REPORT

Validation operation plan

- project deliverable D7.1

Authors: Pilar Orero / Estel·la Oncins / Pau Pamplona

Confidentiality: Public

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 825585.

HELIOS Validation operation plan: Procedures for a coordinated validation

Project name	Grant agreement #
HELIOS	825585

Author(s) & Contributors
Pilar Orero, Estel·la Oncins, Pau Pamplona, Anna Matamala. Elena Parra, Carolina Caride, Chryssanthi Iakovidou, Carlos Alberto Martin, Rolf Nyffenegger, Robin Ribback, Javier Marín, David H. Falagán, Ville Ollikainen.

Pages
1+38

Reviewers:
Katri Grenman (VTT)
Kristina Kapanova (TCD)

Keywords
Coordination, Validation activities, lab tests, trials, pilots.

Deliverable identification code
D7.1

Summary
This deliverable is part of the work performed in WP7- Piloting, Validation and Evaluation. More precisely, it describes the coordination procedures conducted within task 7.1 - Coordination through the first months of the project (M4-M8, so far). The main objective of this task is to coordinate all activities within HELIOS related with the validation of the services, features, technological developments generated in WP3, WP4 and WP5 and integrated in WP6. With this purpose, it has been defined a structure of procedures and resources to coordinate HELIOS partners and external actors (i.e. entities and communities of users) that will be involved in the preparation, execution and validation activities to be performed during the project.

Confidentiality	Public

Barcelona, 30/08/2019

Written by: Pilar Orero (IP); Estel·la Oncins (Researcher); Pau Pamplona (PM)

VTT's contact address
Ville Ollikainen, ville.ollikainen@vtt.fi, +358 400 841116

Distribution
HELIOS project partners, subcontractors, the Project Officer and HELIOS web site

Contents

1. Executive Summary

This deliverable is part of the work performed in WP7- Piloting, Validation and Evaluation. More precisely, it describes the coordination procedures conducted within task 7.1 - Coordination through the first months of the project (M4-M8 so far). The main objective of this task is to coordinate all activities within HELIOS related with the validation of the services, features, technological developments generated in WP3, WP4 and WP5 and integrated in WP6. With this purpose, a structure of procedures and resources has been defined to coordinate HELIOS partners and external actors (i.e. entities and communities of users) that will be involved in the preparation, execution and validation activities to be performed during the project.

The deliverable is structured around 8 sections (9 including this summary). Section 2 provides a definition of concepts that will be used throughout the document to homogenise terminology, and to facilitate its reading and understanding. This is necessary since some concepts might be synonyms in normal contexts, and they might lead to confusion. For instance, trials, tests, and pilots have different meaning here, although they can be used indistinctly as they have similar meaning. Section 3 introduces the project overall structure and WP objectives. WP7 interacts with different work packages to gather information to define the validation activities (WP2, WP6), but also picks up technological developments to evaluate them (WP3-6).

Section 4 gives a summary of the Agile methodology adopted in the whole project, and how this will be integrated in WP7. Section 5 provides a detailed description of the overall validation plan. This section focuses first on the events that will take place, how and when they will happen and how partners will be coordinated. Then, an initial calendar of activities and workflow are also provided, to visualise how WP7 activities are aligned with the overall project progress. The last subsection is called Validation Activities procedures, which explains how validation activities are organised. This subsection is divided into the following three phases - before, during, and after - that are needed to run the validation activities. Section 6 presents the three types of validation activities defined in HELIOS to validate the deployments with users. Lab tests, trials and pilots have been proposed to help in understanding how users perceive the HELIOS social network and how their services, features and interaction can be improved. Section 7 outlines the validation methodologies that will be followed according to a) the characteristics of each validation activity (Lab Test, Trials and Pilots), b) the use case and c) the scope.

Section 8 is a high-level description of the validation methodologies foreseen to be used in WP7. This section was created to not overload the previous one with repetitive explanations or mentions of methodologies which will be common in the different validation activities and UC. A comprehensive description of what they are based on and why they have been selected is provided. Section 9 closes this document with conclusions and further steps on how task 7.1 will be continued.

At the end of the document, three initial templates are provided as annexes. These templates will be crucial to define the validation activities (Annex I), to describe how the evaluation will be conducted on each one of them (Annex II), and to report and provide feedback to the design and development teams (Annex III). Since the project is still in at an early stage, most of this content is susceptible to slight modifications to ensure a correct execution of WP7 and a proper interaction between the design, development and validation partners of the project.

2. Concepts definition

Before reading this deliverable, it is important to clarify the terminology that will be used throughout the document.

Use Case (UC): Actions performed by one or various users. In HELIOS there are three of them and they are used to help in designing the experience, and defining the technologies involved.

Validation activity (VAx): Set of activities organised to validate a Use Case.

UC validation activities: Each activity has a set of actions that users have to accomplish to generate data to provide feedback regarding one or more objectives. These actions are called tasks, and sometimes, a task can have sub-tasks for more specific feedback.

Hierarchy tree:

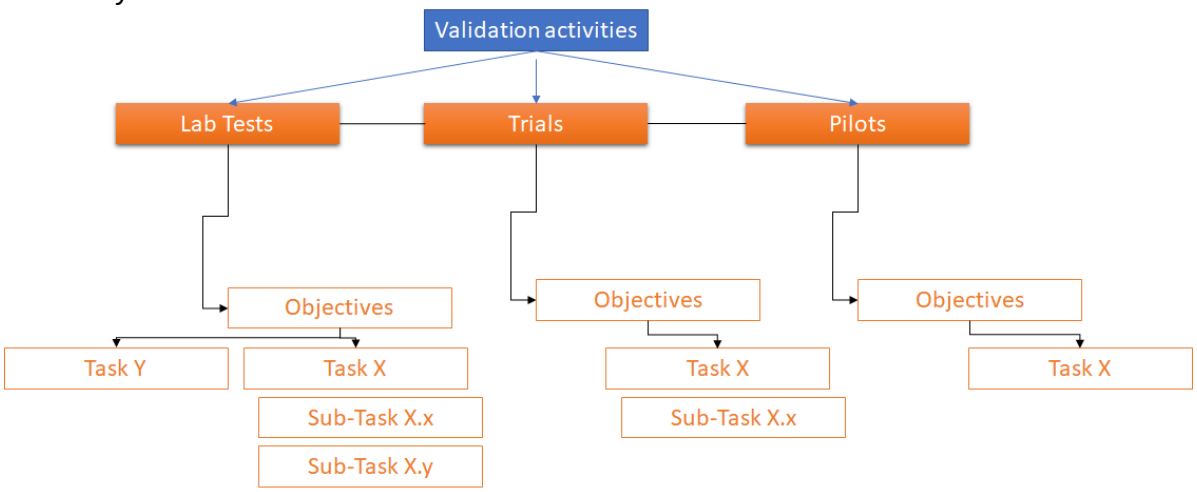

Figure 1. HELIOS validation hierarchy tree[1]

Technical Components: Technology developed (or used) in HELIOS to build the app.

Elements: Characteristics that define a UC in a moment of time.

Working group: Group of partners involved in a particular Validation activity.

Lab test(s), Trial(s), Pilot(s): These are different typologies of activities to evaluate and validate HELIOS technological developments. They differentiate each other in several aspects, like context, volume of users involved, duration, objectives, etc.

[1] If necessary, sub-tasks might be defined.

3. Introduction

Based on the task definition provided in the DoA (or EUGA Technical Annex), **T7.1 Coordination:** *Validation is an ambitious objective for this proposal. It should* ***trigger refinement of user requirements****, help towards a* ***better design,*** *and* ***check on integration and user interaction****. For this project we have decided to identify different validation stages (Tests, Trials and Pilots) to match technology maturity and robustness. The different validation modalities have also been designed to offer quick results, avoiding long development or integration cycles. Finally, to curb efforts when open pilots are organized and coordinated, a low PM partner has taken the lead. This will allow for a coordinated realistic validation in the many and varied university*[2] *activities. This task will* ***focus on the coordination of work among the different tests, trials, and pilots*** *and the smooth collaboration among all consortium partners. It is the responsibility of this task to* ***organize reporting of all issues of T7.1 to T7.4 back to WP6****.*

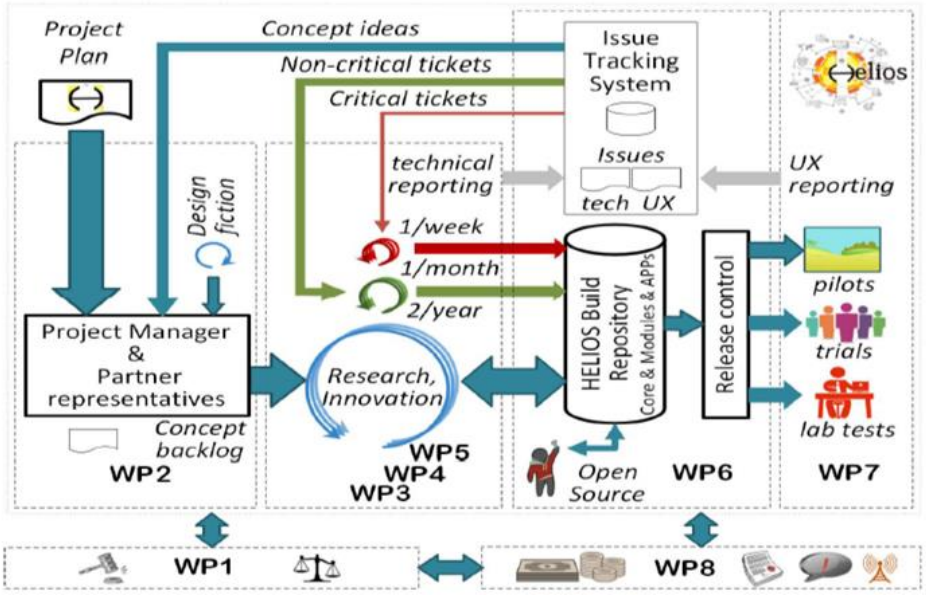

Figure 2. WP workflow and its dependencies within HELIOS

Departing from this description, we can summarise it as follows:

- WP2 will provide the initial Use Case (UC) descriptions to help to understand the activities and context in which HELIOS services must be tested;
- WP6 (the WP integrating all technical components of HELIOS) & WP7 (Validation activities) will get aligned to provide an initial development and deployment calendar to visualise what and when HELIOS services can be tested;
- Later user and technical requirements (WP2) will be used as inputs to refine the evaluation parameters to be used for validation purposes.

As shown in figure 2 of this deliverable, a visual representation of how validation will be coordinated is provided, with first inputs on how it will be conducted. It is complemented with a high-level description of the validation methodology that will be used. Following an agile methodology, all this initial information will be later refined with further and newer inputs, which are necessary to define the specificities of tests, trials and pilots, and the validation material.

[2] During the project progress, it has been decided that trials and pilots will take place in other locations, besides the university.

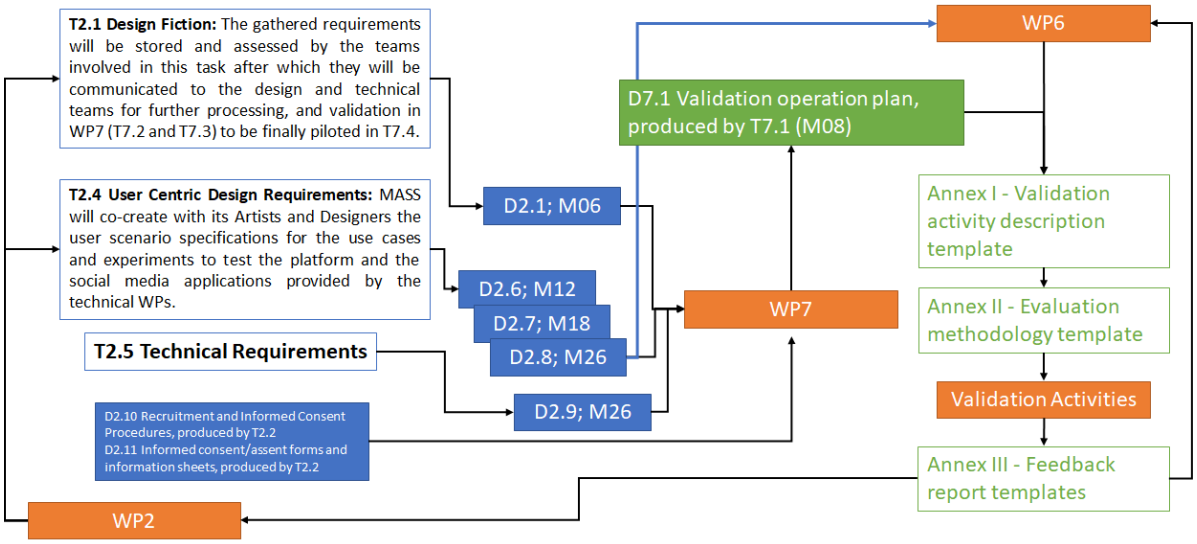

Figure 3. Work packages interaction workflow

At this time of the project there are still details to be defined, for instance, about which technologies will be deployed in each UC, and features and services that will be part of them, etc. For all these reasons and following an Agile approach (see section 4), T7.1 will be constantly interacting with WP2 and WP6 to setup the most adequate processes for HELIOS validation phase, and to contribute providing feedback to improve initial results.

4. The Agile methodology

4.1 Agile in HELIOS

During the past ten years, agile software development methodology has become increasingly popular. By promoting adaptive planning, one of the main advantages of agile development is coping with uncertain environments with fast changing requirements. The methods are no longer used solely in software development, but also in research and other complex processes. Agile methods aim at avoiding pre-defined, overly regulated and micro-managed development. Instead, productive flexibility comes from regular communications between different stakeholders, such as customers/end users, developers and business people, bringing together experts and teams from different fields. Main principles of agile development are collected into a respective Manifesto by Kent et al. (2011).

Since HELIOS is a multidisciplinary project, it is essential to have frequent and efficient interaction between the different approaches. Equally important for a three-year project is to keep ideas fresh, when the world around is rapidly changing: It is not only technology that develops, but also end-users' values, such as attitudes towards privacy and ownership of personal data are likely to evolve. Therefore, it is of paramount importance to keep user studies feeding information to agile platform development.

For this purpose, HELIOS has adopted a SCRUM[3]-like process to:

- provide full transparency between the participants, and
- keep innovation alive throughout the project.

[3] SCRUM is a framework used for software development and is one of the different Agile methodologies. SCRUM fosters frequent and regular communication among teams. https://www.agilealliance.org/glossary/scrum/

Another advantage of an agile process is, in contrast to the traditional waterfall model[4] (requirements -> design -> implement -> evaluate), that it allows flexibility to refocus activities, when necessary. For instance, in practise it often happens that a minor technical feature turns out to be far more complex than originally thought. Pure waterfall does not give any leeway: features, irrespective of their eventual importance, would be implemented whatever it takes. In an agile process, cost/benefit ratios can be evaluated in the course of normal development, enabling efficient use of limited resources.

Figure 2 illustrates the overall process with a feedback loop from 'Piloting, Validation and Evaluation Work Package' (WP7) to technical bug fixing ("tickets" to technical Work Packages 3, 4 and 5) and to 'Concept Development Work Package' (WP2) as "concept ideas". In the middle, WP6 takes care of system integration and operation, as well as the ticketing system, in order to process the feedback efficiently.

In the core of the process is a list of concept related ideas, called a 'Backlog'. Every four months project members prioritize the list for guiding the implementation in the Technical Work Packages. These four-month periods are called 'Research & Innovation Cycles'.

Within each Research & Innovation Cycle, technical Work Packages work on their own, having one-month *Sprints*[5] for actual implementation. These Sprints are transparent to all project members, so that all participants can keep up-to-date on the project status and communicate accordingly.

4.2 Applying Agile in validation activities

Figure 4 below rolls the process into a linear timeline, illustrating how each Research and Innovation Cycle builds knowledge and software on top of previous cycles. It is important to understand that outcomes of these cycles in an agile process are not predetermined. This is once again in contrast to the traditional waterfall design which would keep the results static from the early stages of the project. Cumulative outcomes of an agile process stay fresh.

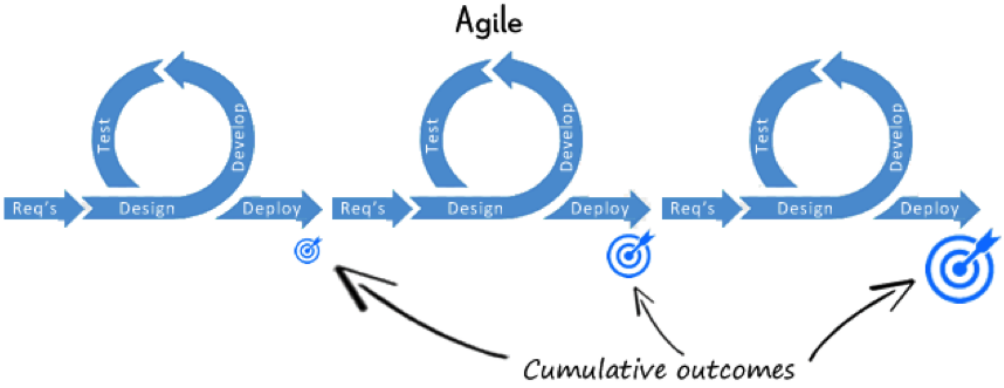

Figure 4. Graphic representation of the Agile Methodology[6]

[4] https://www.softwaretestinghelp.com/what-is-sdlc-waterfall-model/
[5] https://www.atlassian.com/agile/scrum/sprints
[6] http://www.crmsearch.com/agile-versus-waterfall-crm.php

5. Overall validation plan

Three different types of validation activities, lab tests, trials and pilots, will take place to validate HELIOS outcomes in different locations, in different ecosystems (i.e. university campus, art school, tech company, etc.), in different moments, and with different purposes. This diversity is necessary to validate the flexibility and adaptability of HELIOS as a social network and its usability with different target users and their requirements.

Despite this broad approach, a common ground for the evaluation has to be given in order to:

- maximise the efficiency of the activity coordination (i.e. prepare a calendar of activities, engaging users);
- provide coherent feedback to HELIOS design and development teams;
- to minimise risks (i.e. technology not ready to be tested with users).

In the following sections an overall vision of the structure and organisation mechanisms necessary to coordinate these activities is provided. The document is complemented with a description of the workflows to run validation activities (organisation, execution and validation), and how they will be validated (methodologies). It is important to understand that validation involve external users and entities, some of them people with no direct link with HELIOS. This means that WP7 not only coordinates HELIOS partners, but all those involved at some point in any of the activities that will take place between M09 and M34 of the project.

5.1 General approach

The overall process for validating can be summarised as follows: First a set of user requirements have defined the implementation of HELIOS. Once these requirements turned into technological developments, WP7 takes them and organises different activities to validate users' expectations with what has been delivered. The strategy for coordinating all validation activities is provided below.

The whole process starts gathering feedback from WP2 through D2.1 to understand WHAT (use cases) will be tested and HOW (requirements) the testing has to be done. WP2 is responsible for providing the use cases (UC), which are helpful to illustrate a more specific definition of the services to be deployed and consequently, how and where they can be tested. At this point, is it important to make a first comment regarding the scope of the different validation activities: Pilots and trials share a similar approach (i.e. to validate HELIOS with real users), but lab tests have a different objective.

UPV will lead lab tests which are context-agnostic. Their main focus will be in the User Interface (UI), and the tests will be conducted in supervised and closed conditions. Lab tests are fully aligned with the other tasks, and the workflow for preparing, executing and reporting validations actions will remain the same for all partners.

Trials and pilots instead take place in real locations, with real users. Thus, those initial UCs selected in WP2 need to be translated from "fiction" into a new narrative guiding real testing. This translation will also involve technical partners to assess the feasibility of bringing the UCs into the "real world" (from a technical perspective). This joint effort is also useful to identify potential constraints or risks of deployment. Some other considerations are:

- It is important to understand that trials and pilots will try to stay as close as possible to the original UC definition when validating them, although the primary objective is to validate the services and functionalities. UCs are basically "inspirational" stories to guide this process.

- There will be a partner responsible for each validation activity, supported by technical and design teams. The responsible partner will identify the UC defining elements: target users, duration, recommended actions for testing, best evaluation resources to be adopted, etc. Once it is clear that all expected elements are available and the activity is considered as *feasible*, the organisation process of the activity will start.

- Each UC will have services to be tested, thus all participants involved in WP7[7] will support the validation activities at different stages depending on the UC being executed. Although most activities will take place in Barcelona, other locations such as Switzerland, or Finland will be involved (some others might be added in the following months). As the project evolves, a more consolidated calendar of activities will be provided to cope with HELIOS' needs at that moment, with the resources available.

- For each UC there will be a definition of objectives (based on its specifications and requirements) to be validated. Tests, trials and pilots have to provide new inputs to help improving what has been released and re-evaluate it in the following iterations. This deployment-validation-feedback sequence will be done incrementally: first testing basic functionalities and adding new features until achieving more complex and richer features.

- HELIOS is modular, so some developments will be tested with similar or different purposes in the UCs.

- All deployment and execution of validation activities will be reported, first with individual reports to be delivered to WP6, and later compiled in D7.2, D7.3 and D7.4 (all of them to be delivered in M34, intermediate versions will be prepared for reporting/justification purposes).

5.2 Coordination methodology and resources

HELIOS WP7 seeks to provide a testing ground for the validation of technological developments that will be delivered via WP6 (Integration). The results of the evaluation will be reported back to WP6 to iteratively feed them and give new insights to help improve the developments. The technical validation will take place in WP6. Instead, WP7 aims at the validation of the user interface and the experience of the user while using HELIOS (UX, functionalities, usability, etc. see section 7 for further details). The basis of all this validation are the UCs defined in WP2. Three initial UCs have been defined in D2.1 - Concept design to guide the technical definition and implementation. Since HELIOS services will be based, but not limited to these UC, validation will be conducted in similar contexts.

A total of 12 partners are involved in this WP, plus all those from WP6, WP5, WP4 and WP3 (technical teams) with some relevance to WP7, which will get involved when needed.

In this section a description of the approach can be thought as an organisation structure and mechanisms) to align all partners involved is given with these final objectives:

a) to execute all validation activities as planned;

b) to provide relevant feedback to the technical teams;

c) to identify potential issues that risk the execution of the activities and propose remedial actions to overcome them.

With that purpose in mind, Task 7.1 will implement a set of actions to coordinate all activities and players to smoothly (and effectively) execute all validation activities. It is important to

[7] UAB, VTT, ATOS SPAIN SA, CERTH, ISMB, Nagoon AB, SWISS TXT, TCD, UNIPI, UPV, UH and WLI; MASS subcontracted by VTT.

remind that in this document an initial plan of activities will be provided, but since the project is still at a very early stage, this might change and adapt to new conditions not foreseen earlier.

The main bodies and resources set within T7.1 to organise partners and activities are:

1. **WP coordination meetings (in short, WPCM)**. After the submission of this deliverable (M08), monthly meetings will be scheduled for all WP7 partners. When necessary, the frequency of meetings might be adjusted and increased. During these meetings partners will be informed about previous activities, and next steps to be followed. All this information is compiled in action points (APs[8]) to facilitate tracking and progress. This approach is really useful to identify delays, or difficulties that might incur bigger risks of failure or bottlenecks.

2. **Activity coordination meetings (in short, ACM)**. Each validation activity can be understood as a small independent project. Thus, it is necessary to plan activities, define the different roles of partners involved, milestones, execute activities and report them. The general coordination meetings will create sub-working groups per activity. The working group will meet regularly to have specific discussions on the activity. To not overload partners with meetings, WPCMs and ACMs can overlap, taking advantage of the already selected time slots. These meetings will be used to decide the best strategy for the deployment and coordination of efforts.

3. **WP mailing list**. To ease communication among partners a dedicated mailing list will be created just to deal with topics related to WP7. It is critical to minimise the traffic of emails, and to connect only those recipients to whom the information is relevant. It could be dangerous in busy periods of the project to overload partners with too many messages and exchange of information, where this could trigger mistakes and interrupted communications.

4. **Coordination space**. TEAMS by Microsoft is the main platform used in HELIOS to coordinate the different work packages of the project. A dedicated space for WP7 and its different activities will be created. Partners will be reminded to check pending APs and milestones, compiled in an online document. In this space, there will be a calendar of activities and all the reporting templates.

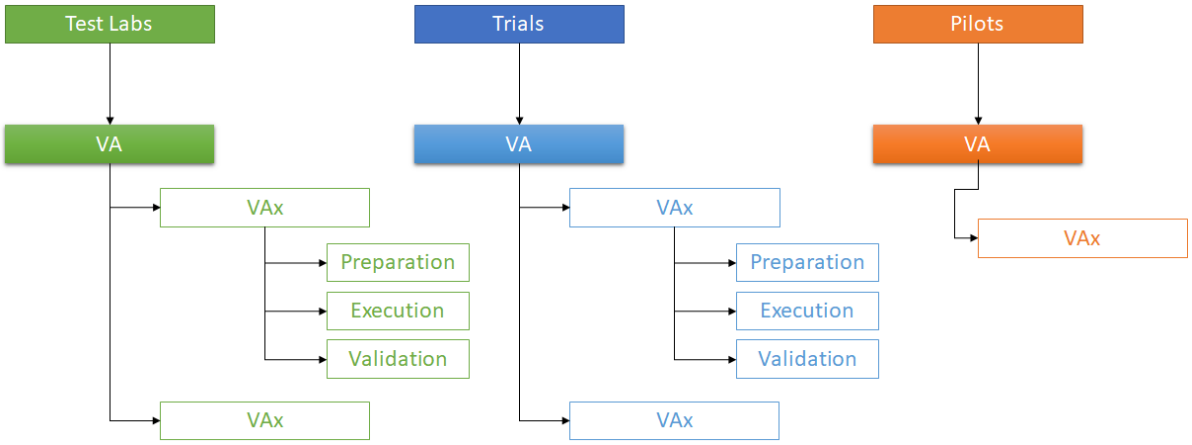

Figure 5. Folder structure to be used in TEAMS.

One of the critical aspects of this task is to ensure regular and effective communication, especially between WP7 and WP6.

[8] Action Points, or APs, are activities that partners have to achieve within a due date.

5.3 Initial calendar of tests, trials and pilot

An initial calendar of activities (see Figure 6) will be provided in this document, including preliminary time periods for possible validation. This calendar will only be used as a base plan to coordinate all activities that will start in M09 with the lab tests. Since HELIOS is still in its early stages, and due to the adopted Agile methodology, some flexibility will be given regarding the time scope to validate each UC. This calendar is based on potential activities, organised by project partners, matching with the UC provided in D2.1. Technical factors such as technology maturity or degree of integration might influence the modification of this calendar. Other factors like users' availability, or slight deviations in the development plan could also have an impact on the final schedule. For all these reasons, WPCMs and ACMs are necessary to ensure that all elements needed for validation are available and in time.

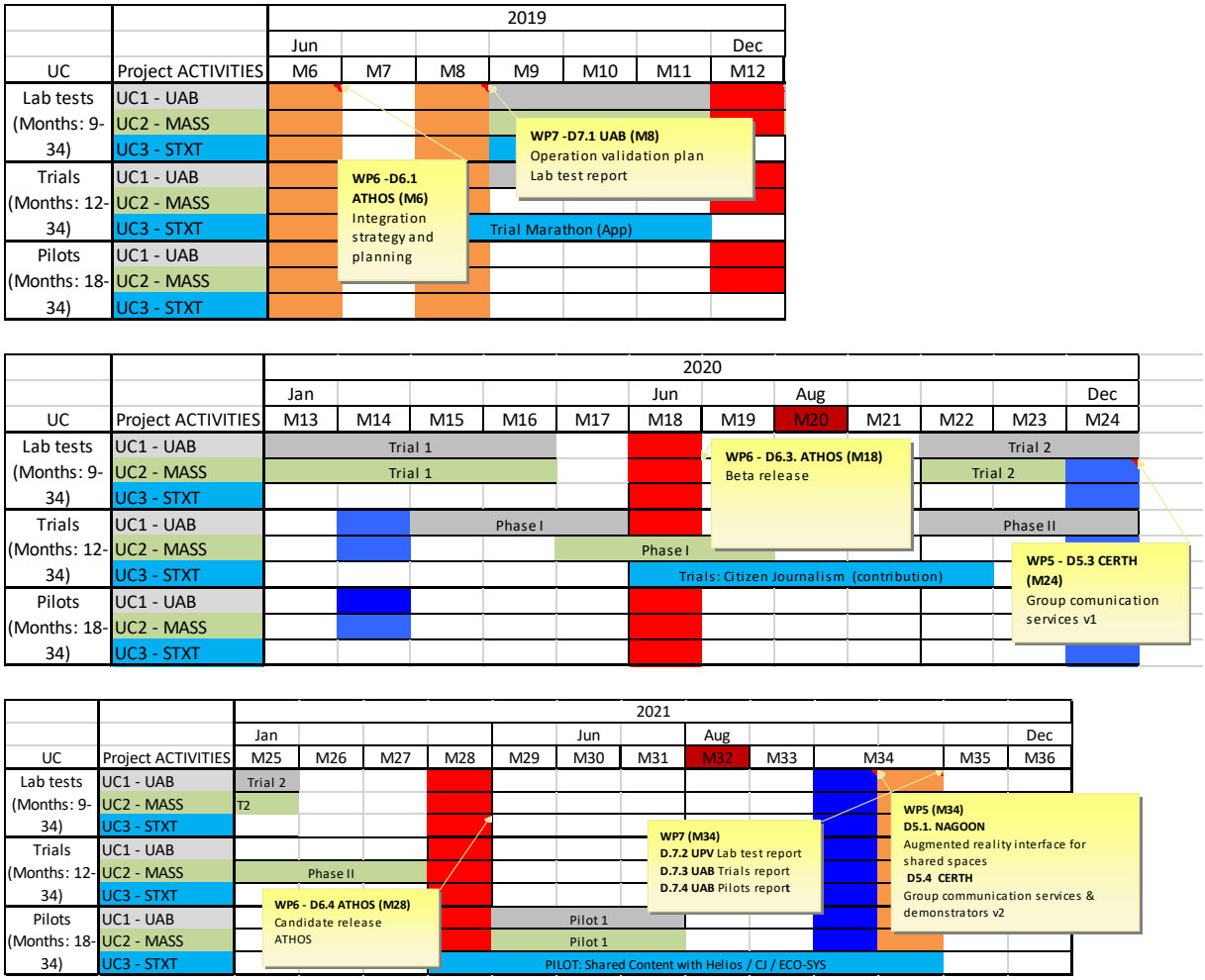

Figure 6. Calendar where possible time slots for validation activities are available

5.4 Validation activities workflow and setup

HELIOS bases the validation of its developments on three types of activities: tests, trials and pilots. Despite each kind of activity having a different approach, a general workflow, which suits coordination needs for all of them, has been designed. Basically, this is a 2-stage (or phase) process (see figure 7).

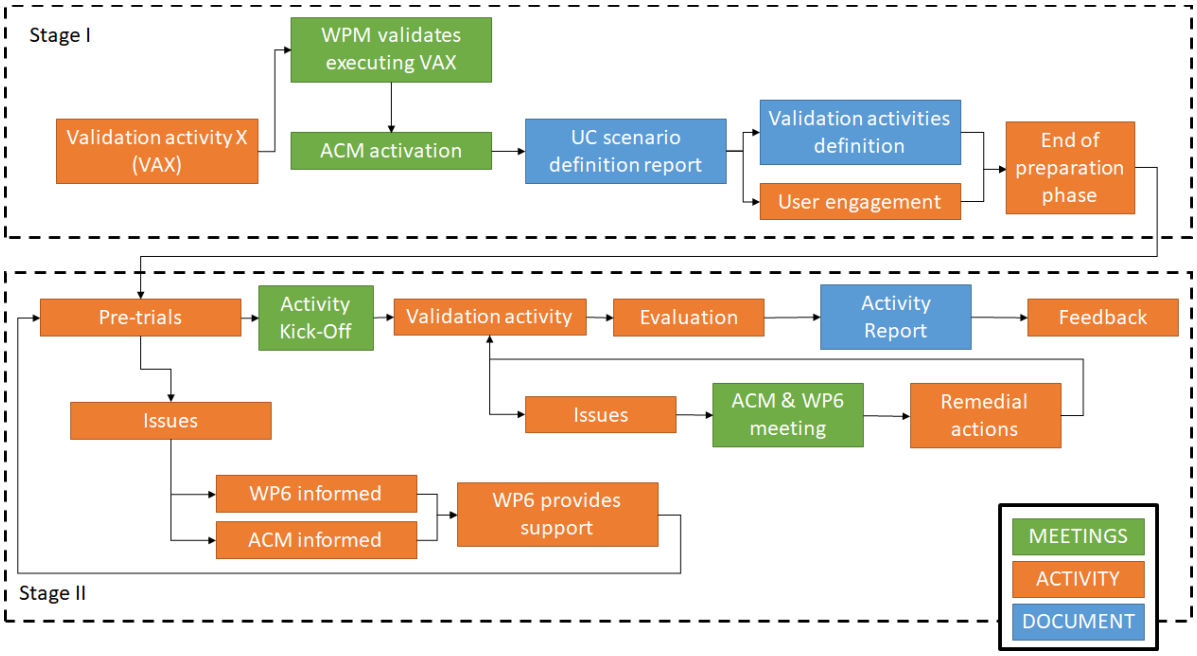

Figure 7. Validation workflow

The first stage aims at the organisation and preparation of the validation activity (VAx). The second one focuses on the execution of the VAx itself and the reporting of results. In section 5.4.1 and 5.4.2 there is a summary of how coordination among partners will take place. Note that section 5.4.3 has been created to put emphasis on the step for reporting and providing feedback, as it is a critical process in HELIOS.

a. Setup phase

This phase will start, at least a month ahead of a planned VAx. If the deployment is complex, i.e. thre is a need to install sensor infrastructures, or give specific training for users, WPCMs will identify it and the setup process might start earlier. The first steps are to meet with the WPCM members in order to assess the feasibility to activate the VAx. It will be necessary to analyse all elements involved in it, starting with the technical maturity of the developments to be deployed and followed by the definition and availability of locations where to run the validation activities. This assessment will determine the possibility to keep the schedule as planned.

If the majority of partners reach an agreement, the ACM will be launched, and all the partners involved will start preparing the UC definition report (see Annex I in this deliverable for further details). The UC definition report contextualises the validation activity and describes the elements to be evaluated, the kind of users that should be involved, and suggests procedures for validation. This document kickstarts the validation process defining WHAT has to be evaluated.

Once the WHAT is defined, partners leading the VAx will proceed to identify WHERE (location, spaces, infrastructure) the VAx will take place and HOW (evaluation methodologies) the VAx should be evaluated (see section 6 for a detailed description of the evaluation methodology). This phase is over when users are recruited to take part. These

steps are straightforward in the case of the lab tests, since the location and environment are the UPV lab. Phase I can now be considered “closed” and the Kick-off of the VAx can start.

Phase I has two main outcomes: a) a document called Validation activity description template (see Annex I); b) Validation methodology definition (see Annex II). The first gathers all the necessary information to summarise and understand WHAT will be tested and how. This information will be used later to define the methodology and tests to validate a specific UC (outcome b), Annex II). This approach is necessary since each UC will have specific features and services, and the different requirements have to be assessed and the methodology adapted.

b. Execution phase

This second phase can be divided into four parts: pre-trials, activity preparation, execution, and finally, evaluation and feedback provision.

Users involved in validation activities will be provided with a usable version of HELIOS. Most of the activities will be based on a mobile App for Android smartphones, although other formats might be considered (i.e. computers connected to an HMD, web browsers). Thus, to ensure the activity can be executed under the right conditions, pre-trials are necessary. Pre-trials, that here do not only refer to task 7.3 - Trials, are *tests* executed in small teams of HELIOS partners (taking a similar approach to the *dogfood*[9] concept).

Here, the objective is twofold. On the one hand, to validate that the technology works, and that users with no or very low experience with the mobile application can use it. Doing this and reporting the findings is very valuable to the development team. On the other hand, it will help HELIOS partners involved in the trials and pilots to better understand how the mobile application works. This internal validation is relevant to partners organising the validation activities in order to understand and interiorise the process to be able to better communicate to the external users involved in the VAx in the future. If no issues are reported, the next phase can start. It is important to highlight that, for a satisfactory outcome of the VAx, functional apps will be mandatory to avoid user frustration. The user, unless a specific VAx is organised with a different purpose, shall be provided with a functional mobile App. All services and technology should have a user-friendly manual (i.e. video or document). Likewise, the minimum and recommended technical requirements to run the HELIOS services (i.e. mobile App for Android smartphones) should be provided to participants to ensure they are valid candidates. If the current release of HELIOS to be evaluated presents any technical problems, the responsible developers will be informed, and the VAx postponed until the release can be tested in proper conditions.

Validation activity scheduling will be defined with a time window to allow the proper mitigation of slight delays (i.e. little developments to solve the issue). If the issue raised is critical, the ACM will organise an urgent WPCM to evaluate whether the VAx should be cancelled, or re-scheduled until it can be resumed in the right conditions. In that sense, delays in one UC should not have a negative impact on the execution of other UC’s VAx, since they are independent, from an organisational perspective. If the issue is related to the core technologies of HELIOS, then the WPCM will have to assess whether other VAx needs to be rescheduled or not.

The HELIOS project is designed to offer several safety systems to minimise the risk of such issues. WP6 will be running internal evaluations of technologies before allowing any Vax to be run. The Agile methodology enables HELIOS to react fast and to be highly flexible in adapting to changing conditions. Following this approach, WP7 is organised and structured so that validation activities are defined upon partners’ agreement on what technologies are ready to be validated. Even if critical issues occur, WP7 will interact with all partners involved to redefine the activities and scheduling.

[9] https://creativemarket.com/blog/ux-dogfooding-what-how-and-why

The next step is VAx setup. It consists of tasks like user recruitment, infrastructure setup, and coordination with third parties involved collaborating with HELIOS, among others. This phase will pose several challenges to partners' abilities to adapt to different conditions. HELIOS will seek, when possible, to run validation activities under a stable validation environment, which at the same time facilitates the repetition of activities (iterations) and brings a higher degree of control on the whole activity.

Before starting the execution phase, a kick-off session with lab tests/trials/pilot participants[10] will be organised. This meeting will be used to inform participants about HELIOS, the services being evaluated, and the overall scope of the VAx. Ethical requirements and signature of consent forms will be administered. It is important to understand that this phase might be different for each UC, since each UC (and also its iterations) has different purposes and objectives, although being complementary among them. After this introductory step, the core of the validation activity will start. Users will be invited to conduct different sorts of actions and HELIOS partners will supervise the evolution of the VAx according to the conditions predefined in Annex I and II.

c. Iterative feedback loop

The final step of each VAx is the evaluation reporting and feedback (Annex III) to technical partners (mainly to WP6 and WP2). The reporting will start during the execution phase by different means, (i.e. interviews, questionnaires, etc.). The results and data gathered at the end of each VAx will be structured and documented to feed both design and technical partners. This is necessary for them to better understand how to later improve the previous deployments and their impact on user's behaviour and interaction. This report will also add observations in order to improve the next VAx for each UC. This will be useful to improve the following iterations of validation activities. At the end of each iteration, reports from each activity will be delivered to design and technical teams via WP6, and distributed onwards to WP2, WP3, WP4 and WP5. All these reports will be compiled at the end of the project in three different deliverables, D7.2, D7.3 and D7.4, to report tests, trials and pilots, respectively. Intermediate compilations will be prepared when necessary (i.e. project reviews).

Interaction with technical teams

WP6 will provide a set of components whose interoperability will be verified by technical teams during its development. These components will be later deployed as a mobile App to be validated in the UCs through the different kinds of VAx.
From the integrated system perspective, two kinds of feedback from the VAx will be provided towards the technical teams:

1. Bugs, that cause a malfunction in the system during test (i.e. user is not able to detect other users connected to the HELIOS platform)
2. Functional and interface improvements (i.e. user reports preference for a bigger size of a button)

While bugs must be debugged, the implementation of improvements and new features will depend on the resources available in the project.

WP6 is responsible for integration, but not for the technical development or debugging, thus the purpose of WP6 is to: *ensure that issues detected in WP7 are quickly and effectively routed to the respective technical teams in WP3, WP4 and WP5 to find a solution*. To ensure this communication, WP6 will offer a set of collaborative tools to be deployed in the project. Currently WP6 is considering two tools for tracking possible bugs or issues:

[10] Participants are not project partners, but the people engaged to take part in the VAx as end-users.

- GitLab and particularly the mechanisms to register and to track issues. This is the tracking mechanism proposed for developers during the deployment.
- JIRA, if it is available for all the partners, since one of the main functionalities of this software is the tracking of errors. JIRA is proposed at a higher level to monitor and control the component refinement.

Improvements, they do not require the creation of issues. The methodology will be different in this case. WP7 will compile and integrate the feedback gathered from users during the validation activities and report it back to WP6 through Annex III.

When suggestions for improvement(s) are clearly identified, it might be necessary to coordinate several partners and tasks. WP6 will identify the respective partners and tasks to prioritise (or not) the implementation of the improvement(s) requested. Here the HELIOS Agile methodology is the key to enable HELIOS to react fast and provide effective solutions. If required specific TelCos with partners of the involved tasks can be organised. The improvements (or the activities to achieve the improvements) will be added to the backlog and prioritised by sprint, following the Agile approach.

Some implementation activities may require issues in the tracking system, but they will be managed by the technical teams and not directly as a result of the validation activities.

Interaction with WP2

The HELIOS design concept is developed in WP2 using fiction as an innovative and creative tool for the visualisation of future user experiences. HELIOS main objectives as a social network are analysed and converted into human-centred design (UCD) requirements. Defining UC has been used to research and define user interactions in the HELIOS social network.

The definition of UCD requirements will be reported in D2.6, D2.7 and D2.8. The requirements report will feed the rest of the tool definition processes and the validation process as well. Two more iterations (M18 and M26) are foreseen in which the exploration can be repeated by adding interface definition conditions. VAx are relevant in order to provide feedback regarding the initial requirements, but also the technological developments delivered by HELIOS. Thus, WP7 will compile and integrate the feedback gathered from users during the validation activities and report it back to WP2 through Annex III. It is important that WP2 and WP7 interact when preparing each VAx to ensure that validation provide relevant answers to support and guide WP2 work.

6. Validation activity procedure definition

The main goal of the validation activities is to assess HELIOS developments and to provide guidance, through structured reports, to the technical (WP6) and design teams (WP2). With this objective in mind, WP7 partners will create a validation strategy for each UC, based on the user requirements (provided by WP2) and technical specifications (provided by technical partners) to measure user performance and to identify potential design elements that could be improved. This will be the basis for HELIOS to increase efficiency, productivity, trust and end-user satisfaction of the services provided by HELIOS. All this process specifically aims at:

a. Identifying design inconsistencies and usability issues within the user interface and content/interactive areas.

b. Assessing the developed mobile application under controlled and uncontrolled test conditions with a representative number of users and end-users. These will be previously defined according to each iteration and UC. Data will be used to define whether objectives regarding an effective, efficient, and well-received user interface and interaction among users have been achieved.

c. Establishing clear benchmarking: user performance and user-satisfaction levels of the user interface and user interaction for future evaluations.

d. Assessing the level of usability, presence, satisfaction and trust. This will be calculated with the NPS, the level of trustiness will be measured in terms of "*How likely is it that you would recommend our company/product/service to a friend or colleague?*" generated by HELIOS social network.

To ensure a consistent approach, each VAx has been divided into three phases: before, during and after the activity. Each phase includes different actions (see description below), that the responsible partners will fulfil when planning and before starting the VAx. This will allow to define a common approach among all the VAx.

6.1 Before the VAx

It is expected that most users participating in the VAx will take part in the test using a mobile device: a smart phone or a tablet. Even if the expected outcomes for each VAx differ, partners involved in each VAx will provide a description of a pre-established set of elements, presented below, which have been identified and defined in the Validation activity description template (Annex I).

Depending on the VAx - lab test, trial or pilot - partners in charge of developing technologies will provide a document (See Annex I) with a technical description of the UC: technical elements deployed, objectives, technical validation activities proposed and technical requirements. This will be crucial to ensure that the developed technologies and services fit with the objectives, and activities of the previously defined UC.

Initial Use Case name and type of VAx

Partners in charge of the VAx will specify the name of the UC and type of VAx (lab test, trial or pilot) they are validating.

Scope/UC iteration

HELIOS as a platform will be developed with three releases, Alpha (early release with just a few functionalities) and Beta versions of the platform will be delivered on M12 and M18. The final version (Release Candidate) of the platform will be delivered in M28. Partners responsible for each VAx in coordination with technical partners, should specify which technical and stable components, released within the platform, shall be validated in each

VAx. This will be crucial to ensure that the developed technologies and services fit with the scope, objectives, and activities of the UC under test. In this sense it would also be interesting to determine or specify the functionalities that the end user needs to validate, and if these are possible and usable from the technical, interaction and interface point of view.

Objectives definition

Depending on the VAx, the iteration or the UC, different objectives need to be defined. These will be used to list tasks/actions that users will have to conduct during the VAx to ensure everything is validated.

Overall objectives

These are general objectives that guide the VAx. The general objectives will be used to structure other more specific objectives (see section 5.3.3.2). These objectives can target user interface related questions (validate interfaces, menus, complexity of usage), user behaviour (exploring all features, facility to connect with users in a certain context) or more technically oriented (proper performance of different features, i.e. AR/VR, in certain devices or scenarios with commercial bandwidth). Other categories might be used when defining each VAx. These objectives will be described in Annex I.

Specific objectives

These objectives will be outlined in the Validation activity description template (see Annex I) and each technical partner will provide the necessary information in collaboration, jointly with partners involved in the VAx, to define them. The next step will be to map the UC activities against these specific objectives to make sure all of them fit and are part of the VAx.

The outcome will be a briefing, based on Annex I, with a description of the UC, the elements within the UC and how they should be validated. This information will be used later to define the validation methodology specific for each UC, and the necessary information to identify the better context to run the validation activities. The UC validation report will also identify the role of each partner in each context.

Example of a task[11] that users might be requested to execute during the VAx: **Objective 1. Peer-to-Peer communications network for mobile devices, which can deliver messages over legacy IP network.** *[Measurable outcome: In a setup of three Android devices, without any centralized server, each device attached to a different point-of-presence, each device can send a message simultaneously to two other devices]*

Session description

Description of the sessions, which will include information about the date and hour of the test, if sessions will be held in a single date or an interval of time, the length of the session (number of minutes), the number of tasks and different parts involved in the activity. Time for welcoming and briefing the participants, time for ethical requirements, as well as, time for any possible delays should be taken into consideration.

Equipment

Equipment consists of as all the devices needed for the execution, at any point, of the VAx. These might be necessary to perform certain actions (i.e. use HELIOS VR), or to generate and collect data. It is expected that devices needed to perform tests, trials and pilots will be smartphones, tablets, computers and HMDs[12]. To gather data, a wider range of equipment is necessary. For instance, in the case of interviews, it might be considered relevant to record participants during the test session, thus cameras or audio recorders might

[11] This example has been taken from the DOA.
[12] Head Mounted Display, used for AR, VR or MR applications.

be necessary. The necessary technical equipment for the collection of psychophysiological information will be determined, specifying galvanic activity of the skin and visual behaviour of the individual during the tests through eye-tracking devices.

Participants (users)

Number and type of users to recruit for each VAx. A description of how users will be recruited, and all documentation related to ethic procedures will be provided.

Another important aspect related with the users in the test is the specification of profiles. In this case these profiles will also be decisive for the definition of the final sample in the test. This definition of profile will be directly linked to UC, the functionality of the App and the technical conditions to be met in each case.

Language

Language that will be used during the tests. Assess whether possible translations might be necessary.

Role

List of the partners involved in each VAx and their role (i.e. responsible for certain features deployed in the VAx). Also, information regarding the roles of each user for each VAx will be specified. Example: In UC3, we can find the role “citizen journalist” (content creator), and the audience (content consumers). Participants might be asked to switch/change roles at a certain point.

User recruitment

In order to run a validation plan for testing the developed technologies, recruiting users is a key aspect. Involving users in research and development projects might be complex. Thus, it is necessary to identify the right target end-users and to create a link of trust. In order to generate this bond, it is mandatory to follow ethical procedures which consist of: information where the project objectives and the tests to be performed are explained. Signature of consent forms will follow the information phase. In the case of HELIOS, a set of tasks to be accomplished for each activity will be provided in order to guide them in the validation phase. UAB will be in charge of ethics, and UP of the legal aspects. Deliverables D2.10 and D2.11 (M6) underlined the ethical issues that the HELIOS project had taken into consideration. Since almost all the information included in that deliverable still applies, only the latest changes regarding data protection regulations are presented. Since UPV (responsible for the lab Test) will follow different approaches for user recruitment, specific ethical issues are provided in the next subsection.

User recruitment for Lab tests

To carry out the lab tests, the intention is to recruit a sample that meets the requirements of the usability study of the HELIOS platform. In order to meet these requirements, a specific profile will be defined and sent to a specialised test recruitment agency. This agency will be in charge of estimating the sample required: (number of users, sociodemographic profile, and experience with different technologies).

Metrics/Methodologies

The use of metrics is essential to monitor the user experience with a product and helps to identify when and where (quantitative metrics) and why (qualitative metrics) issues may be occurring. Therefore, depending on the VAx, the scope and the UC, the following metrics will be collected.

Qualitative metrics/ subjective methodologies

Qualitative metrics include questionnaires that will be used prior to the sessions to gather information regarding the users and their background. Think aloud protocols and semi-structured interviews will be used at the end of the sessions in order to gather valuable feedback for the technical and design teams.

Quantitative metrics/ objective methodologies

Quantitative metrics allow to collect measurable data (e.g., successful completion rates, error rates, time on task). Questionnaires such as SUS (System Usability Scale) and NPS (Net Promoter Score (NPS) structured in Likert-scale questions will be used to allow partners to gather data regarding satisfaction, usability and accessibility.
In the case of lab tests, it is important to mention the psychophysiological metrics that are intended to be used in this case: Galvanic Skin Response (GSR) technologies that allow to evaluate the cognitive and emotional load of the content. In the same way, the analysis of visual behaviour (saccadic eye movement) gathered through eye tracking devices during the use of the App will allow to make a more in-depth analysis of the perception of usability of the technology.

6.2 During the VAx

The partner in charge of each VAx, supported with other partners involved, will provide a description of the test setup. The task descriptions will be reviewed by representatives appointed by technical partners to ensure that content described in Annex I and II are aligned with the activities under validation and the use cases. The same will happen with partners in charge of defining and refining user and technical requirements in WP2. Once the activity is defined, and all partners agree with the terms on how it will be evaluated, the VAx can begin.

Participants will first complete a demographic and background information questionnaire. The facilitator will explain the set of tasks that will be measured at the beginning of each task. The tasks will be defined in the UC description template developed for each UC and they will be identical for all participants of a given user role in the UC.

- Think aloud protocols: In VAx carried out in closed and controlled conditions (i.e. lab tests), the facilitator will instruct participants to ‘think aloud’ so that a verbal record exists of their interaction with the mobile application.
- Psycho-physiological tests: the GSR measurement systems will be placed and the eye tracking system will be calibrated. In the case of GSR it is necessary to establish a baseline maintaining a resting state. In this way it will be able to establish the individual differences that will allow us to carry out the necessary analysis and collect the activity of the user in an appropriate way.

In VAx in open conditions, supervised and unsupervised conditions (i.e. Trials and pilots) video recording might be used for dissemination purposes. Depending on the VAx, after all tasks defined in Annex I have been attempted, the participant will complete the post-test questionnaires described in section 6.3.

6.3 After the VAx

Further activities may follow Vax, namely questionnaires or interviews. For lab tests and trials, post-test questionnaires will be complemented with semi-structured interviews in order to also collect qualitative data to report to the technical and design teams regarding any usability or interaction issues. These will allow to elaborate a feedback report of the results (Annex III), which aims at a) detecting risks, deviations and remedial actions from the initial plan, b) providing quantitative and qualitative usability metrics to evaluate the established

objectives, c) reporting specific usability problems and recommendations for resolution. Data extracted from questionnaires and interviews will help to provide qualitative data in order to collect further recommendations relevant for the development and implementation strategies.

For reporting and feedback delivery purposes, the iterative approach of the VAx in HELIOS is of great importance to support technical and design teams to improve the services developed. Thus, it is important to define clear and lightweight processes for evaluating the services, but also to provide relevant inputs to keep improving the services deployed.

7. Validation methodologies for each VAx

This section provides a brief description of the different typologies of VAx, what is their scope, what the methodologies that will be applied, how will be the internally coordinated coordination (i.e. what processes will follow partners involved to run and execute VAx) and how the way feedback will be provided.

7.1 Organising the Use Case

Lab tests (7.3) will be handled by UPV and the technical leaders of each UC. As stated in section 6.1.8.1, UPV will be supported by a specialised test recruitment agency. UPV will be in charge to of organise organizing the validation activities, execute them and provide feedback to HELIOS partners. Activities, since they are location agnostic, will take place at UPV's facilities in Valencia (Spain).
Trials and pilots (7.4) will take a similar approach in terms of organisation of activities with users. UC1 will be led by UAB with CERTH as technical responsible. UAB has access to students, teachers and the vast network of communities at the campus in Bellaterra (Barcelona), where validation will take place. At this stage of the project (M08), event organisers and communities (i.e. UAB concerts, sport centre, health week organisers, UAB marathon) are being contacted to assess their participation in HELIOS. Thus, these candidates will be invited to join HELIOS partners in an informative and open discussion session, or sessions. Those interested communities interested, will be listed and asked to run a call-to-action to recruit users (most probably through an online form). Each VAx will be coordinated together with the managers of the spaces and communities involved.

UC2 is co-leaded by MASS and UAB, with NAGOON as technical leader. MASS will be in charge of the location and the users. UAB will lead the definition of the validation activities, as well as the involvement of other partners in the preparation and execution of the trials. MASS have direct contact with its students, most of them with art and design background, which make them good candidates due to their cultural and artistic background and awareness of the HELIOS project. VAx are expected to take place in MASS facilities (Design College located at the city centre of Barcelona), as its facilities have plenty of opportunities for experimenting with users, space and content. Organising validation activities should not be challenging in terms of logistics (MASS has a lot of experience in doing this), but engaging users in the VAx might present some difficulties due to the school calendar.

Finally, UC3 is lead, both organisationally and technically, by STXT. STXT has access to a wide network of radio and TV journalists, who can take on the role of "Citizen Journalist" in the validation activities. It is expected to also involve regular citizens during the validation. Part of the validation activities will take place in Switzerland (i.e. Jungfrau marathon, September 2019), but it is currently under consideration to also the deployment of this UC in other locations of Europe for validation purposes. The initial UC describes a Marathon scenario, but to widen its scope and the possibility to validate it, it will be investigated adapting the UC to other contexts will be investigated.

7.2 Providing feedback

At the end of each validation activity a report based on Annex III (Feedback report template) will be generated in order to outline the conclusions and provide recommendations to technical partners. This document will be used to report the activity itself and to feed D7.2, D7.3 and D7.3 (lab tests, trials and pilots activity reports). But, the main utility of this document is to inform partners about the results of the activity. This information will be structured (i.e. Quantitative results, Qualitative results, Conclusions and recommendations) to facilitate understanding and to make a decision on how HELIOS has to react (if necessary). Individual reports are necessary to monitor validation results and distribute them among partners in a rapid and comprehensive manner.

Note that the final structure of Annex III might differ due to the difference between activities, although the structure will be the same.

7.3 Lab tests

Lab tests will be performed at UPV in closed controlled conditions with a small sample of users by UPV. During these tests, metrics from user interaction will be gathered and analysed. The preferred apparatus for data gathering will be sensors from smart phones. Some of the proposed analysis are:

- speech analysis
- voice analysis
- facial coding

Lab tests will also try to gather info regarding:

- contextual information
- participant explicit responses regarding tasks decisions
- perception of usability
- usefulness
- presence
- satisfaction

This layer of metrics will be later analysed and reported to development teams to help to update final versions of technology and services to be validated finally in T7.3 and T7.4.

Validation methodologies

Different tools will be used to gather both qualitative and quantitative data in both objective and subjective methodologies:

Objective methodologies or **implicit processes of user behaviour**:

- Demographics and background questionnaires.
- Execution of tasks: times, routes, loss of control and efficiency.
- Mobile sensors: analysis of user behaviour within the mobile to detect the workflow within a task, number of clicks, times on each screen, routes, etc.

- Eye tracking where we analyse the user's visual behaviour and attention according to the interface and workflow. In the same way it will be used to interpret the attention and interest on the augmented reality of the system.

- Facial coding: Recognition of the emotional expressions of users, both general (positive vs. negative) and points of frustration or confusion during the use of augmented reality.

- Skin conductance and Heart Rate Variability: Recognition of the arousal in each of the tasks, as well as points of stress.

- Satisfaction, usability and accessibility questionnaires (SUS).

- Emotional mood questionnaires before and after the use of App (The state trait anxiety inventory STAI-2)

Subjective methodologies and observation that analyses **explicit processes of user behaviour**:

- Think Aloud: through this technique the user verbalizes what he / she intends to do, what objectives he / she wants to achieve and what he / she does to achieve it. Including in this case comments of losses of control or even victories or positive surprises during the accomplishment of the tasks.

- Ethnography: direct observation by the experimenter or judge of the user execution during the study.

- Semi-structured interviews: in these interviews we expect the user to provide us with a concrete vision of his/her experience based on the elements to be highlighted according to the experimental design (iconography, natural workflows, elements of loss of control, use of RA during the task, information that is missing, what elements are left over, etc...) in the same way, users propose alternatives to the negative aspects of the App.

7.4 Trials and Pilots

Trials will be unsupervised real-life validation in closed controlled conditions. The main aims are to perform validation in a more mature stage of the technology than in Lab tests, and to assess interaction among participants in the VAx.

Pilots will be open and unsupervised validation, with large audiences. Pilots will be the final step of previous testing iterations: focus groups (WP2), neuro tests (T4.6), lab tests (T7.2) and trials (T7.3).

Validation Methodologies

Although we are currently in the initial phases of the project, a list of feasible validation and apparatus can be found in Section 8.

The methodologies in this type of test are on the one hand, objective methodologies or implicit processes of user behaviour, such as:

- Demographics and background information
- Satisfaction, usability and accessibility questionnaires (SUS).
- NPS[13] before and after the use of App

[13] Net Promoter Score (NPS)

In addition, the following subjective methodologies and observation procedures that analyse explicit processes of user behaviour, are relevant:

- Think Aloud: through this technique the user verbalizes what he / she intends to do, what objectives he / she wants to achieve and what he / she does to achieve it. Including in this case comments of losses of control or even victories or positive surprises during the accomplishment of the tasks.

- Ethnography: direct observation by the experimenter or judge of the user's execution during the study.

- Semi-structured interviews: in these interviews we expect the user to provide us with a concrete vision of his/her experience based on the elements to be highlighted according to the experimental design.

7.5 Data protection

As described in D1.1 – Data Management plan, it is expected to generate data during the VAx. In case of handling personal data, HELIOS will under no circumstances publish any data that is personal data, or data that maybe adversely converted into personal data. Decisions about sharing (selected) non-personal datasets will be taken by the Project Coordinator after these datasets are reviewed by the Steering Committee of the project. In general, all activities related with data protection, privacy and ethics, will be assessed by UP and UAB following the principles and procedures described in D1.1 Data Management Plan - initial version, D1.6 Instructions for data retention and deletion, D2.10 Recruitment and informed consent procedures, and D9.2 Ethics Requirements - Protection of Personal Data.

Consent forms and information sheets, as approved by the ethical committee, will be provided to participants, where they will find a description of the activity they participate in, who the data controller is and how and when they can exercise their rights. An ID will be given by data controllers to each participant to anonymise their personal data.

8. VALIDATION METHODS

The objective of the VAx is to understand how users experience and interact with the HELIOS platform. The usability of the HELIOS platform will be validated at different stages of the design and development processes. The different methodologies considered in this section are divided based on a) the stage of the project when it is conducted and b) the type of VAx under test.

There is a myriad of methods to test user experience (Tullis & Albert 2008, Pannafino & McNeil 2017, Goodman et al. 2013) in terms of effectiveness (success), efficiency (time) and satisfaction (ease of use) such as: cognitive walkthroughs, card sorting, icon usability testing, contextual inquiry or online surveys, to name just a few. The usefulness indicates that the system is useful to achieve the specific objective for which it was designed, and the user's objectives were achieved. A system might not be useful for several reasons such as low capacity to create the necessary information, low reliability or even an interface that is too simplistic - very usable but not very useful.

The following section will outline the different validation methodologies that may be used in the VAx to gather user feedback first hand and that allow the design and development teams to take action to implement improvements.

8.1 Demographics and background information

Demographic information[14] seeks to gather basic information about users. It will allow design and development teams to better understand how each person represents the general population. At this stage of project development, relevant demographic information is still to be defined.

Given the GDPR Data minimisation and purpose limitation: any processing of personal data should be taken on board for a specific purpose This should be defined in advance. Further processing of personal data is only allowed when it matches the initial defined purpose. Exceptions to this rule are provided in Article 5.1(b) GDPR and relate to archiving in the public interest, scientific or historical research and statistical purposes. Data minimisation, in turn, requires data processing to be limited to what is strictly necessary to achieve the legitimate aim sought in terms of content and means. To this aim at present only two questions, i.e. age and educational background, have been considered relevant for this section of the questionnaire. Before considering adding further demographic data, it will be assessed whether this information is relevant for the design (WP2) or development (WP6) teams.

Demographic data is mainly useful to forecast market viability for an already existing product, and it does not provide any insights about the reason users may like/dislike a product. Thus, it is necessary to register other sets of information to complement demographics. Background information[15] allows to define the scope of behaviour and to represent the complex set of behaviours, motivations, and goals of real people. Therefore, four background questions based on behavioural attitudes have been added to this questionnaire. These questions are considered relevant for the design and development processes of the mobile application and will enable to a) assess the interaction foreseen in the UC developed in WP2, and b) obtain meaningful insights about user interaction.

Demographic data

The relevance of the data gathered by these questions will have to be taken into consideration prior to any test:

1.- Age range

- 18-24
- 25-34
- 35-44
- 45-54
- 55-64
- +65
- prefer not to tell

2.- Educational background

Highest level of studies reached:

- Lower than primary education
- Primary education
- Secondary education
- Advanced vocational education
- First cycle university education (diploma, degree or graduate studies)

[14] https://www.surveymonkey.com/mp/gathering-demographic-information-from-surveys/
[15] Questions extracted from https://www.surveymonkey.co.uk/r/99CGC3B

- Second cycle university education (master, postgraduate or doctoral studies)
- Prefer not to tell

Background information

3. What kind of social network do you usually use? How much time do you spend on social media per day?

Social Networks:

- Facebook
- Twitter
- Linked-in
- Instagram
- Others (please specify)
- Prefer not to tell

Time of use:

- less than 30 minutes
- 30-60 minutes
- 2 hours
- 2-3 hours
- 3 hours +
- Prefer not to tell

4.- How often do you post on social media?

- never
- every few months
- every few weeks
- weekly
- daily
- multiple times a day
- Prefer not to tell

5.- What do you use social media for?

- keeping in touch with friends and family
- event planning
- buying and selling
- inspiration
- news
- dating
- to meet new friends
- to find employment
- to browse/ time waste
- Prefer not to tell

6. Experience with augmented reality, mixed or virtual technologies.

- high
- medium
- low
- Prefer not to tell

The first question aims to identify the range of age of people taking part in the activity. Potential relation between variables like: AGE and USAGE or INTEREST and TIME OF USAGE and PERMANENCE.

The second question attempts to establish the relation between EDUCATION and TECHNOLOGICAL SKILLS or USE OF TECHNOLOGY in terms of technological capabilities.

The third question allows to gather information related to the user experience. For the usability analysis of a new app the mental model of the user regarding previous experiences has always to be taken into consideration. Users always refer to a previous experience which they use to compare or create a metaphor with the app being tested.

The fourth question aims at differentiating between active and passive social media users. Since different roles from a user perspective have been differentiated in the new mobile app, it is important to consider the activity of the users.

The fifth question provides information regarding the different interests of the users when using a social network. This will also allow design teams to match the interactions described in each UC with the interactions in which the users are interested.

The sixth question is aimed only for the validation activities that will include augmented reality, mixed or virtual reality in order to identify if previous experience with immersive environments have an effect with the developed technologies.

In all questions, the users will have the choice to not answer it in case they do not feel comfortable with giving that information.

8.2 System Usability Scale (SUS)

The SUS is amongst the most popular usability testing tools due to its many advantages, such as its brevity and robustness, as well as it being free of charge (Katsanos *et al.*, 2012: 302; Bangor *et al.*, 2008). Despite its simplicity, Tullis and Stetson (2004) noted that the SUS yielded very reliable results across sample sizes. It has also been successfully applied to a wide range of devices and systems (learning management systems, landline telephones, non-web graphical user interfaces, automated telephone interfaces, web-based interfaces, to name a few), which proves its flexibility and lack of dependence towards the system under study. It was designed by John Brooke in 1986 and it consists of just ten questions, half of which are positive statements, while the rest are negative. These questions are alternated and presented in a fixed standardised order. Informants need to express how much they agree with the proposed statements selecting one of the five options available, ranging from “strongly disagree” to “strongly agree”. Final scores for the SUS can range from 0 to 100, where higher scores indicate better usability.

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

In addition, the following four questions will be annexed to the SUS questionnaire in order to give new insights to help improving the developments:

1. Where you familiar, or have you ever used this kind of service before? If so, which one?
2. Do you think research and development on this area of social media networks should be continued?
    - Yes
    - No
    - Prefer not to tell
3. If you would like to suggest any new functionality or improvement, you can do it now.
4. Is there anything you would like to add?

The first question is considered to be important because previous experience with similar technology could have an impact on the results. Informants will be given some space to write in case they wanted to share the type of technology they had experience with.

The second question asks informants if they think that researching that service is worth pursuing. Its main goal is to know if the services being developed in the project are relevant enough to keep working on them.

The third of these additional questions will allow participants to suggest new functionalities or modifications that they consider could be added to the HELIOS to improve it. It is a non-compulsory open-ended question.

Finally, at the end of the questionnaire, another open-ended question has been added. This one let participants make any other comments that they consider to be valuable to our project. That space can also be used to add information about the informant's demographic profile that has not been requested in the demographic section or inform the developers about issues that are not necessarily usability-related.

8.3 Net Promoter Score (NPS) and KPI

The NPS will also be included in the questionnaire in order to obtain the trust level. This score is calculated based on responses to a single question: "How likely is it that you would recommend our company/product/service to a friend or colleague?". Respondents are asked to rate their response on a 0 to 10 Likert scale, in which 0 means "definitely not" and 10 means "I would recommend it for sure". Those who respond with a score of 9 to 10 are called "promoters"; those who respond with a score of 0 to 6 are labelled "detractors"; and respondents with a score of 7 and 8 are labelled "passives". This question will be followed by an open-ended question in which informants are asked to justify their score.

8.4 Sense of belonging

Sense of belonging is referring to the human need that an individual has to be accepted as a member of a group, to be an integral and relevant part of this group, a basic need of Maslow's pyramid of needs (1968). In this sense, and linked to our social network, our objective here will be to understand if the user really feels that he/she is part of it, that the other users share a certain relationship with him/her and that they take him/her into consideration.

In our social platform, the relationship between the members is not the same as the communication in real life, but at visual, interactive and participative level, the experience could be considered similar in terms of sensations. For example, the possibility offered by a social platform of value information is an unequivocal sample of your relationship with the group and vice versa. To achieve this goal, we use different techniques such as semi-structured interviews, the location of certain elements with which the social platform provides

us information about the feeling of belonging (such as the case of the like button in Facebook) or the realization of adaptations of specific questionnaires about feeling of belonging such as the "belonging scale" (Gambone MA, & Arbreton AJA 1997) or the need to belong scale (Leary et al, 2012, Leary 2013).

8.5 Presence (IPQ Questionnaire)

Presence can be understood as the sense of “being there” (Reeves, 1991), “illusion that a mediated experience is not mediated” (Lombard and Ditton, 1997), “psychological sense of immersion in any mediated environments” (Fryer and Freeman, 2013), “experiential quality metric employed to evaluate broadcast and virtual environment media systems” (Lessiter et al. 2001), “the experience of being engaged by the representation of a virtual world” (Jacobson, 2002) or “perceptual illusion of non-mediation” (Lombard & Jones 2015), to name just a few definitions of this concept. Presence measures allow to assess the entire experience of participants, and, as stated by Walczak & Fryer (2017), 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 (Walczak & Fryer, 2017; Wilken and Kruger, 2016). However, presence is a multi-construct concept that encompasses many dimensions or subcategories, with variations depending on the authors. To measure presence, both subjective and objective measures can be used (van Baren and IJsselsteijn, 2003) but it seems that presence questionnaires are the most common method for multiple reasons: they are specifically aimed to measure the specified concept, they are reliable and valid, they are easy and cheap to implement and non-intrusive, and the results are easy to analyse.

After a review of different standard questionnaires on presence such as tje ICT-SCOPI (Lessiter et al., 2001), the Slater-Usoh-Steed Presence Questionnaire (Slater et al, 1993) or the PQ (Witmer & Singer, 1994)[16], the Igroup Presence Questionnaire (IPQ) (Schubert et al., 2001) has been chosen for various reasons. First of all, it combines previous questionnaires including PQ and SUS) and it was the first one to specifically differentiate between spatial presence, immersion and involvement.

The same methodology will also be used to validate project partner NG’s “shared spaces” social connectivity. In this questionnaire, spatial presence refers to the sense of being there in the virtual environment; involvement refers to the attention to the real and the virtual environment, and realness refers to the reality judgement 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 VEs), including HMD in a laboratory, a situation similar to us. It is available in English, German, Dutch, French, and Japanese. IPQ can be accessed here, and it includes 14 questions, making it an adequate length for experimental purposes: http://www.igroup.org/pq/ipq/download.php

8.6 Preferences and General Opinion

Concerning preferences and general opinion, they will be measured through specific questions in the form of post-questionnaires that will be developed “ad hoc” for each specific test. The questionnaire will rely mainly on open-ended questions with some room for participants to explain their response should they wish to do so.

8.7 The State-Trait Anxiety Inventory (STAI)

The emotional mood, or the emotional state of the moment, provides us with relevant information about how the user is at the time of the study. This information can be evaluated in the sense that our decision-making is filtered by the emotions that are felt at the moment. Our goal with this questionnaire is to establish the emotional state of the subject and see if

[16] http://www8.informatik.umu.se/~jwworth/PresenceMeasurement.pdf

the use of our App has modified this emotional state in some aspect. It will also be interesting to know the influence of certain emotional states on the perception of tool usability, knowing that the level of anxiety of a person affects its effectiveness, and therefore the level of perceived frustration will be different.

The state-trait anxiety inventory (STAI) (Spielberger et al, 1983) is a psychological inventory based on a 4-point likert scale and consists of 40 questions on a self-report basis. the STAI measures two types of anxiety (state anxiety, or anxiety about an event, and trait anxiety, or anxiety level as a personal characteristic.

STAI will be used in the lab test VAx, in which only state anxiety scale will be administered, since the objective is to understand/know how anxiety is located contextually and how the App can affect it.

8.8 Emotional state

Emotional arousal is measured through the galvanic skin response (GSR), also named electrodermal Activity (EDA) and Skin Conductance (SC), refers to changes in sweat gland activity that are reflective of the intensity of our emotional state, otherwise known as emotional arousal. Human body sweating is regulated by the Autonomic Nervous System.

Our level of emotional arousal changes in response to the environment we are in, the experience with the context, and in our case with a new technology context that influences the GSR signal. The GSR signal is very easy to record: in general, just two electrodes put at the second and third finger of one hand are necessary. The variation of a low-voltage applied current between the two electrodes is used as a measure of the GSR.

The objective in the lab test is the use of this type of signal to evaluate the emotional arousal linked to the experience with the App HELIOS. This implicit methodology contributes to increase the validity of the usability study. With GSR measures we can see the cognitive and emotional load the subject interprets of the use of the App.

8.9 Attention

Attention will be measured through eye tracking technology. This refers to the measurement of eye activity. More specifically, eye tracking describes the recording of eye position and movement in an environment based on the optical tracking of corneal reflections to assess visual attention.

This technology has been proposed in lab tests to understand attention and implicit visual behaviour during the use of our App.

8.10 Ethnography methodology

Ethnography is a type of qualitative analysis methodology that aims to study a concrete behaviour within a context and a reality. Despite different scopes that the technique can have, according to the intentions of the researcher with the object of study, the following conditions are characterized:

1. The object of study is approached with a view to understanding and interpreting a reality that interacts with a broader context.

2. The information found in the field – both verbal and non-verbal –must be analysed and interpreted to understand what users do, say and think.

3. An ethnographic study gathers a global vision of the social field studied from different points of view: an internal point of view of the members of the group and an external perspective.

The purpose for using this type of qualitative methodology is to focus the interest in the natural behaviour during the use of the App during the lab test. The natural observation of the user interactions will provide the necessary verbal and nonverbal information to understand user behaviour during the lab test with respect to the HELIOS App.

8.11 Others

Semi-structured interviews and think aloud protocols during and after the VAx are relevant methods where the user can detect not only the utility but also the factors that make it more difficult to use. Both types of subjective methodologies will be used to obtain relevant data to provide feedback to the design teams (WP2) and to the technical teams (WP6).

9. NEXT STEPS AND CONCLUSIONS

This deliverable kick-starts the validation phase, starting in M09 with task T7.2 - Lab tests. Although validation activities will start later in the project, now all the structure will be consolidated and put into motion, making all those arrangements necessary for the proper execution of T7.2, but also later T7.3 - Trial (M12) and T7.4 - Pilots (M18). Despite T7.3 and T7.4 starting later, also in M09, through T7.1 first movements with external entities, involved in the VAx, will be carried out.

In parallel, technological developments will start taking shape, which will feed WP6 and in consequence WP7, contributing to define what will be available for validation. Another important contribution will come with the all the groundwork for the delivery of initial user requirements in D2.6 during M12. These iterative feeds of requirements and validation outputs will keep guiding all the work in WP7 (from mid-September 2019 until the end of the project).

Despite the fact that there are still some questions marks to be answered, preparing this deliverable has brought many interactions among all the partners involved to help understanding better the project, its objectives, how they have to be achieved and what are the expected contributions from WP7 are. This deliverable provides just details and first insights on how all mechanisms for coordinating the validation phase will work, but also the methodologies involved to achieve this objective. It is true that modifications will take place, HELIOS is Agile, and this means also flexible and capable to adapt to changes and unexpected situations. As a result, flexible structures and simple processes are defined to act as small projects when evaluation/validation of deployments is necessary, and feasible. This dynamic should help to improve how it works on every iteration.

10. REFERENCES

Bangor, Aaron; Kortum, Philip T; & Miller, James T. (2008). An Empirical Evaluation of the System Usability Scale. *International Journal of Human-Computer Interaction*, 24(6): 574-594.

Beck, K. et al. (2001) Manifesto for Agile Software Development. Accessed in August 2019 at http://agilemanifesto.org/iso/en/manifesto.html

Brooke, J. (1996). SUS: A quick and dirty usability scale. *Usability Evaluation in Industry, 189*(194): 4-10.

Brooke, J. (2013). SUS: A retrospective. *Journal of Usability Studies, 8*(2): 29–40. Retrieved from http://uxpajournal.org/sus-a-retrospective/ [Last accessed 18 July 2019]

Fryer, L., & Freeman, J. (2013). Visual Impairment and Presence: Measuring the Effect of Audio Description. *ACM International Conference Proceeding Series*.

Gambone, M. A., & Arbreton, A. J. A. (1997). Safe Havens: The contributions of youth organizations to healthy adolescent development. Philadelphia, PA: Public/Private Ventures.

Goodman, E.; Kuniavsky, M. & Moed, A. (2013) Observing the User Experience: A Practitioner's Guide to User Research. Morgan Kaufmann.

Jacobson, D. (2002). On theorizing presence. Journal of Virtual Environments, 6:1

Katsanos, Christos; Tselios, Nikolaos; and Xenos, Michalis (2012). Perceived usability evaluation of learning management systems: a first step towards standardization of the system usability scale in Greek. *Proceedings of the 2012 16th Panhellenic Conference on Informatics*.

Leary, M. R., Kelly, K. M., Cottrell, C. A., & Schreindorfer, L. S. (2012). Individual differences in the need to belong: Mapping the nomological network. Unpublished manuscript, Duke University.

Leary, M. R. (2013). Need to Belong Scale. Measurement Instrument Database for the Social Science. Retrieved from www.midss.ie [Last accessed 31 July 2019]

Lessiter, J.; Freeman, J.; Davidoff, J. B. & Keogh, E. (2001). A Cross-Media Presence Questionnaire: The ITC-Sense of Presence Inventory. *Presence Teleoperators and Virtual Environments*, 10(3): 282-297.

Lombard, M., & Ditton, T. (1997). At the heart of it all: The concept of presence. *Journal of Computer-Mediated Communication, 3*(2).

Lombard, M., & Jones, M. T. (2015). Defining presence. In M. Lombard, F. Biocca, W.A. Ijsselsteijn, J. Freeman, & R. Schaevitz (Editors), *Immersed in Media: Telepresence Theory, Measurement and Technology.* London: Springer.

Maslow, A.H. (1968). Toward a Psychology of Being (2nd ed.). Van Nostrand Reinhold Co.

Pannafino, J. & McNeil, P. (2017). UX Methods: A Quick Guide to User Experience Research Methods. CDUXP LLC.

Reeves, B. (1991). Being there: Television as symbolic versus natural experience. Unpublished manuscript, Stanford University, Institute for Communication Research, Stanford , CA .

Schubert, T., Friedmann, F., & Regenbrecht, H. (2001). The experience of presence: Factor analytic insights. Presence: Teleoperators and Virtual Environments, 10(3): 266-281.

Slater, M. Steed, A. & Usoh, M. (1993). The Virtual Treadmill: A Naturalistic Method for Navigation in Immersive Virtual Environments, First Eurographics Workshop on Virtual Environments, Polytechnical University of Catalonia, ed. Martin Goebel, September 7, 1993, 71-83.

Spielberger, C. D., Gorsuch, R. L., Lushene, R., Vagg, P. R., & Jacobs, G. A. (1983). *Manual for the State-Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press.

Tullis, T. & Stetson, J. (2004). A Comparison of Questionnaires for Assessing Website Usability. Paper presented at the Usability Professionals Association annual conference in Minneapolis. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.396.3677&rep=rep1&type=pdf [Last accessed 18 July 2019]

Tullis, T & Albert, B. (2008). Measuring the User Experience: Collecting, Analyzing, and Presenting Usability Metrics . Morgan Kaufmann Publishers Inc. San Francisco, CA, USA

van Baren, J., IJsselsteijn, W.A., Romero, N., Markopoulos, P., de Ruyter, B. (2003). Affective Benefits in Communication: The development and field-testing of a new questionnaire measure. PRESENCE 2003, Aalborg, Denmark.

Walczak A. & Fryer L. (2017). Creative description: The impact of audio description style on presence in visually impaired audiences. *British Journal of Visual Impairment*, 35:1, pp. 6-17.

Wilken N. & Kruger J. L. (2016). Putting the audience in the picture: mise-en-shot and psychological immersion in audio described film. *Across Languages and Cultures,* (17:2): 251-270.

Witmer, B. G., & Singer, M. J. (1994). Measuring immersion in virtual environments. (ARI Technical Report 1014). *Alexandria, VA: U.S. Army Research Institute for the Behavioural and Social Sciences.*

11. Annex I - Validation activity description template

11.1 UC definition report

UC definition report	
Initial Use Case name:	
Type of validation activity (VAx):	
Scope/ UC Iteration:	
Session description:	
Equipment:	
Participants:	
Language(s):	
Role (Leading partner and partners involved):	
Qualitative metrics:	
Quantitative metrics:	

Initial Use Case description information
Context description:
Target description:
UC description:

Validation activity description
Context description:
Services and features deployed in the UC:
General objectives:
UC validation activities:

User role description
Type of users:
Objectives:
UC validation activities (General and specific tasks description):

Technical description
Technical description of the UC:

Technical elements deployed:
Objectives:
Technical validation activities proposed:
Technical requirements (and additional infrastructure):

12. Annex II - Evaluation methodology template

12.1 Evaluation methodology template

Evaluation methodology template
Description of methodologies used:
Number of users:
Objectives:
Features/services to be tested:

Tests setup and calendar of events
Additional comments:
Annexes (scripts, questionnaires):

13. Annex III - Feedback report templates

13.1 Activity and feedback report templates

Activity and feedback report templates
Use Case summary:
Activity data (start -ending date; location, number of participants, partners involved, partner in charge of the activity):
Activity description (what has been done, narrative):
Risks, deviations and remedial actions from the initial plan:
Quantitative results:

Qualitative results:
Conclusions and recommendations: