Deliverable

Project Acronym:	**IMAC**
Grant Agreement number:	761974
Project Title:	*Immersive Accessibility*

D2.1. User Centered Design

Revision: 1.0
Authors: Anna Matamala & Pilar Orero (UAB)
Delivery date: M3

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement 761974		
Dissemination Level		
P	Public	x
C	Confidential, only for members of the consortium and the Commission Services	

Abstract:
D2.1 documents all activities and their results in the development of ImAc user scenarios and provides detailed use cases for the selected scenarios

REVISION HISTORY

Revision	Date	Author	Organisation	Description
0.1	22-11-2017	Pilar Orero	UAB	Template
0.2	27-11-2017	Anna Matamala	UAB	Template and some sections
0.3	02-02-2017	Francesc Mas	CCMA	Revision and added one sub-section
0.4	02-02-2017	Pilar Orero	UAB	Revision
0.5	04-12-2017	All	All	Consolidation
0.6	09-12-2017	Pilar Orero	UAB	Adding RBB info
0.7	14-12-2017	Pilar Orero	UAB	Adding RNIB & USAL info
0.8	18-12-2017	Enric Torres	ANGLATECNIC	Revised version with comments
1.0	19-01-2018	Pilar Orero	UAB	Final version

EXECUTIVE SUMMARY

ImAc follows a user centered design approach. The project developments must be driven by real user needs, continuously involving these users in every step of the design and implementation. Both professionals and consumers will be consulted on the definition of the solutions as an integral and intuitive part of any immersive content within the convergent media offerings. ImAc will follow these user requirements as they will be further elaborated by the partners involved in the project. The developed solutions will be piloted and validated in public trials, ensuring market relevance and user acceptance. Developed tools will meet the requirements of experienced broadcasters from the start, and tools and services will be thoroughly tested through pilot operations.

This document describes the different stages undertaken during the first three months of the project to define end users, and their interaction with technology giving raise to the user scenarios. These scenarios classified by technological component and user interaction are the foundation of the ImAc roadmap. The document describes the project testing workflow, it identifies the technologies to be tested, end user classification, user testing approaches, and finally the results from the first focus group.

The information from this document was fed to D2.2 which in turn will be fed in D2.3. All these deliverables will be taken into consideration for WP3 and WP4 at the time of development, and will help when setting up WP5 in the D5.1

CONTRIBUTORS

First Name	Last Name	Company	e-Mail
Pilar	Orero	UAB	pilar.orero@uab.cat
Anna	Matamala	UAB	anna.matamala@uab.cat
Doreen	Ritter	Rbb	3nrich.ritter@rbb-online.de
Sonali	Rai	RNIB	sonali.rai@rnib.org.uk
Chris	Hughes	USAL	c.j.hughes@salford.ac.uk
Enric	Torres	ANGLATECNIC	3nrich@anglatecnic.com
Francesc	Mas	CCMA	fmas.z@ccma.cat
Anita	Fidyka	UAB	Anita.fidyka@uab.cat
Belén	Agulló	UAB	Belen.agullo@uab.cat

CONTENTS

TABLES OF FIGURES AND TABLES

LIST OF ACRONYMS

Acronym	Description
AD	Audio Description
AST	Audio Subtitling
CRPD	Convention of Rights of Persons with Disabilities
DTV4ALL	Digital Television for All
HbbTV	Hybrid Broadcast Broadband TV
HBB4ALL	Hybrid Broadcast Broadband Television for All
HoH	Hard of Hearing
HMD	Head-mounted Display
VDPC	Video Disk Control Protocol
VR	Virtual Reality

1. INTRODUCTION

This deliverable documents in detail the end user needs before deciding the approach to develop the technology (WP3 and WP4) and before deciding and defining the scenarios for piloting the services (WP5). Access services in immersive environments may have different behaviours depending on the viewing context, the viewing experience, and the UX with the service itself. Subtitles may be "burnt" or located according to the viewers field of vision, or placed in determined regions within the visual content. Audio description would have to be produced depending on the information/directionality given by the sound, the end user requirements, and the way the audio description is received by the end user: as an atmospheric sound or with an ear piece. There are many technical solutions and it is important to understand those who will allow for mainstreaming accessibility.

Working and testing with persons with disabilities requires specific considerations and a qualitative research approach is the best suited to address the different needs and expectations. Still, organising focus groups with persons with disabilities requires additional preparation [1] [2] with smaller user participation [3] [4]. The reason for this is the time needed to inform everyone [6] and to allow time for presenting the questions and for participants to process information and questions. Also, more time is allowed for participants to express their opinions and thoughts. Balancing time is crucial since participants might experience fatigue [3]. Also, presenting stimulus for persons with disabilities in a focus group is crucial and requires time in preparation [5]. In the case of accessible services for VR –where there is no available solution at hand- is even more challenging. For this reason, during the

focus group session different approaches were taken to deal with available materials and the language of these materials.

This document details the different stages and efforts taken to define end users, and avoiding "professional informers" [6] while looking for industry professionals: a fine balancing act. Given the lack of stimuli and real life simulation a great effort was needed to prepare meaningful focus groups and think of possible scenarios to fulfil a truly user centered methodological approach. These departing scenarios are setting the needs for future deliverables D2.2 and D2.3 and ImAc roadmap. The document describes the project testing workflow, it identifies the technologies to be tested, end user classification, user testing approaches, and finally the results from the first focus groups.

2. USER CENTERED APPROACH IN IMAC

Immersive environments (such as virtual reality and omnidirectional video) are often assumed to be of little or no interest for persons with disabilities. This position reveals a lack of knowledge of the Human Rights obligations (CRPD[1]) ratified by all EU countries, as well as overlooking the versatility of access service and the resourcefulness and creativity of end-users. Immersive environments exist and they should be accessible for all. During ImAc some test cases will be deployed, showing the available accessibility possibilities and offering real life solutions.

During many previously EU funded projects, accessibility services have been tested to adapt to new technologies such as the switch from analogue to digital broadcast (DTV4ALL) and the new connected media ecosystem (HBB4ALL). As this last project has demonstrated, it is of great interest to all stakeholders within the value chain, that accessibility services are tested while defining and developing the new technology. HBB4ALL started with HbbTV1.0 and finished with HbbTV1.5, enabling the project results to be implemented and providing recommendations for future development in HbbTV 2.0. Impact from the project contributed to standardization bodies, creating an exchange of information and common understanding to meet the needs of the manufacturers. The successful deployment of accessibility services by national broadcasters has also proven that it is possible a common shift in attitudes regarding the importance of accessibility services. ImAc is seeking the same success story within the new immersive environment. This time departing from HBB4ALL and adding the expertise from across the ImAc partners in order to achieve a similar situation. As with HBB4ALL the end-user will be at the centre of the project, following a user-centered design. The user being both the audience but also the broadcasters or content developer.

ImAc will follow a bottom-up approach, following a user-centered methodology. The interdisciplinary team has been chosen from complementary partners with a common aim: **making immersive environments accessible for all**. ImAc groups leaders of previous EU funded projects in order to secure an organised and efficient use of resources. Partners have also been chosen to complement their expertise in the value chain, from development to production, distribution and exploitation. The project has an important Social Science element, looking at innovation and impact beyond technical solutions towards a wider social spectrum. ImAc believes in the role played in society by technology, and how good technical solutions can offer massive benefits in terms of social inclusion and cohesion.

[1] http://www.un.org/disabilities/convention/conventionfull.shtml

While media companies are still experimenting a lot with new user experiences, ImAc focuses on the development of strategies and tools for enhancing accessibility services. Particularly investigating the characteristics of 360° experiences. During the project the development of new trends in the immersive media market will be monitored, as this will also affect the corresponding accessibility services. If required, the ImAc project will modify its work plan in line with industry to ensure that the final results fit with the market situation.

ImAc aims at developing and integrating and end to end system to create, distribute and display accessible content as part of omnidirectional video and virtual reality. ImAc development can be grouped in four main sectors where consultation with end users is required: Content Management, Content Production, Content Delivery, and Presentation.

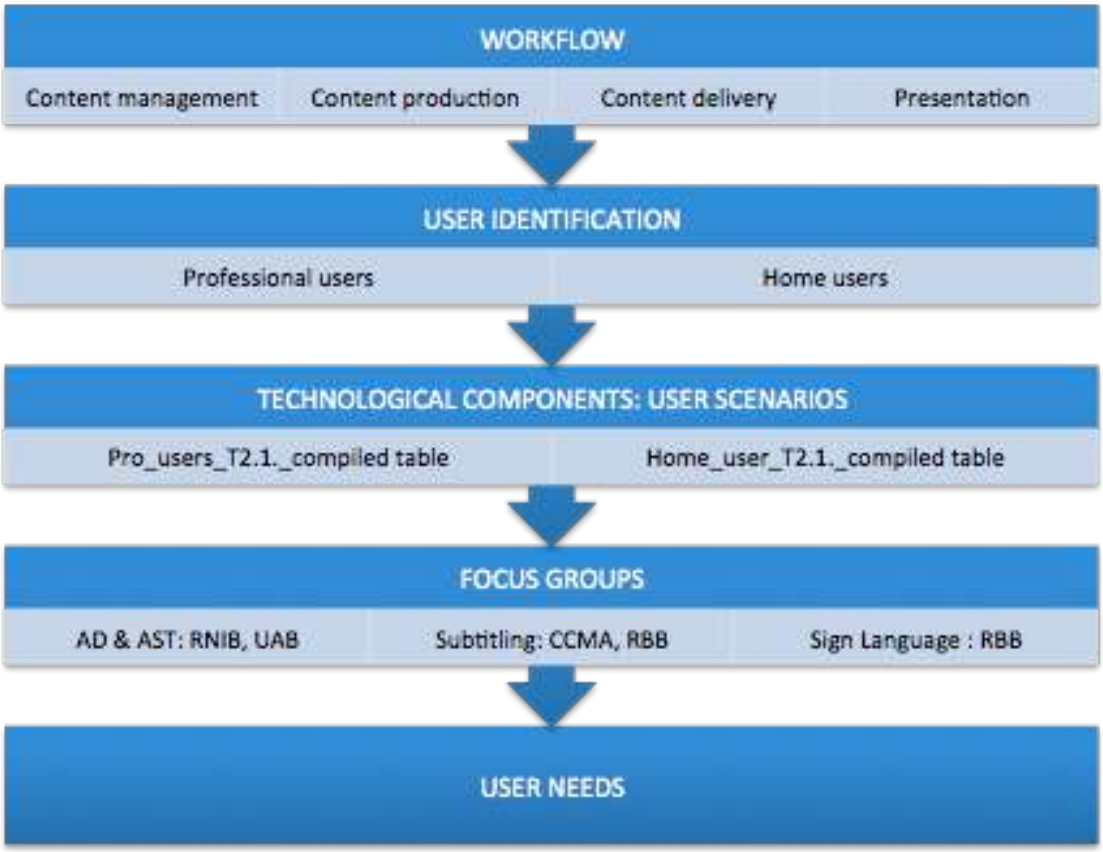

Figure 1 - ImAc user-centered workflow

2.1. Workflow

First of all, in ImAc we defined the four areas where development will take place longitudinally in the life of the project. These areas are the following.

2.1.1. Content Management

In order to handle the storage requirements of the accessibility content and data ImAc will need to deliver a bespoke Content Manager, made available via IP to both direct users and external systems. By hosting the content on a central file system, it will provide control such as

logging and access permissions to the content and related data. A web interface will provide users with a mechanism to interact with the content directly (such as a subtitler who needs to upload the content ahead of broadcast) and a web service will provide automated access to processes (such as streaming platforms and playout systems). The content manager will also make it possible to automate tasks centrally such as notifications and make use of protocols used through the broadcast industry such as Video Disk Control Protocol (VDCP). The content manager will provide an essential backbone to the ImAc platform.

2.1.2. Content production

As broadcasters increasingly integrate web distribution channels, new requirements have to be considered in content production. Additional information and metadata about the content is needed to know how to feed each distribution channel simultaneously. ImAc will address this need, firstly with the specification of suitable formats, and secondly with editor and conversion tools that allow the production of accessibility content for all distribution channels and user platforms.

New formats for subtitle and audio description services will be required, since the current standards only cater for traditional TV broadcast. ImAc will develop solutions that allow for the production and distribution of accessibility content across various channels with the same or minimal extra resources needed on the editor side.

Nowadays, production workflows are designed to supply TV broadcast distribution. ImAc will push towards an inclusion of new services in the production chain. It is important that new services integrate in the current accessibility service workflows to achieve a successful deployment. Format conversion tools developed in ImAc can help when adapting and integrating with existing broadcast systems. Both the developed production tools and the specified formats will meet the requirements defined within the project in order to support the developed services. This follows the understanding that content should be produced only once for economic efficiency. Therefore, all new formats for accessibility data must include enough metadata to supply all distribution channels - new services as well as traditional broadcast services. ImAc will push for rich content annotation during production processes and will provide the tools to do so. For all formats, a very close relation to existing standards is aspired, as it enhances system interoperability and dependability. Both are crucial aspects in broadcasting.

ImAc will evolve production solutions that do not establish additional workflows but instead include enhanced accessibility services for existing system environments. Attractive use cases can then be a driver to develop more enhanced production tools that will support the authoring of enriched content. That content can provide additional value to enhanced platforms while simultaneously serving traditional broadcast playout.

2.1.3. Content delivery

Developments in the project will strongly rely on existing technology and standards, and especially on web technologies which are a central enabler when realizing new user experiences. This is due to high flexibility, the fast-developing market and the cross-platform design of web technologies. Internet streaming has now become a main distribution channel for media content.

Delivering new content types on new platforms require new formats. While audio and video have been in production, accessibility and immersive aspects require extensions to the existing workflows and technologies. This need for new formats is an opportunity to create relevant

requirements for standards and for convergence in formats between the production and the end-user delivery which, for historical reasons, all use different formats. The approach of ImAc is to leverage the latest standards to cover these new use-cases. ImAc will address the openness of the solution by contributing to the standards and will extend or create tools so that the production content is delivered accurately to the end-users, whatever the intended playout.

2.1.4. Presentation

Inclusive TV Content (e.g. with subtitles, audio description, sign language) is usually consumed with a traditional television and via second screen displays. Displaying immersive (virtual or omnidirectional) content has been achieved using players built for specific platforms, mostly Android (Samsung Gear VR, Android DayDream), Windows (HTC Vive, Oculus Rift) or with web players (Facebook, Youtube). ImAc will provide a step forward integrating both experiences from traditional TV viewing and adapted content display. This project will use omnidirectional video enriched with novel techniques for inclusive audiovisual content to deliver immersive content, matching the demands of persons with disabilities when using immersive displays which can be consumed through tablet, mobiles, HMD and traditional TV. A new layer with content adapted to each type of user need (blind, low-sighted, deaf, low hearing, elderly, etc) will provide a personalised experience. For this purpose, we will integrate additional layers into immersive experiences. ImAc will base its developments on ImmersiaTV's player, where it is possible to render an omnidirectional video stream overlapping with complementary and independent augmented layers (including AD), guided with recommendations from the platform itself based upon the user's requirements or by selecting from default presets streams. Audio content (enhanced audio services) will add an additional challenge, as never before have immersive environments been created for blind people. Position based audio fragments and audio indications are the starting point for this innovation.

2.2. User identification

From the four identified areas, it was decided to define who will be the user in the different iteration focus groups, and in the pilots in WP5. In the national pilots (WP5) it was specified in the proposal that at least 80 users will be tested. From previous EU funded projects DTV4ALL and HBB4ALL, where there was an important effort dedicated to testing component, it was easy to define who will be the end user to be consulted in WP5, and how they will be recruited. In ImAc we have a partner RNIB who will not only help out recruiting, but also advising on methodology and accessibility issues when end user interaction is required.
For the focus groups around 10 persons should be consulted. But these 10-people had to be defined further and differentiated from those in the pilots. To this aim we created two different profiles: that of the professional users and that of the advanced home users.

- Professional users included: IT, graphic designer, subtitler, audio describers, sign language interpreters.

- Advanced home users referred, for instance, to deaf, hard-of-hearing, blind, low vision users, the elderly.

It was agreed that the end users to be included in focus group would be experienced or advanced end-users. This meant that, besides their condition of regular lay users, they would

also have some knowledge on the technologies that will be developed in the project. In the focus groups there is a fluent conversation to advance their expectations to match the innovation. It would make no sense to consult end users with no knowledge or experience with neither functional diversity nor technological background since at this stage what we require is not their acceptance of the final service, but issues related to technology development.
Some considerations when drafting the information and designing user requirements were related to the following areas:
While some technological components are language agnostic, others related to user interfaces will be decisive to tag at this stage the language dependency. For this project we have the following language classification, which also defines the end user:

1. Written languages:
    a. Catalan
    b. German
    c. Spanish
    d. English

2. Oral languages:
    a. Catalan Sign Language
    b. German Sign Language
    c. Spanish Sign Language

Hence language definition at this stage is considered to be crucial to understand end user group for testing in the future, the services to be tested and the final (WP5) pilots.
The choice of participants is also dependent of the service for which they will be tested:

- Audio description
- Audio subtitling
- Sign Language
- Subtitling

The end user should also be defined for their sensorial functionality:

- deaf,
- hard of hearing
- blind,
- low vision,

And finally age could also be a factor, since it also often linked to some degree of hearing or vision loss, hence the elderly were also considered.
Further from these considerations, it was crucial to understand from the technological components the type of end user required, which is explained in the next section.

2.3. Technological components: user scenarios

Once we had identified two different end users (advanced home users and professional users), it was required to fine tune end user profiling further and link it to user experience. One table (see Table 1) was created from the previous stage. In this table partners had to identify the technological components to be developed for each WP system side, the users that would be interacting with the component: professional or (advanced home) end-users. And the user scenario they envisaged, i.e. what the already identified user would be experiencing and how.

System side	Technical Component	User: professional user or advanced home user	User scenario: indicate what the previous user will experience and how	input from
Player for mobile phone, TV, head mounted display	ImAc player	Advanced home user	The user will enjoy the experience in the tv, tablet or head mounted display or any combination of these devices. The experience will be synchronized across devices. The player will be available as a web application, so the user won't need to download anything. The tv will start the show, the complementary 360 video is distributed synchronized to the main show, the other devices will show new Audio description and subtitle services adapted to each user impairment The user will access to the contents published in the server and will enjoy an experience adapted to the device where he/she is consuming the content. The user will adapt the device that will store his/her preferences	i2CAT

Table 1 - Sample of table structure

Partners contributed with their expertise to identify technological components, users and related user scenarios, that is, what the already identified users will experience and how for each system side and technical component. Each partner uploaded one document with its input and a first version of the compiled table was created (see Annex I)
The table was discussed in a dedicated meeting. To make the discussion more productive, the previous table was divided into two compiled tables depending on the user profile. The meeting was used to further refine the tables, to which partners added extra input in a second revision.
The revised compiled tables can be found in Annex II for professional users and in Annex III for advanced home users

2.4. Focus Groups: understanding user needs

Once we had the two groups of end users (advanced home and professionals) the technology they will be testing, and the user scenarios, there was the need to define:

- which partner will conduct the focus groups, with

- which end users and how to interact with them, and
- the procedure.

The aim of the focus group was to gather information about user needs, expectations and wishes regarding the various technological components to be defined.
Various methodological tools such as interviews, questionnaires or focus groups can be used to gather user feed-back, but at this stage of the project focus group with experienced users was favoured.

2.4.1. Focus group: partner and service distribution

It was decided that focus groups will be conducted replicating the division made for piloting in WP5. It would make no sense for example for RBB to do a focus group on audio description since they will not pilot the service in WP5.
To this aim the division is by services:

Service	Partner conducting FG	Participating partners
Audio description	RNIB (English) UAB (Catalan)	ANGLATECNIC, CCMA, IRT, RNIB, UAB
Audio subtitling	RNIB (English) UAB (Catalan)	ANGLATECNIC, CCMA, IRT, RNIB, UAB
Sign Language	RBB (German Sign Language)	RBB
Subtitling	RBB (German) CCMA (Catalan)	ANGLATECNIC, CCMA, IRT, RBB, UAB

Table 2 - Focus group by service and participating partner

Beyond the scope of the project, a focus group was also conducted in Poland for audio description and audio subtitling by UAB. Although not initially included in the proposal, it was considered that involving users from a voice-over country such as Poland could yield additional results.

2.4.2. Focus group: end user profiling

Concerning end users, it was agreed that focus group would gather input from both advanced home users and professional users.
Profiling end users was considered necessary, and a specific short questionnaire to be used during the project was developed and translated into the end user languages. The questionnaire includes some basic questions, validated with end users, that can be expanded in future tests but which offer some shared basic information about the users.
Questionnaire was produced first in English (to be found in Annex IV) and translated into German and Catalan.

2.4.3. Focus group: procedure

ImAc produced a shared common procedure for Focus Groups administration and management. A special form was created with general instructions to all partners in English managing a Focus Group to be found in Annex V.
All partners had to write down a procedure following a specific model to be found in Annex VII.

The focus group development was as follows:

1. Welcome participants
2. Information sheet and consent forms to be read and signed
3. Demographic information: short questionnaire administered
4. Group discussion following questions from compiled tables (see Annex VI)
5. Summary of focus group conclusions, to be approved by all participants.

At all times in ImAc testing, piloting or consulting end users follow the Ethical Considerations drafted in a separate document D.1.2 Ethical Considerations and Data Protection Management to be found also in Annex VII

3. RESULTS: USER NEEDS

After the first three months of the project the first action with end users was in place to find out their interests and expectations from the technology and services to be developed in ImAc. The results from this activity, obtained through focus groups, are to be found in this section:

3.1. Audio Description and Audio Subtitling

To mirror the piloting in WP5 two different focus groups were formed to gather information regarding audio description and audio subtitling.

3.1.1. Audio Description in Catalonia

A press release of the Focus Group can be consulted here: http://blogs.uab.cat/blogdtieao/

Figure 2 - Participants in UAB Focus Group

The summary is as follows:

Focus group general information

- Partner responsible for the workshop: UAB.
- Place and date: UAB (Bellaterra, Barcelona), 22.11.17
- Access service(s) discussed: AD and AST.

Participants profile

- Number of home users: 2.
- Number of professional users and profile: 4 (3 audio describers, 1 technical expert).
- Demographics for users:
    1. Sex: 4 “male”, 2 “female”, 0 “other”, 0 “prefer not to reply”.
    2. Age: 25, 25, 30, 34, 51, No Reply.
    3. Main language of the participants: 1 Spanish, 1 Sp/Cat, 3 Catalan.
    4. Level of finished studies: 0 “no studies”, 0 “primary education”, 0 “secondary education”, 1 “further education”, 5 “university”.
    5. (Only for home users): “I define myself as a…” 0 “blind person”, 2 “low vision person”, 0 “deaf person”, 0 “hearing impaired person”, 0 “deaf-blind person”.
    6. (Only for home users): Age in which your disability began: 0 “From birth”, 1 “0-4”, 0 “5-12”, 1 “13-20”, 0 “21-40”, 0 “41-60”, 0 “More than 60”.
    7. What technology do you use on a daily basis? You can select more than one. 4 “TV”, 2 “PC”, 5 “Laptop”, 4 “Mobile phone”, 3 “Tablet”.
    8. Do you have any device to access virtual reality content? 0 “Yes”, 6 “No”, = “I don’t know or I don’t want to reply”.
    9. (Only for home users): Which of the following do you use on your connected devices to access the above content? 1 “Magnification”, 1 “Screen reader”.
    10. (Only for home users): Which of the following controls would you like to use with your screen reader/magnification tool when watching content online? 1 “Identify content”, 2 “Functions such as play, stop, pause, forward, rewind”, 2 “Switch AD/AS on and off”.

Summary of participants’ profile

Six users took part in the focus group (2) home users, 3 audio describers and 1 technical expert). They were 4 males and 2 females, with ages ranging 25-51 and one participant not providing this data. Five participants had university studies and 1 had further education studies. None of them reported having a device to access virtual reality content. A laptop was the most used technology by the participants on a daily basis (5), followed by TV and mobile phone (4), tablet (3) and PC (2). The home users were low vision participants (vision impairment started between 0-4 and 13-20, respectively). One used a magnification tool to access content and the other one, a screen reader. Both participants identified functions such as play, stop, pause, forward, rewind and switch AD/AST on and off as needed to watch content online, and one also considered “Identify content” as needed.

Workshop conclusions: approved by focus group

Conclusions from professional users

Professionals think that AD in immersive environments is challenging because there is more information than in a standard non-immersive film. It is challenging to decide what to audio describe, especially when there is not enough time for the AD. Time management will also be different, because a user can spend more time watching a film.

Professionals think that the space should be described first but it would be interesting to offer simultaneous audio descriptions of different sections of the scene. This will increase the number of audio description units, which will not always be activated, and this may raise some financial issues.

Professionals consider that the type of content will determine what is feasible: a virtual space gives more time for audio description but standard films are already challenging, so immersive films will also be even more difficult.

The audio describer needs to have a general view of the scene (plain view) and the possibility to select different sections (minimum of 4, generally 6, maybe more), which can be opened in new windows.
Audio describers think it would be useful to watch the content with glasses first but do not want to work with an immersive editor, just to check the final outcome.
Concerning current software, audio describers indicate that some functions are only used once and could be deleted from the general view (for instance, voice calibration). Others functions are not used by our professional users ("*thermometers*", "check if mike works").

Conclusions from advanced home users

Concerning the interaction: accessing the services
Users in our focus group consider that voice interaction would be better than video commands, especially if the user is in a private environment, which is where normally this product consumption takes place.

Users in focus group consider that it would be useful that the system identifies user preferences and parameters. Some of these preferences could be transferred between devices but not all of them because the user may have different preferences for a smaller or bigger screen.

Regarding accessing immersive content with or without glasses, users indicate that it depends a lot on the home user.

Concerning the services

Advanced home users consider that 360 immersive content is an interesting technology and implementing accessibility is needed. It is an actual improvement to the content.
Advanced home users consider that enlarging images is a requirement.
Advanced home users think that immersive sound could be useful to position yourself and identify where action is happening. Information could be prioritized according to the volume. Information could be given like “headlines” and then, if you are interested, you turn your head to that area and the volume then increases.
Advanced home users favour the AD position linked to action being described. Users want a main audio description of the main action (so that they can follow the plot), but they also want to be able to choose different secondary audio descriptions for additional action, even if this means the film will last longer. They explain the possibility of watching the film different times and choosing different paths. They also indicate the usefulness of having two different voices: one for the main action and one for the secondary actions.
Advanced home users do not want more information than the one received by persons with no visual impairments, so professionals do not think the action happening outside the view of the user should be described unless specific action is taken by the user (moving the head).
Advanced home users indicate the challenge of moving away from the main action and going back to it. They think a specific sound effect (“beep”) could be good to position themselves and know they are back to the action.
Advanced home users express the need for audio subtitling (spoken subtitles), better than zooming in on the subtitles.

3.1.2. Audio Description in UK

This report concerns the accessibility of 360 degree content for audio description, as tested by a focus group of blind and partially sighted people. The aim of the focus group was to gather feedback from regular users of audio description on viewing and interacting with 360 degree content in an immersive environment.

A review of immersive environments outlined four key areas on which feedback was sought during the focus group. On the form of the audio description track: linear, based on the traditional approach where viewer sees what the director wishes to show on the screen and non-linear, wherein view changes when viewer interacts with the environment, and using audio to enhance the immersive experience. Lastly, issues around accessing the 360 degree content were addressed.

In the absence of any audio described 360 degree content, live description was delivered on a clip available on Youtube in a 360 degree format. We wished to explore how a small group of participants would respond to the material.

Focus group general information

- Partner responsible for the workshop: RNIB and USAL.
- Place and date: RNIB (London, UK) 7 December 2017
- Access service discussed: AD.

Participants profile

- Number of home users: 7.
- Number of professional users and profile: 2 (1 audio describer, 1 technical expert).
- Demographics for home users:
    1. Sex: 4 “male”, 3 “female”, 0 “other”, 0 “prefer not to reply”.
    2. Age: 26, 28, 32, 36, 39, 49, 53.
    3. Main language of the participants: 7 English.
    4. Level of finished studies: 0 “no studies”, 0 “primary education”, 0 “secondary education”, 0 “further education”, 7 “university”.
    5. “I define myself as a…” 5 “blind person”, 2 “low vision person”, 0 “deaf person”, 0 “hearing impaired person”, 0 “deaf-blind person”.
    6. Age in which your disability began: 7 “From birth” (they said the level of sight they had, had deteriorated over the years), 0 “0-4”, 0 “5-12”, 0 “13-20”, 0 “21-40”, 0 “41-60”, 0 “More than 60”.
    7. What technology do you use on a daily basis? You can select more than one. 7 “TV”, 1 “PC”, 6 “Laptop”, 7 “Mobile phone”, 6 “Tablet”.
    8. Do you have any device to access virtual reality content? 0 “Yes”, 7 “No”, 0 “I don’t know or I don’t want to reply”.
    9. Which of the following is your preferred device for watching online video content (i.e., Youtube, Vimeo, Netflix, Amazon Prime, broadcast catch up service etc.)? 1 “PC”, 6 “a combination of smartphone and tablet”.
    10. Which assistive technology do you use? 1 “Magnification (i.e. Zoomtext)”, 5 “Screen reader (i.e. Zoomtext)”, 1 “a combination of magnification and screen reader”.
    11. Which of the following controls would you like to use with your screen reader/magnification tool when watching content online? 7 “Identify content”, 7 “Browse content library”, 7 “Functions such as play, stop, pause, forward, rewind”, 7 “Switch AD on and off”.
    12. Which of the following describes what you are able to see? 1 “well enough to recognise a friend at arm’s length”, 1 “well enough to recognise a friend if they got close to his or her face”, 3 “the shapes of the furniture in a room”, 5 “can tell by the light where the windows are”, 2 “cannot see anything at all”.
    13. Which of the following barriers do you encounter when watching TV? 7 “difficulty seeing buttons on the remote control”, 5 “difficulty seeing the picture on the TV screen”, 7 “difficulty seeing the fine detail on the TV screen”, 7 “difficulty seeing text on the TV screen”, 5 “see the light of the TV screen”, 2 “cannot see anything on the TV screen”, 7 “participants find it difficult to follow what is going on, on the screen”.
    14. How do you currently watch or follow a programme or film on television? 1 “uses residual sight to watch”, 1 “sits closer to the TV screen”, 6 [not the preferred option] “ask their friends or family members to assist by explaining what happens on the screen”, 6 [not the preferred option] “try to pick up as much as they can from the sound of the film or programme”, 6 “use audio description to explain to them what happens on the screen”.

Summary of participants' profile

Seven home users took part in the focus group and 2 professional users (1 audio describers and 1 technical expert). The home users were 4 males and 3 females, with ages ranging 26-53. All home users had university studies. None of them reported having a device to access virtual reality content. TV and mobile phone was the most used technology by the home users on a daily basis (7), followed by laptop and tablet (6), and only (1) used PC. 5 home users were blind from birth and 2 had low vision also from birth (the level of sight they had, had deteriorated over the years). Most home users preferred to watch online video content using a combination of smartphone and tablet (6) and only 1 preferred PC. To access content 1 home user used a magnification tool, 5 a screen reader and 1 a combination of magnification and screen reader. All home users identified "Identify content", "Browse content library", "functions such as play, stop, pause, forward, rewind" and "switch AD/AST on and off" as needed to watch content online. 2 home users cannot see anything at all and 5 can tell by the light where the windows are, 3 can see the shapes of the furniture in a room, 1 can recognise a friend if they got close to his or her face and 1 can see well enough to recognise a friend at arm's length. All home users have difficulty seeing buttons on the remote control, seeing the fine detail on the TV screen and find it difficult to follow on the screen what is going on, of them 5 have difficulty seeing the picture on the TV screen and the light of the TV screen and 2 cannot see anything on the TV screen. Finally, 6 home users use audio description to explain to them what happens on the screen, and also 6 home users, although not the preferred option, ask their friends or family members to assist by explaining what happens on the screen or try to pick up as much as they can from the sound of the film or programme, 1 uses residual sight to watch and 1 sits closer to the TV screen.

Results

Task 1.

Objective: Understand how AD fits with linear storytelling and then how the interactive aspects of 360 degree videos impact the AD track? (Cue: In conventional television we can assume that people are looking directly at a TV screen in front of them and the audio description describes what's on the screen. However, what do we do with the audio description in a 360° experience if the viewer can be looking in any direction?)

Content
Title: Attenborough and the Giant Dinosaur - BBC One
Length: 4 mins (approx.)
Link: https://www.youtube.com/watch?v=rfh-64s5va4
Description delivered live by Describer Roz Chalmers

Feedback
Participants praised the live audio description, it was agreed that the track comprised all those aspects of the video that had lacked audio clues and completed the picture.
"The important thing is the description complemented the narration, it wasn't repetitive and it wasn't overly descriptive."

"I think it managed to describe everything that was on the screen even though there wasn't much time."
Although the description was appreciated in terms of making the story clear, some felt it lacked the elements to build the atmosphere needed for an immersive experience.

"What I didn't get from the description was the ambience – you hear the rustling of the foliage, large clomping feet, what's the weather like, what birds can you see in the sky? This is where your imagination comes in."
"I still can't get a proper visual picture of the dinosaur in my head, so the colour, has it got a spikey back?"
Here the describer pointed out that the brief gaps in narration meant there was a need to prioritise what could be described. "You got what the sighted viewer was getting." It must be noted here that NO ONE wanted the audio description to go over the voiceover.

Task 2
Aim of task 2: Understand how AD fits with the interactive aspects of 360 degree videos impact the AD track? (Cue: In conventional television we can assume that people are looking directly at a TV screen in front of them and the audio description describes what's on the screen. However, what do we do with the AD in a 360° experience if the viewer can be looking in any direction?)

Content
Title: Attenborough and the Giant Dinosaur - BBC One
Length: 4 mins (approx.)
Link: https://www.youtube.com/watch?v=rfh-64s5va4
Description delivered live by Describer Roz Chalmers

In order to emulate a 360 environment, cursor provided within the video was used to shift focus to different views. For example, instead of watching David gaze at the dinosaur during the first break in voiceover, view panned left to the mountains in the distance and consequently the audio description changed as it referred to the on-screen elements now in view. This resulted in participants mostly feeling disconnected with the storyline and found the description to be repetitive and missing important details.

"I thought that sounded disjointed and for me if it is a description of 360 then it would take a lot more than just words. This didn't give me anything really."

While the 360 degree movement seemed to enhance the overall immersive experience visually, it was almost impossible to simulate that experience in audio given the brief gaps in voiceover. Some participants commented that it was difficult to comprehend a 360 view.

"I don't understand, was it vertical now, like a vertical axis, looking down? Was that the view? Was it from the perspective of the dinosaur?"

"For some people who can't see and never have, it is already a challenge to understand 360."

"I don't believe this! Where is the TV screen? I can't get my head around it."

Participants agreed that a lot more would be needed than just audio description to make the environment more immersive for people with significant sight loss and that getting the right balance without information loss would be difficult.

Task 3
Aim of task 3: Use spatial audio to create the illusion of sounds all around the subject thereby creating a 360 degree environment.

Content
Title: Virtual Barber Shop
Length: 5 mins (approx.)
Link: https://www.youtube.com/watch?v=IUDTlvagjJA
Speakers were used to allow the focus group to participate as a group.

Participants responded enthusiastically to this clip, which had no visual content:

“It’s great! I’ve heard it before and it’s a million times better with headphones!”

“For me this was immersive!”

“You’re actually in the room!”

“It was very impressive that last one, I found it very interesting. I know who’s around this table because I’m in the room. Felt as if it was happening around me!”

“Honestly that’s what I need to be completely immersed.”

It was noted that the narration in the clip integrated elements of audio description within it. For example, “now I’m moving to the right.” “Look at my pair of scissors”. This combined with the perception of depth delivered in spatial audio enhanced the experience for the focus group.

Task 4
Establish user preference for tools required to access 360 degree content.

Most participants agreed that a HMD would be unnecessary as audio would be the key feature for them. However technical expert participating in the group pointed out the significance of the headset which may be used to track head movements and subsequently trigger specific descriptions.

The issue of integration with assistive technology tools such as speech readers and magnification was also discussed briefly. It must be noted here that all participants were regular users of video on-demand services such as Netflix, Amazon Videos, BBC iPlayer which are set up to work with assistive technology. Participants reiterated that this was essential to allow independent access.

However in view of the characteristics of the immersive environment that may not allow control via traditional tools such as keyboard and mouse, voice control was considered most appropriate. Once again, members of the focus group have previously used voice control on smartphones i.e., VoiceOver on iOs and Talkback on Android devices.

Discussion

Figure 3 - Focus group in UK

For 360 Image: http://imac.crazysandbox.co.uk/360photo/?img=360_0262

Overall, the focus group agreed that an audio led immersive environment was easier to comprehend than an audio-visual environment for dedicated users of audio description. It was felt that the visual display of a 360 degree environment was somewhat irrelevant without elements of it figuring in the audio description track. Head mounted displays were regarded as unnecessary by the group with the exception of one participant who is not a regular audio description user.

Participants strongly felt that the description track needed to complement the main narrative and any deviation from the primary storyline would lead to unnecessary disorientation. The five key elements of audio description – who, what, why, where and when were prioritised over the description of 360 degree elements by the participants. These were considered more acceptable when offered in the immersive audio environment of the Virtual Barber Shop where sound design was used to set the scene in a 360 degree view but the script followed a single narrative with integral clues on the setting.

It was also felt that since such content was meant to be consumed independently therefore the audio description script could be written in second person i.e., you're only a few paces away from the stage where the band is playing…etc. Professional describer added that this may help pull listeners into the scene and enhance the immersive experience.

The group discussed various factors that could contribute to immersive experience including how many voices would be considered too many in an immersive environment and whether directionality and placement of the audio description would impact the viewer experience for people with sight loss i.e., audio description could be the voice in your ear or coming from somewhere behind you. However, no definite conclusions could be reached in the absence of samples.

On the subject of accessing content, there was a consensus on using a combination of voice control and integration with assistive technology tools – magnification and speech readers. Technical expert in the group however pointed out that immersive experiences are unlikely to support the use of traditional equipment such as keyboard/ mouse therefore voice control may be the only way to access content. It must be noted here that some participants had previous experience of using voice controlled environments such as Amazon Fire TV, Alexa and Google Home.

Further research, which includes specially produced content and a wider focus group comprising of people with different sight levels could clarify the importance of sound in an immersive environment.

3.1.3 Audio Description in Poland

Although not initially planned in the project proposal, a focus group was conducted in Poland as well.

Figure 4 - Focus group in Poland

Focus group general information

- Partner responsible for the workshop: UAB.
- Place and date: Kraków, 28.12.2017, Katedra *UNESCO* do Badań nad Przekładem i Komunikacją Międzykulturową, ul. Czapskich 4
- Access service(s) discussed: AD.

Participants profile

- Number of home users: 3. Number of professional users: 3 audio describers.
- Demographics for users:
  1. Sex: 2 “male”, 4 “female”, 0 “other”, 0 “prefer not to reply”.
  2. Age: 46, 37, 25, 31, 43, 33.
  3. Main language of the participants: 6 Polish.
  4. Level of finished studies: 6 “university”.
  5. (Only for home users): “I define myself as a…” 3 “blind person”, 0 “low vision person”, 0 “deaf person”, 0 “hearing impaired person”, 0 “deaf-blind person”.
  6. (Only for home users): Age in which your disability began: 2 “From birth”, 0 “0-4”, 1 “5-12”, 0 “13-20”, 0 “21-40”, 0 “41-60”, 0 “More than 60”.
  7. What technology do you use on a daily basis? You can select more than one. 1 “TV”, 3 “PC”, 5 “Laptop”, 6 “Mobile phone”, 2 “Tablet”.
  8. Do you have any device to access virtual reality content? 1 “Yes”, 5 “No”, 0 “I don’t know or I don’t want to reply”.
  9. Which of the following is your preferred device for watching online video content (i.e., Youtube, Vimeo, Netflix, Amazon Prime, broadcast catch up service etc.)? You can select more than one. 1 “PC”, 4 “Laptop”, 2 “Smartphone”, 1 “Tablet”, 0 “I don’t watch online video content”, 2: Smart TV, External monitor connected to laptop “Others”
  10. (Only for home users):Which of the following do you use on your connected devices to access the above content? “Magnification”, 2 “Screen reader”, 1 “Both”, 0 “None”.
  11. (Only for home users): Which of the following controls would you like to use with your screenreader/magnification tool when watching content online? 2 “Browse content library”, 2 “Identify content”, 3 “Functions such as play, stop, pause, forward, rewind”, 3 “Switch AD/AS on and off”.

Summary
Six users took part in the focus group (3 end users, 3 audio describers). They were 2 males and 4 females, with ages ranging 25-46. All participants had university studies. One of them reported having a device to access virtual reality content. Mobile phone was the most used technology by the participants on a daily basis (6), followed by laptop (5), PC (3), tablet (2), and TV (1). The end users were blind participants: vision impairment started from birth (2) and between 5-12 (1). Two blind participants reported using screen reader to access content and the third one both magnification and screen reader. All participants identified functions such as "play, stop, pause, forward, rewind" and "switch AD/AST" on and off as needed to watch content online, and two of them also considered "Browse content library" and "Identify content" as needed.

Workshop conclusions: approved by focus groups

Conclusions from professional users

Professional audio describers think that AD in immersive environments is more challenging than AD in standard non-immersive video materials. They consider that immersive environments are interactive, meaning that the user can choose which parts of the visual landscape to consume and that blind users should also be able to consume the contents in such a way.

Professional audio describers consider that the visual scene is much larger. There is much more possible information to convey than in a standard non-immersive video content. From professional point of view, it will considerably increase the amount of work needed. They also consider that providing audio description to immersive content will be much more challenging from financial perspective, as there will be more possible audio description units.

Professional audio describers consider that the question of how to remunerate audio describers writing 360 contents is crucial.

Professional audio describers think that audio description of the main action should be provided as a priority to allow the blind and partially-sighted to follow it, but audio description of the surrounding visual scene should also be offered. During the focus group, three possible ways of implementing AD in immersive environments were explained by professional audio describers:

- To offer audio description of the main action with the possibility to pause the video material and listen to audio descriptions of the surrounding visual scene. It creates the possibility to watch the film many times and listen to different AD units in a personalized way. It will result in different time management, as user can spend much more time watching the film. This is why this possibility will mostly concern video materials consumed at home, not in institutions such as museums. This will also increase the number of audio description units which not always will be activated.
- To offer audio description of the main action with simultaneous audio descriptions of the surrounding scene. It raises the issue of how to come back to the main action.
- Taking into consideration the amount of possible work, the video materials should first be carefully selected and audio description should be produced in close cooperation with the creators of 360 contents. It was suggested that making such video materials

accessible should be an obligation of the content creators. **Visiting paths** should be prepared: Users should have the possibility to choose the option ´guide me´ which would guide them through the main action.

Professional audio describers think that they would not like to work in an immersive editor (because of the difficulties related to writing on keyboard and marking timecodes). A general view of the visual scene would be needed in a plain screen, with a possibility to write audio description in windows linked with different sections of the visual scene.

Text-to-speech module that would allow proof-reading with glasses: As in Poland audio description is recorded and read aloud by a professional, *lektor*, audio describers point to the necessity of having a module with a text-to-speech software in the content manager to check the final outcome of audio description in glasses before it is recorded.

Concerning current programs, professional audio describers use in their daily work Microsoft Office and AD Maker (doc. and srt. files)

Audio describers consider that information could be prioritized according to the volume (audio description of the main action could be played louder). Professional audio describers suggested that two voices could be used: male and female, respectively for main action and surrounding visual scene. Professional audio describers suggest that AD could be given like headlines and then the user can turn their heads towards it, but end users were more interested by being guided by binaural sound.

Professional audio describers consider that they would need to gather information about all possible types of video content (documents, fiction films, simulations) and venues for immersive contents (use at home, museums, training) to gather all challenges and determine what is feasible in each case. It was suggested that, while watching a film, the most relevance is placed upon what the film director wants to convey and not to look around. In simulations or city tours – looking around is important.

Conclusions from advanced home users

Concerning the interaction: accessing the services

Screen readers, screen magnification and voice commands should be implemented in immersive technology to allow blind and partially-sighted users to access video materials. End users should be able to mark them in check-boxes.

Users in the focus group consider that their preferences and parameters should be replicated between devices.

Regarding accessing immersive content with or without glasses, users indicate that it depends on the end user.

Concerning the services

Users and audio describers consider that implementing access services to 360 immersive content is needed: if this technology is developed, access services should also be offered.

Binaural sound: users think that it could be useful to orient yourself in the scene. It is also needed to know where to turn your head to receive audio description. They consider that it deepens the sensation of being in the centre of the action. They expressed an opinion that object-based audio description could be very interesting, as it would deepen the experience of being inside the film.

Users favour the AD position linked to action being described. They are also interested in looking around and trigger secondary audio description tracks (e.g. to turn around and listen about what is behind them).

Returning to the main action: both end users and audio describers indicate the challenge of moving away from the main action and going back to it. They think that the information concerning the main action could be given (1) by the different voice, or (2) by a different volume of the sound.

Users (and audio describers) are strongly in favour of dubbing instead of voice over or subtitles when consuming foreign-language film.

Sensor with eye-tracker: End users think that this technology should respond when a user closes his or her eyes (or moves his or her head because the content provokes disgust or anxiety): there should be a sensor that detects it and audio description should be then stopped automatically.

Both professional audio describers and home users are more interested in using this technology in museums or planetariums (e.g. in form of simulations of travels or diving in an ocean) and not at home. The reasons behind it are following: (1) watching films is mostly a social activity and spending time in glasses keeps are removed in our own worlds – it is impossible to share the experience of watching film with others (2) this technology is more adequate for watching short-duration video materials (e.g a city tour) as it could be too cognitively tiring to watch full-length film (over-stimulation).

End users are interested in using this technology for educational materials (that are also meant to be entertaining) or for training purposes (they said that *it feels as if such a content is real* which is the potential of this technology). End users are interested in watching materials no longer than 5-15 minutes.

Physical effort: Participants consider that they would not always like to turn their heads while consuming content, especially if they used this technology at home. They are much more interested in turning themselves around and turning their heads around while consuming 360 video contents in museums or other institutions.

3.2. Subtitling

Two subtitling focus groups were performed. They correspond to the two broadcasters CCMA and RBB who will later perform the pilots. In each group, the way to represent VR to users was different. In Catalonia CCMA has its own produced content in VR to present the possibilities of VR and elicit answers. Rbb on the other hand could not show Catalan content because it was not translated into German, so they resorted, as can be seen in Figure 5 to sketches on the whiteboard.

3.2.1. Subtitling and Signer in Germany RBB

Focus group general information

- Partner responsible for the workshop: RBB
- Place and date: RBB (Potsdam, Germany), 28.11.17
- Access service(s) discussed: subtitles, signer

Participants profile

- Number of professional users and profile: 4 (2 subtitle editors, 2 experts for 360° videos).
- Number of home users: 5
- Demographics for home users:
    1. Sex: 2 “male”, 3 “female”, 0 “other”, 0 “prefer not to reply”.
    2. Age: 37, 2x40, 52, 62
    3. Main language of the participants: 4 “German sign language”, 1 German
    4. Level of finished studies: we omitted this question
    5. “I define myself as a…”: 0 “blind person”, 0 “low vision person”, 2 “deaf person”, 3 “hearing impaired person”, 0 “deaf-blind person”.

    5.1 level of hearing impairment according to WHO: 5 “profound hearing loss (over 81 db)”

    6. Age in which your disability began: 1 “From birth”, 3 “0-4”, 1 “5-12”, 0 “13-20”, 0 “21-40”, 0 “41-60”, 0 “More than 60”.
    7. What technology do you use on a daily basis? You can select more than one. 5 “TV”, 1 “PC”, 5 “Laptop”, 5 “Mobile phone”, 2 “Tablet”.
    8. Do you have any device to access virtual reality content? 0 “Yes”, 5 “No”, = “I don’t know or I don’t want to reply”.

Summary of participants’ profile

Eight users took part in the focus group (5 home users, 2 subtitle editors and 1 expert for 360° content). They were 3 males and 5 females. The home user ages are ranging from 37 to 62. None of them reported having a device to access virtual reality content. Laptop, TV and mobile were the most-used technology by the participants on a daily basis (5), followed by tablet (2) and PC (1). The home users were hearing impaired participants. According to the WHO grades of hearing impairment, all participants have a profound hearing loss (hearing impairment started from birth to 5-12).

Approach

RBB's focus group followed the same procedure as all other focus group. In preparation of the session with both user groups it was clear that demonstrating only 360° content in different devices and ask questions might not be sufficient to get the desired results concerning the accessibility services.

Consequently, we identified four use cases to break up the complexity. We used paper prototypes to help to envisage possible solutions. One scene from this video was illustrated with sketches on the whiteboard (see Figure 5). According to the open discussion by the users we realised the results with paper stripes for the subtitles or a sketch of the signer. The users agreed after each use case on the results.

At the end of the session we addressed open questions coming from the collecting of the user scenarios table.

Workshop conclusions: approved by focus group

Conclusions from professional users

Subtitles

In the separate session with the professional users elaborated the subtitle production workflow and then defined any necessary enhancements for positioning in 360° content.

The current subtitle workflow depicts the necessary workflow steps for the production part.

1. Content is available (video, manuscripts and/or subtitles)
2. Import/open content
3. In the subtitle software
    a. Check integrity
    b. Format text in "frames"
    c. Define time code base
    d. Set time code for frames
    e. Check
    f. Replay (high quality preview, less authoring)
    g. Acceptance
4. Export/save/send

The workflow has to be enhanced with additional tools to enable the following additional features:

- The existence of additional spatial information doubles the effort required to place subtitles within media items
- Time-based PLUS angle-based navigation with the help of shortcuts, scroll wheel and input field
- Time-based PLUS angle-based subtitle definition
- Time-based PLUS angle-based editing
- Time-based PLUS angle-based preview
- Time-based PLUS angle-based replay
- Preview in low-res flat angle view
- Replay and acceptance in high-res flat angle, flat unfolded and HMD view

Signer

Here, the workflow is obviously quite similar to that of the production of subtitles. Sign language video is pre-produced and needs to be added to the omnidirectional main video, depending on positions in the 360° field:

- Content is available (signer video)
- Import/open content
- Cut signer video if necessary
- Time-based PLUS angle-based signer definition
- Name, information of the speaker
- Time-based PLUS angle-based editing and positioning of textual and/or graphical notices
- Time-based PLUS angle-based preview
- Time-based PLUS angle-based replay
- Preview in low-res flat angle view
- Replay and acceptance in high-res flat angle, flat unfolded and HMD view

Conclusions from advanced home users

Figure 5 - Explaining immersive subtitles in RBB

a) Concerning the interaction: accessing the services

They would like to have a similar approach to the one currents used in RBB's web player for video on demand and the TV player used for catch-up services. These offer a dedicated button with the abbreviation UT (for subtitles) allowing the user to switch the subtitles on or off and to access the settings. The adjustment wheel icon for accessing all available settings is also very common and was mentioned. Summarized they would like to see a solution based on existing settings similar to those used for the catch-up TV service (position, background, size).
The usage of the user interface in a HMD was identified as a possible challenge.

b) Concerning the services

The users saw a 360° video of RBB's news magazine Abendschau showing a "behind the scenes" story of the show's production. The video was shown in a desktop-based web browser and in a HMD (Oculus rift, VR glasses with smartphone).

One issue that was not obvious before the focus group discussion is that people with hearing loss have balance disorders and the usage of a HMD can potentially cause motion sickness. The users were sitting during the service consumption and our observation was that the advanced home users only made tentative head movements.

One scene from this video was illustrated with sketches on the whiteboard. The blue frame symbolised the users' field of view.

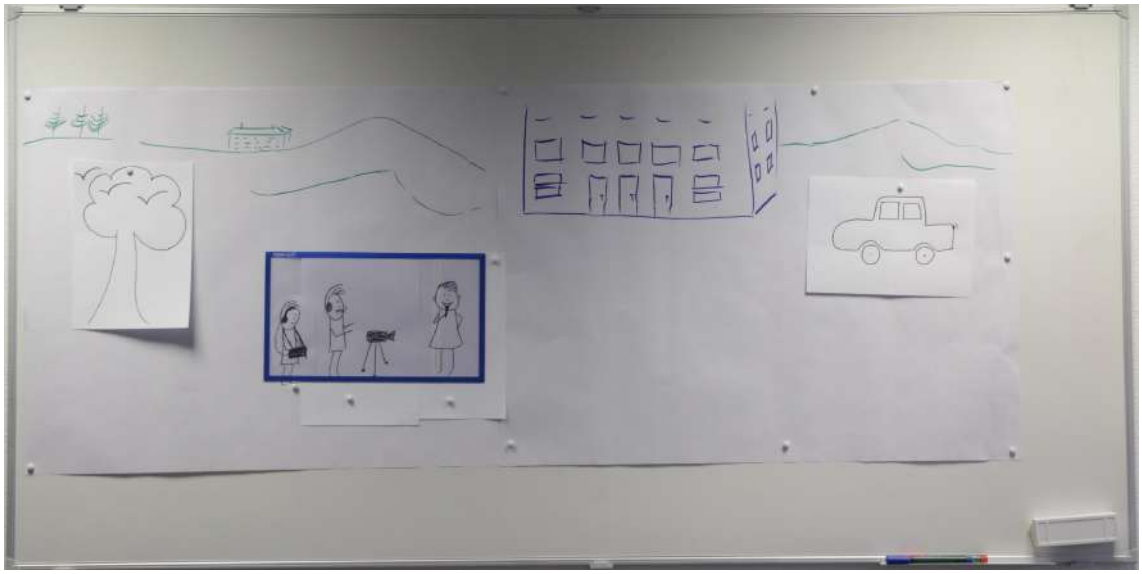

Figure 6 - Basic scene

1st Use case: off-speaker, subtitles

The speech bubbles symbolised that somebody is speaking. In the first case, illustrated in Figure 7, an off-screen speaker explains something and the user is looking around in the content.

The question was: Where should the subtitles been positioned? The red stripe symbolised the subtitles.

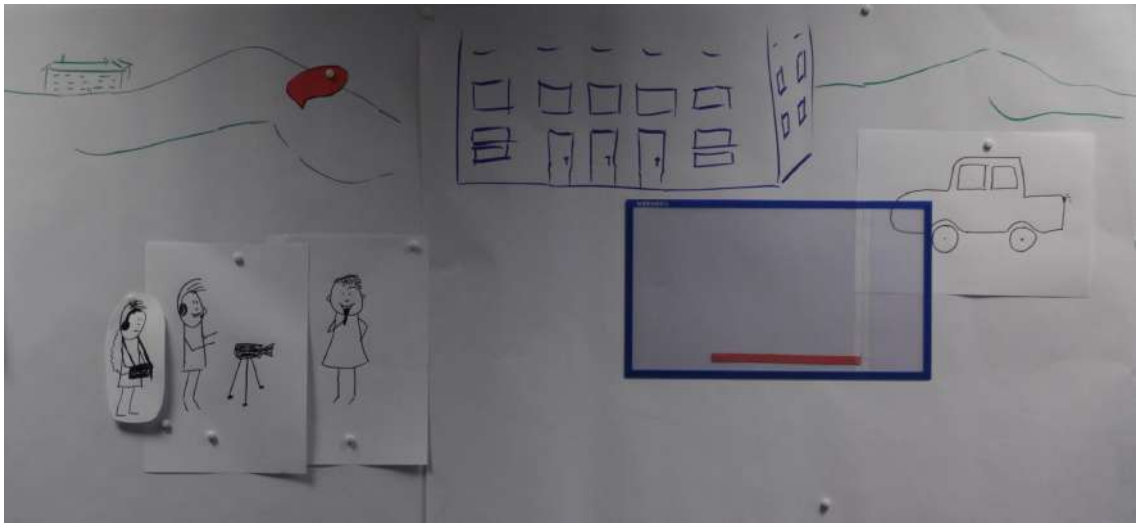

Figure 7 - Off-scene speaker

Conclusion:
The subtitles should be in a fixed position in the user's field of view according to the current standards, two-lined and each speaker gets its own colour. It is important for the users to get the same information as hearing people.

2nd Use case – one speaker, subtitles

In this example a presenter is speaking and different acoustic information such as singing birds in the tree, a car door slamming and a camera team preparing for a live recording is visible and audible.

The questions to the users were: Do you want to have a description of this audible background noises in addition to the subtitles for the speaker? If yes, how should they be displayed?

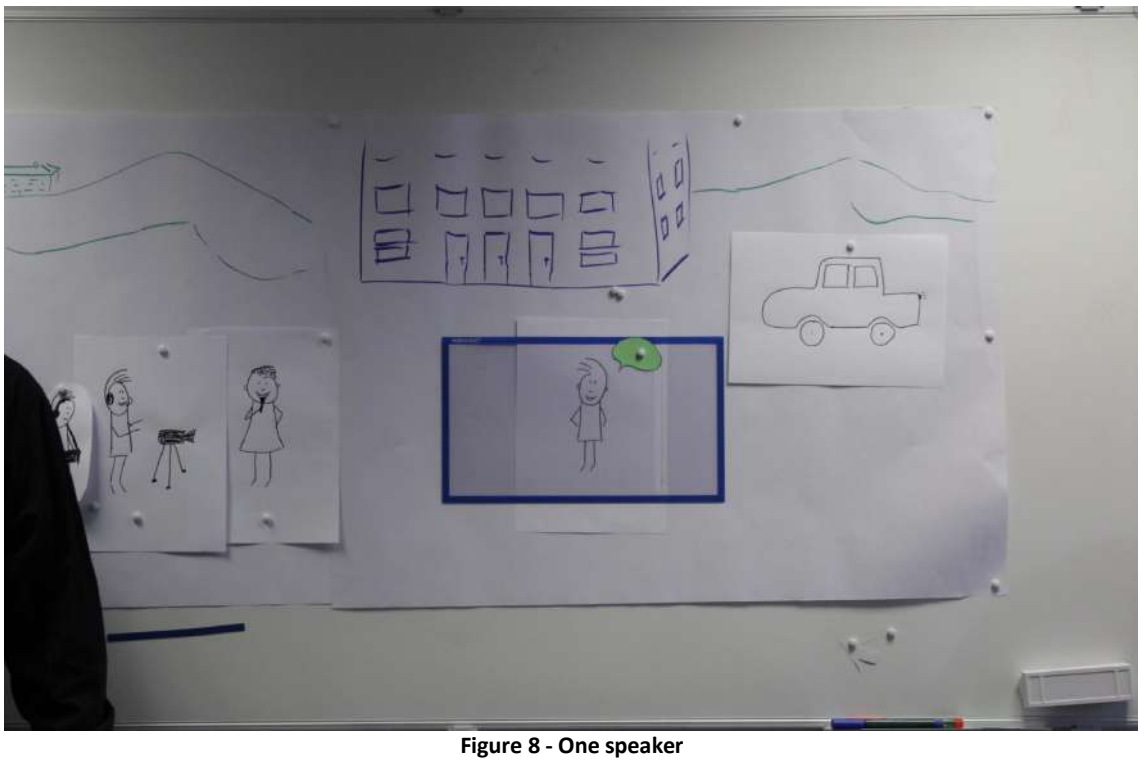

Figure 8 - One speaker

Conclusion:
In contrast to the above statement that all information i.e. speech and background noises should be available for people with hearing loss, one user was not sure if these noises are important enough to be mentioned.

One of our RBB team members asked again concretely: How should a user get the information that something important was happening outside of the field of view?

The user maintained his opinion and wanted to explore the content by himself. At the end of the discussion all agreed that it could be important in fictional films or documentaries to inform the user about background noises for a better understanding of the dramaturgy. The manuscript authors or subtitle editors responsible should decide which information is important for the advanced home user.
The conclusion was that information about background noises should be clearly distinguishable from subtitles. The users didn't express a clear preference if this information should be textual or graphical. They would like to test both uses cases in a pilot and also have an option to switch the information on or off.

In general the 360° content consumed with a HMD is more difficult for people with hearing loss as they don't get any acoustic feedback if something dramatic happens. An example was an approaching herd of horses. They won't hear them and will be surprised when they suddenly appear in their field of view.
Vibration feedback is one of the familiar feedback mechanisms. Testers suggested delivering such feedback via a combination of head-mounted displays and controllers.

3rd Use case – two speakers, subtitles

In the 3rd use case one presenter is speaking and appears in the field of view. He stops talking and goes into the building. Simultaneously another presenter starts to speak outside of the field of view.

The question to the users was: How does the user know who to attribute the new subtitles to and where the presenter is in the 360° space?

Conclusion:
The first idea of the users was to add an arrow to the subtitles to indicate the position of the speaker, depending on the actual viewing angle. The arrow should disappear as soon as the user reaches the respective field of view, and speaker and subtitles aligned.
One of the subtitle editors proposed an alternative solution. The positioning of subtitles on the left or right edge of the current field of view indicates the direction in which the speaker is talking and the subtitles will go to the middle as soon as the user reaches the respective field of view.

4th use case – two speakers, signer

The 4th use case has the same starting situation as the 3rd use case. The difference is the type of accessibility service.

The questions to the user were: where should the signer be positioned? How does the user know where the second presenter is?

Conclusion: The signer should be positioned as usual top right of the speaker or if no speaker is visible top right of the field of view. The signer should always be visible in the field of view.
The users see difficulties if two speakers are talking in parallel, especially when both or only one of them is in the field of view. Usually, the signer indicates with a gesture the direction of the speaker. This can be done in a studio situation for 2D video. In 360° video content the signer does not know where the speakers are positioned in relation to the user's field of view.

Different approaches were discussed:

1. An arrow indicates the position of the speaker directly under the signer window and/or the name or description of the speaker is displayed.
2. The field of view will be changed by the video player ("forced perspective") so that the user will see the speaker when a conversation between two persons starts and can align the translation of the signer to the speaker. Afterwards the user decides the field of view self-sufficient.
3. The signer image has a fixed location and the user decides where he/she wants to look.

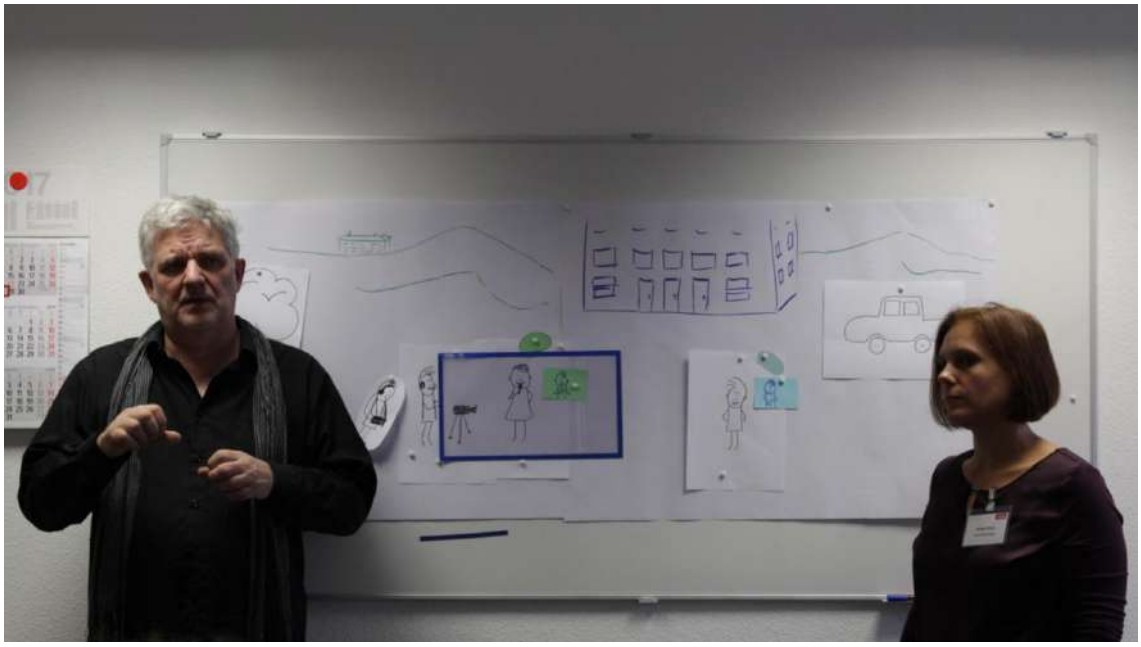

Figure 9 - Focus group - RBB

Additional questions:

Personalisation options

1. Subtitles - the same settings as in catch-up service:
   a. switch on or off subtitles
   b. position (top, bottom)
   c. size (three different: small, medium and large)
   d. background (semi-transparent box, outline, frame around TV picture)
   e. switch on or off notices
2. Signer:
   a. position
   b. size
   c. switch on or of notices
   d. switch on or off forced perspective on the speaker

Parallel usage in a group of people – synchronisation between devices

- the parallel usage in a group is a scenario that is conceivable but more in front of the TV
- the tablet as a second device was not preferred as the display is too small for the consumption of subtitles and signer
- the head mounted displays were not preferred – they prevent the communication in sign language and a usage alone was more preferred
- a HoloLens approach was discussed, where users can see each other and at the same time have a sign language translator positioned next to the TV-based end device.

3.2.2. Subtitling in Catalonia

The whole document with notes from the focus group can be found here https://drive.google.com/file/d/1mVo3wkJ8TSVS6cougZB5zopJdrCeBI1s/view?usp=sharing
A press release of the Focus Group can be consulted here: http://www.ccma.cat/premsa/la-ccma-participa-en-el-projecte-imac/nota-de-premsa/2824511/
And tweets to the press release:

https://twitter.com/CCMA_cat/status/936617166408953856 (Catalan)
https://twitter.com/CCMA_cat/status/936624649286438912 (English)

Figure 10 - Participants in CCMA focus group

Figure 11 - Participants in CCM focus group and Sign Language interpreter

The summary is as follows:

Focus group general information

- Partner responsible for the workshop: CCMA.
- Place and date: CCMA (Sant Joan Despí, Barcelona), 28.11.17
- Access service(s) discussed: subtitling and sign language interpreting.

Participants profile

- Number of home users: 10 (4 Sign Language, 6 Oralists).
- Number of professional users and profile: 6 (2 subtitlers, 3 technical experts, 1 user association member –not deaf–).
- Demographics for users:
    1. Sex: 8 "male", 8 "female", 0 "other", 0 "prefer not to reply".
    2. Age: 25, 34, 38, 38, 41, 46, 47, 47, 49, 53, 53, 58, 61, 62, 65, 66.
    3. Main language of the participants: 1 Spanish, 1 Sp/Cat, 10 Catalan, 3 Catalan Sign Language, 1 Spanish Sign Language.
    4. Level of finished studies: 0 "no studies", 0 "primary education", 3 "secondary education", 4 "further education", 8 "university". One person did not answer.
    5. (Only for home users): "I define myself as a…" 0 "blind person", 0 "low vision person", 8 "deaf person", 2 "hearing impaired person", 0 "deaf-blind person".
    6. (Only for home users): Age in which your disability began: 4 "From birth", 5 "0-4", 0 "5-12", 0 "13-20", 0 "21-40", 1 "41-60", 0 "More than 60".
    7. What technology do you use on a daily basis? You can select more than one. 14 "TV", 10 "PC", 12 "Laptop", 16 "Mobile phone", 8 "Tablet".
    8. Do you have any device to access virtual reality content? 3 "Yes" (VCR?, Glasses, PC), 13 "No", = "I don't know or I don't want to reply".
    9. Which of the following is your preferred device for watching online video content (i.e., Youtube, Vimeo, Netflix, Amazon Prime, broadcast catch up service etc.)? You can select more than one. 7 "PC", 7 "Laptop", 5 "Smartphone", 3 "Tablet", 0 "I don't watch online video content", 3 "Others" (TV)

Summary of participants' profile

Sixteen users took part in the focus group (10 home users –4 Sign Language, 6 Oralists–, 2 subtitlers, 3 technical experts and 1 user association member –not deaf–). They were 8 males and 8 females, four participants with ages ranging 21-40, eight participants with ages ranging 41-60 and finally four participants with more than 60. Three participants had secondary education studies, four participants had further education studies, eight had university studies and one person did not reply to this question. Three of them reported having a device to access virtual reality content (VCR, glasses and PC, respectively). Mobile phone was the most used technology by the participants on a daily basis (16), followed by TV (14), laptop (12), PC (10) and tablet (8). The home users were deaf (8) and hearing impaired people (2), most of them having acquired the disability from birth (4) or with ages ranging 0-4 (5), and one person with ages ranging 41-60. The preferred devices for watching online video content was PC (7) and laptop (7), followed by smartphone (5), tablet (3) and TV (3).

Workshop conclusions: approved by focus group

A first session was carried out with all the focus group participants. The excellent turnout of home users in the focus group resulted in a slower development. It was therefore agreed to focus on questions of interest for both home and professional users, and leave the specific questions addressed to professional users for a second meeting only for professional users. This is why the information is reported in the changed order. First, we report about the results of the first focus group, in which both home and professionals users took part, and then we report on conclusions from the professional users.

Conclusions from advanced home users

a) Concerning the interaction: accessing the services

Regarding interaction with the interface, both professional and advanced home users consider that it would be positive to be able to customize the interface once and then that the interface would remember those parameters.

Users suggest that this customization should be transferred from one device to the other (importing profile) and they request the possibility to create more than one profile. They also consider the possibility of transferring a profile from your device to another external device (for example, at your friend's home).

Regarding interaction with access services, users positively value the possibility of alternative interactions (for example, voice commands), although they do not find it necessary in their case and they indicate that the implementation costs should be taken into account. However, they think that, if this is to be developed for other profile types anyway, it could be an additional resource.

Regarding companion screens, users like the possibility of using the smartphone to move the screen in a tactile way (like a "mouse") and to customize their preferences. A user even suggested the possibility to include a sensor in the finger that allows users to see their own fingers on the image. There are different opinions regarding the need of reproducing the content in the mobile too, since the smartphone is often used as an element to access additional contents.

When accessing audiovisual content with other people, users do not want the subtitles to be consumed on a different screen (for example, a smartphone), but on the same screen as the other users.

b) Concerning the services

Users suggest that subtitles should always appear in a fixed position in relation to the users' field of view.

Users suggest that subtitles in immersive media should be based on approved subtitling rules (for example, UNE rule in Spain) and, if necessary, improvements might be implemented to adapt existing rules to the new needs posed by immersive environments.

Users state that it is necessary to keep colour coding to identify characters.

Users require that basic subtitling elements that have been already approved in the regulations (for example, how to indicate music) need to be kept. However, they accept that new technologies may bring new possibilities.

Users require that subtitles include all the information, both the information present on the screen and also the information off the screen (ON and OFF).

Users suggest that it is necessary to include non-speech information (that is, sounds, extra linguistic information, etc.) and directions.

Users state that it is difficult to know where to look for the character who is speaking. The subtitle must indicate where you need to move your head (four directions). It is suggested that a wind rose or a compass is drawn to indicate where the sounds come from.
It seems that users prefer that icons or similar elements are always located in the same position. Some users prefer at the top, others at the bottom closed to the subtitle (dialogue), and others would like to move them. All in all, it seems that customization is the solution.
Users suggest the possibility to use a closed list of icons in order to illustrate non-speech information. For example, a lightning to indicate the sound of a storm.
There are different opinions regarding how to include non-speech information in the subtitles: with icons or text. In that sense, customization should be prioritized.
Users positively value the possibility of personalization, that is, having different layers that one can activate or not, depending on their needs. Like a "menu of options". However, there are some elements that do not need customization, such as the position of the subtitles – always at the bottom – except from specific cases, such as football matches, etc.
It becomes clear that there are different needs among users and, consequently, subtitles must be adapted to different profiles. For example, there could be different levels of speed (faster/slower).
Additionally, sign language must also be considered. Sign language would appear simultaneously to the person speaking on the screen and it would satisfy the needs from other users.
Users consider that summarized or simplified subtitles that do not reproduce exactly what is being said word by word do not generally help deaf people, because this type of subtitles make it more difficult to follow the audiovisual content. However, they admit that simplified subtitles may be useful for users with other type of needs (for example, people who need easy-to-read texts). In that sense, simplified/summarized subtitles can be an alternative.
Regarding sign language, users require that the sign language interpreter is always located at the same fixed position in relation to the users' field of view and with a background. They also prefer that each user has the possibility to customize the position of the SL interpreter.
Users raise doubts regarding the success of these new technologies and they believe that we need to be ready, but they are afraid that it happens the same as happened with other services, such as 3D cinema

Conclusions from professional users

- These are the answers or proposals expressed by professional subtitling producers: Vertical positioning of subtitles is not a main need in 360 subtitles, but it must be interesting to use different vertical positioning to separate different subtitles or non-speech information. Anyway home users must be able to decide or setup where they want to have the subtitle text.
- As agreed with end-users, there is not demand for supporting 'moving positions' (e.g. a subtitle following a person in 360º), and agree with expert users about the need to keep a subtitle fixed position in relation to the users' field of view.

- During the production of subtitles, producers prefer to have an on-screen display (player) showing one dynamic angle of the 360 view, so they can choose which angle to see through cursors or mouse movements. They don't need to have a 2D distorted panoramic view showing full 360º.
- For testing purposes of contents, producers think it should be done with HMD and on-screen display, so they can test both results, using HMD or directly from a display (PC screen or smartphone).
- A 360 web subtitling editor could be very similar to, for example, the Anglatècnic tool they use nowadays, but it should add the 360 displaying and the possibility to add 'emoticons' and text messages to show 'sound actions' that take place in parallel to dialogue subtitles.
- The Web Editor tool must offer original 360 immersive audio, because it is important to work correctly with subtitles and notifications showing where the sound comes from. The tool would add a wind rose item that will help the home user to locate the position of different sounds or dialogues.
- About the possibility to have a Web Editor tool for Sign Language, it is observed that Sign Language would be produced in a similar way as it is done nowadays for classic 2D audiovisual contents. The position for the Sign Language Picture in Picture box would be configured by the home user from player interface, and would be a fixed position in the visual area.

4. REFERENCES

[1] Kitzinger, J. (1995). Qualitative research: Introducing focus groups. *British Medical Journal*, *311*, 299-302.

[2] Morgan, D. L. (1997). *Focus groups as qualitative research* (2nd ed.). Thousand Oaks, CA: Sage.

[3] Barret, J & Kirk, S. (2000) Running focus groups with elderly and disabled elderly participants. *Applied ergonomics*, 31, 621-629.

[4] Seymour, J., Ingleton, C., Payne, S., & Beddow, V. (2003). Specialist palliative care: Patients' experiences. *Journal of Advanced Nursing*, *44*, 24-33.

[5] Harris, J., Foster, M., Jackson, K., & Morgan, H. (2005). *Outcomes for disabled service users.* Retrieved March 22, 2006, from http://www.york.ac.uk/inst/spru/pubs/ccatreps.htm#2005

[6] Kroll, T., Barbour, R. and J. Harris (2005) Using Focus Groups in Disability Research. *Qualitative Health Research* Volume 17 Number 5 May 2007 690-698 © 2007 Sage Publications 10.1177/1049732307301488

5. ANNEX I – T2.1. PROFESSIONAL USERS

	Technical Component	User scenario: indicate what the previous user will experience and how	input from
Content Management		User will use Web GUI Interface to manually control the ImAc platform, and API to automatically access & control the ImAc platform from broadcaster OTT related workflows: ● Access the platform ● Upload/Download video, audio, subtitle, sign language, AD contents. ● Monitor ImAc contents ● Download content to broadcaster premises ● Export into different formats a processed ImAc file (ST, sign language video or AD) via a save-file GUI ● Select an existing ImAc file via mouse click ● Manually assign an ImAc file to a video asset ● Check integrity of ImAc file ● Check synchronisation between main and ImAc content, alarm if necessary ● Report status of available data and its content type **Content Management - Input for focus groups:** ❑ Is there a specific content management system for your work today? ❑ What file formats do you currently work with (for video, ST, AD, SL) ❑ Do you like open-file-menus? ❑ Do you like drag and drop? ❑ Do you check integrity today, if yes how? ❑ What feedback do want from the content manager? number of emissions, … ❑ ...	RBB, CCMA, ANGLA
Content packaging and distribution		User can: ● Ingest an ImAc setup from the Content Management described by: ○ 1) output from production (to retrieve content and authoring infos) and ○ 2) user experience (player). ● Trigger/monitor automatic, open and closed (as nowadays subtitles) packaging of ST, SL,	RBB, MSE, ANGLA

	Technical Component	User scenario: indicate what the previous user will experience and how	input from
		AD enhanced omnidirectional media ● Trigger/monitor automatic, open and closed (as nowadays subtitles) distribution of ST, SL, AD enhanced omnidirectional media ● Monitor status of content packaging and distribution ● Retrieve the encoded and packaged ImAc result (to be consumed/previewed by a player). **Content packaging and distribution - Input for focus groups:** ❑ How do you monitor the processing of your results today? How does it look like? ❑ Do you need a preview here? ❑ …	
Interface: the signalisation of accessibility services in the content stream.		Accessibility services will be automatically signaled when they are present. Users can: ● Monitor signalisation of ImAC content ● Configure signalisation of ImAc content ● Preview signalised ImAc content **Signalisation of accessibility services in the content stream - Input for focus groups:** ❑ Do you like drag and drop? ❑ Do you check integrity today, if yes how? ❑ How would you signalize accessibility in broadcast contents? ❑ How would you signalize accessibility in broadband contents?	RBB, CCMA
Interface: handling presentation settings from different layers (like content stream, operating system, player),		Depending on the player and its operating system a user can ● pre-define the presentation settings (fixed for the end user or offering him a selection of different options) for every targeted content stream from a given number of presentation variants **Interface: handling presentation settings from different layers (like content stream, operating system, player) - Input for focus groups:** ❑ Which presentation settings do you see? Which options should the user have to personalize the accessibility services? ❑ Which presentation variants for subtitles, audio description and signer do you want to offer? ❑ Do you want to offer a selection of variants to choose from?	RBB

	Technical Component	User scenario: indicate what the previous user will experience and how	input from
		❑ For ST and SL: ❑ What do you think about viewing angle feedback mechanisms (for instance "ST follows view", "ST is fixed on objects")? ❑ What do you think about event notifications ("there is something happening/somebody speaking behind you")?	
Integration and testing		(Is related with content management, packaging and distribution) Professional user would use this module to: ● integrate IMAC cloud platform with internal broadcaster systems & workflows (CMS, MAM, Accessibility content management systems…) **Integration and testing - Input for focus groups:** ❑ Which systems of the broadcasters environment should be interact with the ImAc platform? ❑ ...	CCMA
Player for preview in editor tools		The users of the web tools and editors need to: ● Use the player to position and preview the result (angle, frame jump, slow speed, etc.) when they create the individual items (subtitle, audio description segment or sign language segment) ● See the final result over the video using the player. **Player for preview of content - Input for focus groups:** ❑ Is it sufficient to have a spheric preview in a plain web player? ? And for monitoring the final result? ❑ Is it necessary to have a preview in a head mounted display and/or smartphone with VR glasses? And for monitoring the final result? ❑ ...	ANGLA
Audio production tools		The ImAc content producer will use the web tool (in stage 1) or the professional editor (in stage 2) to: ● Perform file operations such as import/export audio track ● Preview and control the main video (angle, frame jump, slow speed, etc.)	ANGLA, CCMA, IRT

	Technical Component	User scenario: indicate what the previous user will experience and how	input from
		● Create plain (object-based see below) audio description segment with text, TCs, audio, segment metadata (attenuation, position, etc.) ● Edit AD data (key shortcuts, segment operations such as insert/delete/test, text edition such as cut/paste and search/replace) ● Compose object-based audio scenes and monitor the rendering in real-time for various output formats (e.g. 2.0, 5.1, 22.2, headphones, …) ● Export an ADM (Audio Definition Model) file which can be used for the distribution of the content ● Add certain accessibility related features, e.g. ○ Audio Description objects ○ Interaction or adjustment ranges for speech/dialogue related objects ● Preview the result with the video ● Verify the quality with the help of graphics to help the user (vumeter, waveform, thermometers, time left during the recording, etc.), also let the user set the parameters (audio parameters, speed, time restrictions, windows setup, default values, etc.). ● Verify synchronisation with main video ● Adapt synchronisation	

	Technical Component	User scenario: indicate what the previous user will experience and how	input from
Audio reception tools		A framework written in JavaScript which can be used for rendering of object-based audio content in Web Browsers. The professional user (Web developer, content provider, etc.) can use the framework to: ● receive and render object-based audio content in ADM format. It may be integrated in any HTML page. ● personalize and / or interact with the program: ○ Adjust the level of speech/dialogue within the allowed range ○ Enable / disable additional Audio Description tracks ○ Adjust the level of the Audio Description ○ Navigate within the audio scene and adapt head movements (especially useful for headphone listening) ○ Render the program for various output formats (e.g. 2.0 and headphones)	**IRT**
Subtitling Tools		User can ● Perform file operations such as import/export ST files ● Preview and control the main video (angle, frame jump, slow speed, etc.) ● Preview the authoring outcome in low res video (plane or 360°) ● Produce ST by inserting text with the keyboard ● (re-)structure text input into ST frames (e.g. two lines of 36 characters each) ● (re-)assign ST frames to positions in time, height and viewing angle ● (re-)define presentation styles of ST frames ● (re-)define which presentation style will be offered to the end user ● (re-)define the ST feedback (fixed absolute position | fixed viewing position | confinedly following viewing angle | added graphical hints where ST are located) ● Verify the result regarding synchronisation and defined presentation parameters **Subtitling tools - Input for focus groups:** ❑ What end user requirements for the presentation? ❑ What degree of freedom should be supported (e.g. should vertical positioning of subtitles be	**RBB, IRT, CCMA, ANGLA**

	Technical Component	User scenario: indicate what the previous user will experience and how	input from
		supported?) ❑ It is not planned to support "moving positions" (e.g. a subtitle following a person in the video) – Is there a demand for it? ❑ Notes: I would suggest asking professional subtitle editors how they could imagine to present subtitles in 360° or VR to the user. VR also adds the question where a subtitle is positioned in 3D. These are the more interesting questions we need to answer in phase 1, I think. The requirements for the editor software (e.g. regarding UI) are probably similar to those in standard editing environments for most parts. The main question regarding the editor would be, if an editor prefers to work in a 360° environment (HMD?), or if he would edit in a usual 2D interface and would just check the results in 360°/VR. ❑ During the production of subtitles, do you prefer a 360 view presented in one 2D panoramic view (although distorted) or show only the 2D view in one direction (not distorted view) and use the arrow keys to move around?	
Sign Language Editor		User can ● Perform file operations such as import/export SL files ● Preview and control the main video (angle, frame jump, slow speed, etc.) ● Add sign language video items (SL) to the omnidirectional video ● Split SL files into clips ● (re-)assign SL clips to positions in time, height and viewing angle ● (re-)define presentation styles of added SL clips (lighting, contrast etc.) ● (re-)define the SL feedback (fixed absolute position | fixed viewing position | confinedly following viewing angle | added graphical hints where SL are located) **Sign Language Editor Tool - Input for focus groups:** ❑ Same questions as in ST tool	**RBB, ANGLA**
Content production		These user scenarios are described above in "Editor tools (WP4)" for each type of web tool or editor. Users can produce ImAc content ● ST enhanced media ● SL enhanced media	**ANGLA, RBB,CC MA**

6. ANNEX II – T2.1. ADVANCED HOME USERS

	Technical Component	User scenario: indicate what the previous user will experience and how	input from
Interface: providing access to accessibility services	1. Screen reader integration (JAWS NVDA, Voiceover, Talkback) 2. Speech Recognition API 3. Compliance with the requirements in WCAG 2.0 (video playerS)	User uses existing tools which supports him controlling the player and the accessibility interfaces 1. Audio feedback: Depending on the platform, users with sight loss will be able to use audible feedback from JAWS/ NVDA/ Voiceover/ Talkback to navigate the player and control functions such as play, forward, rewind, stop, pause, volume, skip etc. 2. Voice commands: A voice controlled ecosystem will allow users with sight loss (also with mobility issues) to navigate the player by voice and choose any of the controls mentioned above 3. Screen magnification users will be able to navigate the player using magnification softwares such as Zoomtext. Bigger buttons & bigger font letters for people with vision difficulties. ❑	RNIB, CCMA
Multiplatform Player for desktop, mobile phone (cardboard supported, gyroscope sensor based) , TV, head mounted display (e.g. Oculus, Playstation VR,	Web application	User can ● start, pause, resume and stop playback of ImAc enhanced omnidirectional media in a plain video view (mobile phone, TV) or in a connected HMD or as a combination of these devices (TV + smartphone + HMD, TV + smartphone, TV + HMD, TV, HMD, smartphone, smartphone + HMD) ○ Automatic adaptation to different interfaces and platforms. ○ Automatic detection and presentation of accessibility contents. ● The experience will be synchronized across devices. ○ complementary 360 video is distributed synchronized to the main show, the other devices will show new audio description and subtitle services adapted to each user impairment ● The user will adapt the device that will store his/her preferences for the accessibility interface and the player.	RBB, i2CAT, CCMA

	Technical Component	User scenario: indicate what the previous user will experience and how	input from
Vive) WP3: Immersive platform, T3.5 player		**Input for focus groups - define use cases for synchronization between different devices:** ❑ Do you want to consume together with a group of people 360° content? The 360° videos is playing on the TV. The user can navigate in the video by using the arrow keys on the remote control AND/OR is it possible to use a second device like tablet or smartphone to control the viewing direction? This case will imply the synchronisation between the devices. ❑ You need to consume the 360° content with accessibility services. You will use a second device like a smartphone with glasses or a head mounted display. Do you see the necessity to consume the content in your device synchronously to the content on the TV or with other devices?	
Accessibility interface: Subtitles **WP3: Immersive platform, T3.5**		User can ● switch off and on ST presentation ● watch ST in omnidirectional media ● choose from different presentation styles ● choose from different ST feedback styles **Input for focus groups:** ❑ Possible scenarios to present subtitles ❑ Subtitles are presented always in front of you (slightly below eye line) and follow the head position, independently to the content ❑ The 360° space is divided into several parts and subtitles are available for each section on a fixed position (slightly below eye line) ❑ Subtitles are presented always in front of you (slightly below eye line) and follow the head position, independently to the content, additionally ❑ The user gets notices about the position of speaker - clarify in the focus group the presentation (another font?, graphics?, background?) ❑ Should it be possible to switch off the notices? ❑ The subtitles are presented in the spherical location where the speaker is ❑ Which settings want the user have to personalize the subtitles? Size, font, background, position (slightly below eye line, besides speaker...?)	**RBB, CCMA**

	Technical Component	User scenario: indicate what the previous user will experience and how	input from
Accessibility interface: Audio Description / Audio Subtitles **WP3: Immersive platform, T3.5**		Users can ● Hear different audio description/audio subtitles depending on the current angle visualization. (different use cases are defined below, users of the focus groups will give us feedback on that) User preferences: ● Users with sight loss will have the ability to control the audio description/audio subtitle track in relation to the main media ○ volume of the track and type – object based, surround, stereo ● Depending on the authoring metadata within the ADM file or stream and the graphical user interface, the end user has certain possibilities to personalize and / or interact with the program: ○ Adjust the level of speech/dialogue within the allowed range ○ Enable / disable additional Audio Description/Audio subtitle tracks ○ Adjust the level of the Audio Description/audio subtitle ○ Navigate within the audio scene and adapt head movements (especially useful for headphone listening) ○ Render the program for various output formats (e.g. 2.0 and headphones) ● User will have access a customised experience - player to have the ability to remember user preferences i.e., user with sight loss should have audio description/audio subtitles switched-on as default. **Input for focus groups: use cases for AD** ● The first is around the placement of audio description of which I can see three main strategies: ○ A, AD position linked to action being described, ○ B, AD position fixed in relation to scene (as if they are sitting on the sofa next to you as you turn your head), ○ C, AD position fixed in relation to users head (like a devil sitting on your shoulder as you turn your body to follow the action). ● B and C then invite a further question on where the audio description should be placed. Do users prefer it off to one side, in front of them, behind them or potentially even centred on the user. A tool to allow users to place the AD and record their preference could give us some great data. ● B and C also raise the question of how to describe in AD where action is happening if the	**CCMA, RNIB, IRT**

	Technical Component	User scenario: indicate what the previous user will experience and how	input from
		placement of the AD doesn't communicate that. Do you pair the audio description to a sound effect "[scraping sound placed within the scene] Daniel takes a knife from the block", can you get away with just the AD (or just the sound effect) or do you need to describe where the action takes place "a shadow moves at the back of the room"? ● This is potentially a far trickier piece of work because it requires a variety of content and different storytelling techniques. Since the intention is to build an audio description capable 360 video player then if the tool created in the project can be made available for further research then it would allow this question to be explored further.	
Accessibility interface: Sign Language **WP3: Immersive platform, T3.5**		User can ● switch off and on SL presentation ● watch SL in omnidirectional media ● choose from different SL feedback styles ● choose from different SL feedback styles **Input for focus groups:** ❑ Possible scenarios to present a signer ❑ Signer is presented beside the speaker and/or on a fixed position and follow the head position, independently to the content ❑ The 360° space is divided into several parts and signer are available for each section on a fixed position (beside speaker and/or on a defined place) ❑ Signer is presented in front of you beside and/or on a fixed position and follow the head position, independently to the content ❑ The user gets notices about the position of speaker - clarify in the focus group the presentation (another font?, graphics?, background?, signer?) ❑ Should it be possible to switch off the notices? ❑ Which settings want the user have to personalize the signer? Size, background, position (fixed position, beside speaker...?)	**RBB, CCMA**
Pilots, WP5 Content consumption		Users can consume ImAc content (rough summary, details are in the lines above concerning accessibility interfaces) ● ST enhanced media	**RBB, CCMA, RNIB**

	Technical Component	User scenario: indicate what the previous user will experience and how	input from
		• Signer enhanced media • SL enhanced media • Audio described media (in different audio formats) +audio subtitles [for the cross national pilot] • Media with 3D audio and no audio description but with audio subtitles [for the cross national pilot]	

7. ANNEX III - GENERAL QUESTIONNAIRE IN ENGLISH

General questions

1. Sex:
   a) Female
   b) Male
   c) Other
   d) I prefer not to reply

2. Age: _____

3. Main language: _____

4. Level of finished studies
   a) No studies
   b) Primary education
   c) Secondary education
   d) Further education
   e) University

5. I define myself as a…
   a) Blind person
   b) Low vision person
   c) Deaf person
   d) Hearing impaired person
   e) Deaf-blind person

6. Age in which your disability began:
   a) From birth
   b) 0-4
   c) 5-12
   d) 13-20
   e) 21-40
   f) 41-60
   g) more than 60

7. What technology do you use on a daily basis? You can select more than one.
   a) TV
   b) PC
   c) Laptop
   d) Mobile phone

e) Tablet

8. Do you have any device to access virtual reality content?
   a) Yes (If yes, which one? _____)
   b) No
   c) I don't know or I don't want to reply

9. Which of the following is your preferred device for watching online video content (i.e., Youtube, Vimeo, Netflix, Amazon Prime, broadcast catch up service etc.)? You can select more than one.
   a) PC
   b) Laptop
   c) Smartphone
   d) Tablet
   e) I don't watch online video content.
   f) Others (if so, please specify:_____)

10. **(only for visually impaired users)** Which of the following do you use on your connected devices to access the above content?
   a) Magnification (i.e. Zoomtext)
   b) Screen readers (i.e., JAWS, VoiceOver, TalkBack)
   c) Both
   d) None

11. **(only for visually impaired users)** Which of the following controls would you like to use with your screenreader /magnification tool when watching content online?
   a) Browse content library
   b) Identify content
   c) Functions such as play, stop, pause, forward, rewind
   d) Switch AD/AS on and off

8. ANNEX IV – FOCUS GROUP INSTRUCTIONS

T2.1. USER CENTERED DESIGN: INSTRUCTIONS ON HOW TO DEVELOP FOCUS GROUP

The aim of the focus group is to gather input from users (both end-users and professionals) on the needs, expectations and requirements for access services in immersive accessibility.

Previous to the focus group

1. Focus groups will ask questions departing from Pro-user-T2.1. compiled table and End-user_T2.1._compiled table. If any partners wants to add questions, please add them on Monday 21st November at the latest.
2. To prepare for your focus group, read the recommendations available here and write down a similar procedure: https://drive.google.com/drive/folders/1JRbnZegeccVuRJsBrqD_E3s4wh5WckEZ
3. Number of participants: a minimum of 5. Number of facilitators: 2.
4. Have the documentation concerning consent form and information sheet available in your language (see step 1 below)

During the focus group, please follow these steps:

1. Welcome participants and ask them to read (or read them out) the consent form and the information sheet, available here: https://drive.google.com/open?id=1G-qpAqo3tdFE34voFPY-chOPDlWMjkAq
   Version in English: T2.1._IMAC INFORMATION SHEET and consent_final.docx
   Version in Catalan: T2.1._IMAC INFORMATION SHEET and consent_final_CAT.docx
   Version in Spanish: T2.1._IMAC INFORMATION SHEET and consent_final_ES.docx

2. Ask participants to sign consent forms and information sheets.
   (Please keep the signed consent forms and hand in to Pilar in next meeting in Barcelona February 2018)

3. Ask participants to fill in the demographic information. Please notice that two questions do no apply to professional users without disabilities.
   https://drive.google.com/open?id=1G-qpAqo3tdFE34voFPY-chOPDlWMjkAq
   Version in English: T2.1. IMAC_GENERALQUESTIONNAIRE_FINAL.docx.
   Version in Catalan: T2.1. IMAC_GENERALQUESTIONNAIRE_FINAL_CAT.docx.
   Versión in Spanish: T2.1. IMAC_GENERALQUESTIONNAIRE_FINAL_SPA.docx.

4. Start focus group, following suggested structure and taking into account the list of questions provided. Perhaps you can have a photograph taken, or interview the participants for one of the the ImAc films we have to provide.

5. Summarise the results of the report and read them aloud to the participants, who have to approve them.

After the focus group

6. Send your focus group report approx. 01/12/2017. It must follow the template available here: https://drive.google.com/open?id=1TUjkFCnK3Dcx1x1HP2zoEB5UOJHs6lSL

   (T2.1.focusgroupreporting_v1, under FOCUS_GROUP_ALL folder)

7. Provide original consent forms to UAB (next February face-to-face meeting or send by snail mail).

9. ANNEX V – FOCUS GROUP TEMPLATE

Focus group template

1. Focus group general information

- Partner responsible for the workshop: *please indicate your acronym.*
- Place and date: *please indicate place and date.*
- Access service(s) discussed: *please indicate if SDH, SL, AD or AST.*

2. Participants profile

- Number of home users: *please indicate only number of home users (persons with disabilities).*
- Number of professional users and profile: *please indicate only number of professional users and profile (subtitlers, audio describers, sign language interpreters).*
- Demographics for users (*please notice that questions 5 and 6 may not be applicable): (please add your numbers where an X is found).*
    1. Sex: X "male", X "female", X "other", X "prefer not to reply".
    2. Age: *please specify the different ages*
    3. Main language of the participants: *please specify the different languages*
    4. Level of finished studies: X "no studies", X "primary education", X "secondary education", X "further education", X "university".
    5. "I define myself as a…" X "blind person", X "low vision person", X "deaf person", X "hearing impaired person", X "deaf-blind person".
    6. Age in which your disability began: X "From birth", X "0-4", X "5-12", X "13-20", X "21-40", X "41-60", X "More than 60".
    7. What technology do you use on a daily basis? You can select more than one. X "TV", X "PC", X "Laptop", X "Mobile phone", X "Tablet".
    8. Do you have any device to access virtual reality content? X "Yes" (*and please indicate which ones they name: XXXX*), X "No", X "I don't know or I don't want to reply".
    9. Please add results for questions 10 and 11 if you had blind and visually impaired users in your testing.

3. Workshop conclusions

3.1. Conclusions from professional users

Please summarise your results here, indicating user needs, expectations and recommendations, taking the compiled table as a reference.

3.2. Conclusions from advanced home users

a) Concerning the interaction: accessing the services

Please summarise your results here, indicating user needs, expectations and recommendations, taking the compiled table as a reference.

b) Concerning the services

Please summarise your results here, clearly identifying the access services you refer to, and identifying user needs, expectations and recommendations, taking the compiled table as a reference.

10. ANNEX VI – FOCUS GROUP QUESTIONS

Audio description activity:

1. visual input

We are going to show a 360º video. We kindly ask the audio describers in the room to produce a live audio description in turns addressed to the final users in the room.

Possible videos: http://www.uab.cat/web/tour-virtual-1345727862214.html

https://www.youtube.com/watch?v=FySF4ezUwqE

The aim of this activity is that audio describers indicate us the main challenges and how to address them in terms of production; end users indicate us the main challenges in terms of consumption, and both groups suggest us ideas on how virtual reality could be made accessible.

2. Sound possibilities

You are now going to listen to a specific type of audio. Close your eyes and listen carefully. At the end, we would like you to discuss if this type of audio could be used, and how, in providing audio description for virtual reality.

Possible audio input:

English
https://www.youtube.com/watch?v=IUDTlvagjJA

Spanish
https://www.youtube.com/watch?v=VoaUy0CiyCw

https://www.youtube.com/watch?v=s5UYmGQYZtQ

Spanish Audio subtitles
https://www.youtube.com/watch?v=XgcUyCKurHA

3. Questions
From the table below ask the following questions:
"What do you like about the 360?"
"If you could generate a 360 AD from scratch, what would you put in to make a better one?"
"What would make 360 AD more appealing to your peers?"
"Do you have any suggestions about 360AD?"

11. ANNEX VII – ETHICAL FORMS

INFORMATION SHEET

Project: ImAc (Immersive Accessibility)

Main researcher: Sergi Fernández (i2Cat)
Ethical adviser: Pilar Orero

The aim of the tests is to get feed-back on how access services can be implemented in immersive media. This will allow us to identify the needs of diverse audiences and research how the quality of experience and the quality of the service can be improved.

During the test, which can take various forms (experiment with questionnaire, focus groups, interviews, etc.), you will be asked to provide some demographic data. Then, you will be asked to watch an input, perform a task or give your opinion on various aspects. If needed, objective data will be recorded during the session. The researcher will give you more details of the specific test assigned to you and the data collection methods. Please ask as many questions as needed to clarify the procedure.

Virtual reality may produce some sort of discomfort such as virtual reality sickness when visualizing virtual reality contents, information will be provided and appropriate measures will be taken to guarantee the participants' safety and well-being. Immersive environments are not recommended for individuals with claustrophobia, heart conditions, back conditions, a history of seizures, epilepsy, and/or sensitivity to flashing lights. Also participants thought to be unstable or under the influence of drugs or alcohol will not be admitted.

In the case that some physiological or eye-tracking apparatus are used to gather data, you will not experience any discomfort, since the apparatus used are the latest generation and are not invasive.

Now please read the consent form.

CONSENT FORM (written version)

Project: ImAC (Immersive Accessibility)

Your participation in the tests is absolutely voluntary.

You can discontinue your involvement in the study at any time without prior justification. This shall have no repercussions or negative consequences of any sort.
In the case that some physiological or eye-tracking apparatus are used to gather data, you will not experience any discomfort, since the apparatus used are the latest generation and are not invasive.
Virtual reality may produce some sort of discomfort such as virtual reality sickness when visualizing virtual reality contents, information will be provided and appropriate measures will be taken to guarantee the participants' safety and well-being. Immersive environments are not recommended for individuals with claustrophobia, heart conditions, back conditions, a history of seizures, epilepsy, and/or sensitivity to flashing lights. Also participants thought to be unstable or under the influence of drugs or alcohol will not be admitted.
The information you provide will be used in the project but it will remain anonymous.
ImAc is a European project led by Sergi Fernández, from the company i2Cat. The ethical adviser responsible of ethical procedures is Pilar Orero. You can contact Pilar Orero at pilar.orero@uab.cat and ask for more information about the project and the project results.

The researcher administering the test is ((NAME and SURNAME)).
If you are willing to participate, please confirm the following statements by signing at the end of this document.

- I have read and understood the information given for this research or have had the information read to me,
- I have had the opportunity to ask questions about the research.
- I consent to take part in the research sessions.

_____

Name of the participant Date Signature

_____

Name of the researcher Date Signature

Signed by Pilar Orero (UAB IP ImAc)

<END OF DOCUMENT>