Intelligent Verification/Validation for XR Based Systems

Research and Innovation Action

Grant agreement no.: 856716

D1.1 – First Project Report

iv4XR – WP1 – D1.1

Version 1.9

December 2020

Project Reference	EU H2020-ICT-2018-3 - 856716
Due Date	31/12/2020
Actual Date	30/10/2020
Document Author/s	Marta Couto (INESC-ID), Rui Prada (INESC-ID) Pedro Fernandes (INESC-ID), Joseph Davidson (GA), Wishnu Prasetya (UU), Tanja Vos (UPV), Fernando Pastor (UPV), Fitsum Meshesha Kifetew (FBK), Ilda Ribeiro (INES-ID), Angelo Susi (FBK), Ian Saunter (GW), Manuel Lopes (INESC-ID)
Version	1.9
Dissemination level	Public
Status	Final

This project has received funding from the European Union's Horizon 2020 Research and innovation programme under grant agreement No 856716

Document Version Control			
Version	**Date**	**Change Made (and if appropriate reason for change)**	**Initials of Commentator(s) or Author(s)**
1.0	29/10/2020	Initial document structure and contents	RP
1.1	19/11/2020	Added WP1 and WP4	MC, PF
1.2	27/11/2020	Added WP5 part	JD
1.3	04/12/2020	Added WP2	WP
1.4	06/12/2020	Some changes in WP4 summary	RP
1.5	11/12/2020	Added WP3	TV, FP
1.6	12/12/2020	Added WP6	FK
1.7	21/12/2020	Overall minor changes	RP
1.8	21/12/2020	Added financial information	IR
1.9	30/12/2020	Final arrangements for submission	RP, MC

Document Quality Control			
Version QA	**Date**	**Comments (and if appropriate reason for change)**	**Initials of QA Person**
1.0	30/10/2020	First comments	RP
1.5	06/12/2020	General comments	RP
1.6	17/12/2020	Review	AS
1.6	17/12/2020	Review	IS
1.8	22/12/2020	Review	ML

Document Authors and Quality Assurance Checks		
Author Initials	**Name of Author**	**Institution**
RP	Rui Prada	INESC-ID
RP	Rui Prada	INESC-ID
MC	Marta Couto	INESC-ID
PF	Pedro Fernandes	INESC-ID
JD	Joseph Davidson	GA
AS	Angelo Susi	FBK
IS	Ian Saunter	GWE
WP	Wishnu Prasetya	UU
TV	Tanja Vos	UPV
FP	Fernando Pastor	UPV
FK	Fitsum Meshesha Kifetew	FBK
IR	Ilda Ribeiro	INESC-ID
ML	Manuel Lopes	INESC-ID

TABLE OF CONTENTS

EXECUTIVE SUMMARY

This deliverable presents the progress made in the first year of the Iv4XR project. Due to covid-19 and the lockdowns across Europe, we asked for a three-month extension which was granted. Therefore the current document reports the progress made in the first 15 months of the project (M1 - M15). The document reports the progress made by each work package and a financial summary for all partners and work packages.

1. OVERVIEW OF THE PROGRESS

1.1 OBJECTIVES

With the increasing complexity of Extended Reality systems (XR), the industry finds itself between XR's fine-grained and high level of interactivity and its realism, making such systems very hard and expensive to test.

Iv4XR aims to build a novel verification and validation technology for XR systems based on AI techniques to provide learning and reasoning over a virtual world.

With this technology, XR developers can deploy powerful test agents to automatically explore and test their virtual worlds' correct parameters as they iteratively develop and refine them. Given the importance of user experience for XR systems, we will also develop socio-emotional AI to enable test agents to conduct an automated assessment of user experience quality and parameterization by different demographics and socio-economic types.

1.2 EXPLANATION OF THE WORK CARRIED OUT PER WP

1.2.1 WORKPACKAGE 1

WP1 is responsible for all project management activities, covering the project's execution, monitoring progress, controlling the budget, managing risks, coordinating reviews, the consortium agreement, communication between the partners, and liaising with the EC. To facilitate this, we created a management structure at the first consortium meeting, which includes a management board (MB), the quality assurance (QA) team, and the Science and Technology Committee (S&T).

The Management Board meets during the consortium meetings and aids the project coordinator with administrative and financial tasks. So far, we have had three MB meetings; the first was in October 2019, during the Kick-off meeting of the project; the second was in January 2020 at Utrecht, and the third one in July 2020, in a virtual meeting.

The Science and Technology Committee focuses on R&D and executive decisions and meets once a month, using teleconference technology (on Skype), to discuss and solve possible

conflicts or problems regarding technical issues, identify necessary changes to the direction in which the project is going concerning technical content development, and share progress.

The QA team is responsible for ensuring that the project members deliver high-quality outputs in all project phases. All project deliverables are subject to quality control by at least one member from the QA team.

We have a series of channels to facilitate communication between the project's members. We have a Slack channel for all members, mailing lists for all members, for the S&T Committee and the Management Board. To increase communication between the members, the project's coordinator also promoted the creation of working groups. Each WP coordinator sets up monthly meetings to discuss their progress and foster collaboration between the different partners. The meetings are open to all iv4XR members who want to join.

The current pandemic imposed constraints to most partners. We requested an amendment to extend the project for three months. The amendment was approved in July 2020.
All deliverables were submitted on time, taking in consideration the extension approved. The first project milestone, “MS1 - Basic iv4XR Framework”, was met. The github page[1] of the project contains the related software.

Plans for 2nd year

The next year should follow as planned. We will work to meet the second project milestone and reach an intermediate framework with testing agents with mature capabilities and running in the case studies. We will coordinate the preparations for the first project review.
We expect to organise two project consortium meetings.

1.2.2 WORKPACKAGE 2

WP2 is progressing according to the plan. The first prototype of the iv4XR agent-based testing Framework is available and operational. A summary of its features is given in deliverable D2.1 - “1st prototype of iv4XR Framework”. All components mentioned in tasks T2.1 and T2.4 (agents’ runtime system, World Object Model, information collection and reporting) are present in the prototype, though varying levels of completion. With the basic features that it has now, the

[1] https://github.com/iv4xr-project

Framework is operational and has been tested with some in-house examples/pilots. For further information, see the document D2.1.

Plans for 2nd year

We aim to deliver the 2nd prototype, which would integrate new capabilities developed in WP3 (e.g. reinforcement learning, search-based testing, model-based testing, integration alternative testing tools such as TESTAR) and WP4 (socio-emotional test agents).

1.2.3 WORKPACKAGE 3

Work Package 3 focuses on the development and integration of Functional Test Agents (FTAs) intended to verify if the behaviour of the XR systems responds correctly according to the interaction of users. A first prototype of different FTAs allows the navigation over LabRecruits System Under Test (SUT) and the interaction with different WOM entities while verifying the correctness of the functionality. Document D3.2 - "1st prototype of functional test agents (FTAs)" includes an explanation of the first prototypes resulting from WP3.

We distinguish two types of FTAs:

1. *FTA focused on SUT Goal Solving*

   Using the Test Specification Language (TSL) developed in Task 3.1 (see D3.1 - "Test Specification Language"), agents can take deliberations about which strategies will allow them to navigate and interact with different entities to solve their goals. Solving goals to test specific tasks, means that dealing with hazardous elements (Task 3.3) will not be very different from dealing with other XR application elements. Consequently, the goal solving FTA can already handle these hazardous elements which completes Task 3.3 and its inclusion in the FTA prototype.

2. *FTA focussed on exploration of the SUT*

   This is a type of agent that does exploration and the validation of the robustness of the SUT. Decision making that executes non-sequential actions allows detecting abnormal and unexpected SUT behaviors.

Plans for 2nd year

Over the next year, the integration of FTAs on the XR systems of industrial partners will begin. Also it is planned to research and develop a second prototype that will include:

- Reinforcement learning (RL) algorithms and strategies to diversify executions on existing FTAs. We will consider RL rewards that are general and can improve exploration as well as domain specific RL rewards for optimal goal solving (This is conform with Task 3.2).
- Improve the action/tactic selection and maximize goal test coverage (This is conform with Task 3.4).

1.2.4 WORKPACKAGE 4

Work Package 4 focuses on developing socio-emotional test agents (SETAs) to aid the systematic assessment of user experience of XR systems while minimizing the manual effort.
The current implementation of the SETA (task 4.1) is based on the two-dimensional core affect theory, modelling the pleasure and arousal emotional dimensions. This implementation will be improved and validated in the following months through user-testing based on physiological measurements (e.g., electrodermal activity).

WP4 has studies in progress to build intelligent agents capable of simulating user-experience during interaction with XR systems (task 4.2), including:

1. Testing a model to predict the cognitive load imposed by a game. This is the first step towards developing a model that can tell designers the amount of cognitive load that their systems are imposing on the user.

2. Testing a machine learning model that automatically assesses if a scenario promotes a particular behaviour of its users, for example, cooperation. This is done by using machine learning to learn a policy to solve the scenario, and comparing the policy learn to a set of policies that define the baseline we want to test.

A first implementation of the SETA has already been integrated with the iv4XR Framework (Task 4.5) and used to test several maps of the Lab Recruits game. Further details on this integration are provided in document D4.1- “1st prototype of SETAs” deliverable.

WP4 meets on the last Monday of the month to discuss the socio-emotional test agents (SETA). Partners present the ongoing work and share research plans or results, which are then discussed with the attendees (these discussions have been linked to tasks 4.1 and 4.2). The second part of the meeting is usually about test coverage and what it means for SETA (affective coverage - task 4.3). These discussions are fruitful not only for the project development but also because they allow students to get more feedback on their work.

Plans for 2nd year

The current pandemic delayed plans for some user studies, namely those that require the collection of physiological data. We are currently preparing those studies that should take place in the coming year. We will expand and improve the model to predict the cognitive load imposed by interacting with a system. We will also expand the automatic assessment of scenarios to different behaviours beyond cooperation and we are validating our model with real users. We have started to explore machine learning approaches to assess the difficulty of scenarios and possibly estimate the experience/capabilities of users interacting with the system.

1.2.5 WORKPACKAGE 5

Work Package 5 focuses on the preparation of pilot programs by the industrial partners in order to showcase how iv4XR can be used. These pilots are a 3D game, a simulation scenario depicting the infiltration of a nuclear power plant, and a sensor suite for the monitoring of civil engineering projects.

In the first year, the work has focused on the integration of the iv4XR system with the various pilots. The original plan was to have completed the second stage of the integration of iv4XR with the pilots, however COVID-19 has delayed that progress for some of the partners, necessitating the 3 month extension to the project.

Task 5.1 - “Preparing the Pilots” concerns the preparation of the pilots - this effort is primarily reflected in the progress towards deliverables D5.1 - D5.3 reporting the integration of the pilots. The pilot programs require some adaptation to allow for interfacing with the framework, and for the case of the power plant the ability to greatly speed up the simulation to make it suitable for reinforcement learning methods. Deliverable 5.1 - “Basic integration of the pilots” was produced

in this period which resulted in public code releases of the initial integrations of the pilots on the project's github page[2].

Task 5.2 - "Pilot Environment" is about building the testing scaffold for the pilots so that an iv4XR test suite can be executed with minimal human intervention. Some of this work has been started on the industrial partner side with their integration efforts including features such as the automatic loading of testing scenarios without the need to select them from inside the pilot itself.

Task 5.3 - "Demos" is about packaging the pilots into a form suitable for the public. The initial stages of this have happened as a result of D5.1 with the launching of the public repositories.

Plans for 2nd year

The next year will see the completion of D5.2 and D5.3, and will mark the end of the pilot integration with iv4XR. We also anticipate that coordination between the industrial and academic partners will increase as the functionality to support the test cases is completed and agents hosted by the iv4XR system start to be developed.

Now that each of the pilots has some basic connection to the ivXR framework, similarly to the other work packages, there will be a regular meeting scheduled for WP5 where progress and requests can be shared. This will be crucial in permitting the coordination mentioned above and will allow the industrial partners to share their test cases, and the academic partners to share their requirements for the interfaces between the pilots and the framework.

1.2.6 WORKPACKAGE 6

The tasks in WP6 have been executed according to the original plan:

- The website is in place https://iv4xr-project.eu/ (D6.1)
- The data management plan has been delivered (D6.2). It includes an assessment of the various data generated and maintained in the project by all the partners and outlined the project's management plan with respect to the storage and dissemination of the various data produced in the scope of the project.

[2] https://github.com/iv4xr-project

- The 1st Dissemination and exploitation plan has been delivered (D6.3) and it describes more detailed plans for dissemination and the results for the first year for Tasks 6.1, 6.2 and 6.3.

Plans for 2nd year

For next year, the task will continue to carry out dissemination and communication activities as outlined in D6.3 and we will start to work on exploitation. One of the important tasks next year will be the organisation of the first project workshop.

1.2.6 WORKPACKAGE 7

Workpackge 7 has been included to address ethical questions and guideline development of the project. Its main goal was to develop deliverable D7.1 - “POPD - Requirement No. 1” and assure that its principles and guidelines are followed.

We created the Ethics Advisory Board (POPD requirement No.1) to maintain information on all studies conducted with humans within the project. The Ethics Advisory Board (EAB) is responsible for ensuring that all studies that involve humans are submitted to an Ethical Review Committee (ERC) for approval. The EAB will also help the partners compile the required materials for ERC submissions, inform all partners about the decisions, and help with any modifications requested by the ERC's. This Advisory Board will also be tasked with assisting all partners in complying with the principles defined in D7.1.

2. FINANCIAL SUMMARY

The following tables detail the planned and allocated resources from M1 to M15

	WP1 - Project Management		WP2 - The iv4XR Framework		WP3 - Functional Test Agents		WP4 - Socio-emotional Test Agents		WP5 - Pilots and Validation		WP6 - Dissemination & Exploitation			
	PLANNED PMs	ACTUAL PMs	PLANNED PMs	ACTUAL PMs	PLANNED PMs	ACTUAL PMs	PLANNED PMs	ACTUAL PMs	PLANNED PMs	ACTUAL PMs	PLANNED PMs	ACTUAL PMs	TT PLANNED PMs	TT ACTUAL PMs
INESC-ID	4.62	6.23	4.92	6.23	8.77	10.42	11.69	13.40			1.44	1.95	31.44	38.23
U.UTRECHT	1.54	0.61	17.78	16.37	9.83	10.90	0.69		1.63	1.73	1.16	0.66	32.63	30.27
FBK	1.54	1.93	1.77	2.81	11.96	14.97	3.33	5.25			1.96	3.47	20.55	28.43
UPV	1.54	1.64	8.88	4.09	16.60	5.76	2.40	0.04	1.75	0.39	1.57	0.82	32.75	12.73
GWE	1.54	1.68	8.08	10.53	1.15	1.01	0.18		10.89	9.12	1.77	2.01	23.61	24.35
GA	1.54	0.60	7.08	2.09	3.95	0.25	0.18	0.00	6.32	2.89	1.44	0.52	20.50	6.35
THALES-SIX	0.77	0.83	1.38	2.00	5.16	6.45	0.74	1.58	0.92	1.10	0.87	0.74	9.84	12.70
THALES-AVS	1.15	1.15							6.21	8.51	0.58	0.05	7.95	9.71
UMU	0.19	0.30	0.51	0.06	0.55	0.39	0.96	0.66			0.15	0.09	2.36	1.50
	14.42	**14.97**	**50.40**	**44.18**	**57.97**	**50.15**	**20.18**	**20.93**	**27.73**	**23.74**	**10.94**	**10.31**	**181.64**	**164.28**

Table 1: Effort per partner and WP

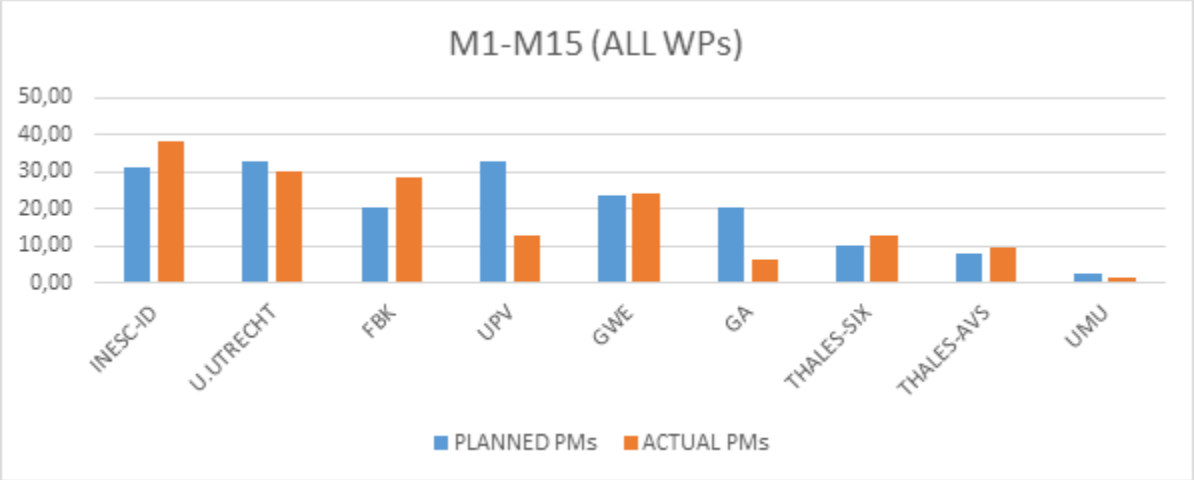

Figure 1: Personnel costs per partner (all WP)

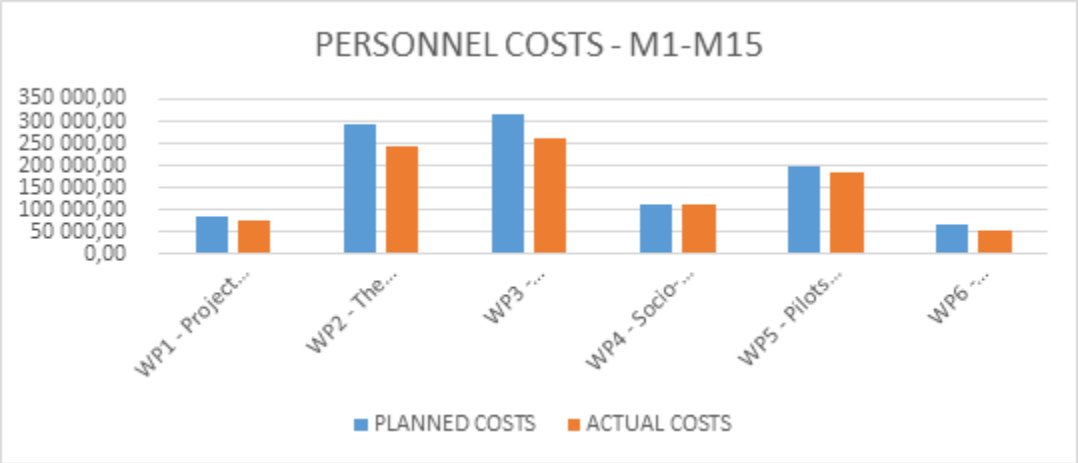

Figure 2: Personnel costs per WP

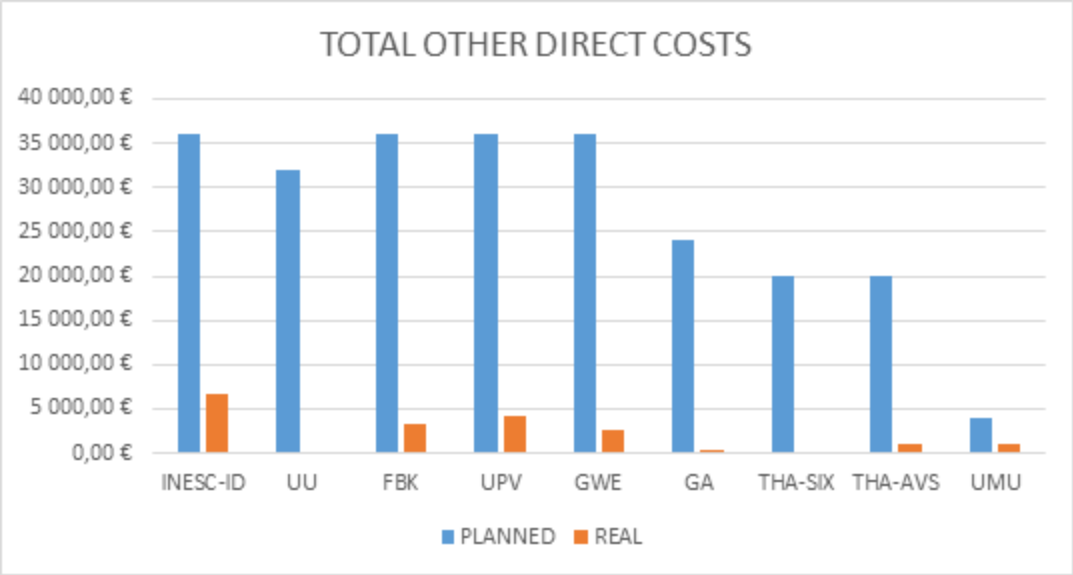

Figure 3: Total of other direct costs per partner (all WP)

Overall the actual costs are lower than the estimation. This difference arises from the fact that one project meeting was held online due to the current pandemic. Participation in conferences and workshops was also moved online, which decreased a lot of travel costs. The delay in data collection using physiological measures accounts for lower spending with equipment. Aside from these minor issues, the budget is also on track, and we have no significant deviations to report.

3.CONCLUSIONS

This deliverable provides an overview of the work carried out in each work package throughout the first year. As should be clear from the document, we consider the project work to be well on track and all deliverables have been submitted on time. All outputs from the project can be found on Zenodo[3], Github[4], Project Website[5] and on Social Media: Twitter[6] and Facebook[7].

[3] https://zenodo.org/communities/iv4xr-project/?page=1&size=20
[4] https://github.com/iv4xr-project
[5] https://iv4xr-project.eu/
[6] https://twitter.com/iv4xr
[7] https://www.facebook.com/iv4xr