This is an opportunity to inform the performing arts sector by bringing an awareness of what is possible with these Immersive and AI technologies. This use case pilot is motivated to inform future directions for other art pieces such as theatre, opera, movement Repeatability in terms of content use is also considered a key component, e.g. capture content may be (re) adapted towards various new pieces. The result is new business opportunities with the creation of new types of employment in art sector and the resale of tickets for additional experiences.

Professional users are performers (Artistic Director, Actors, Dancer, Opera Singer, Musicians, Producer, Costume Designers, Set Designer, …) while the end-users are the audience. Currently, the audience is predominately aged over 55 years old upper socio-economic background groups. Over 90% attending from a white ethnic background. In TRANSMIXR, via the performing arts use case pilot, we want to disrupt this limitation, increasing accessibility to new audiences via immersive technologies.

"Tale Weaver" is an immersive theatre experience that blends physical and Mixed Reality (MR) elements to create a multi-layered and dynamic narrative. This form of theatre is distinguished by its interactive and engaging nature, drawing the audience into the story's world rather than having them merely observe from the outside. This is a live performance.

The concept is an immersive theatre experience, featuring in-real-life live actors (Violet, Chris, and restaurant occupants from the synopsis bellow) and Mixed Reality Actors (only visible through MR). In addition to the actors, 24 interactors, each wearing a

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

headset, play a crucial role. These headsets not only identify each interactor but also allow them to see additional scenes and actors within the MR world. This enables interactors to engage with both the live actors and the MR elements. The result is the creation of a multi-dimensional and interactive narrative. For example, imagine that 4 actors can at any given point puppeteer the MR actors (through Mocap), thus making the interactors not know if they are with an NPC, an actor, or a pre-recorded movement. This immersive setup blends live performances with MR technology, creating a dynamic and multi-layered storytelling environment, offering a unique exploration of the emotional and social dynamics.

The synopsis of the show is:

"Violet is waiting at the restaurant. She has been there for 20 mins and is starting to believe she will be stood up. When Chris arrives, he apologizes, saying he had a mishap. He tries to break the ice with a cheesy pickup line that he thinks is charming, but it falls flat, leading to an awkward silence that he attempts to dispel with a self-deprecating joke.

Christian tries to impress by ordering food in a foreign language, but due to a mispronunciation, he accidentally orders something completely unexpected, leading to a surprising and humorous reveal when the dish arrives.

At the end of the night, Christian reaches for the bill only to discover he's left his wallet at home, he describes how his friends told him to wear these trousers and his wallet must have been left in the other trousers. This leads to an impromptu adventure as they figure out a creative solution, showcasing his resourcefulness and their growing connection." Again, the specifications below are a work in progress informed by the ongoing requirements and design workshops of WP1.

4.2 Needs and Specifications

The specifications outlined in Table 4 are a work in progress informed by the ongoing requirements and design workshops of WP1.

Table 4: Enriching performing arts specifications.

Need	Associated specifications	Envisioned technical solutions
Create an immersive and engaging experience where the user is mentally and physically involved	Blend of physical and MR elements to offer a rich and layered storytelling environment.	MR actors with humanoid avatars. MR scenes.
	Dynamic and multi-layered narrative.	Unfolds through a mix of scripted scenes,

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

		improvisation, and audience interaction.
	Interactivity with both live actors and MR elements.	Some MR actors can be puppeteer by live actors through Mocap.
	Experience various emotional and Social Dynamics.	Mix of scripted and spontaneous moments.
Audio-Visual enhancement	Immersive experience with sensory cues.	Spatial audio and lightning.
	Augment the physical set with added details or effects.	Changing ambiance settings or introducing fantastical elements.
	Possible to receive textual information.	AR.
Seamless integration of real and virtual	The user can choose if they want to access extra XR layer (HMD) or not.	Easy access to XR layer. To know which technology to use.
	The user can choose if they want to actively participate or just observe.	Let the choice. The scenario do not necessarily requires user action.
	Ability to have hybrid (online – offline) experience.	Let the choice.
Freedom of action	Ability to customize the XR environment for the directors.	Change the look and feeling, add different audio files.
	Ability to move and change objects within the XR environment.	For Choreographers and Directors to change their mind. Use of USD.
	The user can choose between physical and online.	Option.
	Ability to share this experience in group or alone.	Since it is a theatrical experience, it can be attended alone or with a group.

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

	No constraint about where to look at. Similarly, performers shall be able to move as on a normal stage.	360° performance with spatial cues but no constraint about where to look.
Clear understanding of the environment	Ability to collaborate with their team.	Collaboration cues. Synchronous movements.
	Gather audience feedback.	Enhanced communication.
	Ability to be aware of their own body movement when integrating XR.	See their own body, AR.
Strong ethical consideration. Security and safe feeling	GDPR and privacy compliant.	The system will not save the details of the audience at the theatre. For viewers at home, they can opt out.
	To not feel alienated or overwhelm with their surroundings.	Let user some space, freedom of action.

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

5 NISV – Immersive and Data-Driven Cultural Heritage Experience

5.1 Scenario

This use case pilot focuses on introducing a replicable immersive storytelling format for cultural heritage organisations. It aims to deliver innovation in three areas:

1. Using social XR capabilities, streamline communication between design companies and cultural heritage organisation to facilitate collaboration during the production process,
2. Introduce ways to use cultural heritage metadata for storytelling in an immersive environment,
3. Democratise immersive storytelling in the cultural heritage sector by creating reusable templates that make such experiences much more accessible to organisations.

Figure 5: Cultural heritage characteristics.

This use-case aims to introduce a customisable and reusable format for immersive storytelling with heritage data that is powered by social XR for collaboration between professional users during pre-production and post-production. We are bringing the huge amounts of data that are present in heritage organizations into XR. By not only using individual heritage objects but also metadata, we create a new tool for storytelling that curators can use to show connections across heritage data. Audiences gain a new way to experience heritage collections where XR tools allow them to experience the large scale and scope of collections that would not be possible to access in a physical space. Professional users of this use case are the staff/curators of cultural heritage organisations and design entities creating immersive for / with them, while end-users are visitors of cultural experiences aged 8-

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

14.[1] The end-users attend locally whereas the professional users can work together remotely.

As a cultural heritage professional, the aim is to create an engaging immersive reality experience that (1) harnesses the rich metadata available in digital archives to tell stories, (2) can be co-created between multiple experts using social XR, (3) can be iteratively updated with new data and stories, and (4) is economically sustainable for the organisation.

As a visitor, the aim is to have an educational and fun museum experience so that I have a nice visit and learn something new, whilst interacting with others and sharing experiences. The workflow of the cultural heritage use case pilot is as per Fig. 6.

[1] Note that this age group is specific to the context of the use case partner executing the pilot. It could be adapted to include other end users from other age ranges.

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

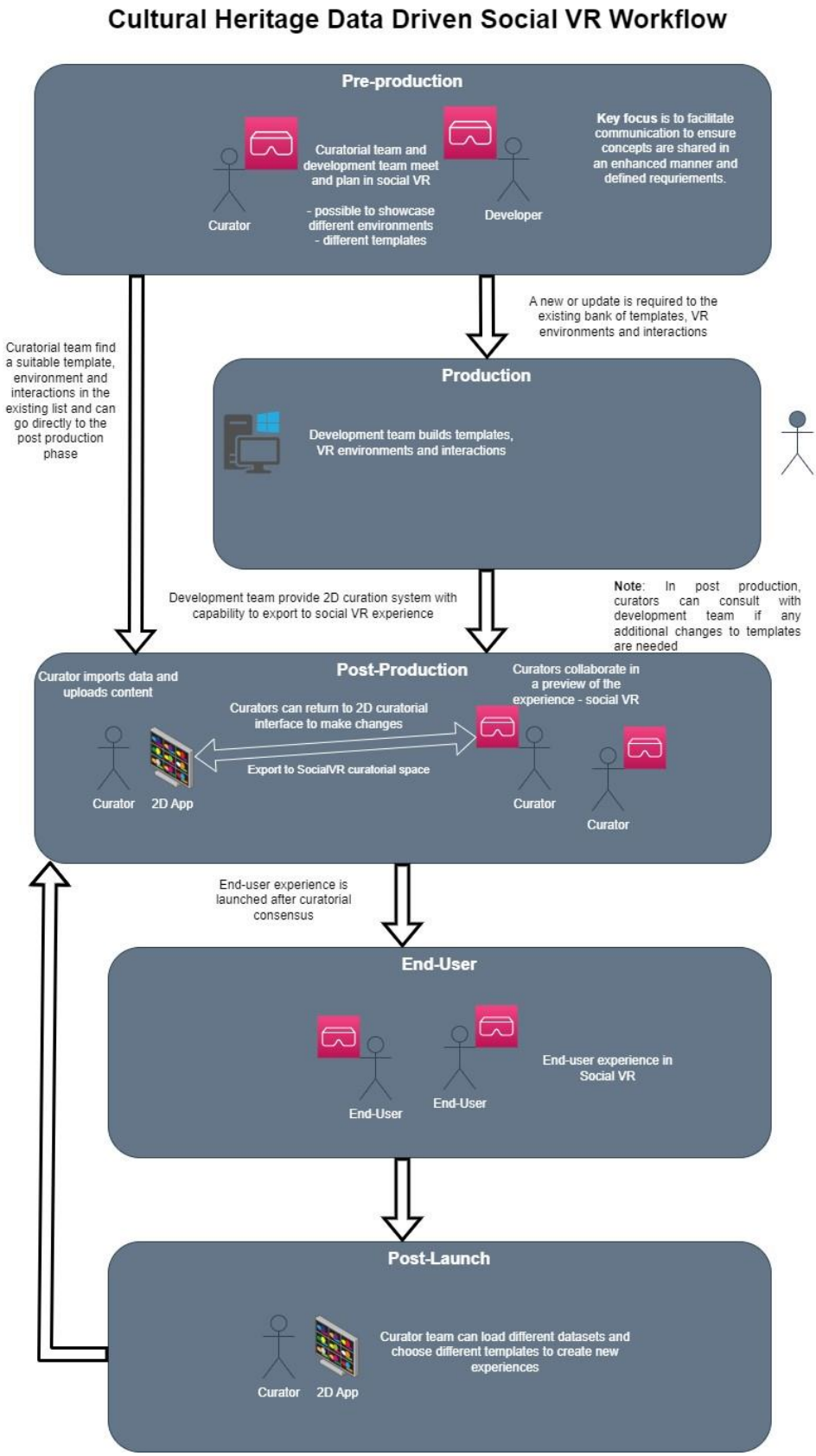
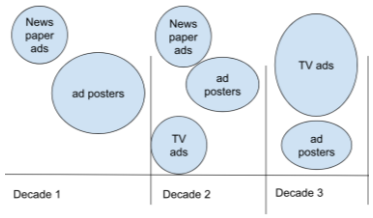
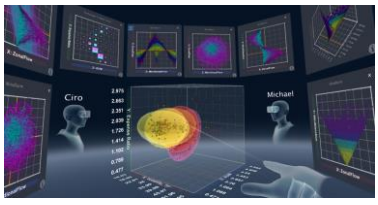
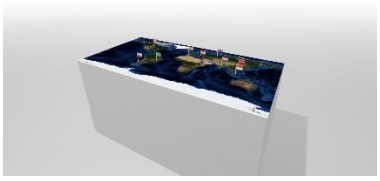
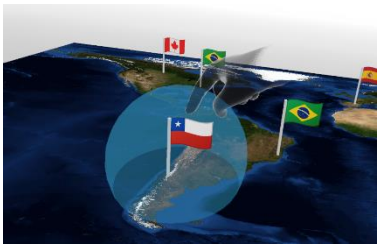


Figure 6: Cultural Heritage Data Driven Social XR scenario.

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

Examples of storytelling concepts currently in being defined in WP1 (requirements and design workshops) and WP5 (use case pilots) are presented in Table 5.

Table 5: Cultural heritage storyboards.

AR	VR	XR
A young museum visitor becomes an investigative journalist to uncover the bigger story that lies behind the official exhibition thanks to the archives through AR. She gets the instructions from a tablet, following an 8-steps quest. 	Two museum visitors equipped with VR headsets. They find themselves surrounded by visualization with each frame representing a topic. Puzzle to solve through metadata.  	Discovering the climate change worldwide thanks to media collection. On a map table, pins and relevant connections to answer quiz.  

5.2 Needs and Specifications

The needs and specifications in Table 6 are a work in progress informed by the ongoing requirements and design workshops of WP1.

Table 6: Cultural heritage specifications.

Need	Associated Specifications	Envisioned Technical Solutions
Enable professional users to interact seamlessly within the virtual environment during pre-	Include indirect and complementary interactions.	Use head movements for steering, use handhelds
	Optimized data transfer and minimized latency.	Good network connection.

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

/post-production processes.		
Make the data-driven storytelling experience both educational and engaging for end-users.	Include elements of gamification.	Reusable templates for gamified interactions with data in the immersive environment
	Facilitate real-time communication and collaboration among trusted end users (e.g. siblings or friends visiting the experience together).	Integrate voice chat and possible avatars that can interact
Possibility for different heritage organisations to reuse the experience format multiple times to tell different stories.	A portfolio of pre-defined immersive data-driven storytelling elements available for heritage organisations	Templates for data visualisations and gamified interactions with data in an immersive environment available for choosing during the pre-production process
	A formalised approach to deliver data in the right format for the experience.	Predefined workflow and formats to upload heritage data for the experience using a browser-based UI

About the pre-identified technical components of the use-case:

- **Dashboard** – In our first prototype, heritage data will be accessed via APIs from a data repository (e.g. akin to the Weblyzard Dashboard). This retrieved data is viewed in XR. It is challenging to make the experience both gamified and educational without it requiring too much development resources. Also, scalability is a challenge (not creating a 1-off). Indeed scalability of such

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

creation to support the cultural heritage sector is a key motivation and is being realized via the TRANSMIXR Templates outlined below.

- **Social XR** – Social XR will play a key role in supporting communication between designers and heritage professionals during the pre-production and post-production stages of designing the experiences. Additionally, this use case will explore whether elements of social XR could be introduced in the end user experience (e.g. joining the immersive experience together, interacting with content created by previous visitors, etc.).
- **Templates** – We create templates for data visualizations as well as templates for interacting with data. Organizations can mix and match these and insert their own data to create a new storyline. Using the templates makes the concept more scalable reducing development costs and time with each iteration.

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

6 TG4 – Control Room of the Future

6.1 Scenario

This use case pilot targets control room with the goal of offering a sandbox for a small broadcaster. To do so, it proposes a workflow including a method to interface virtual cameras and to exploit existing 360 video. It allows to news broadcasters to explore volumetric capture as a new format.

Figure 7: Control room characteristics.

This section introduces the specifications for live media fostered during a workshop with TG4 professionals in Immersion. They met for a day of demos and activities allowing us to gather a bench of inputs for the scenario. They experienced state-of-the-art XR devices including different ranges of XR HMD (HoloLens, Meta Quest 3, Varjo XR-3 with and without Chromakey), stereoscopic displays (L-Screen, Looking Glass), haptic gloves (We-Art) and Volumetric Capture. These demos aimed to open the minds of participants before the ideation session. They also experienced the virtual control room Immersion prototyped and first presented the General Assembly 2023. The participants valued the functionality of a virtual control room for remote operations to reduce the need to send too many people into the field. They highlighted the need of quicker interaction. Moreover, participants expressed the feeling that virtual reality is going to be a breakthrough.

Professional users are journalists on live media production and diffusion (hereafter called Presenter, Journalist and Production Teams) while end-users are the Audience.

- The Audience is receiving the immersive information. As our end-users (consumers) the platform seeks to influence and increase engagement.
- The Presenter is in charge of leading the Audience through the experience, from a controlled studio and recorded with volumetric capture.
- The Journalist stands in the reported environment (out of the studio) and is sharing facts to the Presenter and the Audience. The Journalist is the experience guide on site, recorded in 360 Video.
- The Production Teams will be a backstage operating system, in charge of installations, maintenance, filming, assets production, broadcasting… Two directors will be present: one in the studio instructing the studio's production

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

team by accessing the AR virtual studio platform and another in the field instructing the on-site production team. The studio's director will manage the broadcasted content and set virtual cameras for phone and web users.

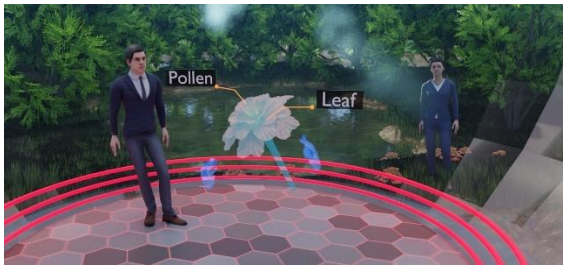
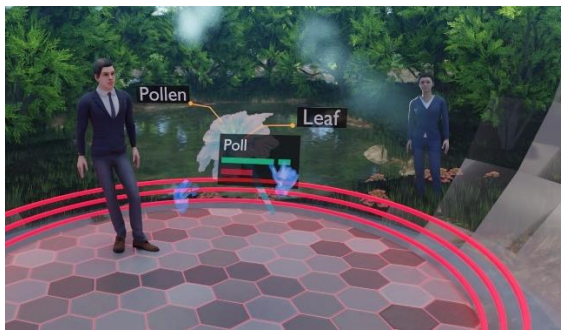
The expressed demand was to create a platform that is accessible by all, that allows to share news information in an immersive way for an enhanced experience. Everyone should receive information in the clearest and most transparent way, so Audience can reach the platform with the most accessible hardware they can provide, ranging from phones, web services to XR headsets for the most immersive experience. The respectively provided Hardware will only define the level of stimulation and depth of immersion in the experience, thus the information transmitted will remain as true on all platforms. We focus on the XR end-users because of the higher complexity of the setup.

The use-case storyboard from the Audience would be the following:

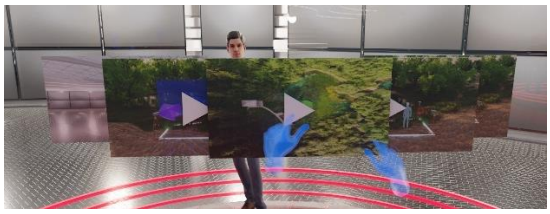
Table 7: Control room storyboard.

Sean just got back to work. When scrolling on his phone he receives a notification from his XR news app with attached QR Code inviting him for a live reporting of a newly discovered flower.	
A Presenter projected in 3D into the virtual studio welcomes the Audience. Sean is free to navigate in the defined red platform. The Audience is able to see other people sitting on benches behind them to give an enhanced immersive feeling. The Presenter introduces the subject and asks the Journalist in the field.	

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

The Journalist starts reporting and invites the Presenter and the Audience to join by teleporting into the live streamed 360 environment of the real site to follow the reporting in a more immersive way.	
The Journalist talks about the discovery of a new flower. The Presenter shows the Audience a live example of the flower that he has in the studio. The Audience can approach the 3D capture flower and inspect it up close.	
Then the Presenter shows an enhanced 3D model of the flower and indicates its essential characteristics to reinforce the Journalist's reporting.	
The Presenter launches polls to get feedback from the Audience so that he can tailor the reporting experience based on their curiosity and understanding. Sean answers the poll.	

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

The reporting done, the Presenter and the Audience are back in the studio where the Presenter presents infographics to the Audience and interacts with them by initiating debates on the gathered data.	
The session ends and Sean decides to stay in the virtual world and search for another show to watch from his personalized lobby.	

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

The stakeholders interact together as described in the service architecture bellow demonstrating the functional core process:

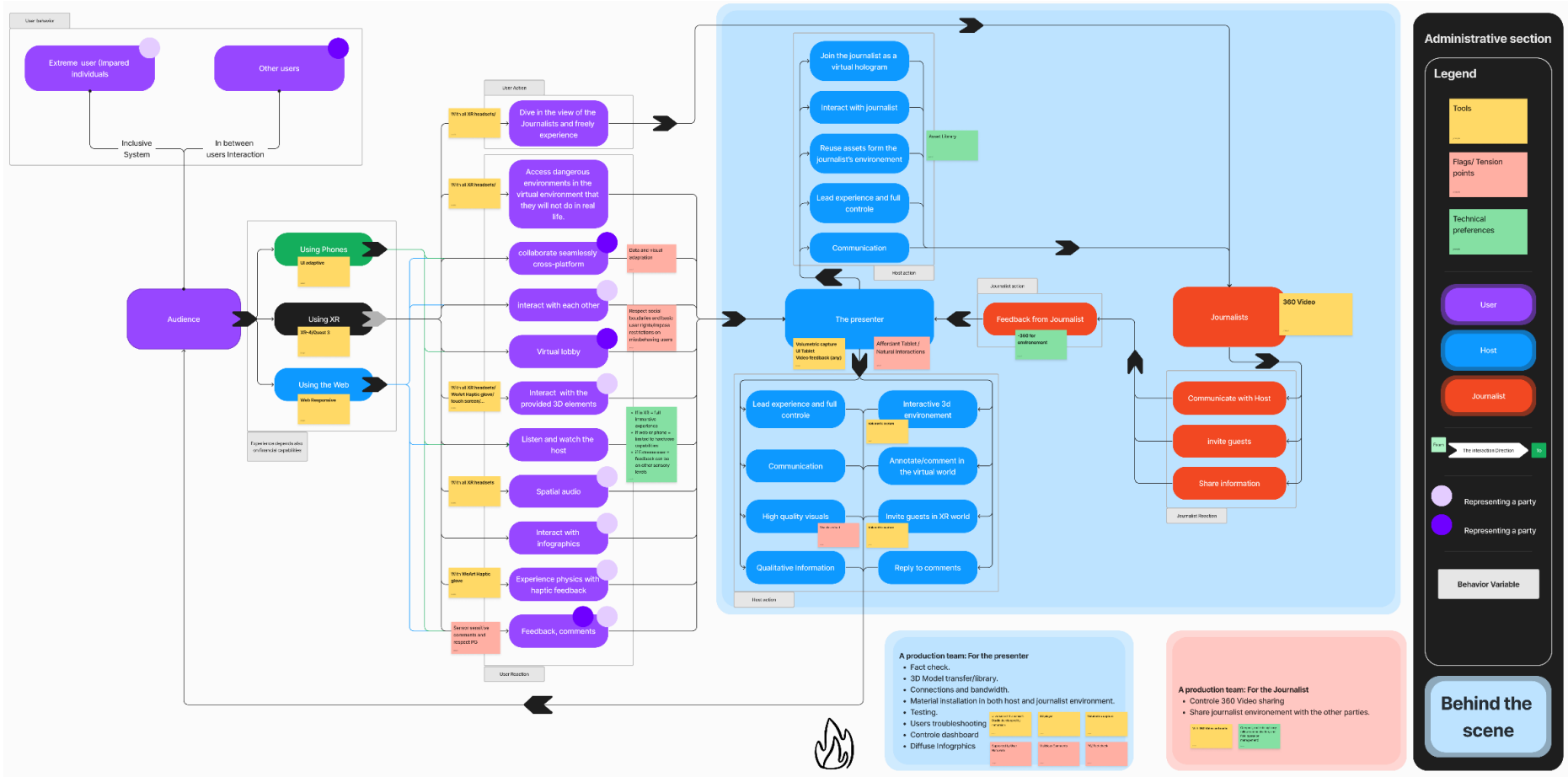


Figure 8: Control room functional core process.

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

6.2 Needs and Specifications

Table 8: Control room specifications.

Need	Associated Specifications	Envisioned Technical Solutions
Control room needs to be easy to handle. Functional	Facilitate the interaction gestures.	Include a learning stage.
	Reduce the number of clicks.	Include voice command.
	Contact the remote team members.	Include calls in a future version. This will not be handled in the demonstrators.
Culture of innovation, grasp new technology seamlessly.	Accessible anytime / anywhere / on any platform.	Phone, web or XR access.
	Limit the number of intermediaries for connection, make it direct.	App sends QRCode directly to the Audience.
Generate curiosity and engagement even among younger audience.	Provide different perspectives (format, age, localization etc.).	Personalized lobby.
	Strong interaction thanks to social networking.	Include comments with various levels of access (anonymous, friends, public) and moderation.
	Gamification for unique experience.	Polls, stats and feedback.

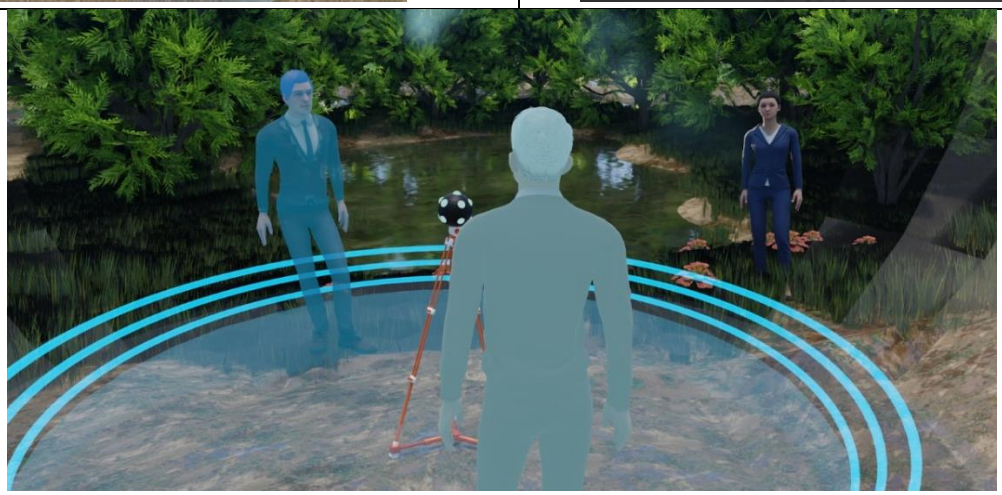
This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

	Allow the Audience to freely explore an aera. Immersive experience.	Red platform representing the volumetric capture area where the Audience can move. 360° video capture.
Informative and interactive content.	Feed the Presenter with cues to adapt the rhythm of the presentation.	Use user comments and polls as feedback. Dashboard screens on the Presenter side with all the needed information.
	Include 3D models and operations such as exploding, pointing...	Cloud and local libraries.
Strong values (sustainable, accessible, trustworthy, secured, reliable, qualitative, innovative).	Allow the user to understand the project and how it works.	Opening sequence and first episode about TRANSMIXR project.

The envisioned setup includes two main complementary technical components:

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

Table 9: Control room live capture components.

360° Video Capture of the Environment	Presenter Volumetric Capture
The Audience can look around them and see the Journalist as part of the immersive environment. They cannot walk out of the predesigned platform in red that will be present during all of the experience.	The Audience can move around the Presenter and (if) objects of interest which are volumetrically captured, it is a more active experience.
	
	

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

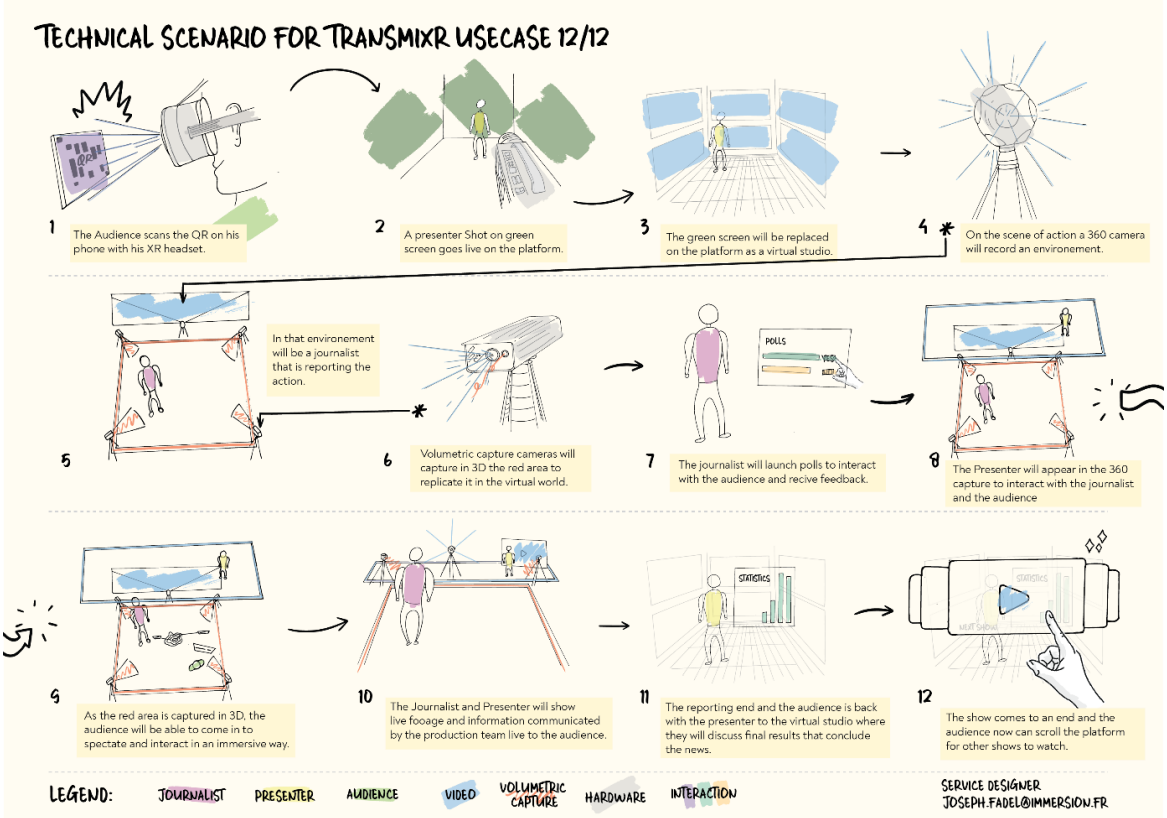


Figure 9: Control room technical scenario.

Let notice that the Presenter and Journalist have no headsets while Production Team in the studio has access to the distributed control room in AR similarly to the previous demonstrator built by Immersion. It will be revised in accordance with the feedback gathered during the workshop.

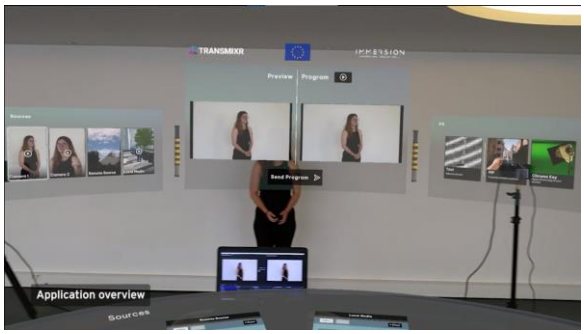


Figure 10: Control room broadcasting component

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

7 Conclusion and Future Work

From the different use cases each offering an entire pre-production, post-production, delivery and post-release pipeline – also illustrated through our work packages understanding (WP2), creation (WP3) and experience (WP4), we created a common framework in order to cover various modalities of XR media.

While bringing the use cases together, this document compiles the outcomes of initial workshops with industry professionals. It meticulously delineates specifications derived from pinpointed requirements and constructive feedback, offering a comprehensive view of selected use case scenarios. These scenarios vividly illustrate practical applications and potential challenges, providing a clear roadmap for development.

Based on it, upcoming future work is to iteratively design and build demonstrators towards the pilots. This process is instrumental in guiding the creation of the deliverable demonstrator for WP4. Envisioned as a tangible representation of our project's progress, it will serve as a pivotal tool for evaluating the applicability and efficacy of our solutions in real-world settings. Inherent in its design is the flexibility for continual refinement, ensuring that the demonstrator evolves through successive iterations of testing and feedback.

This iterative approach, aligned with best practices in project development, allows for adaptive modifications, enhancing the demonstrator's relevance and effectiveness as new insights and requirements emerge from ongoing evaluations and stakeholder interactions. It will lead to a minimum viable product for each use case.

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

8 Appendix – State of the Art

This section sets the current state of the art of Social XR first by introducing some definitions and characteristics, then exploring the existing solutions with opportunities and obstacles.

8.1 Definitions

8.1.1 XR and Social XR

Extended Reality (XR) is a global term which gathers various definitions. One of the most common one is based on different levels of virtuality such as Augmented Reality (AR), Mixed Reality (MR) and Virtual Reality (VR) (MILGRAM & KISHINO, 1994; Rauschnabel et al., 2022; U. S. Government Accountability, 2022).

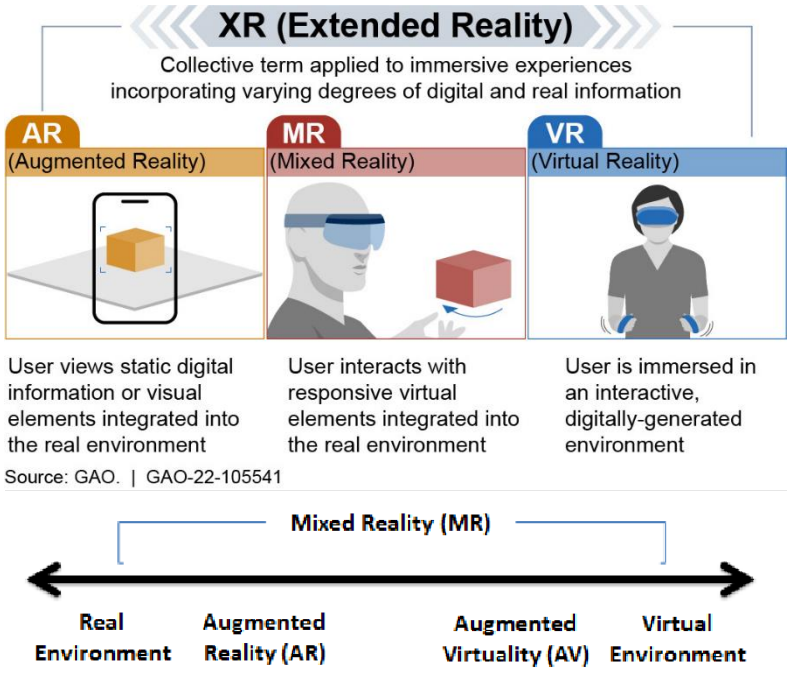


Figure 11: Definition of XR.

XR solutions are to be used in industry, education, entertainment, healthcare, psychology, art, and many other fields. More and more of these XR solutions are made for groups of users and collaboration, which raises many questions about the social aspects of XR.

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

"Social XR allows several people or a group of people to experience social presence and co-presence by engaging in activities and real-time conversations with each other." (Görlich, 2022)

8.1.2 How Humans Interact with Each Other

To unleash the potential of Social XR, it is first necessary to understand the rules of human interactions. Human interaction is complex and have many characteristics (Levinson, 2006) such as:

- Humans respond to actions or intentions, not to behaviours.
- Interaction implies mutual stimulation with reciprocity of roles (speaker-addressee, giver-taker) and their alternation over time.
- Interaction implies a constantly modulating participation structure.
- Interaction sequences are governed by expectation, not by rule. On a close timing.
- Interaction appears to have detailed universal properties such as turn taking, adjacency pairs, greetings, repairing misunderstandings, etc.
- Although human interaction is dominated by the use of language, language does not actually code the crucial actions being performed. Besides, interaction is characterized by multimodal signal streams (visual, auditory, and haptic at the receiving end, and kinesics, vocal, and motor at the producing end), specifically linked across time and modality.

In addition, (Pelikan & Broth, 2016) highlighted several organizational speaking processes that occur during interaction:

- Turns at talk are not simply following one after another but they project back on what has been said before and create expectations about relevant next turns.
- It relates not only to the current conversational context, but also to its addressee (regarding knowledge, motives, and expectancies).
- There is a set of normative rules that regulates turn-taking for conversation, minimizing gaps and overlaps. Sequencing occurs due to "adjacency pairs" such as greetings.
- Some speaking, hearing or understanding mistakes can occur and may be repaired in various ways. A self-repair is usually initiated before the next speaker's turn.

(Mondada, 2012) points out the large diversity of acoustic language (prosody, phonetics, syntax, lexicon, etc.) and body visual dimensions (gestures, facial mimics, posture, head movements etc.). Humans possess innate senses and high perceptive, integrative, and interpretive capabilities for various sensorial information. Multimodal

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

information can be either alternative or complementary depending on the performed task (Rakkolainen et al., 2021). Hand or finger pointing have been used since childhood as an intuitive gesture. Pointing tools are universal and necessary in human-human and human-machine interaction as location solver. They can easily be completed by language (deictic information).

As highlighted above, even if interaction is dominated by the language modality, the main part of information is differently coded. Indeed, non-verbal communication would represent 70% of intrapersonal communication (Wang, 2020). Social XR could represent non-verbal information in real time.

Many authors investigated multimodal interaction with different multimodal input and output interfaces that are necessary to design in order to provide best answer to user communication (Frohlich, 1992; Nigay & Coutaz, 1996; Rakkolainen et al., 2021).

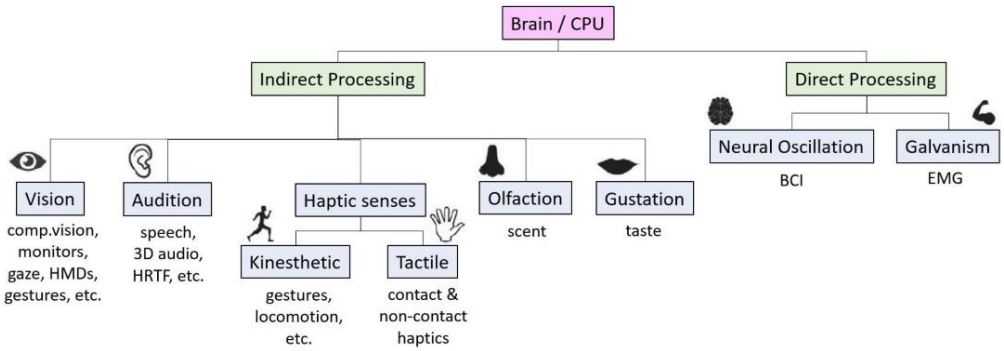


Figure 12: (Rakkolainen et al., 2021)'s taxonomy based on human senses and classifying both input and output devices and technologies for multimodal interaction.

8.2 Reproducing Human-to-Human communication mechanisms

8.2.1 Represent the Users

In order to have multiple users interacting naturally remotely, the representation of each other becomes an important issue. (Lee & Yoo, 2021) suggest a representation of user engagement in the real-virtual continuum:

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

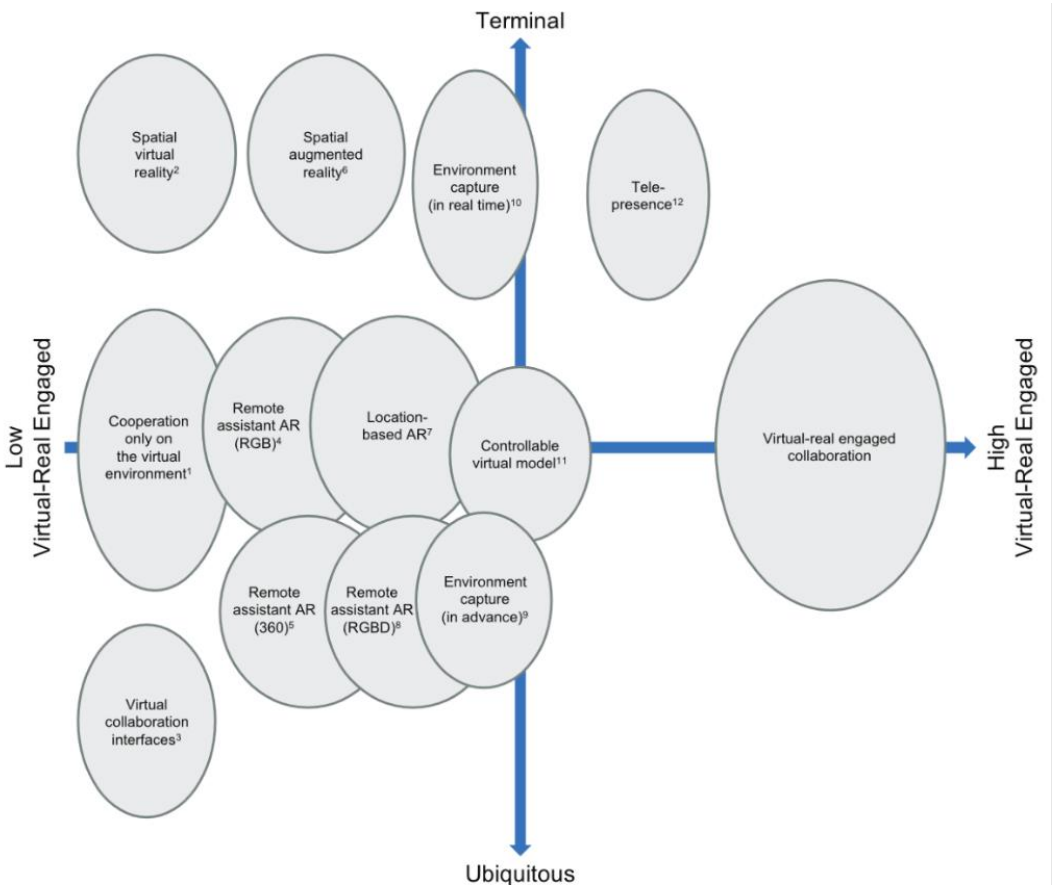


Figure 13: The virtual-real engagement and ubiquitous computing continuum by (Lee & Yoo, 2021)

Some studies suggest that ideal avatars created from strong personalization might enhance user identification and entertainment, therefore increasing brand engagement and loyalty, hence purchase (Jeon, 2023).

While some existing solutions prefer to leave it up to the user, others major social applications in XR the choice of representation falls into three categories: default avatars, cartoon avatars and realistic avatars.

8.2.2 Default Avatars

"Default" avatars are minorly present in today's applications. This category includes avatars with a simple, impersonal face. Most often these are robots, silhouettes, or holograms. The problem with these avatars is that they are often emotionless and allow less natural social interaction due to the lack of a reference point, like a human face. This type of avatar can be found in applications such as Nvidia Holodeck[2] or in

[2] https://www.nvidia.com/fr-fr/design-visualization/technologies/holodeck/

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

the avatars at the launch of VRChat[3]. However, this type of avatar can sometimes be used in some applications to transfer a message, as in Where thoughts go, where users are invited to reveal their secrets, fears or dreams via recorded voice message responses. In this game, players who have left a message are represented by small white balls with black eyes, all identical, thus making the people who left messages unidentifiable.

Figure 14: Nvidia Holodeck (Greenstein, 2017).

8.2.3 Cartoon Avatars

The second type of avatar is currently the most common: cartoon avatars. They can be found in the majority of social applications in XR, and the choice of representation is generally made by creating an avatar via a creation menu in the application.

The choices are generally:

- **Gender:** nowadays the choice of gender is less and less requested when creating an avatar, and we rather go towards a "type" of body, with more masculine or feminine features,
- Skin complexion,
- **The face:** the main features of a face can be modified to a greater or lesser extent: mouth, nose, eyes, eyebrows. We can often modify the facial hair and make-up or some marks on the face (wrinkles, freckles, ...),
- Clothing and accessories.

Users can create cartoonish avatars that look like them or what they want. With fairly human features, these avatars offer a realism in interactions that works very well. By adding eye and face tracking technologies, we end up with expressions that are close to reality. This type of avatar is most often found in applications on virtual reality headsets. The choice of a cartoonish avatar therefore makes it easier to integrate it

[3] https://hello.vrchat.com/

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

into the application's environment and to find a uniformity in the experience, as in AltSpaceVR[4] from Microsoft or Horizon[5] from Meta. Moreover, technically and in terms of hardware, a cartoon avatar is much easier to integrate than a realistic avatar, which is why it is preferred even in certain AR applications such as Microsoft's Mesh[6].

Figure 15: Horizon Worlds (Rozier, 2022).

For both default and cartoon avatars, the user is generally represented only by the upper body, as the hardware does not allow the reproduction of lower body movements. However, some applications simulate the lower body to increase the realism a bit.

8.2.4 Realistic Avatars

And finally, the last type: realistic avatars. A realistic avatar is an avatar that represents the user with high fidelity. Among the solutions to create a realistic avatar, there are more or less simple solutions to represent the user. For example, Spatial.io[7] uses a photo of the user which it then integrates into a 3D avatar, a technique that remains rather limited for the fidelity of facial expressions. Other scanning techniques are used, such as a 3D scan through photogrammetry of the head or the whole body (Reimat et al., 2021), Microsoft's Holoportation[8], or a dynamic point cloud capture via Kinect (Reimat et al., 2022). RGBD capture is developing and related works (Gunkel et al., 2023) contribute to pave its path.

According to the studies of Pakanen et al, 2022, users would prefer a full photorealistic avatar, both in AR and VR. A realistic representation in real time would enhance the

[4] https://www.theverge.com/2023/1/21/23565188/altspace-vr-shutting-down-microsoft-layoffs

[5] https://www.meta.com/fr-fr/experiences/2532035600194083/

[6] https://adoption.microsoft.com/fr-fr/microsoft-mesh/

[7] https://www.spatial.io/

[8] https://www.microsoft.com/en-us/research/project/holoportation-3/

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

sense of presence of the person in front and thus improve comfort in social interactions.

However, building users' realistic representation often implies some hardware setup issues as a complex equipment is required and not available to all users. Some of these installations are to be seen in Starline Project and Mediascape XR (Lawrence, 2021; Reimat et al., 2022).

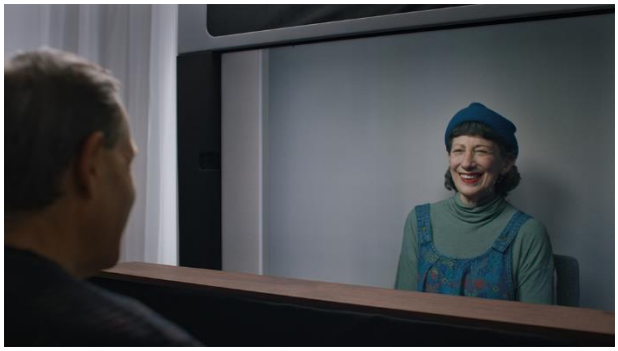


Figure 16: Starline Project (Bavor, 2021).

For realistic avatars, with the complexity of the systems in place, full body capture is largely possible and therefore users are often represented in their entirety. Freedom of movement depends on the devices used, in the case of a device similar to Mediascape XR (Reimat et al., 2022), the solution allows the user to move, but in solutions like Project Starline (Lawrence, 2021), the user must remain in front of the device.

To overcome the complexity of the set ups while ensuring a realistic approach, Krome et Kopp, 2023, propose an AI-based gesture generation relying on speech to allow access from desktop PCs or smartphones. Produced gestures must be convincing and believable, hence real-time. They set requirements for a co-speech gesture generation to achieve online interaction. The evaluation reveals a promising approach.

8.2.5 Existing Social XR Apps

Existing social XR applications can be divided into two categories:

- Public apps which are accessible to all users. They allow for joining virtual worlds created by other users to interact, play, and socialize with people from all around the world. This is getting close to the metaverses definition.
- Private apps which are accessible to a group of invited users. They allow for small group to join a private space to communicate, share content and collaborate with other users.

As some applications are multi-platform, the functionalities may differ depending on the hardware used. Particularly with VR/AR headsets, some headsets allow more or

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

less precise tracking of the user (hand tracking, upper body tracking, etc.), allowing new types of interaction.

We listed the main existing social XR apps in the following table:

Table 10: List of the main existing social XR apps.

Application	Goal	Virtuality	Interactions
VRChat, 2014	Creation of worlds in which one can interact with other users.	VR / desktop mode	Public, Text chat, voice chat, emotes (gestures, expressions), emoji, face tracking (lip and eyes),
RecRoom, 2016	Join or create rooms to interact and play with other users.	VR / desktop mode	Public, Text chat, voice chat, emotes (gestures and face expressions)
Altspace, 2015 (shut down March 2023)	Organized into worlds created by developers or individual players to meet and play with other users.	VR / web navigator	Public, Voice chat, text chat (only in private message), emoji
OrbusVR, 2019	An open world MMORPG where players can battle or join forces to visit this world.	VR (room scale)	Public, Voice chat, text chat (only in private message)
Where Thought Go, 2018	Wander from chapter to chapter and discover users recorded voice responses to intimate/secret questions and record an answer yourself.	VR	Public, Hand tracking, no direct interaction, only voice recorded messages
SteamVR Home, 2017	Home on the SteamVR loader that can be opened to other users to interact with them.	VR	Public, Voice chat

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

High Fidelity, 2016	Create, deploy and visit virtual worlds with other users.	VR	Public, Voice chat, Text chat, emote, face tracking
VTime, 2015	To meet, share content or chat with users from all over the world.	VR / mobile	Private, Voice chat, text chat (private message), emote, face tracking
Bigscreen, 2016	A place to hang out with friends, collaborate with coworkers, play games, watch movies ...	VR	Private, Voice chat
Holoportation, Project ongoing, techno more than app	Real-time holographic communication.	AR	Private, Voice chat
Mozilla Hubs, 2018	Create 3D spaces and collaborate by sharing images, videos, 3D models, ..., up to 25 users.	VR / navigateur web	Private, Voice chat, text chat, emoji
Microsoft Mesh, 2022 (preversion)	Collaboration platform in the form of private rooms, to share content and chat with other users. Allows mixing local and remote users.	AR	Private, Voice chat
Horizon Worlds, 2021	Create and join virtual worlds to play, chat, with other users. You can also attend live events and concerts.	VR	Public, Voice chat, text chat (friend only, via oculus menu), emoji
Horizon Workrooms, 2021	Remote meeting application, to collaborate with colleagues in a virtual meeting room.	VR / navigateur web	Private, Voice chat, text chat (only in browser), emoji

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

However, those solutions are not suitable for our media needs thus in TRANSMIXR project we will explore our own solution.

8.2.6 Reproduce Multimodal Interaction

There are many papers about multimodal human-machine interaction. However, only a few of them focus on human-to-human multimodal interaction in XR. (Poller et al., 2021; Schütt et al., 2019) provided works about semantic and physical interaction (pointing, gaze) with virtual objects in VR and with real objects in AR through 3D mapping from HoloLens.

(Rakkolainen et al., 2021) identified core multimodal interaction techniques which are quite mature and already taking integrant part of XR developments, in opposition to the emerging and future multimodal UI prospects:

Table 11: List of core multimodal interaction techniques according to (Rakkolainen et al., 2021).

Modality	How it is used	Existing solutions
Auditory interfaces	Many existing forms: ambient, directional, musical, speech, sonic. Simulate sounds from the virtual space with its proper characteristics. Act as element of interaction. Serve as augmentation (warnings). Gaining efficiency in traditional interactions such as button click.	Most XR solutions utilize headphones. There are bone conduction headsets. Spatialization of the sound. Sonification: nonspeech sound to render complex data. Auditory icons and earcons: elements of interface. Music: media for emotion mostly but can be used for communicating information also.
Speech	Speech recognition (speech to text), speech synthesis (text to speech), speaker recognition and identification and emotion recognition from speech.	Most AR and VR headsets include audio input and all have audio output. Voice assistants: Apple's Siri, Google Assistant, Microsoft's Cortana, Amazon's Alexa. AI.

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

	Voice input allows both hands and eyes free usage. Voice input is efficient and expressive. Especially powerful when communicate abstract concepts and relations. To be combined with gesture to communicate about direction, distances, spatial relations. Speech output is particularly useful in XR due to text rendering challenge and eyes being focused on other tasks. Challenges: noisy environments and error management which is critical.	Robust speech recognition and sophisticated dialogue modelling techniques. As it requires data-intensive and substantial development efforts, it works better with widely spoken languages.
Gesture recognition technologies	Widely used is HCI, XR, security, video surveillance, film / TV / game studios, etc. Hand, body and head gestures can be useful in various contexts such as deaf sign language. User and object tracking is essential for XR. Motion capture technologies for real-time tracking. Besides, hand and finger tracking are important in human gesture recognition application as an essential part of human natural communication. It is challenging due to their 27 degrees of freedom, fast and precise movements.	Tracking software and hardware improved a lot. There are many position and orientation trackers (sensor fusion, inertial measurement unit, optical / magnetic / acoustic tracking). Modern HMDs embed many sensors (position, orientation, motion, and gesture tracking) for gesture / hand tracking e.g. HoloLens 2, Magic Leap 1, Oculus Quest 2, DecaGear. External devices (e.g., HTC Vice base stations) can be used to improve accuracy. Hand-held controllers, data gloves or full-body VR suits

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

	Gestural interaction in XR can be split into mid-air gestures, gestures with hand-held devices, and touch based gestures.	can also track the user's movements and possibly give some tactile feedback in addition. Some advanced hand-held controllers (Valve Knuckles) include grip force sensor and finger tracking. Motion capture solutions can provide high-resolution data with good data rate. They require calibration and body-worn equipment. Kinect depth camera and Ultraleap tracker popularized gesture. Depth sensors is used on many technologies. Some academic and commercial superwide field of view are emerging. Computer vision methods can track arms, hands, and fingers. Deep learning methods are promising for hand pose estimation.
Locomotion interfaces	Enable the user to feel the movement while moving within the virtual space. In opposition to existing techniques that often deal with motion sickness such as teleportation, astral body projection, blinking etc. Locomotion by walking supports user's spatial understanding and enhance immersion.	Flight simulator, pneumatic platform, treadmill, … Specific VR motion platforms: Cybershoes (Oculus Quest), Virtusphere VR hamster ball, Kat walk, Infinadeck, Omnideck, Stewart platform, 3dRudder, Audi Holoride. Hip based tracker for better body tracking, easier inverse

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

		kinematics, reduced motion sickness.
Gaze	Eye tracking requires various sensor technologies to identify the gaze locations and their sequences. Eye video analyzing (video-oculography), sensors touching the skin close to eyes (electro-oculography). Dual use that the system needs to distinguish: identify users' interests (perception) and provide specific commands (control).	Tobii range of solutions for VR, screens etc. Tobii and Varjo are providing integration and analyzing tools for gaze data in VR. HMDs including eye tracking sensors (e.g., HTC Vive Pro Eye, Fove, Magic Leap 1, Varjo, HP Reverb G2, Pico Neo2 Eye). A common method to differentiate viewing and gaze-control is inducing a brief delay ("dwell time"). There are also blink and wink detection. It can be combined with other control methods (gesture, audio) to activate selection. Other common methods to use gaze are rapid sequence of eye movements (gaze gesture) and continuous movement of gaze while tracking a specific target (smooth pursuit).
Haptics	Integral part of human-real world interaction and therefore important interface in human-machine interaction. Tactile and kinesthetic feedbacks that can greatly improve immersion, performance, and quality of the interaction experience in XR.	The technology is still limited. Global vibrotactile feedback: hand-held controller, gloves, bandages, full-body suits, HMDs. Promising technologies: - non-contact interaction (ultrasound / pneumatic)

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

	- surface-based interaction such as smartphone or projection (vibrotactile feedback) - wearable device interaction (force / tactile feedback) - Multi-device and full-body interaction (force / tactile feedback) - dynamic physical environments (physical forces from real objects)

To reproduce those human-human interaction within XR environments, some authors highlight the necessity to treat the entire XR continuum with fluid transitions from one state to another. Consistency of behaviours and interaction techniques should reduce cognitive workload and learning effort (Riegler et al., 2020).

On the other hand, others argue the need for new interaction techniques designed especially for XR (McVeigh-Schultz et al., 2021; Vasilchenko et al., 2020). Some related works focused on the transformation of nonverbal behaviours during interaction such as Proteus effect which shows altered behaviour due to transformed self-representation (McVeigh-Schultz & Isbister, 2021; Yee & Bailenson, 2007).

Indeed, usual interaction rituals are disrupted within virtual environment. For example, in VR, sudden moves (teleportation, fly) and rapid avatar changes interrupt interpersonal interactions.

8.2.7 Going Beyond Traditional Mechanisms

Instead of only aiming to mimic and reproduce the physical world and its mechanisms, XR systems can also propose new types of visualisations and interaction to support communication and social exchanges.

XR designers explored other ways of moving within virtual environments, like teleportation. This kind of “superpowers”, often mediated with controllers, is specific to XR and not reproductible in our real environment. Besides, it also (at least partially) solves or limits cybersickness issues.

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

In their review, (Rakkolainen et al., 2021) make a list of emerging and future multimodal UI prospects : facial expression interfaces, scent, taste, exhalation interfaces, tongue interfaces, brain-computer interfaces (BCI), biometric signals interfaces, but there are not mature yet.

(McVeigh-Schultz & Isbister, 2021) take a more radical point of view and aim to create new types of social coordination with new embodied capabilities, new social artefacts, and new environment functionalities. During their research they analysed new ways of expression that appeared in VR. Affection proves, attention geometry and nonverbal emotional reactions indeed change a lot. Visual and sound feedback act as principal cues. For example: kissing sounds express greetings, “feeding” another avatar involves appreciation and sharing, “mirror cuddling” shows affection by allowing users to have visual and auditory feedback instead of usual haptic one, emojis illustrate phatic reaction of the users, meaning being used for general purposes of social interaction rather than to convey specific information.

Social XR also involves new geometries of attention. In VR for example, the difficulty of not seeing the speaker can be fixed with artefacts (e.g., a cat sitting on the side of the speaker). XR brings the opportunity to mediate the interactions using a visualisation of behaviours not directly visible in the real world such as gaze and speech duration.

8.3 Opportunities & Obstacles

8.3.1 Opportunities

8.3.1.1 Empathy Machine

According to (Milk, 2015; Paananen et al., 2022) virtual environments can be considered as “the ultimate empathy machine”. Indeed, XR offer the opportunity to create new experiences and allow a more immersive empathy by embodied users into a reality of which they can experience the issues. It augments the environment with information the users could not perceive otherwise.

This embodiment has a strong psychological impact on the users, especially through the manipulation of themselves and others (Mcgill & Khamis, 2021)

XR therefore foster the capability for a user to identify themselves to another by feeling the same experience. This is empathy.

8.3.1.2 Enhanced Understanding

XR technologies allow for supporting both physical and cognitive aspects of learning. They increase user motivation, content understanding and long-term memory

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

retention. Moreover, social XR enhance this potential through its collaborative characteristic which strong social interactivity benefits learning (Vasilchenko et al., 2020).

Beyond pedagogical purpose, XR also ease understanding in research and industry fields allowing a more explicit content. It can be an asset in Big Data understanding regarding their spatial and quantitative characteristics to create visual patterns (Riegler et al., 2020). Besides, by the augmentation of physical space with other users' and environments' information, XR technologies provide a more rapid and explicit communication than traditional 2D UIs (Paananen et al., 2022).

8.3.1.3 A logical follow-up

Social XR can allow sharing abroad while limiting physical contacts, in the continuity of videoconference's solutions, habits that have drastically increased due to COVID19 crisis. (McVeigh-Schultz & Isbister, 2021).

New emerging technologies such as 5G offer stable and necessary network to allow the large number of data required by XR. New use cases are expected to develop then (i.e. European Horizon project Evolved5G).

Social XR may be used in various fields already raised in this document. Another example is dating applications which have been explored in related work since 2008 (Shanker & Zytko, 2022). As XR technologies could reduce some dangers allowing people to better know each other's before meeting in real life, it might also raise new contexts of harms.

8.3.2 Obstacles

8.3.2.1 Ethical Challenges

XR technologies also bring new ethical issues:

- Increasing amount of collected personal data (facial mimics, speech, retinal pattern, etc.) (Shanker & Zytko, 2022; Vasilchenko et al., 2020; Wang, 2020).
- Besides, creating immersive experiences for sensibilization might be intrusive and seeing as exploitation of vulnerable populations (Paananen et al., 2022). XR increase data privacy challenges.
- A frequent use of XR might be harmful for mental health and social isolation (Vasilchenko et al., 2020).
- New harassment risks regarding verbal and physical fierceness on avatars and a potential rise of violence within a virtual environment (McVeigh-Schultz & Isbister, 2021; Shanker & Zytko, 2022; Wang, 2020).

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

The society will need an acceptance and adaptation time for those technologies. Population's lack of knowledge on these technologies can make it anxious about their effects. Legislative, legal and societal preconditions must be created to protect users and their data (Görlich, 2022; Shanker & Zytko, 2022; Vasilchenko et al., 2020; Wang, 2020).

Potential abuses also raise the question of moderation tools within those new environments. (Mcgill & Khamis, 2021) underline the tension between safety and privacy of users. They propose two types of moderation: an allowlist and control (more intrusive) or an automatic journal of key sensitive events to offer support (don't prevent aberrant behaviours).

(Görlich, 2022) highlights the weak accessibility of social XR by opposing the term of "Societal XR". Indeed, actual social XR often requires body worn devices which is not accessible to elderlies, young children nor people with disabilities.

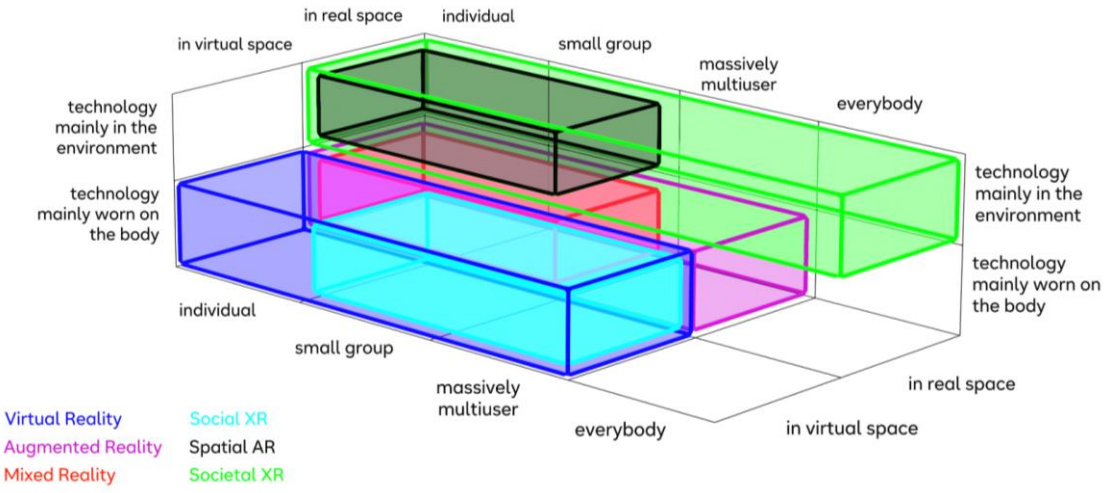


Figure 17: Three-dimensional categorization of XR paradigms including the vision of Societal XR proposed by (Görlich, 2022).

Finally, due to the drastic change of paradigm in XR environment, users usually have more difficulties to transform their intentions into actions with XR UIs than traditional 2D UIs (Paananen et al., 2022).

8.3.2.2 Technical Limitations

XR technologies have been used for 28 years in the industries with engineers and experts. It is used in daily task to share 3D mockup of products or working area. It is already useful and offered ROI to industries in the world. But XR technologies are also not fully mature yet and might slow down the massive deployment of Social XR. AR has a tight Field-Of-View (FOV) and poor display contrasts ; Virtual reality is isolating and lacks haptic feedback (Riegler et al., 2020). They are not adapted to public spaces

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.

such as outdoor or subway (Vasilchenko et al., 2020; Wang, 2020). Besides, they are still expensive and cumbersome.

Another issue is the data transmission delay as XR requires reliable and efficient network (e.g. deep learning, facial and hand recognition, 3D mesh of the scene …) (Wang, 2020).

This state of the art proves the high potential of XR technologies and Social XR. As technologies are growing fast, many new possibilities are to come soon.

This project has been funded by the European Union as part of the Horizon Europe Framework Program (HORIZON), under the grant agreement 101070109.