

Deliverable

Project Acronym:	IMAC
Grant Agreement number:	761974
Project Title:	<i>Immersive Accessibility</i>



D5.2-Pilot evaluation methodology and plan

Revision: 1.0

Authors: Anna Matamala and Pilar Orero (UAB)

Delivery date: M21 (28-06-2019)

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

Abstract:
 This document defines for each iteration the evaluation methodology to be followed, based on the requirements. It establishes what elements are evaluated and the calendar of evaluation activities.

REVISION HISTORY

Revision	Date	Author	Organisation	Description
0.1	22-01-2018	Pilar Orero	UAB	Table of Content
0.2	21-03-2018	Anna Matamala	UAB	Updated Table of Content
0.3	10-04-2018	Anna Matamala	UAB	First draft merged version
0.4	16-04-2018	Anna Matamala	UAB	Draft version incorporating partners' input
0.5	18-04-2018	Anna Matamala	UAB	Final draft version for review
0.6	27-04-2018	Anna Matamala	UAB	Final version after review by i2Cat and RBB (1 st iteration)
0.7	18-05-2019	Anna Matamala	UAB	Updated ToC for second iteration and first draft
0.8	13-06-2019	Anna Matamala	UAB	Final draft version for review (2 ⁿ iteration)
1.0	28.06.2019	Anna Matamala	UAB	Final version

Disclaimer

The information, documentation and figures available in this deliverable, is written by the IMAC – project consortium under EC grant agreement H2020-ICT-2016-2 761974 and does not necessarily reflect the views of the European Commission. The European Commission is not liable for any use that may be made of the information contained herein.

Statement of originality:

This document contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.

EXECUTIVE SUMMARY

This document describes the evaluation methodology for the pilots to be developed under WP5 (Demonstration pilots), namely the German Pilot (T5.3.), the Spanish Pilot (T5.4), and the Cross-national Pilot (T5.5), dealing with subtitling, sign language, and audio description/audio subtitling. Easy-to-read is also added as an innovative access service.

In WP2, User requirements (D2.2) and Platform specifications (D2.3) were described and defined. They are the departing points for this deliverable, enriched by the information provided by the reports from the technical development achieved in WP3 and WP4.

This deliverable specifies the methodology for the ImAc project pilot actions: the criteria for defining tests, their procedure, and the evaluation plan. It also identifies the elements to be tested by the two very different target users: the home user, and the professional user.

On the one hand, it defines the evaluation methodology to be followed in home user actions. It differentiates between tests performed during development --dealing both with interface interaction and service presentation modes-- and final pilot actions. The specificities of each pilot action are considered. On the other hand, it defines the evaluation methodology to be followed in professional user actions.

Three key concepts, developed in this document, sustain the methodological approach, which is mainly based on qualitative data: usability, presence, and preferences.

CONTRIBUTORS

First Name	Last Name	Company	e-Mail
Sven	Glaser	RBB	sven.glaser@rbb-online.de
Francesc	Mas	CCMA	fmas.z@ccma.cat
Anna	Matamala	UAB	Anna.matamala@uab.cat
Mario	Montagud	i2Cat	mario.montagud@i2cat.net
Pilar	Orero	UAB	Pilar.orero@uab.cat
John	Patton	RNIB	John.Paton@rnib.org.uk
Àngel	Quero	CCMA	aquero.f@ccma.cat
Sonali	Rai	RNIB	sonali.raai@rnib.org.uk
Doreen	Ritter	RBB	doreen.ritter@rbb-online.de
Zora	Schaerer	RBB	zora.schaerer@rbb-online.de

CONTENTS

Revision History.....	2
Executive Summary.....	3
Contributors.....	4
Tables of Figures and Tables.....	6
List of acronyms.....	7
1. Introduction.....	8
1.1. Purpose of this document.....	8
1.2. Scope of this document.....	8
1.3. Status of this document.....	9
1.4. Relation with other ImAc activities.....	9
2. Criteria for user testing.....	10
3. Evaluation plan.....	12
3.1. Elements to be evaluated.....	12
3.1.1. Services: home users.....	12
3.1.2. Tools: professional users.....	16
3.2. Summary.....	17
4. Evaluation methodology.....	18
4.1. Methodology to evaluate services during development: presentation modes.....	20
4.2. Methodology to evaluate services during development: interface.....	21
4.3. Methodology to evaluate services: implementation (German and Spanish pilots) ...	22
4.4. Methodology to evaluate services: implementation (cross-national pilot).....	22
4.5. Methodology to evaluate professional tools: development and implementation.....	23
5. References.....	24
Annex I – Demographic questionnaires.....	26
Annex II – System Usability Scale.....	40
Annex III – Igroup Presence Questionnaire (IPQ).....	41
Annex IV – Post-questionnaire: preference and opinions.....	43

TABLES OF FIGURES AND TABLES

Image 1: Diagram of relation between work packages, and its cycles (iterations).	8
Image 2: Diagram of tasks and its outcomes (deliverables).	9
Image 3: WP2-WP5 relationships.....	12
Image 4: WP5 pilots' diagram.	17

LIST OF ACRONYMS

Acronym	Description
AD	Audio description
AST	Audio subtitling
CAVE	Cave Automatic Virtual Environment
CCMA	Corporació Catalana de Mitjans Audiovisuals
D	Deliverable
ER	Easy-to-Read
HMD	Head-mounted display
HUR	Home User Requirements
ImAc	Immersive Accessibility
ICT-SOPI	ICT-Sense of Presence Inventory
IPQ	Igroup Presence Questionnaire
IRT	Institut für Rundfunktechnik
M	Month
RBB	Rundfunk Berlin-Brandenburg
RNIB	Royal National Institute of Blind People
SL	Sign Language
SUB	Subtitling
SUS	System Usability Scale
T	Task
UAB	Universitat Autònoma de Barcelona
USAL	University of Salford
VR	Virtual Reality
WP	Work Package

1. INTRODUCTION

This introduction describes the purpose of this deliverable, its scope, status and relationship with other ImAC activities.

1.1. Purpose of this document

This document establishes the evaluation methodology and plan for the demonstration pilots to be developed under Work Package 5 (WP5). Pilot actions are understood as any test in which ImAc services and products are demonstrated and feedback from users is gathered.

Defining a sound methodological approach and a feasible timeline is crucial to achieve the project aims satisfactorily. More specifically, defining an evaluation plan and a calendar of activities for the pilots (for each of its iterations) and for the cross-national is one of the WP5 objectives, tightly linked to the other project WPs (Image 1).

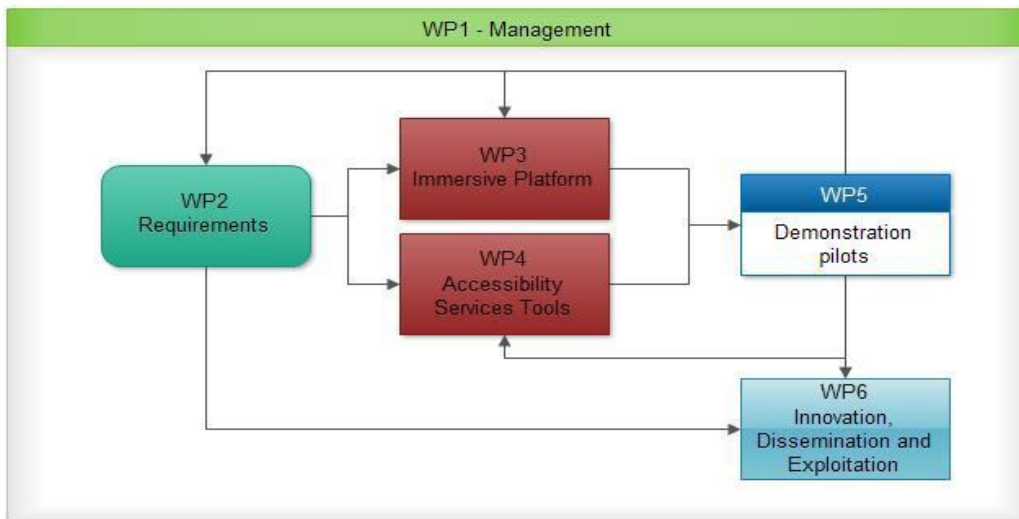


Image 1: Diagram of relation between work packages, and its cycles (iterations).

1.2. Scope of this document

ImAc follows a user-centric design (D2.1), which means that technological components, services and tools are developed based on user requirements. These developments are evaluated where appropriate in different iterations, guaranteeing the participation of end users at different stages in the process.

ImAc follows a methodology inspired by Harte et al. [1], which is based on:

- An explicit understanding of users, tasks, and environments.
- A permanent involvement of users throughout the design and development.
- A design driven and refined by user-centric evaluation activities.
- An iterative approach.
- A multidisciplinary team with different skills and perspectives.
- A clear but at the same time flexible approach to user testing.

This deliverable explains how user testing was derived from user requirements developed under WP2, differentiating between home user requirements, linked to services, and professional user requirements, linked to the tools. Criteria for selecting the specific tests are forward in Section 2. Section 3 presents the evaluation plan. On the one hand, it describes the

elements to be evaluated, differentiating between services for home users and tools for professional users. On the other hand, it provides a summary of all the phases.

Section 4 presents the specific methodology to be used when testing during development and in its final implementation, considering the specificities of each pilot and the different approaches to end user services and professional user tools.

1.3. Status of this document

This is the second iteration of D5.2 with delivery foreseen in M21. This document designs a broad methodological framework that needs to be adapted to each specific pilot action, based on the actual content produced. Specific instructions will be developed and added as annexes in D.5.4, so that both the final methodology for each action and the results can be consulted in the same document.

This document presents the methodological framework for the Spanish and German pilots and for the cross-national pilot. As indicated in the Grant Agreement, the Spanish and German pilots include two pilot phases, whereas the cross-national has one phase. For professional tools, two phases are generally considered regardless of the pilot they are assigned to.

1.4. Relation with other ImAc activities

D5.2. originates from T2.2. User Requirements, which feeds into T5.1. Execution and Evaluation Plan, and impacts on all WP5 tasks, as shown in Image 2.

D5.2. is closely related to D5.1. Pilot Operation Plan, and is the basis for the future T5.5. Evaluation, which will result in D5.4. Pilot evaluation report.

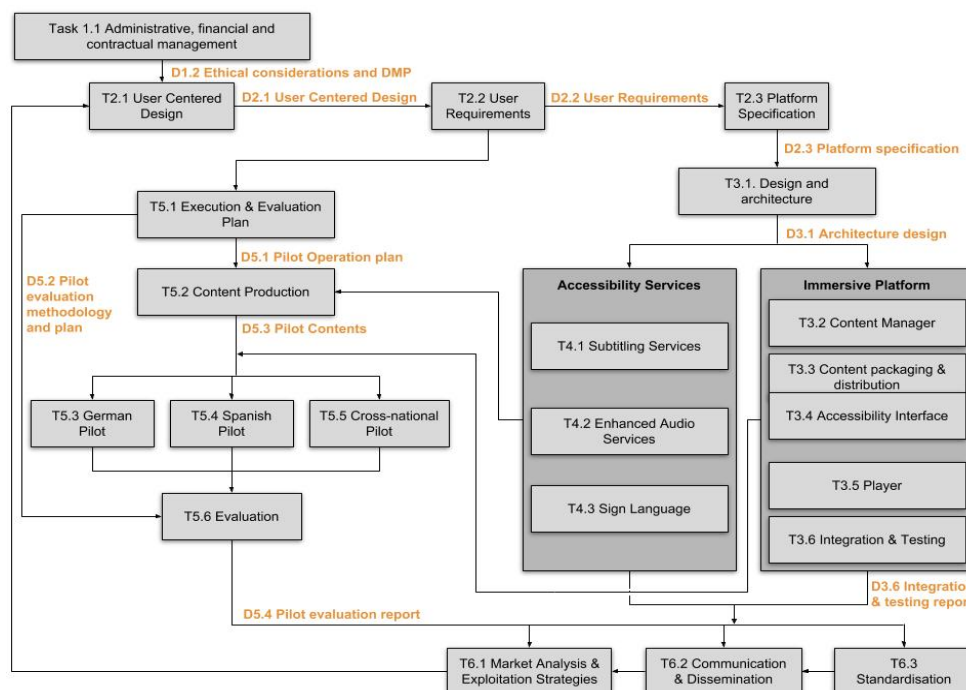


Image 2: Diagram of tasks and its outcomes (deliverables).

2. CRITERIA FOR USER TESTING

The user requirements established in ImAc in the interactions with users are described in detail in D2.2 User Requirements, and are the basis for defining testing criteria. User requirements are grouped under two conceptual categories:

- a. home user requirements concerning the services, and
- b. professional user requirements concerning the professional tools.

The home user requirements for all services can be clustered into four groups: 1) access to and control of service, 2) presentation modes, 3) personalisation options, and 4) content-related requirements.

The professional user requirements can be clustered into six groups 1) access to content, 2) signalization, 3) packaging/distribution, 4) ST (subtitling) editing tool, 5) SL (sign language) editing tool, 6) AD (audio description) editing tool.

The main criteria established to define user testing are described next.

1. Requirements referring to **professional tools**: they should be implemented as indicated by the professional users, according to the priorities given, if technically feasible. Tests for the AD editing tool, the SL editing tool and the ST editing tool should be planned, as well as for the Accessibility Content Manager (ACM).
2. In requirements referring to the **services**, there are two possibilities:
 - a) When users suggest only one implementation option, it should be implemented directly, if technically feasible (see also the prioritization assigned to the requirements that is described in D2.2).
 - b) When users suggest more than one implementation option, testing should be planned if technically and methodologically feasible.

A series of pre-pilot actions should be performed prior to the two main pilots, to narrow down the features to be implemented and tested.

3. Two main categories are identified for services testing purposes: **presentation modes** and **interface personalisation and interaction** (including access to and control of service, and personalisation options).

Testing for these two main broad categories happens at different stages of the project, depending on the service:

- **Subtitling**: presentation modes are tested in pre-pilot actions 1 (which is part of WP2). Within WP5, they are tested in pilot phase 1 and in pre-pilot actions previous to pilot phase 2. Interface personalisation and interaction is tested in pilot phase 1 (access) and in pre-pilot 2 (improved interface and personalisation). Pilot phase 2 tests the outcome of previous tests globally, including the improved presentation modes and interface features identified in the previous pilot actions.
- **Sign language**: presentation modes are tested in pre-pilot 1 (WP2) and pre-pilot 2 (WP5). Pilot phase 2 tests the outcome of previous tests globally, including the resulting presentation modes and interface features identified in the previous pilot actions.

- **Audio description and audio subtitling (AST)**, both in terms of presentation modes and interface personalisation and interaction are tested in the cross-national pilot, preceded by pre-pilot actions 1 (WP2) and 2 (WP5).

Regarding content-related requirements, small pilot actions can be planned where technically and methodologically relevant, following the methodological approach for the “presentation modes” category.

4. When technical partners identify that further specifications are needed, user testing may also be planned, if technically and methodologically feasible, following these criteria.

5. When requirements have already been tested in previous projects (for instance, colours, position, font size, font type in subtitling), no testing should be planned.

The demonstration pilots are divided into:

- **German Pilot** (M10-M28), executed by Rundfunk Berlin-Brandenburg (RBB) in cooperation with Institut für Rundfunktechnik (IRT). The focus is on subtitling and sign language. Regarding services, it includes one phase with small-scale lab tests with a dedicated group of users (pilot phase 1), and a second phase with large-scale deployment of services (pilot phase 2). Pilot phase 2 also includes an action with a group of selected users. Pre-pilot actions have been added. Regarding professional tools, it includes two phases, the first one evaluating a preliminary version of the prototype and the second one evaluating the final implementation.
- **Spanish Pilot** (M10-M28), executed by Corporació Catalana de Mitjans Audiovisuals (CCMA) in cooperation with i2CAT and Universitat Autònoma de Barcelona (UAB). The focus is on subtitling. Regarding services, it includes one phase with a panel of users (pilot phase 1), and a second phase that is part of a large-scale open pilot (pilot phase 2). Pilot phase 2 also includes an action with a group of selected users to assure a minimum feedback on Head-Mounted Display (HMD) use and accessibility functions usability. Pre-pilot actions have been added. Regarding professional tools, it includes two phases.
- **Cross-national Pilot** (M11-M28), executed by RNIB, in cooperation with University of Salford (USAL), UAB and CCMA, in which audio description (including audio subtitling where relevant) is tested. It is a fully collaborative pilot in English, Spanish and Catalan in which tests are conducted with user panels in a semi-open environment in one single phase, preceded by a pre-pilot action to test the methodology. Regarding professional tools, it includes two phases (except in the United Kingdom pilots where the professional tools are not being tested) as with the other ImAc tools.

Image 3 summarises the relationship between WP2 user requirements and WP5 testing.

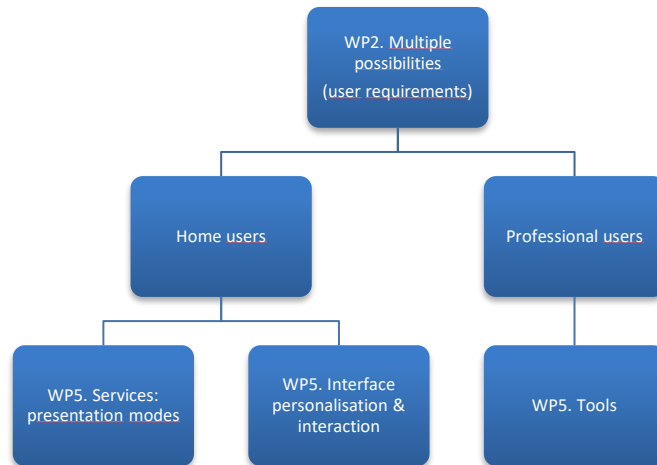


Image 3: WP2-WP5 relationships.

3. EVALUATION PLAN

The evaluation plan follows a qualitative approach, with quantitative indicators where relevant.

3.1. Elements to be evaluated

Based on the input from WP2 and the criteria established above, the following elements are identified as candidates for testing at different stages of the project development.

3.1.1. Services: home users

A number of different services are tested in the German, Spanish and cross-national pilots, as defined for the following service categories.

3.1.1.1. Subtitling

Subtitling is tested by home users at RBB as part of the German pilot and by CCMA as part of the Spanish pilot, namely in four distinctive set of actions: pre-pilot 1, pilot phase 1, pre-pilot 2, pilot phase 2. UAB also contributes with some additional intermediate actions, not initially planned.

Pre-pilot 1

As described in Section 2, a qualitative pre-pilot test was considered necessary before pilot phase 1, in order to narrow down the features that needed to be implemented and evaluated in pilot phase 1. Pre-pilot tests 1 considered three different presentation modes identified in the focus groups with home users, described below. It was carried out within the framework of WP2 with a small user group at RBB and CCMA in April and May 2018. The user requirements included in pre-pilot 1 were the following:

1. basic presentation mode (Home User Requirements HUR.02.27.0): the subtitles are presented in six field of view levels in a 16:9 ratio with the aim of identifying the comfort field of view to consume subtitles.
2. position icons with arrow and compass (HUR.02.31.1) and angular-based position mechanism (HUR.02.34.0): the user is guided to the speaker through three strategies: arrow, compass, and sided text.

The aim of pre-pilot 1 actions was to get feedback for improvements, an early rating of the different approaches and the definition of the comfort field of view to consume the subtitles. Pre-pilot tests also aimed to identify how the different requirements could be grouped in pilot phase 1, where the two best-rated modes were considered. The results of the pre-pilot 1 are reported in D2.1.

Pilot phase 1

The purpose of the actual pilot (German pilot and Spanish pilot) was to introduce a panel of users to the developed solution for consuming fully accessible subtitled 360° contents and, at the same time, to gather qualitative measurements and feedback about the user experience when consuming those services in an immersive environment.

These qualitative measurements took into consideration user feedback regarding:

- subtitle presentation modes, more specifically arrow versus radar, which were the ones that were best rated by users in pre-pilot 1 (HUR.02.31.1), and
- interface personalisation and interaction (i.e. access and control), when using the traditional menu in HMD and on a tablet (more details in D.3.5).

During pilot phase 1, the focus was on presentation modes and interface interaction (access to services).

Pre-pilot 2

The aim of pre-pilot 2 actions, led by RBB and CCMA, was to further refine some user requirements and further technological developments through qualitative testing, focusing on:

- subtitle presentation modes, more specifically on:
 - always visible subtitles with (enhanced version of the) arrow versus fixed positioned subtitles with (enhanced version of the) arrow (HUR.2.27.1), and
 - non-speech information presentation: emojis versus text (HUR.3.3.0);
- interface personalisation and interaction, namely the updated version of the traditional menu on the ImAc player, both in HMD and on a tablet.

An additional intermediate action led by UAB was also included in order to compare the current solutions found during the state-of-the-art review (subtitles implemented by *The New York Times* or BBC) and the ImAc solutions regarding subtitling. This action had two stages. In the first one, a reduced number of users tested subtitle presentation mode (always visible versus subtitles equally spaced in a fixed position) and guiding mechanisms (arrow versus autopositioning). Based on the results of this first stage, a second action aimed to test with a wider number of users the solutions developed in ImAc compared to the current solutions in the market and confirm that ImAc solutions are more immersive and preferred by users. More specifically:

- subtitle presentation mode: always visible subtitles (ImAc solution) versus subtitles equally spaced by 120° in a fixed position (current solutions found in *The New York Times* and BBC).

- subtitle presentation mode: arrow (ImAc solution) versus radar (solution found in The New York Time Virtual Reality (VR) Player (<https://www.nytimes.com/video/the-daily-360>)).

Pilot phase 2

During pilot phase 2, the panel of users will be wider. This will enable an open pilot to be run, where a bigger user group will have access to the contents and will enjoy the accessible immersive contents. Pilot phase 2 will allow not only gathering qualitative feedback, but also quantitative data about the user behaviour while using the services. It is planned to test the improved presentation modes and the player's user interface together with the personalisation options identified in the previous actions.

In the Spanish Pilot, four scenarios are planned:

- Open pilot for testing the consumption of immersive content with subtitles (device selected by user).
- Open pilot for testing the consumption of immersive content with subtitles on companion screens synchronously with standard content on a HbbTV 2.0.1 compatible SmartTV.
- Open pilot for testing the consumption of immersive content with subtitles in companion screens synchronously with standard content on a web-browser on a PC.
- Semi-open pilot with a group of selected users that will test the same content available in the open pilot in a controlled environment. The aim is to assure a minimum feedback on HMD use and on accessibility functions (subtitles).

In the German Pilot, two scenarios are planned:

- Open pilot for testing the consumption of immersive content with subtitles (device selected by user).
- Closed pilot with a group of selected users to assure a minimum feedback on accessibility functions (subtitles).

3.1.1.2. Services: Easy-to-Read subtitles

ImAc has proposed to transfer the guidelines of Easy-to-Read Language into subtitling in order to implement the simplification required by users under HUR.03.04.0. This innovative access service, not originally planned, has been tested with a panel of users made up of elderly people. Two types of subtitles have been tested:

- Subtitles as proposed in the other tests, without any additional language adaptation.
- Easy-to-Read subtitles: subtitles with a simplified language structure and lexicon, which aim to enhance comprehension.

3.1.1.3. Services: sign language interpretation

The sign language interpretation service will be tested by home users at RBB as part of the German pilot phase 2. Two set of pre-pilot actions have been conducted:

Pre-pilot action 1

As with the subtitling services, a pre-pilot test was carried out for the sign language interpretation service with a small group of home users at RBB in April 2018. The focus was on the evaluation of the different presentation modes and the definition of the comfort field of

view to consume the sign language interpretation. The user requirements included in the pre-pilots were the following:

- Basic presentation mode (HUR.02.18.0): the signer video is positioned at the bottom right in six field of view levels in a 16:9 ratio with the aim to identify the comfort field of view to consume sign language interpretation.
- Position icons (HUR.02.21.0), angular-based position mechanism (HUR.02.23.0), and forced perspective (HUR. 02.21.1): three strategies were tested to guide the user to the speaker. In the first one, the user is guided to the speaker with an arrow. In the second one, the user is guided by the position of the signer video window on the left or right edge. In the last one, the field of view is automatically changed by the video player so that the user can see the new active speaker. Afterwards, the user can change the direction he/she wants to look.

The results of this pre-pilot test are reported in D2.1.

Pre-pilot action 2

Pre-pilot actions with a reduced number of users were carried out for sign language interpretation service aiming at gathering users' feedback on different presentation modes for Sign Language, namely:

- presentation modes. Display of signer video: continuous versus non-continuous. The continuous signer is visible all the time while the non-continuous signer video is faded out when no text or dialogue was translated by the signer.
- presentation modes: Sign Language plus subtitles versus Sign Language only (suggestion by users in pilot 1).
- presentation modes: textual versus emoticons for speaker representation (HUR.2.21.0).

Interface personalisation and interaction was not tested in this pre-pilot action because this was part of the subtitle pre-pilot 2 test (see above) conducted with more than 20 users.

Pilot phase 2

During pilot phase 2, as with the subtitling test, the panel of users will be wider, and an open pilot will be run. A larger user group will have access to the contents and will enjoy the accessible immersive contents. Pilot phase 2 will allow gathering qualitative and quantitative data about the user behaviour while using the service. More specifically there will be:

- Open pilot for testing the consumption of immersive content with sign language (device selected by user)
- Closed pilot with a group of selected users to assure a minimum feedback on accessibility functions (sign language)

3.1.1.4. Services: audio description and audio subtitling

Audio description and audio subtitling will be considered as part of the cross-national pilot, in the UK (Royal National Institute of Blind People, RNIB) and Spain (UAB). As with the other services, pre-pilot tests were conducted with a small group of home users at RNIB and UAB.

Pre-pilot 1

The focus was on the evaluation of three different presentation modes, specifically related to the placement of the AD, in order to gather initial feedback. The user requirements included in the pre-pilots were the following:

1. AD placed on the action (HUR.03.40.1): audio description placed on or in the direction of the action, which after some terminological discussion has been termed “Dynamic AD”;
2. AD anchored to the soundscape (HUR.03.42.1): audio description anchored in the scene, in which a first-person narrative is used, which has been termed “Static AD”.
3. AD anchored to the head position, which is the standard procedure: audio description in which the “voice of God” is used, finally termed “Classic AD”.

The results of the pre-pilot test have been reported in D2.1.

Pre-pilot 2

The aim of this pre-pilot 2 was to test the methodology for the cross-national pilot concerning the following tests:

- Interface interaction, namely enhanced menu and voice interaction.
- AD presentation modes: Classic, Standard, Dynamic.
- Audio subtitling presentation modes: Classic versus Dynamic.

Cross-national pilot

The purpose of the pilot will be to introduce the target group, in this case persons with sight loss, to the proposed solutions for consuming fully accessible 360° content and gather feedback. It will aim to further understand user experience in relation to:

- using the ImAc player to access, play and control content (interface interaction and personalisation), and
- watching 360° content with AD and AST (presentation modes).

Tests will aim to elicit whether the ImAc player meets the core functional user requirements of blind and partially sighted people, and how accessible the interface is for people with varying degrees of sight loss, with specific emphasis on independently accessing, playing and controlling content. In this context, specific feedback on the significance of audible menus, voice control, and default settings within the ImAc environment is expected to be gathered. The cross-national pilot also aims to research how blind and partially sighted people evaluate the immersiveness of the test material in 360° content, and which type of AD of the ones presented in the pilot is preferred to maximize the quality of experience in a 360° environment. The expected outcome of this test is a technical evaluation of the accessibility of the ImAc player and insights into how people with sight loss respond to content in 360° when it is delivered with different types of audio description and also audio subtitling, if necessary.

3.1.2. Tools: professional users

The professional tools are tested at two stages of the project.

Tools are developed and enhanced as part of WP4, which aims to develop and provide the technological basis for the content production of accessibility services. More specifically, it aims to investigate and enhance existing tools dedicated to three main access services: a ST editing tool (D4.1), an AD editing tool (D4.2), a SL editing tool (D4.4), and an ACM (D3.2).

Tests are planned for:

- ACM, to be tested in a preliminary version by RBB and CCMA as part of pilot phase 1 and in its improved version as part of pilot phase 2.

- AD professional tool, to be tested in a preliminary version by UAB as part of pilot phase 1, to gather input on its development, and once it is fully developed as part of pilot phase 2.
- ST professional tool, to be tested in a preliminary version by UAB as part of pilot phase 1, to gather input on its development, and to be tested by broadcasters RBB and CCMA once it is fully developed as part of pilot phase 2.
- SL editing tool, to be tested by RBB as a professional tool in one single iteration, once it is fully developed.

3.2. Summary

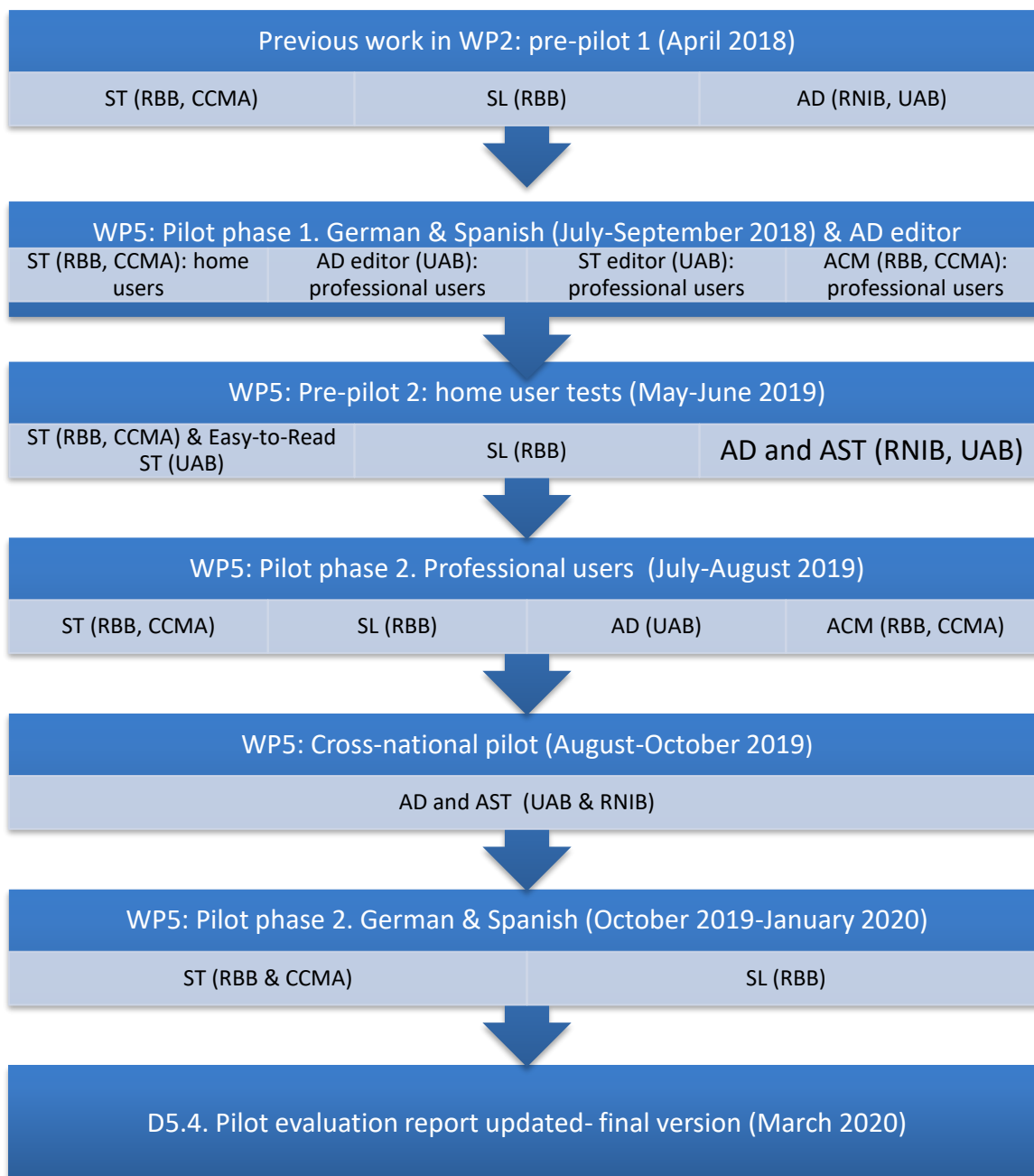


Image 4: WP5 pilots' diagram.

4. EVALUATION METHODOLOGY

The evaluation approach differentiates between:

- Methodology for home user actions:
 - During development: service presentation modes (section 4.1.),
 - During development: interface personalisation and interaction (section 4.2.),
 - In its final implementations, differentiating between the German/Spanish pilots (section 4.3) and the cross-national pilot (section 4.4), as they have different stages.
- Methodology for professional user actions: development and implementation (section 4.5).

For each scenario, the following items are defined: measures, participants, materials, experimental protocol, and reporting.

The methodological approach in pre-pilot actions and pilot phase 1 is formative, meaning that the testing is not aimed to assess a finished product or service, but to assess an on-going technological development by means of qualitative testing and/or to test the methodology.

Methodologies in ImAc follow a user-centric approach. The approach in WP5 is both behavioural and attitudinal, meaning that participants will perform tasks in which different measures will be assessed through testing, but they will also be asked for opinions or preferences. In all cases, the emphasis will be put on qualitative data, rather than on quantitative approaches.

Some aspects in all evaluation processes are shared, so they are included here.

1. This document provides a general framework, but specific instructions will be issued per pilot action/test. These instructions will be included as annexes in D5.4.

2. All testing should follow ethical procedures, already approved under WP1 in the project. This includes signing an information sheet and consent form in an accessible format.

3. Prior to each pilot action/test, the planned methodology and instructions should be tested by the partner responsible for carrying out the pilot action/test.

4. Testing in a controlled environment is recommended for pilot phase 1, pre-pilots and cross-national pilot.

5. Regarding participants, for each specific pilot action/test, recruiting criteria will be established. However, their profiling will be carried out through a questionnaire (see Annex 1), which will gather data on three main aspects:

- Personal characteristics linked to demographic information.
- Behavioural categories linked to participants' actual usage of the technologies and /or services being tested.
- Attitudinal categories linked to participants' opinions prior to the test on the technologies and /or services being tested.

This questionnaire can be shortened where relevant for a pilot action.

6. The measures that will be generally considered are **usability, presence, and preferences/opinions**, with differences depending on each pilot action/test.

There are a myriad of methods to test user experience [2], [3], [4] in terms of effectiveness, efficiency and satisfaction, such as cognitive walkthroughs, icon usability testing, contextual inquiry or online surveys, to name just a few. ImAc focuses on **usability**, understood as the ability of the user to use a thing or to carry out a task successfully, and uses one of the widest known scales, namely the System Usability Scale (SUS), available here: <https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html>

SUS includes 10 items (see Annex II). It is easy to administer and provides reliable results with small sample sizes and has been validated and used in many investigations as a standard self-reported metrics [5].

The second measure considered in ImAc is **presence**, sometimes also referred to as immersion, which can be understood as the sense of “being there” [6], “perceptual illusion of non-mediation” [7], “psychological sense of immersion in any mediated environments” [8], “the experience of being engaged by the representation of a virtual world” [9], “experiential quality metric employed to evaluate broadcast and virtual environment media systems” [10], to name just a few definitions of this concept. Measuring presence when users are faced with different presentation modes in virtual environments is especially relevant, as it is possible to evaluate whether a presentation mode enhances or decreases immersion.

Presence measures allow assessing the entire experience of audiences [11] and they also prove to be a successful tool for measuring the emotional response of target users in previous research conducted in the field of Media Accessibility [12]. However, presence is a multi-construct concept that encompasses many dimensions or subcategories, with variations depending on the authors. To measure presence, questionnaires are the most common method for multiple reasons: they are specifically aimed to measure the specified concept, they are reliable and valid, non-intrusive and cheap to implement, and results are easy to analyse.

After a review of different standard questionnaires on presence such as the ICT-Sense of Presence Inventory (ICT-SOPI) [10], Slater-Usuh-Steed Presence Questionnaire [13] or the Igroup Presence Questionnaire (IPQ) [14], [15], IPQ [16] (see Annex III) has been chosen for various reasons. First of all, it combines previous questionnaires and it was the first one to specifically differentiate between spatial presence, involvement and experienced realism. In this questionnaire, spatial presence refers to the sense of being there in the virtual environment; involvement refers to the attention to the virtual environment, and experienced realism refers to the reality judgment of the virtual environment. The questionnaire has been validated in different forms of virtual environments (users of VR or CAVE-like systems, desktop VR, players of 3D games and text-based virtual environments), including HMD in a laboratory, a situation similar to ImAc. It is available in English, German, Dutch, French, and Japanese. IPQ includes 14 questions, making it an adequate length for experimental purposes, especially in pre-pilot actions and pilot phase 1. It can be accessed online: <http://www.igroup.org/pq/ipq/download.php>

The third measure is **preferences and general opinion**, for which specific questions in the form of post-questionnaires will be developed “ad hoc” for each test. The questionnaire relies mainly on open-ended questions with some room for participants to explain their responses. A sample of possible closed and open questions is presented in Annex IV.

This general framework needs to be adapted to each specific test. Additionally, the analysis of behavioural data recorded through data logs will be considered in the final open tests.

4.1. Methodology to evaluate services during development: presentation modes

This sub-section provides a summary of the methodological framework used to evaluate presentation modes in accessibility services. Please notice that testing for presentation modes and the interface personalisation and interaction can be combined in one single testing session in the specific instructions that were developed, depending on the final stimuli and experimental conditions.

Measures: presence, preferences, and usability.

Participants: end users, specific profile to be defined depending on the access service and their needs.

A within-subject approach is favoured, meaning that all participants are exposed to all conditions in their specific test. This has an impact on the stimuli for the experiments.

Materials: HMD (alternatively, smartphone and VR glasses, depending on the test), player (together with all the involved server-based and delivery technologies), and two comparable clips (A, B) with 2 conditions (1, 2) for each presentation mode tested, so that participants can watch both test conditions in different but comparable clips to avoid a learning effect.

When different presentations modes need to be assessed in one test, there should be different clips used in each one to avoid a learning effect by users.

Content needs to be suitable for the test conditions and was produced/acquired as part of T.5.2. A minimum length was sought in line with existing immersive content. Furthermore, the material needed to meet specific requirements depending on the presentation modes tested. For example, in the case of the guiding mechanisms, the content needed to involve several speakers at different angles. Taking these considerations into account, possible test material for the subtitling tests was: “I, Philip” short film, “Romeo and Juliette” opera excerpt, “Desconcert” recordings, “Rapzember” radio interview at Radio Fritz, RBB. For AD and AST tests, “Opera” and “Holy Land” were considered. The public version of the player where some of these materials are available can be found here: <https://imac.gpac-licensing.com/player/>

Experimental protocol:

- Welcome the participants.
- Ethical procedures: information about the project, signing informed consent form.
- Pre-questionnaire administration.
- Actual testing: watching stimuli (information on order of presentation below).
- Post-questionnaire in digital format.
- Discussion and thanks.

The order of presentation of the stimuli should be balanced across participants. Two sample Latin square protocols are presented next, but specific ones will be adapted for the instructions of each test.

Example 1. Latin square protocol in which one presentation mode is tested with two conditions (for instance, guiding user to speaker with arrows versus guiding user to speaker with compass)

Stimuli: clip A1, clip A2, clip B1, clip B2.

Order of presentation (please repeat up to the number of agreed participants):

- Participant 1: A1-B2
- Participant 2: A2-B1

- Participant 3: B1-A2
- Participant 4: B2-A1

Example 2. Latin square protocol in which two presentation modes are tested with two conditions each (for instance, guiding user to speaker with arrows versus guiding user to speaker with compass + comfort view X versus comfort view Z).

Stimuli: clip A1, clip A2, clip B1, clip B2// clip C1, clip C2, clip D1, clip D2. Order of presentation (to be repeated under the number of participants is reached):

- Participant 1: A1-B2 / C1-D2
- Participant 2: A2-B1 / C1-D2
- Participant 3: A1-B2 / C2-D1
- Participant 4: A2-B1 / C2-D1
- Participant 5: B2-A1 / C1-D2
- Participant 6: B1-A2 / C1-D2
- Participant 7: B2-A1 / C2-D1
- Participant 8: B1-A2 / C2-D1
- Participant 9: A1-B2 / D2-C1
- Participant 10: A2-B1 / D2-C1
- Participant 11: A1-B2 / D1-C2
- Participant 12: A2-B1 / D1-C2
- Participant 13: B2-A1 / D2-C1
- Participant 14: B1-A2 / D2-C1
- Participant 15: B2-A1 / D1-C2
- Participant 16: B1-A2 / D1-C2

Reporting: results will be included in a reporting template that will be developed once the final instructions per test are available.

4.2. Methodology to evaluate services during development: interface

Measures: usability and preferences.

Participants: end users (same profile as in section 4.1., when evaluating presentation modes).

The number of participants per usability test is not pre-established and is highly dependent on the testing conditions. It is recommended that users taking part in the evaluation of the presentation modes are also involved in the interface evaluation, which can include interface personalisation and/or interaction (access), depending on the pilot and the pilot phase. A single testing session may be planned to optimise resources.

Materials: HMD (alternatively smartphone and VR glasses, depending on the test), player (together with all the involved server-based and delivery technologies), and content to be viewed.

Experimental protocol: users will be asked to perform certain tasks and then report on the usability, using the agreed questionnaire. The specific tasks will be included in the instructions.

- Welcome to the participants.
- Ethical procedures: information about the project, signing informed consent form.
- Pre-questionnaire administration.
- Actual testing: performing tasks.

- Post-questionnaire in digital format.
- Discussion and thanks.

Reporting: results will be included in a reporting template that will be developed once the final instructions per test are available.

4.3. Methodology to evaluate services: implementation (German and Spanish pilots)

Measures: usability and preferences.

Participants: for the German and Spanish pilots: open pilots are addressed to all types of users. Although users who are consumers of access services will be reached, access services are expected to benefit other types of users, hence the profile will be wider, including persons with diverse capabilities.

A pilot in a controlled environment will also be executed in both German and Spanish pilots with a reduced group of users to assure a minimum feedback on HMD use and on accessibility functions.

Materials: for the German and Spanish pilots, the final ImAc player needs to be integrated in the CCMA/RBB websites, with accessible 360° videos. Additionally, an ImAc subsection in the CCMA HbbTV application to access to the ImAc contents, and a specific webpage -with the final ImAc player integrated- to play synchronized 360° contents on the second screen.

Experimental protocol: German and Spanish pilots: users will be asked to use the ImAc services and then fill out a short questionnaire. The general framework is as follows:

- Welcoming page, with ethical procedures (consent form) and brief pre-questionnaire (demographics).
- ImAc services player use, during which quantitative usage data will be obtained.
- Post-questionnaire to gather qualitative feed-back on usability and preferences.

The specific instructions and questionnaires will be developed once the final version is implemented.

Reporting: results will be included in a reporting template that will be developed once the final instructions are available.

4.4. Methodology to evaluate services: implementation (cross-national pilot)

The cross-national pilot takes a different approach from the German and Spanish pilots, as it is only developed in one stage as a semi-open pilot in a controlled environment. It aims to take, as indicated in the Grant Agreement, a first approach towards using audio description and audio subtitle in immersive environments.

The aim of the pilot actions will be to test the presentation modes for audio description, for audio subtitling, and the user interface.

Measures: usability, presence, and preferences.

Participants: the aim is to involve 30 users with sight loss per country to obtain statistically relevant results.

Materials: for the cross-national pilot, the user interface should be available both with voice interaction and enhanced menu. Content which allows to test the three AD presentation modes (Classic/Static/Dynamic) and the two AST presentation modes (Classic/Dynamic), both in the UK and in Spain, is needed.

Experimental protocol:

For the user interface test, users will be asked to perform certain tasks and then report on the usability, using the agreed questionnaire. The specific tasks will be included in the instructions.

- Welcome to the participants.
- Ethical procedures: information about the project, signing informed consent form.
- Pre-questionnaire administration.
- Actual testing: performing tasks.
- Post-questionnaire in digital format.
- Discussion and thanks.

For the presentation modes tests, the procedure will be as follows:

- Welcome the participants.
- Ethical procedures: information about the project, signing informed consent form.
- Pre-questionnaire administration.
- Actual testing: watching stimuli (information on order of presentation in 4.1.)
- Post-questionnaire in digital format.
- Discussion and thanks.

The previous actions can be combined in one testing session to facilitate user participation.

Reporting: results will be included in a reporting template that will be developed once the final instructions are available.

4.5. Methodology to evaluate professional tools: development and implementation

The tools are evaluated in two iterations. Once WP2 requirements are implemented, a first round of tests should be performed on a preliminary version of the tool. The feed-back gathered should then be used to improve the editors, which should be tested in a second and final round. Previous to these final tests, broadcasters (CCMA and RBB) should perform an internal evaluation action to test whether the tool is fully functional and ready for a test with external professional users.

Measures: usability, preferences.

Participants: professional users. The project aims at involving around 30 professional users. In the second iteration the same professional users will be contacted, and new ones will be added to reach as many users as possible.

Materials: editing tool ready, user guide, and 360° videos to use in the tasks.

Experimental protocol: users will be asked to perform certain tasks and then report on the usability and preferences, using the agreed questionnaire. The tasks, and corresponding preference questions, will be developed once the editors are available.

Reporting: results should be included in a reporting template to be developed once the final instructions per test are available.

5. REFERENCES

1. System Usability Scale (SUS). [Online]. n.d. <https://www.usability.gov/how-to-and-tools/methods/system-usability-scale.html>
 2. Igroup presence questionnaire (IPQ). [Online]. n.d. <http://www.igroup.org/pq/ipq/download.php>
- [1] R. Harte, L. Glynn, A. Rodríguez-Molinero, P. M. Baker, T. Scharf, L.R. Quinlan, & G. ÓLaighin, "A Human-Centered Design Methodology to Enhance the Usability, Human Factors, and User Experience of Connected Health Systems: A Three-Phase Methodology", *JMIR Human Factors*, 2017. <http://humanfactors.jmir.org/2017/1/e8/>
- [2] T. Tullis & W. Albert, *Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics*, Morgan Kaufmann, 2013.
- [3] J. Pannafino & P. McNeil, *UX Methods: A Quick Guide to User Experience Research Methods*, CDUXP LLC, 2017.
- [4] E. Goodman, M. Kuniavsky & A. Moed, *Observing the User Experience: A Practitioner's Guide to User Research*, Morgan Kaufmann, 2012.
- [5] J. Brookes, "SUS: A Retrospective", *Journal of Usability Studies*, vol. 8, n. 2, 2013, 29-40.
- [6] B. R. Reeves, "Being there: Television as symbolic versus natural experience", Unpublished manuscript, Stanford University, Institute for Communication Research, Stanford, CA, 1991.
- [7] M. Lombard & T. Ditton, "At the Heart of It All: The Concept of Presence", *Journal of Computer-Mediated Communication*, vol. 3, n. 2, 1997.
- [8] L. Fryer & J. Freeman, "Presence in those with and without Sight: Audio Description and its potential for virtual reality applications", *Journal of Cyber Therapy & Rehabilitation*, vol. 5, n. 1, 2012, 15–23.
- [9] D. Jacobson, "On theorizing presence", *Journal of Virtual Environments*, vol. 6, n. 1, 2002.
- [10] J. Lessiter, J. Freeman, E. Keogh & J. Davidoff, "A Cross-media Presence Questionnaire: The ITC-Sense of Presence Inventory", *Presence: Teleoperators, and Virtual Environments*, vol. 10, n. 3, 282–297, 2001.
- [11] A. Walczak, "Immersion in audio description. The impact of style and vocal delivery on user's experience", Phd thesis, Universitat Autònoma de Barcelona, 2017.

[12] N. Wilken & J. L. Kruger, "Putting the audience in the picture: mise-en-shot and psychological immersion in audio described film", *Across Languages and Cultures*, vol. 17, n. 2, 2016, 251–270.

[13] J. Van Baren & W. IJsselsteijn, "Measuring Presence: A Guide to Current Measurement Approaches", Deliverable of the OmniPres project IST-2001-39237, 2004. <http://www8.informatik.umu.se/~jwworth/PresenceMeasurement.pdf>

[14] M. Slater & M. Usoh, "Representations systems, perceptual position, and presence in immersive virtual environments", *Presence*, vol. 2, n. 3, 1993, 221–233.

[15] B. G. Witmer & M. J. Singer, "Measuring Presence in Virtual Environments: A Presence Questionnaire", *Presence Teleoperators & Virtual Environments*, vol. 7, n. 3, 1998, 225–240.

[16] T. Schubert, F. Friedmann & H. Regenbrecht, "The experience of presence: Factor analytic insights", *Presence: Teleoperators and Virtual Environments*, vol. 10, n. 3, 2001, 266–281.

ANNEX I – DEMOGRAPHIC QUESTIONNAIRES

		General questionnaire – Professional users
<i>This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 761974</i>		

Pre-questionnaire addressed to professional users

(Only for the second iteration) 0. This is the second ImAc subtitle editor test. Did you take part in the first ImAc subtitle/AD editor test?

- a) Yes
- b) No

1. Sex

- a) Female
- b) Male
- c) Other
- d) I prefer not to reply

2. Age: _____

3. Main language: _____

4. Please describe your current job: _____

5. Have you ever subtitled/audio described/interpreted a 360° video?

- Yes
- No

6. For how long have you been working in the field of audio description/subtitling/sign language interpretation? _____

7. How many hours of audio description/subtitling/sign language interpretation have you produced in your professional life?

- a) Less than 50 hours
- b) 51-150 hours
- c) 151-300 hours
- d) More than 300 hours

8. In what language or languages do you normally audio describe/subtitle/interpret? _____

9. What software do you normally use when subtitling / writing AD/ interpreting for films, TV programs or other videos?

ONLY FOR AD: Do you record your AD?

- a) Yes
- b) No
- c) Other

10. Please indicate your level of studies.

- a) Primary education
- b) Secondary education
- c) Further education. Please specify _____
- d) University. Please specify _____

11. If you have received specific training on audio description/subtitling/sign language interpretation, please indicate it here. _____

12. What devices do you use on a daily basis? Multiple replies are possible.

- a) TV
- b) PC
- c) Laptop
- d) Mobile phone
- e) Tablet
- f) HMD
- g) Other: _____

13. How often do you watch virtual reality content (for instance, 360° videos)?

	Never	Occasionally	At least once a month	At least once a week	Every day
In smartphone					
On a tablet					
On a PC					
In smartphone plugged to HMD					
In HMD					

14. If you have never used virtual reality content such as 360° videos or only occasionally, please indicate why. Multiple answers are possible.

- a) Because I am not interested.
- b) Because it is not accessible.
- c) Because I have not had the chance to use it.
- d) Other reasons. Please explain: _____

15. Please state your level of agreement with the following statement: "I am interested in virtual reality content (such as 360° videos)."

- a) I strongly agree
- b) I agree
- c) Neither agree nor disagree
- d) Disagree
- e) Strongly disagree

16. Do you own 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

18. If you replied "yes" to the previous question, please specify which device(s).

		T1.2. General questionnaire – deaf and hard-of-hearing (oralist)
<p><i>This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 761974</i></p>		

Pre-questionnaire addressed to home users (subtitling)

1. Sex:

- a) Female
- b) Male
- c) Other
- d) I prefer not to reply

2. Age: _____

3. Main language: _____

4. Please indicate your level of studies.

- a) No studies
- b) Primary education
- c) Secondary education
- d) Further education
- e) University

5. I define myself as a...

- a) Deaf person
- b) Hearing impaired person
- c) Deaf-blind person
- d) Other: _____

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 devices do you use on a daily basis? Multiple replies are possible.

- a) TV
- b) PC
- c) Laptop
- d) Mobile phone
- e) Tablet
- f) Head Mounted Display (HMD)
- g) Other: _____

8. How often do you watch virtual reality content (for instance, 360° videos)?

	Never	Occasionally	At least once a month	At least once a week	Every day
In smartphone					
On a tablet					
On a PC					
In smartphone plugged to HMD					
In HMD					

9. If you have never used virtual reality content such as 360° videos or only occasionally, please indicate why. Multiple answers are possible.

- a) Because I am not interested.
- b) Because it is not accessible.
- c) Because I have not had the chance to use it.
- d) Other reasons. Please explain: _____

10. Please state your level of agreement with the following statement: "I am interested in virtual reality content (such as 360° videos)."

- a) I strongly agree
- b) I agree
- c) Neither agree nor disagree
- d) Disagree
- e) Strongly disagree

11. Do you own 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

12. Do you like watching the following types of content on television or online?

	I like it very much	I like it	Neither like it nor dislike it	I don't like it	I don't like it at all
News					
Fiction (series, films)					
Talk shows					
Documentaries					
Sports					
Cartoons					

13. When subtitling is available, do you activate it for the following type of content?

	Always	Sometimes	Rarely	Never
News				
Fiction (series, films)				
Talk shows				
Documentaries				
Sports				
Cartoons				

14. If it is available and you do not activate it, please select the reasons why

- a) Because the interface is not accessible.
- b) Because I don't want subtitling in all the content, only in certain types of content.
- c) Other: _____.

15. How many hours a day do you watch subtitled content?

- a) None
- b) Less than 1 hour

- c) 1-2 hours
- d) 2-3 hours
- e) 3-4 hours
- f) 4 hours or more

16. What do you use subtitles for?

- a) They help me understand
- b) They are my only way to have access to the dialogue
- c) I use them for language learning
- d) Other: _____



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 761974

Pre-questionnaire addressed to home users (sign language)

1. Sex:

- a) Female
- b) Male
- c) Other
- d) I prefer not to reply

2. Age: _____

3. Main language: _____

4. Please indicate your level of studies.

- a) No studies
- b) Primary education
- c) Secondary education
- d) Further education
- e) University

5. I define myself as a...

- a) Deaf person
- b) Hearing impaired person
- c) Deaf-blind person
- d) Other: _____

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 devices do you use on a daily basis? Multiple replies are possible.

- a) TV
- b) PC
- c) Laptop
- d) Mobile phone
- e) Tablet
- f) Head Mounted Display (HMD)
- g) Other: _____

8. How often do you watch virtual reality content (for instance, 360° videos)?

	Never	Occasionally	At least once a month	At least once a week	Every day
In smartphone					
On a tablet					
On a PC					
In smartphone plugged to HMD					
In HMD					

9. If you have never used virtual reality content such as 360° videos or only occasionally, please indicate why. Multiple answers are possible.

- a) Because I am not interested.
- b) Because it is not accessible.
- c) Because I have not had the chance to use it.
- d) Other reasons. Please explain: _____

10. Please state your level of agreement with the following statement: "I am interested in virtual reality content (such as 360° videos)."

- a) I strongly agree
- b) I agree
- c) Neither agree nor disagree
- d) Disagree
- e) Strongly disagree

11. Do you own 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

12. Do you like watching the following types of content on television or online?

	I like it very much	I like it	Neither like it nor dislike it	I don't like it	I don't like it at all
News					
Fiction (series, films)					
Talk shows					
Documentaries					
Sports					
Cartoons					

13. When sign language is available, do you activate it for the following type of content?

	Always	Some times	Rarely	Never
News				
Fiction (series, films)				
Talk shows				
Documen taries				
Sports				
Cartoons				

14. If sign language interpretation is available and you do not activate it, please select the reasons why:

- a) Because the interface is not accessible.
- b) Because I don't want sign language in all the content, only in certain types of content.
- c) Other: _____.

15. How many hours a day do you watch sign language interpreted content?

- a) None
- b) Less than 1 hour

- c) 1-2 hours
- d) 2-3 hours
- e) 3-4 hours
- f) 4 hours or more

16. How many hours a day do you watch SL interpreted content?

- a) None
- b) Less than 1 hour
- c) 1-2 hours
- d) 2-3 hours
- e) 3-4 hours
- f) 4 hours or more

17. What do you use sign language interpretation for?

- a) It helps me understand
- b) It is my only way to have access to the dialogue
- c) I use it for language learning
- d) Other

18. What do you use SL interpretation for?

- a) They help me understand
- b) They are my only way to have access to the dialogue
- c) I use them for language learning
- d) Other

		T1.2. General questionnaire – blind and visually impaired
<p><i>This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 761974</i></p>		

Pre-questionnaire addressed to home users (AD/AST)

1. Sex:

- a) Female
- b) Male
- c) Other
- d) I prefer not to reply

2. Age: _____

3. Main language: _____

4. Please indicate your level of 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) Partially-sighted person
- c) Other: _____

6. Age at which you started losing your sight:

- a) From birth
- b) 0-4
- c) 5-12
- d) 13-20
- e) 21-40
- f) 41-60
- g) more than 60

7. What devices do you use on a daily basis? Multiple replies are possible.

- a) TV
- b) PC
- c) Laptop
- d) Mobile phone
- e) Tablet
- f) HMD
- g) Other: _____

8. How often do you watch virtual reality content (for instance, 360° videos)?

	Never	Occasionally	At least once a month	At least once a week	Every day
In smartphone					
On a tablet					
On a PC					
In smartphone plugged to HMD					
In HMD					

9. If you have never used virtual reality content such as 360° videos or only occasionally, please indicate why. Multiple answers are possible.

- a) Because I am not interested.
- b) Because it is not accessible.
- c) Because I have not had the chance to use it.
- d) Other reasons. Please explain: _____

10. Please state your level of agreement with the following statement: "I am interested in virtual reality content (such as 360° videos)."

- a) I strongly agree
- b) I agree
- c) Neither agree nor disagree
- d) Disagree
- e) Strongly disagree

11. Do you own 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

12. Do you like watching the following types of content on television or online?

	I like it very much	I like it	Neither like it nor dislike it	I don't like it	I don't like it at all
News					
Fiction (series, films)					
Talk shows					
Documentaries					
Sports					
Cartoons					

13. How many hours a day do you watch audio described content?

- a) None
- b) Less than 1 hour
- c) 1-2 hours
- d) 2-3 hours
- e) 3-4 hours
- f) 4 hours or more

14. What do you use to access online content?

- a) Magnification (for example, Zoomtext)
- b) Screen readers (for example, JAWS, VoiceOver, TalkBack)
- c) Both
- d) None

ANNEX II – SYSTEM USABILITY SCALE

© Digital Equipment Corporation, 1986.

	Strongly disagree				Strongly agree
1. I think that I would like to use this system frequently	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	1	2	3	4	5
2. I found the system unnecessarily complex	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	1	2	3	4	5
3. I thought the system was easy to use	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	1	2	3	4	5
4. I think that I would need the support of a technical person to be able to use this system	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	1	2	3	4	5
5. I found the various functions in this system were well integrated	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	1	2	3	4	5
6. I thought there was too much inconsistency in this system	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	1	2	3	4	5
7. I would imagine that most people would learn to use this system very quickly	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	1	2	3	4	5
8. I found the system very cumbersome to use	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	1	2	3	4	5
9. I felt very confident using the system	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	1	2	3	4	5
10. I needed to learn a lot of things before I could get going with this system	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
	1	2	3	4	5

ANNEX III – IGROUP PRESENCE QUESTIONNAIRE (IPQ)

This is the official version of the IPQ. Its formatting will be simplified to make it easier for end users to answer it.

Number	PQI/II Nr. (internal)	IPQ item name	shortcut	loading on ...	English question	English anchors	Copyright (item source)
1	s62	G1	sense of being there	PRES	In the computer generated world I had a sense of "being there"	not at all--very much	Slater & Usoh (1994)
2	s44	SP1	sense of VE behind	SP	Somehow I felt that the virtual world surrounded me.	fully disagree--fully agree	IPQ
3	s30	SP2	only pictures	SP	I felt like I was just perceiving pictures.	fully disagree--fully agree	IPQ
4	s28	SP3	not sense of being in v. space	SP	I did not feel present in the virtual space.	did not feel--felt present	
5	s31	SP4	sense of acting in VE	SP	I had a sense of acting in the virtual space, rather than operating something from outside.	fully disagree--fully agree	IPQ
6	s33	SP5	sense of being present in VE	SP	I felt present in the virtual space.	fully disagree--fully agree	IPQ
7	s64	INV1	awareness of real env.	INV	How aware were you of the real world surrounding while navigating in the virtual world? (i.e. sounds, room temperature, other people, etc.)?	extremely aware--moderately aware--not aware at all	Witmer & Singer (1994)
8	s37	INV2	not aware of real env.	INV	I was not aware of my real environment.	fully disagree--fully agree	IPQ
9	s40	INV3	no attention to real env.	INV	I still paid attention to the real environment.	fully disagree--fully agree	IPQ
10	s38	INV4	attention captivated by VE	INV	I was completely captivated by the virtual world.	fully disagree--fully agree	IPQ

11	s48	REAL1	VE real (real/not real)	REAL	How real did the virtual world seem to you?	completely real--not real at all	Hendrix (1994)
12	s7	REAL2	experience similar to real env.	REAL	How much did your experience in the virtual environment seem consistent with your real world experience ?	not consistent--moderately consistent-very consistent	Witmer & Singer (1994)
13	s59	REAL3	VE real (imagined/real)	REAL	How real did the virtual world seem to you?	about as real as an imagined world--indistinguishable from the real world	Carlin, Hoffman, & Weghorst (1997)
14	s47	REAL4	VE wirklich	REAL	The virtual world seemed more realistic than the real world.	fully disagree--fully agree	IPQ

ANNEX IV – POST-QUESTIONNAIRE: PREFERENCE AND OPINIONS

Some examples of open and closed questions are provided next. They will be further developed for each test in the specific instructions to be developed. The final instruction will be included in D.5.4.

Tests with professional users:

1. What did you like most about the (tool name)?
2. What did you like less about the (tool name)?
3. What do you think could be improved, and how?
4. Did you miss any functionality? If yes, which one?
5. If (specific feature) easy to use? Explain why.
6. Was (specific feature) useful for you? Explain why.
7. Do you think it will take you longer to (subtitle/interpret/AD) videos in 360°? Why?
8. Do you think 360° videos will impact your work as an (audio describer/SL interpreter/subtitler)?
9. Other comments

Tests with home users: presentation modes

1. What system/presentation mode do you prefer for (condition being tested)? a/b. If more than 2, ranking.
2. Please, explain why you prefer the above indicated option/ Please explain why you ranked the AD modes in this way.
3. Please explain why you did not choose the other option in question 1).
4. What do you think (service/feature) could be improved, and how?
5. How easy was it to understand/find the subtitles/find the speaker (etc.) in (condition)? (Likert scale)
6. Did you like (specific feature)? Yes/No.
7. What system would you prefer? A/B/C.
8. Explain your choice in the previous question.
9. Other comments:

Tests with home users: interface interaction

1. What do you like the most about the menu interaction?
2. What do you like less?
3. What do you think could be improved and how?
4. Other comments:

<END OF DOCUMENT>