



D1.5: Final Periodic Report

Dissemination level: Public

Document type: Report

Version: 1.0.0

Date: August 31st, 2020



This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement #769553. This result only reflects the author's view and the EU is not responsible for any use that may be made of the information it contains.

Document Details

Project Number	769553
Project title	Council of Coaches
Title of deliverable	Final Periodic Report
Due date of deliverable	August 31 st , 2020
Work package	WP1
Author(s)	Jorien van Loon (CMC), Harm op den Akker (RRD), Tessa Beinema (RRD), Randy Klaassen (CMC), Oresti Banos (CMC), Sofoklis Kyriazakos (iSPRINT), Konstantina Kostopoulou (iSPRINT), Mark Snaith (UDun), Catherine Pelachaud (SU), Reshmashree Bangalore Kantharaju (SU), Sita Ramchandra Kotnis (DBT), Álvaro Fides Valero (UPV), Marian Hurmuz (RRD), Kostas Konsolakis (CMC), Stephanie Jansen-Kosterink (RRD)
Reviewer(s)	Harm op den Akker (RRD)
Approved by	Coordinator
Dissemination level	PU: Public
Document type	Report
Total number of pages	78

Partners

- University of Twente – Centre for Monitoring and Coaching (CMC)
- Roessingh Research and Development (RRD)
- Danish Board of Technology Foundation (DBT)
- Sorbonne University (SU)
- University of Dundee (UDun)
- Universitat Politècnica de València, Grupa SABIEN (UPV)
- Innovation Sprint (iSPRINT)

Abstract

This document contains a summary of all results achieved in the Council of Coaches, with a focus on the final 18 months of the project, covering the period from the first review of the project in February 2019 until the end of the project in August 2020.

Table of Contents

1	Introduction.....	7
2	Summary of Progress.....	8
2.1	Summary.....	8
2.2	List of Achievements	9
2.3	Overall project choices	10
2.3.1	Demonstrator explanation.....	11
2.3.2	Explanation on UI Input Modalities.....	11
2.3.3	Consent dialogue	12
3	Project Objectives.....	13
4	Work Packages	18
5	WP1: Management	19
5.1	Planning	19
5.2	Main Achievements.....	19
5.2.1	First Period (M1-M18).....	19
5.2.2	Second Period (M19-M36)	19
5.3	Individual Tasks.....	19
5.3.1	T1.1: Coordination, general management.....	19
5.3.2	T1.2: Administrative, Legal and Financial Management	22
5.3.3	T1.3: Technical and Scientific Coordination	23
5.3.4	T1.4: Continuous Quality Control, Risk Identification and Mitigation Strategies.....	24
5.3.5	T1.5: Innovation Management	24
6	WP2: Responsible Research and Participatory Design	26
6.1	Planning	26
6.2	Main Achievements.....	26
6.2.1	First Period (M1-M18).....	26
6.2.2	Second Period (M19-M36)	26
6.3	Individual Tasks.....	27
6.3.1	T2.1: Developing a Shared Vision of RRI for the Council of Coaches	27
6.3.2	T2.2: Stakeholder Engagement Process	29
6.3.3	T2.3: User Needs and Continuous Evaluation.....	31
6.3.4	T2.4: Socio-Technical Integration	34
7	WP3: Coaching Strategies and Knowledge Base	35
7.1	Planning	35
7.2	Main Achievements.....	35
7.2.1	First Period (M1-M18).....	35
7.2.2	Second Period (M19-M36)	35
7.3	Individual Tasks.....	36
7.3.1	T3.1: Definition of Tailored Coaching Strategies.....	36



7.3.2	T3.2: Definition of Coaching Actions and Content	37
7.3.3	T3.3: Development of Shared Knowledge Base.....	39
8	WP4: User Behaviour Sensing, Modelling and Analysis.....	41
8.1	Planning	41
8.2	Main Achievements.....	41
8.2.1	First Period (M1-M18).....	41
8.2.2	Second Period (M19-M36)	41
8.3	Individual Tasks.....	42
8.3.1	T4.1: Inference of Short-Term Behaviours from Sensor Data.....	42
8.3.2	T4.2: Identification of Long-Term Behaviours from Short-Term Information.....	43
8.3.3	T4.3: Automatic Detection of Behaviour Changes.....	44
8.3.4	T4.4: Evaluation of Holistic Behaviour Analysis Framework	45
9	WP5: Dialogue and Argumentation Framework (UDun).....	47
9.1	Planning	47
9.2	Main Achievements.....	47
9.2.1	First Period (M1-M18).....	47
9.2.2	Second Period (M19-M36)	47
9.3	Individual Tasks.....	48
9.3.1	T5.1: Theoretical Model.....	48
9.3.2	T5.2: Abstract Model	48
9.3.3	T5.3: Computational Model.....	49
9.3.4	T5.4: Evaluation of Framework	50
10	WP6: Human-Computer Interfaces	51
10.1	Planning	51
10.2	Main Achievements.....	51
10.2.1	First Period (M1-M18).....	51
10.2.2	Second Period (M19-M36)	51
10.3	Individual Tasks.....	52
10.3.1	T6.1: User Interface Framework for Home and Mobile.....	52
10.3.2	T6.2: Design and Creation of Virtual Agents.....	54
10.3.3	T6.3: Modelling Group Interaction	56
10.3.4	T6.4: Engagement Model	57
10.3.5	T6.5 Simulation of Council of Coaching Sessions	58
11	WP 7: Continuous Integration and Demonstration.....	60
11.1	Planning	60
11.2	Main Achievements.....	60
11.2.1	First Period (M1-M18).....	60
11.2.2	Second Period (M19-M36)	60
11.3	Individual Tasks.....	61
11.3.1	T7.1: Technical Requirements, Specifications and Maintenance of APIs	61

11.3.2	T7.2: Development and Integration of the Functional Prototype.....	62
11.3.3	T7.3: Development and Integration of the Technical Prototype.....	64
11.3.4	T7.4: Final Demonstration.....	65
12	WP 8: Dissemination and Exploitation.....	67
12.1	Planning.....	67
12.2	Main Achievements.....	67
12.2.1	First Period (M1-M18).....	67
12.2.2	Second Period (M19-M36).....	67
12.3	Individual Tasks.....	68
12.3.1	T8.1: Dissemination, Communication and Pre-Marketing Activities.....	68
12.3.2	T8.2: Exploitation and Business Planning.....	69
12.3.3	T8.3: Ecosystem Building of the Open Agent Platform.....	70
12.3.4	T8.4: Standardization Activities.....	71
12.3.5	T8.5: Training activities.....	71
13	WP9: Ethics.....	73
13.1	Planning.....	73
13.2	Main Achievements.....	73
13.3	Individual Tasks.....	73
13.3.1	T9.1: H- Requirements No. 1.....	73
13.3.2	T9.2: POPD- Requirements No. 2.....	74
14	Bibliography.....	76

List of figures

Figure 1: Work-in-progress screenshot of the Council of Coaches functional demonstrator with integrated speech recognition.	12
Figure 2: Set of coaches implemented in the Council of Coaches VCS (Virtual Coaching System).	13
Figure 3: Screenshots of the Council of Coaches Virtual Coaches System.	14
Figure 4: The WOOL dialogue for Carlos' (Peer Coach) "Experience Sampling Dialogue", an example of a coach-as-a-sensor implementation for sensing user's mood data (screenshot taken from the WOOL Dialogue Editor).....	15
Figure 5: Screenshot of the Agents-United.org website, the entry point for the community surrounding the Agents United Platform.	17
Figure 6: Four appearances modelled to fit authoritative and peer agent role.....	32
Figure 7: Initial design of the knowledge base.....	39
Figure 8: Operation flow of the HBAF.	42
Figure 10: Mobile UIs inspired by WhatsApp and chatbox styling.....	53
Figure 10: A Unity scene with multiple cameras, resembling video conferencing systems like Skype or Microsoft Teams.....	54
Figure 11: Character designs, from left to right Helen, Hank, Owen, Alexa, Francois, and Melissa.	55
Figure 12: The agents subjected to testing, differing in gender, age and role.	56

Symbols, abbreviations and acronyms

AAL	Ambient Assisted Living
ASAP	Articulated Social Agents Platform
CMC	Centre for Monitoring and Coaching
COUCH	Council of Coaches
D	Deliverable
DBT	Danish Board of Technology Foundation
DoA	Document of Action
EC	European Commission
H2020	Horizon 2020
HBAF	Holistic Behaviour Analysis Framework
ISPRINT	Innovation Sprint
KB	Knowledge Base
M	Month
MS	Milestone
PM	Person Month
RRD	Roessingh Research and Development
SU	Sorbonne University
UDun	University of Dundee
UPV	Universitat Politècnica de València
UT	University of Twente
VCS	Virtual Coaching System
WP	Work Package

1 Introduction

This deliverable reports on the progress of the Council of Coaches project in the final 18 months. Section 2 provides a summary of the project including an overview of the work performed and a description of the major achievements of the project. Additionally, we explain important project choices that have been made during the first period and the follow up of these choices in the final period. In Section 3, we describe the objectives of the project and report on how we have achieved them. After that, in Section 4, we describe for each Work Package the progress, any deviations from the project's Document of Action (DoA), and the main achievements per work package. Some tasks have reported all of their progress in their respective deliverables already. In this case, we refer to those deliverables, no additional technical progress will be reported here.

This Deliverable 1.5: Final Periodic Report serves as the basis for the official project Periodic Report, to be submitted to the European Commission. In the periodic report there will also be an overview of the finances and the justification on the PM's and budgets. Given the public nature of the document no financial information is reported here.

2 Summary of Progress

2.1 Summary

Society is aging at a rapid pace. Our parents and grandparents are getting older than any generation before them, but in the meanwhile, they develop many age-related impairments. They visit their doctor's office frequently but would probably be happier not to. The doctors on the other hand spend many hours providing general care at huge costs, which could in many cases be taken over by technology. This would lead to care with more attention for the patient at lower costs.

Since everyone recognizes these benefits, many coaching applications have been developed in which an end user is being coached by a virtual coach for different kinds of ailments. In these kinds of applications, the success rate depends heavily on the relation you, as the end user, will establish with the coach. The coach will need to build a personal connection with you and must know how to handle all kinds of tricky obstacles you can throw at the coach. If you are not really willing to change, you could place all kinds of linguistics roadblock, for which the coach is not equipped of dealing with.

Coach Anton: *"Let's take a look at your step count"*

User: *"No, I don't want to".*

You have clearly ended the discussion right there and then.

Researchers within the Council of Coaches project were wondering if there was a way around these roadblocks in coaching dialogue. How can we develop a virtual setting that can equip a coach to handle such cases? The solution can be found in a Council of Coaches. In this council, multiple virtual agents, each with their different personality, character, and speciality, take turns in a dialogue, while interacting with you. In our project we aim to develop a supportive virtual coaching team of two to five coaches that can help you with a range of health relation issues. A conversation with the coaches might now look like this:

Coach Anton: *"Hi, shall we take a look at your step count?"*

User: *"No, I don't want to."*

Coach Betty: *"Ok, I understand, let's go back to Anton later. Can you tell me what you have been doing lately?"*

Bringing multiple virtual coaches to a table also introduces a lot of challenges. First of all, the interaction between the coaches should feel natural and realistic to the end user. Therefore, the Council of Coaches project is developing tools to be able to adjust the body language of the coaches, such as gazing and gesturing, to reflect natural discussions. For the coaching to have a convincing dialogue, the speech patterns used by the coaches should be adapted to the personalities of the coaches and the number of coaches as well.

Secondly, for the coaches to be convincing to the end user, the coaches will need to be responsive to user's individual preferences. What these preferences are, and how we can develop a virtual agent reflecting these preferences is part of our research as well. Making one coach likeable to the end user is easy, making a virtual team likeable is another story. For example: Do we need four likeable coaches, or do we need one bad cop, to make the other ones look good?

Thirdly, how can we develop coaching strategies that can be applied by multiple coaches. Each coach will add something to the discussion, but who is in charge? And what if they contradict each other? When we send reminders to a mobile phone, do we need it from one coach specialized in the topic, or do we need an intervention on a larger screen by the entire council? From a technical point of view, the challenge is: can we unobtrusively gather data on your behaviour, so the council doesn't have to bother you with asking all kinds of questions every time you meet. And can we then use this data as information in coaching strategies and dialogues between the coaches and the end user.

So far, the Council of Coaches project has been working on solving these questions, and many more. In our attempt to drive innovation in this critical domain of health and wellbeing, the project has embraced an approach based on principles of Responsible Research and Innovation – one in which we constantly

reflect on the potential impact of our work. In **Deliverable 2.7: Report on ethical, social and legal considerations and implications** and **Deliverable 2.8: COUCH Participatory Innovation Process** we have reported on how we have implemented this responsible research and innovation approach, and discuss the major outcomes of following this process.

Broadly speaking, the research and work performed in the project can be divided into two main tracks. On the one hand, the project has aimed to answer “fundamental” questions on how having multiple coaches in a room effects the interaction with the end-users: does it work? Is it fun? Do people listen to what the coaches have to say? To answer these type of questions, the project has created a very stable, and easily deployable demonstrator – it runs in your web browser, and you can check it out at www.council-of-coaches.eu/beta – and tested this demonstrator with a large number of older adult end-users (in the project plan we aimed for 50, but we ended up with more than 90). The results of this evaluation are reported in **Deliverable 7.7: Final Demonstration Results**, given us more than enough inspiration to create further improvements to the Council of Coaches product, and continuing on our path towards product exploitation, as described in **D8.8: Final Report on Standardisation and Exploitation**.

The second track is focused on the more technical challenges: how can we create a platform in which you can “plug” a group of (more or less) autonomous virtual coaches. Can we support different technology implementations for the individual coaches, while still allowing them to collaborate on the same team (and gaze at each other, respond to each other, etc...). We strongly believe that Council of Coaches is just the beginning of a whole field of research on multiple embodied conversational agents interacting with users to achieve some goal (like coaching, but not limited to it). Therefore, it’s important to lay the groundwork for our colleagues in research and industry to start experimenting with these novel technologies. This work has resulted in a fully open source embodied conversational agents’ platform: Agents United – a platform that is being maintained by the newly founded Agents United Alliance (with initial members of the consortium partners, but already expanding). Find out more about the platform on the website: www.agents-united.org and check out some of the more technical deliverables for additional background information (e.g. **Deliverable 6.6: Final virtual coach design and model**).

2.2 List of Achievements

The aim of the project has been to develop a coaching application that covers multiple coaching areas such as psychical, mental, social and cognitive aspects, suitable for elderly people with age related impairments, which is still very engaging to the user and will encourage them to keep on using the system in order to improve their quality of life. Our solution is a virtual coaching platform, where multiple coaches will coach you in an interactive and engaging way. Coaches you can bond with and coaches you trust with your sensitive data. Below, we will guide you through all the achievements towards this goal.

Firstly, several different subsystems have been developed and integrated. Since the different partners worked with different systems to visualize the coaches, develop dialogues and measure sensor data, the different systems needed to be integrated. More specifically, the University of Twente’s ASAP platform has been integrated with Sorbonne University’s GRETA platform, as well as University of Dundee’s DGEP platform, RRD’s Knowledge Base and WOOL platform (specifically developed for Council of Coaches) and the UT’s Holistic Behaviour Analysis Framework. Together these systems have been the base of the technical demonstrator, which has been further developed in the Agents United Platform.

For the entire technical demonstrator multiple mobile interfaces have been developed (WhatsApp / Chatbot / Skype) as well as virtual reality interfaces. This way, the coaching can take place through multiple channels, depending on the preference of the end-user. A speech-based interface has been implemented as well (and will be evaluated in the next months). For multiple coaches in the same platform to be convincing, they require realistic body language and interaction. Multiple studies have been conducted on body language, turn taking, gesturing, as well as conflict resolution and other non-verbal behaviour to create coherency in a group.

The coaches not only interact through body language but through dialogues as well. To have a realistic dialogue is almost as important as the body language. However, we do not want to write all dialogues by hand. Steps towards solving this problem has been taken by implementing the utterance generator. This generator will find the next “move” in a dialogue that will make sense in a natural conversation.

For the coaches to be able to give specific feedback the integration with the Holistic Behaviour Analysis Framework has been completed. This HBAF will send information based on the sensing of physical, emotional, social and cognitive aspects of an end user to the Knowledge Base, which stores all information regarding the end users, dialogues, coaching strategies etc. From the KB, the information on e.g. steps can be used as coaching content in the dialogues. For the HBAF to collect information, the Coach-as-a-Sensor concept has been further developed to be able to gather information from the end user on both their emotional as well as cognitive wellbeing. The possibility to detect changes in short term as well as long term behaviour changes are now in place.

The technical demonstrator has been further developed into the Agents United Platform. This platform hosts the open source software necessary for other user groups to develop their own multi agent coaching systems. The Agents United Platform has its own website, newsletter and mailing list as well as a GitHub page where all components are downloadable, documentation has been setup and the community will be further established. Training material have been developed for different components within the project. The future of the platform will be coordinated through the Agents United Alliance, a newly established cooperation with the sole purpose of maintaining the uptake and expanding the reach of the platform.

The project has developed a highly robust, flexible and simple to use demonstrator to test the fundamentals of multiple coach interactions. This demonstrator has in the second half of the project been further extended with more suitable dialogue generation as well as an automatic topic selection module. A new standard has been developed, WOOL, enabling content authoring by domain experts. The knowledge base has been integrated more fully. Furthermore, an enormous task was to develop enough coaching content to be used during the six weeks final evaluation period featuring multiple coaches. The functional demonstrator has been extensively tested (under new and challenging circumstances) to evaluate the engagement of the end user with the platform, the engagement with the different coaches and different coaching strategies for different end user’s motivations. The functional demonstrator forms the basis for the “Coach as a Product” line of exploitation, for which a business case and further exploitation plan has been described in D8.8.

During the project’s technical development, the RRI issues have been kept on the radar through dedicated RRI workshops and visits from the RRI team to researchers working in their labs. External stakeholders have been asked for advice on the developments within the projects through a dedicated RRI workshop. A final overview has been documented in D2.7 on the ethical, social and legal implications that have arisen during the project and or that may arise during exploitation. A dedicated Participatory Innovation Process for RRI has been developed within the project, documented in D2.8.

The entire project has been communicated and disseminated through our website, social media channels as well as a final Workshop with external experts. Additional communication efforts have been set up to start building the Agents United Community. In total 24 papers have been published in journals and in conferences, while a couple papers are still under review. However, the dissemination does not end with the project, two more workshops and a press release are planned for the future, as well as submitting papers of work currently still being performed.

Overall, the project has executed the Proposal as originally intended. We are exceptionally proud of all researchers who have worked extremely hard during the entire project, but especially the last months under the challenging new circumstances. The two main exploitable outcomes of the project are the result of great cooperation between the different partners and we stand by them in the future growth and development.

2.3 Overall project choices

In the paragraphs below we will describe some bigger overall project choices we have made. In the first period, some big project discussion, concerning the demonstrators, UI input modalities, and the consent

dialogues have been made. In the second period, most of these choices have been further integrated in the project, or have, due to renewed insights been abandoned. We provide more details below on our “Demonstrator” pathway (Section 2.3.1), User Interface Input Modalities (Section 2.3.2) and the concept of Consent Dialogues (Section 2.3.3).

2.3.1 Demonstrator explanation

As explained in the first progress report, we have two pathways for the development of the demonstrators; the functional and the technical demonstrator. Both demonstrators have been expanded in the final period and new developments have been incorporated. Both demonstrators have been extensively tested. Two clear and distinct roadmaps for both the functional and the technical demonstrators have been developed, resulting in the two core technical achievements of the project. The functional demonstrator is a first prototype of “*Council of Coaches as a Health and Wellbeing Product*”, of which we have demonstrated its potential in the project’s final demonstration trial. For the technical demonstrator, the Agents United Platform has been developed – an open source platform allowing researchers and developers to create their own multi-agent dialogue systems for various domains. Both demonstrators and their testing have influenced further development in the other demonstrator, resulting in platforms that have been mutually beneficial. We fully expect that this will continue to happen in the future: findings from the Agents United Platform research and development activities are meant to find their way in the “Council of Coaches product”, while the real user feedback from the “product” track should prove to be useful for the Agents United research community.

In order to create a coherent prototype, there is much more to do than just the functional and technical demonstrator. In modern development, integrating stakeholders’ wishes, creating solutions for ethical issues such as privacy and thinking about exploitation even before any development starts, are key to developing a prototype that end users will actually use when developed into a product. In the Council of Coaches, we have tried to incorporate all those elements from the start of the project. All researchers have been aware of the importance of these aspects and have been taking them actively into consideration while developing their solutions.

2.3.2 Explanation on UI Input Modalities

During the first project period (M1-M18), we have chosen not to incorporate the free speech and language recognition as an input modality. Upon request from the reviewers, we have taken up the work to implement speech recognition and perform additional studies to find how to include some form of speech recognition in the demonstrators without negatively impacting usability and user experience. In Figure 1 below, we show the work-in-progress version of the Council of Coaches functional demonstrator with integrated speech. In the top right you see additional buttons for starting/stopping speech input, and a “recording” icon to indicate that the microphone is currently actively listening for input. In this implementation the system listens for relevant “keywords” in the user’s speech, and attempts to map those onto the reply options that are available to the user. Initial experiments indicate that such a system could work with high accuracy, if manual keywords are added to the dialogue scripts (a feature supported by the dialogue framework: WOOL). Initial experiments also showed that having speech input (user talks to system) is very strange without having speech output (coaches talk to user). Therefore, work is underway to implement a realistic set of voices for the different virtual coaches in order to conduct a decent experiment to test the efficacy of text input.

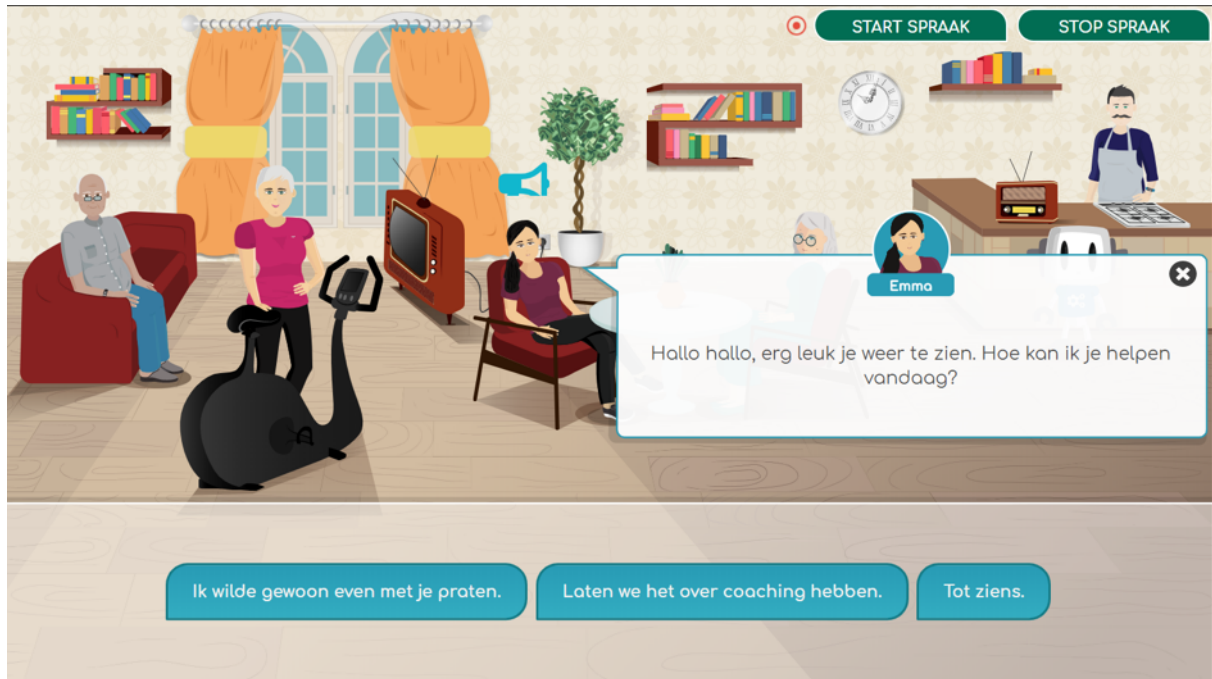


Figure 1: Work-in-progress screenshot of the Council of Coaches functional demonstrator with integrated speech recognition.

2.3.3 Consent dialogue

During the first project period, plans were developed to implement a consent dialogue. However, after one of the RRI workshops in Brussels, experts advised us that implementing such a dialogue was not as straight forward as we thought it was. From a legal perspective, you have to have participants to agree to the legal version and not an explanation of the consent form. Since the consent in our case was mainly about the trials and those could be handled very well by the researchers in the room. Additional consent on all things related to data management, would be something to be handled in the future by the developing party of the actual application. Data Management in those cases will be completely different from data management in our research environment. Therefore, we have chosen not to further implement the consent dialogue.

3 Project Objectives

Below we outline the objectives of the Council of Coaches project as set out in the Document of Action, and as proposed in the original project proposal submitted to the European Commission in January of 2017. The objectives (highlighted in bold), are followed by a commentary describing how the project has achieved those objectives.

Objective 1 – The project will design, develop and evaluate a Virtual Coaching System (VCS) targeted at older adults that are undergoing life changing events that have a potential negative impact on physical, cognitive, mental, and/or social well-being as well as elderly users suffering from NCDs, specifically Diabetes or Chronic Pain. The main success criteria for the VCS will be its ability to aid the user in preventing expected negative effects on targeted domains (physical, cognitive, mental, social) and disease specific behaviours (medication intake, diet, coping with pain) measured using Huber’s holistic model of health (Huber, et al., 2016).

The VCS that was developed in Council of Coaches is a unique approach to eHealth coaching: using a virtual council of embodied conversational coaches. Where most coaching applications provide help to the end user on a one to one basis, Council of Coaches has multiple coaches readily at hand to assist the end user on the different domains. In order to demonstrate the proof-of-concept, the project has implemented a large, but manageable set of coaches to cover various domains: nutrition, cognition, social, (peer) support, physical activity, chronic pain, and diabetes (see Figure 2).

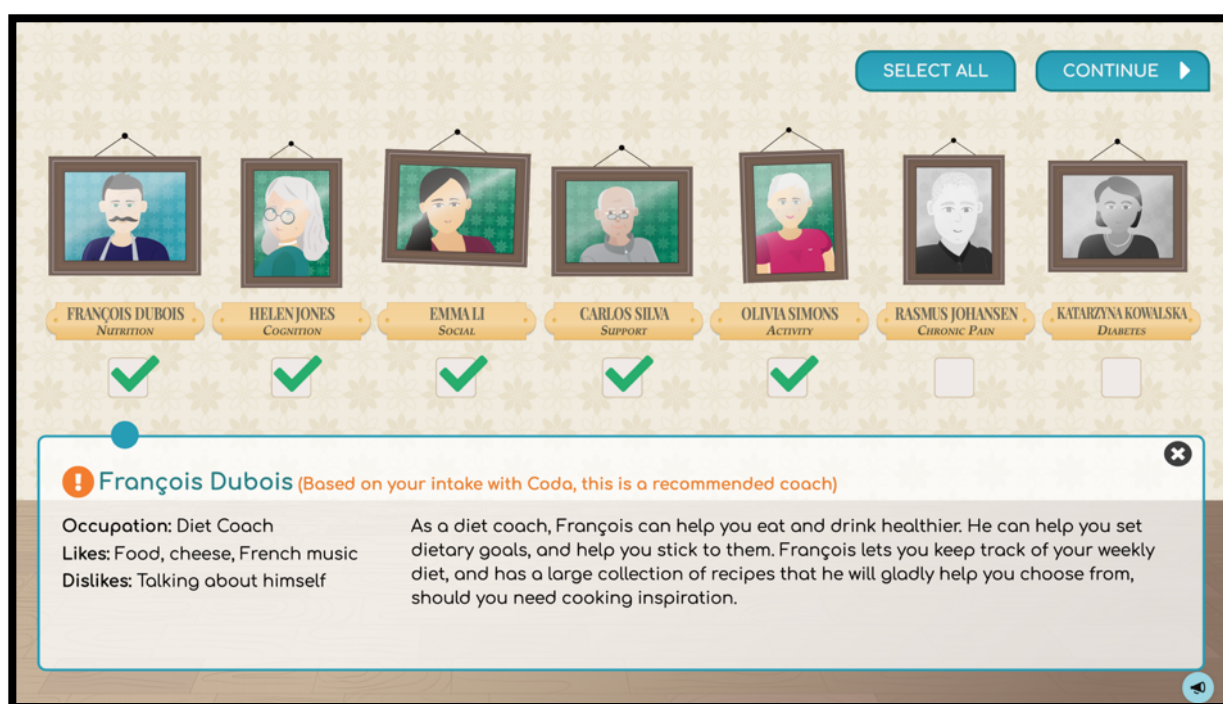


Figure 2: Set of coaches implemented in the Council of Coaches VCS (Virtual Coaching System).

This high TRL application was co-developed with a large number of stakeholders (N=82) and end-users over the course of the project, and can run as a website on the user’s tablet or desktop computer (see Figure 3 below). The final version was evaluated in a 2-phase, 2-site long-term evaluation (5 to 9 weeks) with a total of 92 older adult participants (planned: 50). Although it has become clear that in order to be successful, such products need a very large amount of coaching content embedded, almost half of the users enjoyed using the application so much that they continued using it in the facultative phase of the evaluation.

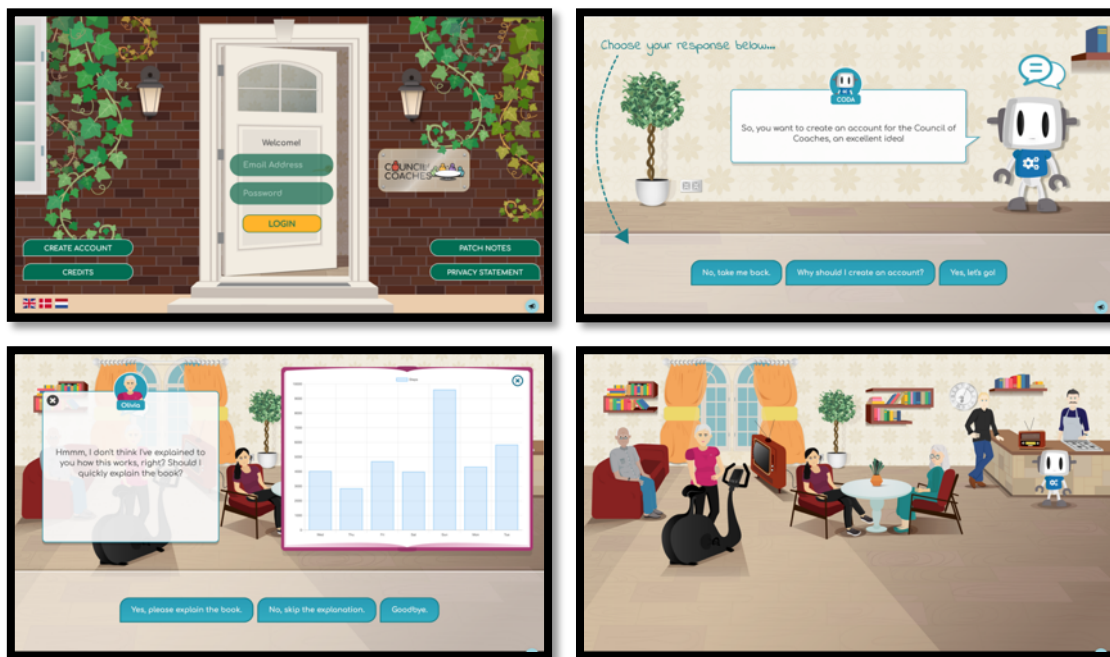


Figure 3: Screenshots of the Council of Coaches Virtual Coaches System.

Relevant details related to the fulfilling of Objective 1 can be found in the following documents:

- D7.5: Final Council of Coaches Technical Prototype – Also contains a description of the final “functional” prototype used in this evaluation.
- D7.6: Demonstration protocol, Ethical approval – Describes in detail the evaluation protocol followed during the final demonstration study (protocol was also published in (Hurmuz, Jansen-Kosterink, op den Akker, & Hermens, 2020)).
- D7.7: Final Demonstration Results – Describes the execution, collected data and results analysis for the final demonstration study.

Objective 2 – The project will introduce the new coaching paradigm of the Council of Coaches. In this paradigm, the virtual coach is manifested in a group of virtual characters that each represents a different knowledge domain of the coach. Coaching sessions consist of an interactive group discussion between the user and several members of the council, in which the user and his behaviours are the main topic of discussion. This paradigm’s success criteria are to significantly increase the engagement of the user with the system, and his willingness to actively participate in the coaching sessions.

This objective has been reached: from the first functional demonstrators delivered by the project in M9 and M15, all the way through the demonstrator used in the final project demonstration, the Council of Coaches paradigm has been demonstrated in an interactive manner.

Objective 3 – In order to provide tailored communication and personalized/targeted coaching actions, the virtual coach must be able to gather a large amount of knowledge regarding the primary user and his context. Traditional, unobtrusive on-body and in-home sensing technologies provide a basic level of information, but have limitations in the type of information that can be obtained. In order to sense the required psychological user related information, the project introduces the concept of the coach-as-a-sensor. Through the use of interactive dialogues, the coach will obtain knowledge about the primary user by engaging in interactions with the user. The success of the coach-as-a-sensor concept is based on its ability to obtain relevant information about the user in a way that is not achievable by state-of-the-art unobtrusive sensing technologies.

This objective has been achieved. The embodied virtual coaches of the Council of Coaches can autonomously (in the technical demonstrator, through the Dialogue and Argumentation Framework), or in a scripted manner (using the WOOL Dialogue Platform) personalize their content by accessing up-to-

date information from the Knowledge Base. This Knowledge Base contains, as planned, a mixture of information obtained through more “traditional” sensing methods (integration with step counters, weight scale, etc...), and information obtained directly through the dialogues. The coach-as-a-sensor concept was worked out in the context of WP4, and integrated within the project’s functional demonstrator, and finally demonstrated and evaluated in the final demonstration under WP7 of the project (see Figure 4).

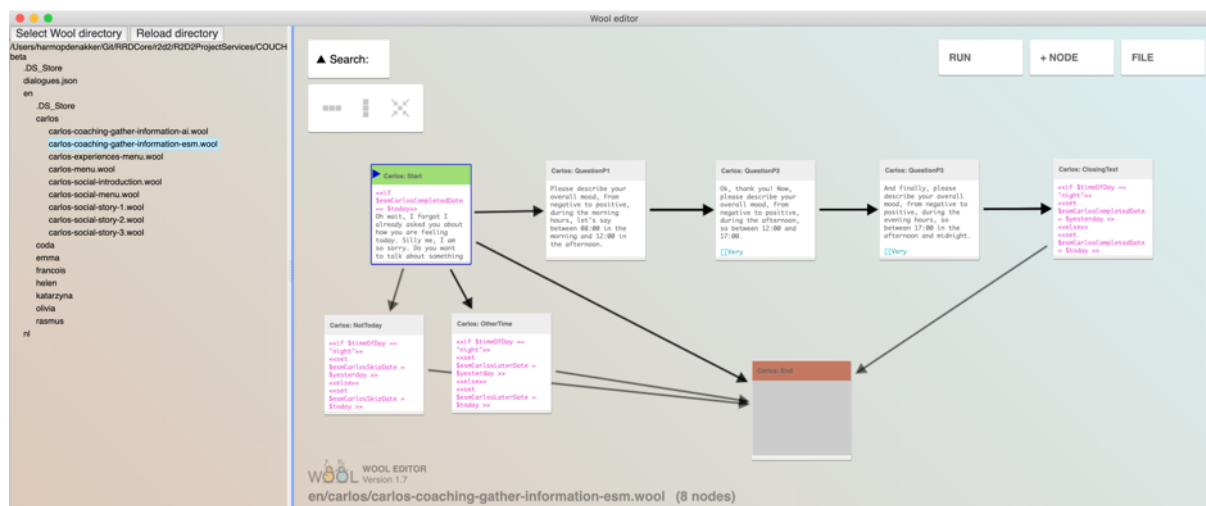


Figure 4: The WOOL dialogue for Carlos' (Peer Coach) "Experience Sampling Dialogue", an example of a coach-as-a-sensor implementation for sensing user's mood data (screenshot taken from the WOOL Dialogue Editor).

Objective 4 – The project will emphasize deep character design based on sound theories from the video game and film worlds. The added value in terms of engagement will be measured and existing methods and guidelines (Isbister, 2006) (Sloan, 2015) will be adopted to the successful design of virtual characters in coaching contexts that have an impact beyond mere entertainment.

The design of characters is a creative process. From the research into methodologies of character design it can be concluded that there is practically little to work with, especially in the vastly novel area of virtual (health) coaches. Principles from video game design often do not apply to the Council of Coaches setting, making it hard to draw from the sources mentioned in this objective. Having said that, the project has addressed character and application design from a creative point of view. In the second period of the project, a professional graphic designer was brought on board of the project team which has led to a great boost in the professional look and feel of the application and the characters. Besides the creative efforts being performed, the project has delivered multiple fundamental studies on how various character attributes are being perceived by an end user – the results of which will be used to tailor the specific coaches to the end users in a personalised approach, therefore improving the engagement of the end user. The experiments performed, and numerous studies still ongoing are planned to be compiled into a “hand book” on how to design virtual coaches in the eHealth domain.

Objective 5 – The project will combine smart multimodal sensing technologies in order to seamlessly and opportunistically measure and model the user behaviour in a comprehensive fashion, including physical, cognitive, mental and social aspects. This holistic sensing and modelling approach not only aims at registering, analysing and inferring each determinant of behaviour in a user-centric manner but also mining the interactions among users and with their physical and virtual environment.

This objective has been attained via the development of one of the core engines of COUCH, the so-called Holistic Behaviour Analysis Framework (HBAF). In a nutshell, the HBAF consists of a set of AI-powered methods that process and analyse heterogeneous data collected via smart devices in order to broadly identify human behaviour at different time scales. To do so, the HBAF combines behaviour data registered via physical devices (namely commonplace sensors available on smartphones and wearables like accelerometers or phone logs, which passively measure some user's explicit actions such as steps taken or number and duration of phone calls) and virtual environments (here

operationalised via the virtual council, leading to the novel "Coach-as-a-Sensor" concept, which actively measures some user's thoughts and observations such as mood state or type and duration of the undertaken cognitive tasks, information that is otherwise not measurable to date via hardware sensors). Moreover, the HBAF enables us to infer human behaviour continuously at two levels of granularity or time scales. The first level, here termed as "short-term" behaviours, define the user's feelings and actions in the range of minutes, hours or even days. As an example, the framework allows us to identify the amount of active and passive physical activities for a user per minute, hour and day. This information constitutes the building blocks for the second level, here termed as "long-term" behaviours, which continuously aggregates the short-term behaviour data to generate time series in the order of weeks and potentially months to identify specific trends or tendencies (as for the example given before, it allows us for example to categorise the user as a sedentary or active person), as well as changes in such patterns which might call for the council's attention. The developed framework has been mainly designed to identify and analyse behaviour in an individual manner, i.e. based on the capacities or functions naturally expressed by a given user in their daily living. This happens to be the norm for inferring physical behaviours, cognitive and emotional behaviours, which are innerly attached to every individual. Nonetheless, the interactions with other people have been also considered to determine the nature and extent of other components of behaviours, namely those of a social nature. Hence, social interactions are analysed based here again on the data collected via physical devices (ambient noise or contact traces from the microphone and Bluetooth smartphone data respectively) and virtual environments (number, type and satisfaction with the encountered daily social interactions communicated during virtual coaching sessions).

Objective 6 – In order to accelerate open innovation and future developments in the domain of virtual coaching the project will develop the council of coaches on the principles of a generic coaching framework that builds upon and extends existing platforms such as FIWARE and universAAL to provide a set of protocols and definitions for generic, interchangeable virtual coaching modules. The generic coaching framework will be delivered as an open-source extension to existing platforms and the project will lay the groundwork for providing future support in the dissemination of this platform.

The consortium set out, early in the project, to investigate the potential of FIWARE and universAAL as platforms to support the delivery of the innovative multi-party embodied conversational agent platform that was foreseen in the Council of Coaches project. It was concluded early on that there was little to gain from following the FIWARE pathway (a framework focused on supporting IoT applications) and universAAL integrations may provide benefits only at the traditional sensing side of the Council of Coaches – e.g. focusing on sensor hardware integrations. With regards to the FIWARE and universAAL platforms, this is reflected in the outcome: where FIWARE has not proven useful, and universAAL is used in the integration of various sensors.

The main focus of objective #6 though is that the project aimed to “accelerate open innovation and future developments in the domain of virtual coaching”. That objective has certainly been reached. Practically all of the technology innovation that has taken place in the course of the project has now been released as open source software, free to use and expand upon by the community. The WOOL Platform (www.woolplatform.eu) was developed within the project (from scratch, although partly based off existing open source software) to enable non-technical experts to write scripted dialogue, and to parse and execute these dialogues within web or mobile applications. The WOOL platform has proven to be a huge success, and is already being adopted in several research projects at European level (H2020-SmartWork, H2020-Bionic, AAL-LEAVES) and within a Dutch project (PACO), thus having secured future development and support for the coming 3 years.

But WOOL is only a small part of the larger open source platform release that has come out of the Council of Coaches project: The Agents United Platform (www.agents-united.org) is a fully open source platform, consisting of different agent embodiment platforms (GRETA, ASAP), the Dialogue and Argumentation Framework (DAF), the WOOL Platform, and Holistic Behaviour Analysis Framework (HBAF).

Objective 7 – COUCH will develop an ecosystem of services, along with a community of healthcare services providers (including health professionals) and innovators around the project’s open coaching platform, as a means of sustaining and commercializing the project’s results.

The Agents United Platform has been launched as open source software on GitHub (see <https://github.com/AgentsUnited>), free for all researchers and interested parties to download, modify, and use. In order to ensure support of the platform and to promote its uptake in the years to come, the *Agents United Alliance* has been set up as the official organisation to support the platform. The Agents United Alliance (www.agents-united.org – see Figure 5) was founded by the Personalized eHealth Technology programme at the University of Twente which will support the platform, both in budget and in personnel. Official members of the AUA are some of the original consortium members: The University of Twente, Roessingh Research and Development, the University of Dundee, Sorbonne University, and the Polytechnic University of Valencia, and the first external expression of interest to join the alliance has already been obtained. The project has not simply slapped an “open source” label on some software, but has set up the groundwork for a professionally maintained and impactful open source project. Additional details on the Agents United exploitation path are provided in **D8.8: Final report on standardisation and exploitation**, while an overview of all training material and documentation that was generated in support of the platform is provided in **D8.9: Training Material**.

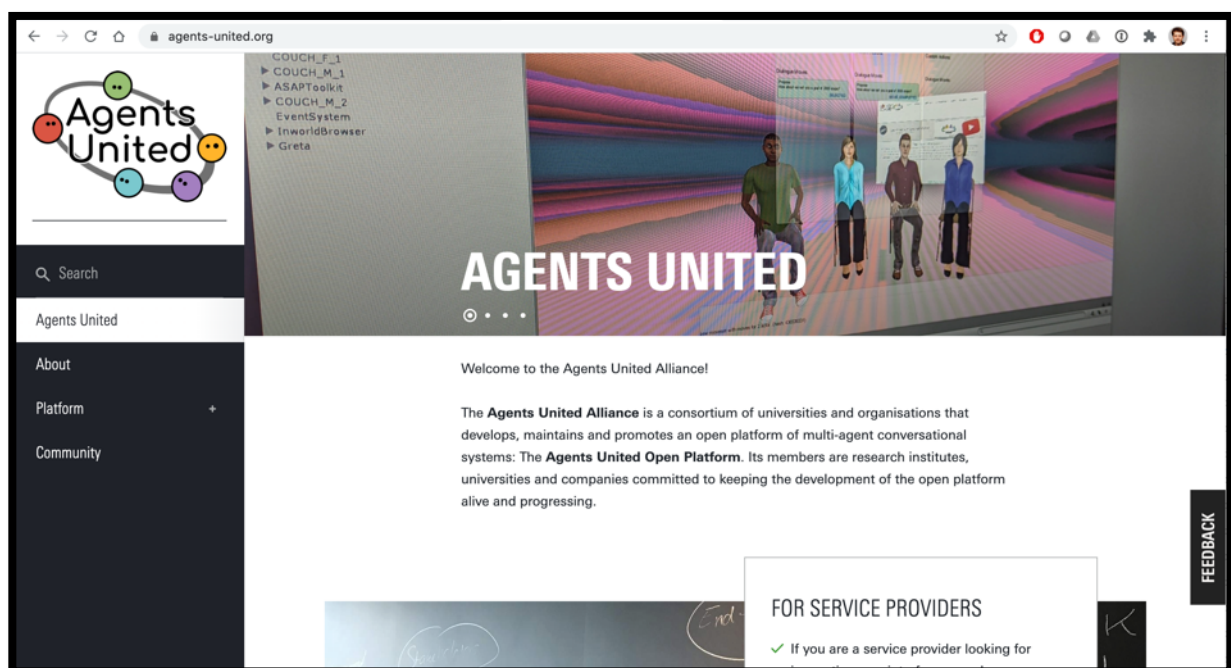


Figure 5: Screenshot of the Agents-United.org website, the entry point for the community surrounding the Agents United Platform.

4 Work Packages

In the following sections, the nine different work packages of the Council of Coaches will be addressed. An index is given below for convenience:

- WP1: Management (§5)
- WP2: Responsible Research and Participatory Design (§0)
- WP3: Coaching Strategies and Knowledge Base (§7)
- WP4: User Behaviour Sensing, Modelling and Analysis (§8)
- WP5: Dialogue and Argumentation Framework (§9)
- WP6: Human Computer Interfaces (§10)
- WP7: Continuous Integration and Demonstration (§11)
- WP8: Dissemination and Exploitation (§12)
- WP9: Ethics (§0)

Each work package progress overview is structured as follows:

- Planning and general information of the Work Package
- Main Achievements
 - First Period (M1-M18)
 - Second Period (M19-M36)
- Overview of all individual Tasks:
 - Objectives of the Task
 - Progress in the First Period (M1-M18)
 - Achievements in the First Period (M1-M18)
 - Progress in the Second Period (M19-M36)
 - Achievements in the Second Period (M19-M36)
 - Deviations from the Document of Action (DoA)

5 WP1: Management

5.1 Planning

WP / Task	Title	Participants	Months
WP1	Management	CMC	M1-M36
T1.1	Coordination, general management	CMC, RRD, DBT, SU, UDun, UPV, iSPRINT	M1-M36
T1.2	Administrative, Legal and Financial Management	CMC	M1-M36
T1.3	Technical and Scientific Coordination	RRD, CMC	M1-M36
T1.4	Continuous Quality Control, Risk Identification and Mitigation Strategies	CMC, RRD	M1-M36
T1.5	Innovation Management	iSPRINT, CMC, RRD	M1-M36

5.2 Main Achievements

The following main achievements result from this work package:

5.2.1 First Period (M1-M18)

- All procedures are put in place: Administrative, financial, communication, quality control, risk management. Excellent communications throughout the project.
- Technical progress has been established as is reported in the deliverables. All milestones have been reached in time.
- Distinctive project identity, both graphically and personally.
- Connection with other PM-15 project, standardization bodies and networks of excellence are established.
- Two Innovation workshops have been organised to create awareness of innovation and exploitation among the Council of Coaches researchers. Innovation Guidelines have been established.
- Consortium Agreement signed by all parties.

5.2.2 Second Period (M19-M36)

- Successful completion of the project.
- Managing the risks during Covid-19.
The project held a very interactive and engaging remote plenary meeting in June 2020.
- In order to keep engagement with the project and general motivation high, the project increased its regular teleconference schedule back to being bi-weekly.
- Business Model workshop, Brussels, July 2019.
- Sprint workshop, Paris, February 2020.
- Handover of exploitation potentials to WP8.
- Updated templates and procedures

5.3 Individual Tasks

5.3.1 T1.1: Coordination, general management

5.3.1.1 Objectives

This task groups all the activities to be carried out by the Project Coordinator and includes:

- a) maintaining contractual documents;
- b) issue of internal and contractual periodic reports;
- c) organising project launch (establishing procedures, project management methods and tools, organising project kick-off meeting) and follow-up (periodic project boards meetings, progress review, and conflict resolution);
- d) co-ordinating the timely production of deliverables and reports and maintaining the project archive;
- e) coordination of financial and administrative issues: establish and maintain financial records, coordinate financial statements submission by all project partners and audit certificates as needed, administer the EC financial contribution, and distribute partner shares according to the rules defined in the Grant Agreement and the Consortium Agreement;
- f) executing and controlling global expenses;
- g) providing assistance to individual partners on specific administrative issues. The task will centralize control of the project progress by ensuring administrative and contractual relationships both within the consortium and with the European Commission.

The Project Coordinator will assume responsibility for contacting the Project Officer, formulating propositions for possible modifications of the work plan, supervising contacts with all external organisations, and delivering all types of deliverables. An important key for project success will be ensuring effective collaboration and easy communication between partners and towards external entities, as well as ensuring that the project's RRI principles are executed. To this goal, this task will setup and manage several collaboration tools (e.g. internal collaboration and document sharing platform, mailing lists, virtual conferencing).

5.3.1.2 Progress (M1-M18)

Communication – The project management is besides the regular responsibilities of the financial and administrative coordination, also the responsibility of effective collaboration and easy communication between the partners. This has been a focal point from the beginning, since communication is the oil in the project's machinery.

As the project management you lead by example. We aim for open, timely, and frequent communication in the project. This way we try to ensure nothing comes as a surprise and people will have enough preparation time for deadline and deliverables. We encourage new initiatives and try to make everyone feel heard. To facilitate this, we actively encourage everyone's participation in the different meetings. Since cooperation within the project is necessary on the most basic levels, we aim for all researchers communicating directly, not only through the Work Package leaders.

To establish good communication within the project several face-to-face meetings have been organised in the first year, starting with a 2 day Kick off meeting in November 2017 in Enschede. Further consortium meetings were held in March 2018 in Valencia, June 2018 in Dundee and in November 2018 in Enschede again. During the meetings, the partners all got to know each other and were eager to share their knowledge. Also, the technical integration weeks, where researchers got together to successfully integrate the different platforms (what has always been considered a big risk in the project) provided an efficient and effective way of collaboration. During these technical meeting there has always been an RRI "intervention" in which the RRI issues were further discussed and could be taken into account while developing the prototypes further. The technical integration weeks have been held in April 2018 and October 2018 in Enschede, as well as in March 2018 jointly with the plenary meeting in Valencia. In addition to the consortium meetings and the technical integration weeks, an additional RRI workshop and exploitation workshop have been organised by the responsible parties. At both these workshops several consortium members were present as well.

Since many tasks are assigned to many partners, and vice versa, many partners are assigned to a lot of tasks, we have planned bi-weekly overall project telco's where everyone can inform the rest of the project about their work and can ask questions. This has proven to a highly successful, and easy way to take away barriers between partners (the bi-weekly call has occurred continuously since the beginning of the project, only interrupted by Christmas breaks). Furthermore, many work packages have initiated their own telco's on regular intervals, depending on the perceived necessity.

Mailing lists have been set up, to mail the different groups easily. A Slack channel has been opened for internal technical communication and for day to day business during meetings, a WhatsApp group has been created.

An online collaboration platform has been established which ensured a swift sharing of documents, video's, and other data. We are aware of the different limitations online platforms have in terms of privacy and have therefore chosen a variety of means to cooperate. The day to day cooperation goes through Dropbox. However, since privacy is limited, any kind of datasets or other information that contains personal data cannot be shared through this medium and has to be shared through a SharePoint platform hosted by the University of Twente, specifically designed for privacy purposes. After a year we ran into the limitation of Dropbox in combination with some universities administrative policies. The 2 GB limit was reached and people needed to buy a business account, which was not reimbursed by some administrations. To be able to keep using Dropbox, we started an archive of non-sensitive data on google drive. This way, all partners still have an easy to use, accessible way of sharing information, which respects the privacy of the end users and the financial situations of the researchers.

Deliverables – Another aspect of successful project management is submitting all deliverables in time. From the beginning this has been a focus point and the procedure for reviewing the deliverables has been communicated multiple times. So far, five deliverables have been submitted late. The first two deliverables were ready on time, but due to inexperience with submitting documents in the portal, were submitted three weeks late. Another two deliverables were late due to some very sad personal circumstances, whereas project management we chose for a more personal approach, instead of clinching to a hard deadline from which nobody would benefit. The last deliverable that was late (exploitation and standardization), was originally planned for May 2018 and has had a delay of six months, due to the content of the deliverable being not ready at such an early stage in the project. Furthermore, all deliverables were submitted in time.

Milestones – So far, all milestones have been reached in time. With the earlier described split between functional and technical demonstrator, even when time was really not on our side, we managed to deliver the demonstrators in time.

Project Presence – A very distinct project graphical style has been developed and is consequently used throughout the project. This has created a light and playful atmosphere, but at the same time a kind of identity for the researchers to feel at home with. So far, we have seen all researchers taking over the style and adding their own touches to it, from sign posts to baby gifts to a collection of new pictures.

In each project there are some rules and regulations that should be engraved in the researchers' memories, for instance, the dissemination disclaimer and the review procedure. For all those small but important titbits of information, we have developed a so called 2-page cheat sheet, where this information is gathered and can be easily found for researchers.

At the kick off meeting, all researchers received this cheat sheet and a limited-edition Council of Coaches coffee mug, that due to a lack of coffee, misspelled the word "coffee" Both can be put on your desk and will serve as visual reminders for researchers working on the project.

Organisation – As described in the project proposal, the project management consists of the following persons. Professor Hermie Hermens acts as coordinator for the project. Due to health reasons, professor Dirk Heylen has taken over the coordination of the project for some months, until Professor Hermens returned in January 2019. The day to day management of the project is done by Harm op den Akker from Roessingh Research and Development for the technical coordination and the administrative and process coordination is done by Jorien van Loon from the University of Twente.

5.3.1.3 Achievements (M1-M18)

Up till this point in the project, the following achievements have been reached by the project management.

- Excellent communications throughout the project;
- An amicable and friendly atmosphere during meetings and in general;
- Active participation by all project partners, and through all levels of researchers;

- Timely submitting deliverables and milestones reached;
- Distinctive project style;
- Cheat sheet with important information.

5.3.1.4 Progress (M19-M36)

Communication – The communication structure that was set up during the first period has shown fruitful in managing the project. It has been adapted over time. After April 2019 bi-weekly meetings have been decreased to a six-week interval, since all participants knew how to find each other (regular bilateral remote meetings were taking place), and we did not want to overburden them with meetings. However, once Covid-19 came around, we increased the frequency back to bi-weekly, to make sure everyone was doing alright and keeping the project on track during a challenging situation. The online communication platform has been extended from Dropbox to Google Drive, Zoom, GitHub and Slack during our adapted (online) technical integration weeks, online workshop and online consortium meetings.

Deliverables – We have rescheduled some deliverables with approval of the PO, since the deadline did not fit the timeline of the expected research results.

Milestones – All milestones have been reached in time.

Project presence – The project presence has been continued and extended during the second period as well. The project identity can also be found in the new logo for the Agents United Platform. The identity has proven effective, not only in bringing a smile to audience's faces, but also in assisting researchers in communicating what might be difficult explanations in a fun and easy way.

Organisation – No changes in organisational structure have taken place during the second period.

Covid-19 – As project management, not often do you get challenges as unexpected as the Covid-19 pandemic. We have identified the risks, set up mitigation plans and increased the communication frequency for the remainder of the project. It became evident early on that project partners did not want, and in some cases could not afford an extension. Therefore, it became urgent to keep the project going in a new way that would produce the same results. An inventory was sent out to all project partners to see which risks and challenges they were running into. These ranged from closed labs, working from home (with children), reduced access to elderly test subjects to mentally staying strong to face the crisis and reduced working hours to take care of relatives. In response, we have increased the communication frequency, to make sure we kept in contact with everyone, and to discuss plenary how to tackle the challenges. We are extremely proud of our consortium on how they handled the entire situation. Plans were re-arranged, new online technologies and communication platforms were quickly embraced, working hours were changed and everyone put in more effort than what could reasonably expected.

5.3.1.5 Achievements (M19-M36)

In addition to the achievements of period 1, of which almost all of them are ongoing during the second period, we want to add the following achievements for Period 2:

- Successful finishing of the project.
- Facilitating the participants in a re-orientation of tasks during Covid-19.

5.3.1.6 Deviations from DoA

There have been no deviations from the DoA.

5.3.2 T1.2: Administrative, Legal and Financial Management

5.3.2.1 Objectives

This task comprises all tasks associated with the administrative and financial management of the project, including regular reporting to the EC, as well as monitoring of the progress and evolution of the various tasks. Furthermore, it will deal with IPR management issues, through maintaining and updating an IPR list and regulating IPR shares of the various contributors to the COUCH product, based on the

level of their contributions and the background components that will contribute in the scope of the product's development and integration.

5.3.2.2 Progress (M1-M18)

So far, the first reporting period is coming to a close. The administrative and financial management procedures are put in place to support the reporting in a timely fashion. The IPR list has been setup, but so far, no exploitable IPR has been generated. However, an overview has been created of possible exploitable outcomes of the project and this could lead to other IPR registries as well. Considering the legal management, the consortium agreement has been signed. The only legal issue that is still open is a data processing agreement between the partners as mentioned in the progress description on the ethics work package.

5.3.2.3 Achievements (M1-M18)

The following achievements have been reached:

- The IPR list is being created.
- All financial and administrative procedures are put in place.
- Consortium Agreement signed.

5.3.2.4 Progress (M19-M36)

The administrative and financial management has been kept up. An additional financial update was asked from the partners in November 2019, to keep spending in line with the results. As we have waited for Brexit, which finally came around in February 2020, and negotiations still going forward, we have continued with the data processing as defined earlier under the GDPR.

5.3.2.5 Achievements (M19-M36)

- The IPR list has been updated
- All financial and administrative procedures have been followed.

5.3.2.6 Deviations from DoA

There have been no deviations from the DoA.

5.3.3 T1.3: Technical and Scientific Coordination

5.3.3.1 Objectives

The task includes all the activities to be carried out by the Technical Manager and includes:

- Monitoring of the general scientific and technological evolution;
- Monitoring of the progress of scientific and technological developments carried out in COUCH;
- Identification and troubleshooting of technical problems;
- Liaison with related projects, standardization bodies and networks of excellence. The task includes also review and approval of technical reports and deliverables.

5.3.3.2 Progress (M1-M18)

- As the general communication and collaboration in the project is going well, so is the technical and scientific progress. The project has achieved major technological and challenging feats, mainly related to the integration of complex technical systems developed by various University partners. The project has a very high scientific output in terms of reporting the work done in the form of articles and conference papers.
- Besides having clear work structures set out in the Document of Action, the project is focused towards the development of its main two technical solutions: The Functional- and Technical demonstrators. Early in the project, a clear vision for the functional demonstrator has been set out – a vision that is shared throughout the consortium and that is collectively refined at each of the plenary project meetings.

5.3.3.3 Achievements (M1-M18)

The following achievements have been reached during the first reporting period of the project (M1-M18):

- Efficient collaboration and communication between technical partners.
- Provision of a clear “shared vision” of the Council of Coaches.
- Effective technical integration work, delivering rapid results.
- Effective scientific output, with strong collaboration between partners.

5.3.3.4 Progress (M19-M36)

- Based on critical observations during the first project review, we have updated our deliverable template documents, and improved the quality control procedures, ensuring that all deliverable documents are reviewed in a timely manner, and receive a final quality check before being submitted.

5.3.3.5 Achievements (M19-M36)

- Improved deliverable templates
- Improved quality control procedures

5.3.3.6 Deviations from DoA

There have been no deviations from the DoA.

5.3.4 T1.4: Continuous Quality Control, Risk Identification and Mitigation Strategies

5.3.4.1 Objectives

The objective of this task is to continuously monitor the project’s technical progress, identify risks incurred through the project’s internal or external processes and define and update the risk mitigation plans. The task will establish and apply a number of quality and risk management processes, towards ensuring the quality of the project’s results (i.e. software, prototypes, services, documents), while at the same time mitigating risks and undertaking relevant contingency actions.

5.3.4.2 Progress (M1-M18)

The project management has implemented the quality assurance procedures as described in the *Deliverable 1.1: Quality Risk and IPR management* on how to handle the quality control aspects within the projects. A review procedure has been implemented and is being used for all deliverables. The Risk log is being updated by the management team and mitigation measures are being put in place if necessary. So far, there has not been a necessity to use them, fortunately.

5.3.4.3 Achievements (M1-M18)

- Review procedure has been put in place and is being used.
- Risk management procedures are implemented and a Risk Log is being kept and updated.

5.3.4.4 Progress (M19-M36)

The risk log was updated with the inventory of the Covid Risk evaluation we sent out to all consortium partners. As described, mitigation plans were made and all partners have gone out of their way to host technical integration weeks, workshops and consortium meetings online, instead of face to face.

5.3.4.5 Achievements (M19-M36)

Proper risk management during the Covid pandemic.

5.3.4.6 Deviations from DoA

There have been no deviations from the DoA.

5.3.5 T1.5: Innovation Management

5.3.5.1 Objectives

The main objective of the Innovation Management activity in the COUCH project is to monitor and control the process of creation of novel results with strong such as new ideas, algorithms, concepts, methods, products, services or applications that can be exploited through effective monitoring and

controlling processes. In order to control and monitor the progress of the innovation management, an innovation management table is being maintained with detailed information on each exploitable deliverable. This activity will be supported by the exploitation management team of WP8.

5.3.5.2 Progress (M1-M18)

In this reporting period, task T1.5 the project has delivered in M6 the Innovation Management Guidelines (D1.3), which describes the processes to monitor and support innovation outcomes that can lead into exploitation. In addition, innovation management telco's and two workshops in Dundee and Enschede have been organised, during project plenary meetings. In Dundee the workshop has focused on a pitching session, to shape and focus the exploitable outcomes of COUCH, where in Enschede the partners worked on the business modelling of the major outcome of the project, the COUCH system.

5.3.5.3 Achievements (M1-M18)

The main achievements of T1.5 have been:

- The deliverable D1.3, Innovation Management Guidelines.
- Two workshops organized.

5.3.5.4 Progress (M19-M36)

During this reporting period, task 1.5 has delivered two physical workshops and numerous calls around Innovation potentials of COUCH project. The first physical workshop took place in Brussels in July 2019, in which the partners were split into teams and brainstormed about potential business models for the two main exploitation potentials, namely the Open Agent Platform and the "COUCH as a product". Later, in February 2020, partners met in Paris and worked on the Product Design Sprint process, to optimize the offering of both exploitation potentials. Furthermore, the project has maintained and IPR and Innovation Management tables and strongly collaborated with WP8 for the "handover" of the potentials for exploitation plans.

5.3.5.5 Achievements (M19-M36)

The main achievements of T1.5 have been:

- Business Model workshop, Brussels, July 2019.
- Sprint workshop, Paris, February 2020.
- Handover of exploitation potentials to WP8.

5.3.5.6 Deviations from DoA

There have been no deviations from the DoA.

6 WP2: Responsible Research and Participatory Design

6.1 Planning

WP / Task	Title	Participants	Months
WP2	Responsible Research and Participatory Design	DBT	M1-M30
T2.1	Developing a Shared Vision of RRI for the Council of Coaches	DBT, CMC, RRD, SU, UDun, UPV, iSPRINT	M1-M24
T2.2	Stakeholder Engagement Process	DBT, CMC, RRD, UDun, iSPRINT	M1-M30
T2.3	User Needs and Continuous Evaluation	RRD, CMC, DBT, SU, UDun, UPV, iSPRINT	M1-M24
T2.4	Socio-Technical Integration	DBT, CMC, RRD, SU, UDun, UPV, iSPRINT	M7-M24

6.2 Main Achievements

The following main achievements result from this work package:

6.2.1 First Period (M1-M18)

The following main achievements result from this work package:

- The Council of Coaches RRI vision was established.
- Five RRI workshops organised.
- A framework of higher-level requirements, aimed to ensure responsible research and to prioritize the research areas.
- Established the STIRRING methodology as documented in the internal reports from the RRI workshops (D2.1).
- Presentation about RRI issues in the COUCH project at the COMMA conference during Warsaw Argumentation Week, *Argumentation & Society 2018*, Warsaw, September 2018
- Elaborate patient journeys per patient group describing how the Council of Coaches technology can support patients with a chronic disease in multiple stages in their daily life and the roles of virtual coaches during these stages.
- Insights in user needs and requirements derived from multiple and mixed research methods, for all three primary end-user groups, leading to an initial set of user requirements for the technology innovation process
- User evaluation of the first functional prototype that serves as input for the design cycles.
- Additional user feedback on alternative functional prototypes (Low-fidelity demonstrator and a 2D/Web demonstrator) for improving and supporting the innovation process.
- Categorization of positive and negative aspects of the characters, background stories and roles of the virtual coaches, based on users' thoughts and opinions which are derived from focus groups

6.2.2 Second Period (M19-M36)

- The *second stakeholder workshop* (under T2.2) was carried out in Brussels, February 26th, 2019
- Results and recommendations from the second stakeholder workshop were introduced to the consortium and debated in the *fifth RRI-workshop* held in Twente, March 28th, 2019

- The *third stakeholder workshop* (under T2.2) was divided in three sessions and carried out online on May 11th, 12th and 20th, 2020
- The STIR process was finalized
- D2.7, D2.8 were submitted
- User evaluation of the second functional prototype that serves as input for design cycles (D2.5).
- Categorization of usability issues encountered during the evaluation of the second functional prototype, classified in four categories (Navigation & Structure, Content & Information, Design & Presentation, and Other). Recommendations for improvement for each of these categories were written, and served as input for the updated requirements (D2.5).
- Using the same prototype as the user evaluation, an online experiment was conducted that investigated the tailoring of motivation strategies to users' type of motivation (D2.5). A paper about this experiment has been submitted to the Computers in Human Behavior journal and is currently under review.
- User evaluation of the third functional prototype that serves as input for design cycles (D2.6).
- User feedback on newly added system elements (activity book, coach selection screen, Fitbit connection, account creation process) and redesigned user interface. Feedback was gathered through usability issues extracted from interviews as well as using evaluations scales (SUS and TAM) (D2.6).
- Insight in target population user performance of common tasks in health applications (e.g. account creation, Fitbit connection) using five different task metrics (D2.6).
- Both the second and third usability evaluation did not only provide insight into negative usability issues, but also provided insight into the positive usability 'issues'; thus clarifying which elements of the prototype were appreciated by end users in our target population (D2.5, D.6).
- Both evaluations were conducted with end users from the Netherlands, Scotland and Denmark. These end users were all part of the project's target population. (D2.5, D2.6)

6.3 Individual Tasks

6.3.1 T2.1: Developing a Shared Vision of RRI for the Council of Coaches

6.3.1.1 Objectives

In this task, the objective is to develop a shared understanding and vision of what RRI means for the COUCH consortium as well as how this vision should be achieved in order for RRI to become fully embedded in the COUCH R&I processes. Imposing a pre-defined RRI-understanding top-down will neither induce sufficient ownership amongst the partners nor is such an approach suitable to accommodate project-specific requirements and conditions. It is a central consideration to ensure that the research and innovation process in the project follows the principles of Responsible Research and Innovation, implementing the framework laid out in the European Responsible Industry project (EU-FP7-609817). It is also a central consideration to ensure that the tools and coaching methods, as well as the actual R&I processes in COUCH, contribute effectively to the needs of future users and society at large and are aligned with societal values and ethical considerations.

6.3.1.2 Progress (M1-M18)

The consortium laid the groundwork for the COUCH RRI Vision in a **two-day internal workshop** in Copenhagen in January 2018 by applying the RRI "Co-construction Method" (www.responsibility-navigator.eu), a deliberative methodology designed to create upstream reflection on research and innovation, and to facilitate related debate, negotiation and learning on how to implement RRI in the specific R&I processes. Rather than trying to deduce in a top-down manner the responsibilities that fall on the shoulders on the Council of Coaches consortium members from existing legal and ethical frameworks, the workshop attempted to stimulate bottom-up reflection within the consortium about which responsibilities arise naturally out of the ambitions of the project. At the workshop, the consortium members decided on an initial list of priorities and a rough plan for how to deal with them within the project. The inputs for this process of reflection included an overview of the RRI debate in general, suggestions for professional and ethical frameworks that might be relevant to the project, and

inputs from stakeholders about the social and ethical concerns that might prove salient for the project. This workshop also provided the framework for *Deliverable D2.1: The COUCH RRI Vision* which provides the internal guidelines and the strategy for how the principles of RRI will be implemented throughout the project.

On March 5th and March 6th, 2018, the consortium held their first technical integration week in order to integrate the different platform systems. During this week, the Danish Board of Technology (DBT) facilitated **the first RRI-workshop** to support reflection on the four main RRI issues during this early-stage system development sprint. The workshop consisted of facilitated group work with the aim to integrate the RRI-issues further in the mindset and discussions of the technical partners. A small presentation was made by DBT to sum up on the different RRI-issues and stakeholder concerns for the people who were not present when the issues were identified in Task 2.1. The group was then divided into smaller groups, all with at least one member who was there when the RRI-issues were originally identified and engaged in debate around focused questions. Lastly, each group presented their issue(s) and view upon how to deal with the issue(s), to the rest of the technical experts and DBT. The presentations were discussed in plenum and the discussion was recorded.

The second RRI-workshop held at the University of Twente in Enschede on April 25th, 2018, in connection with COUCH's second technical integration week. In the RRI team we were working on a sketch for the overall RRI procedure for the whole project period. Also, we were developing and advancing the individual interventions to take place along the way and planning 'what to do next based on what happened last' in a systematic manner. Since the process is empirically attentive and responsive, the exact content of each intervention needed to be planned in accordance with other developments and needs in the project, but the strategic map of where we were at and where we wanted to end up by month 36 was indeed possible to externalize. Methodically, we consider the course of the interventions as a *mechanics* – as punctual occasions for reflection where tacit knowledge can be put into words and said out loud in plenum. At this second workshop the RRI-process followed the developments-history and caught on to other COUCH tasks. Through the work carried out in Task 2.1, four major RRI-issues and a list of 'sleepers' were identified by the consortium and listed in D2.1. In Valencia the RRI-partners facilitated a structured reflection upon the RRI issues, which fed in to the first internal RRI report from Task 2.4. At this workshop we asked participants to reflect on a different set of issues that are in fact equally relevant to the RRI profile of the COUCH-project: The sleeper-issues. The name 'sleeper-issue' came into being as the structure of the initial RRI-workshop in Copenhagen in January (where the RRI-issues and sleeper issues were identified) did not facilitate enough time to dive deeply into the whole spectrum of issues that came up. Furthermore, at the consortium meeting in Valencia, many 'sleepers' were mentioned and debated as important at the present moment in time. Therefore, in Enschede the sleeper issues were brought back into the picture. In addition to reflecting on the issues, we also asked participants to also consider: What are the remaining sleepers at this point?

The third RRI-workshop was held in Dundee on June 28th, 2018 as part of the 3rd Council of Coaches plenary consortium meeting. The reason for this internal brief was to gather the information provided in a ready-to-hand form for use in deliverable 2.8 on the Council of Coaches Participatory Design Process. The workshop was designed to provide to serve two purposes at once. The primary purpose was to gauge the impact of the RRI activities so far on the decision-making processes within the Council of Coaches consortium. Such input would be necessary to ensure that the ongoing method development in task 2.4 was taking place from a well-informed starting point. A secondary purpose was to stimulate reflection within the Council of Coaches consortium on said processes, including reflection about the relative impacts of internal vs external factors on the decisions being made in the course of prototype development. In methodological terms, the bespoke method developed for this workshop was designed to elicit responses from consortium members about the factors influencing decision-making in their day-to-day work towards the objectives of Council of Coaches. In terms of scope, we decided to focus on the prototyping process, which at the point in time had just completed its first round. The first functional prototype had been completed, and planning for the next was high on the agenda of the consortium meeting. Focusing on the prototyping process would thus provide us with a discrete slice of project activity to look at. The exercise took place and was designed against the background of the method development work in task 2.4. In that context, the ancillary benefit of having this workshop

would be to get a reality check on a perceived 'prioritization vacuum' in the RRI method being developed by our team.

October 8-12th, 2018, marked the week of the third technical Integration Week, which was hosted at the University of Twente. At this event the RRI-team rolled out **the fourth RRI-workshop**, where the focus was a beginning process of letting stakeholder perspectives inform the hardest issues at the given moment in time, as experienced from inside the project. As we were just beginning to prepare for our **second stakeholder workshop** (under T2.2 which will be held in Brussels, February 26th, 2019), we asked the participants to scrutinize their impressions of the state of RRI issues right now, hard nuts to crack, identification of what is needed and who may help. They were asked to imagine which persons, professions or organisations would be most valuable for them to draw into their decision-making and solution-work at the time, and why. Also, they were invited to reflect on which exact questions they would like to discuss with these people/professions/organisations. It is our objective to make all this happen at our second stakeholder workshop in February 2019.

6.3.1.3 Achievements (M1-M18)

The following achievements result from this task:

- Background brief for building the Council of Coaches RRI vision.
- Deliverable D2.1 The COUCH RRI Vision.
- Internal report from the first RRI-workshop in Valencia, March 2018.
- Internal report from the second RRI-workshop in Enschede, April 2018.
- Internal report from the third RRI-workshop in Dundee, June 2018.
- Internal report from the fourth RRI-workshop in Enschede, October 2018.
- Presentation about RRI issues in the COUCH project at the COMMA conference during Warsaw.
- Argumentation Week, *Argumentation & Society 2018*, Warsaw, September 2018.

6.3.1.4 Progress (M19-M36)

Most of the work in this task was performed in the first period. In the second period the RRI Vision work was finalized and D2.8 was submitted.

6.3.1.5 Achievements (M19-M36)

- Internal report from the fifth RRI-workshop in Enschede, March 2019.
- D2.8 submitted

6.3.1.6 Deviations from DoA

D2.8 was submitted later than planned in the DoA, as approved by the PO.

6.3.2 T2.2: Stakeholder Engagement Process

6.3.2.1 Objectives

The objective of this task is to better align COUCH R&I with societal demands and values through a series of deliberative, co-constructive multi-stakeholder workshops. A series of altogether three interactive and participatory stakeholder workshops will be conducted. As part of the integration of outside perspectives, interviews with key stakeholders in three different countries are held prior to the workshops, in order to also allow for different cultural and social settings. The workshop concept allows for multi-stakeholder engagement and interchange of viewpoints from the consortium, industry, care providers, health professionals, policymaking and government, civil society and patient organizations. Issues to be discussed and negotiated include identification and assessment of risks and benefits, of conflicting values and interests, potential barriers to use, hazard identification, ethical, legal and social implications, professional needs, etc. In addition to the briefing notes from the workshops which have the purpose to inform the consortium, each workshop will be followed-up with an internal, lessons-oriented de-briefing of the consortium members in order to facilitate meaningful and effective integration of the workshop results in COUCH.

6.3.2.2 Progress (M1-M18)

The first RRI workshop was held in Copenhagen in January 2018 (M5). It focused on the needs, problems, requirements, and possible social and ethical implications, which will provide value input and requirements for the design of functionalities.

The second RRI workshop will be held in Brussels on February 26th, 2019 (M18). The purpose of this workshop is to obtain stakeholder views on the responsibility of the project's tech and design choices. The focus of the workshop is action-oriented and will revolve around how the RRI issues are being tackled in the project. The participants will be informed about the project's aims and its RRI vision and how the project has sought to implement the RRI vision in the prototypes and the architecture. Responses will be gathered about: a) the substantial tech and design choices made and whether they hold up to stakeholders' expectations of a societal responsibility profile and b) the underlying approach for achieving responsibility and how it could be improved. The workshop will also provide feedback on the usability of the prototypes, and further discuss professional needs, ethical, legal, and regulatory aspects that need to be taken into consideration in the innovation, or in the use of the service.

6.3.2.3 Achievements (M1-M18)

At this point, the following achievements result from this task:

- Background brief for building the Council of Coaches RRI vision.
- Deliverable **D2.2 Report on user and stakeholder needs and expectations**.
- Interviews with key stakeholders in 3 countries for the first workshop (Denmark, Holland, UK).
- Interviews with key stakeholders in 3 countries for the first workshop (Germany, Norway, Ireland).

6.3.2.4 Progress (M19-M36)

- The *second stakeholder workshop* (under T2.2) was carried out in Brussels, February 26th, 2019
- Results and recommendations from the second stakeholder workshop was introduced to the consortium and debated at the *fifth RRI-workshop* held in Twente, March 28th, 2019
- The *third stakeholder workshop* (under T2.2) was divided in three sessions and carried out online on May 11th, 12th and 20th, 2020

6.3.2.5 Achievements (M19-M36)

- Report from *second stakeholder workshop on February 26th, 2019*
- Report from *third stakeholder workshop on May 11th, 12th 20th, 2020*.

6.3.2.6 Deviations from DoA

The timing of the workshops was to be aligned so that the inputs could be used for the delivery of the three functional prototypes in M9, M15 and M21. However, in order to gather sufficient material for the participants in the second workshop, the timing of this workshop was postponed to M18, and the inputs fed into prototype three instead.

The third RRI workshop, due to the impact of COVID19, was converted to an online format and carried out in May 2020.

The theme of the third stakeholder workshop 'responsible exploitation' had the purpose of evaluating and further developing guidelines towards a practical approach to responsible prototyping. The workshop was planned to be launched in March 2020 in the shape of a physical event in Brussels. However, due to the incident of COVID-19, the event was first postponed and after some time redesigned into an online format (using the platform Zoom). This makeover gave us the opportunity to expand our geographical coverage regarding participants as well as the scope of themes. The workshop themes and related stakeholder-groups were designed to complement a desire for high diversity in stakeholder perspectives.

6.3.3 T2.3: User Needs and Continuous Evaluation

6.3.3.1 Objectives

The main objective of this task is to gain insights in the needs of targeted (primary) end-user groups that serve as input for the technology innovation processes in WP3 to WP6. This task consists of two goals. The first is the elicitation of user needs and requirements from the three main patient groups (older adults with age-related impairments, patients with chronic pain and diabetes type 2). The second is a continuous evaluation process of the functional prototypes of the Council of Coaches system. There are three evaluation rounds, where the focus of the evaluation gradually shifts from technology acceptance (with low-fidelity prototypes) to usability and user experience (with high-fidelity prototypes), to the human-computer interaction in the final functional prototype. These evaluation rounds also serve to update user requirements and to collect redesign input.

6.3.3.2 Progress (M1-M18)

This task runs from the beginning of the project (M1) until August 2019 (M24). Four deliverables result from this task, namely:

- D2.3: Initial user requirements (M12)
- D2.4: Evaluation results of the first functional prototype and updated requirements (M12)
- D2.5: Evaluation results of the second functional prototype and updated requirements (M18)
- D2.6: Evaluation results of the third functional prototype and updated requirements (M24)

Deliverables D2.3 and D2.4 were finished on schedule in August 2018 (M12). For the elicitation of user requirements, a longitudinal diary study was set up in which chronic patients (age-related impairments, chronic pain, diabetes type 2) were followed for four weeks. In Scotland, it was decided to utilize the National Diabetes Foundation for the recruitment of participants, as this would result in a high-quality sample. As a result, medical ethical approval needed to be obtained twice (internally and externally); a process that was confronted with several delays. In the end, these delays made the consortium decide to abandon this option for recruitment and to use a snowball/convenience sample instead, so as not to delay recruitment further. Additionally, we were confronted with an unexpectedly high degree of participant drop-out. As a result, the study did not reach the minimally required sample size. Additional recruitment and data collection was done to reach the necessary number of participants. At the time of writing D2.3, data collection for the patients with chronic pain and diabetes type 2 was therefore not yet completed. The deliverable therefore provide only a brief summary of the main results of the diary study. During M13 to M16 we completed the full analysis of the diary study. Updated results and requirements are provided in D2.5.

Besides the diary study, D2.3 provides results of studies that the consortium performed in addition to the DoA to further strengthen our understanding of health coaching and to elicit user requirements from different perspectives and research methods. Sorbonne University conducted a quantitative study on the effects of virtual agent's gender, role and focus on user's perspective.

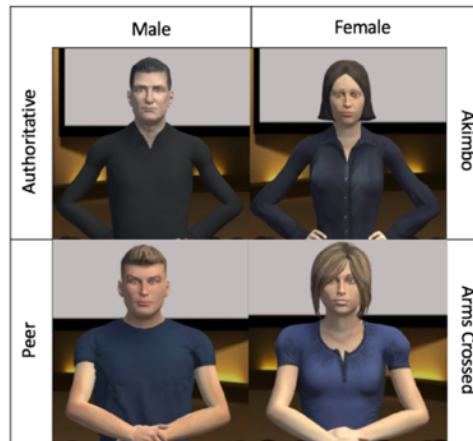


Figure 6: Four appearances modelled to fit authoritative and peer agent role.

Also, at the University of Dundee a role-playing video study was set up in which actors simulated an interaction between a patient and multiple medical practitioners. Last, D2.3 reports on the elicitation of the high-level requirements process.

Because of the aforementioned challenges in the diary study, D2.5 will also suffer a slight delay. This delay will not affect the development process of the technology, as the results were reported to the consortium in a timely manner.

6.3.3.3 Achievements (M1-M18)

The following achievements result from this task:

- An initial set of user requirements for the technology innovation process (D2.3).
- Insights in user needs and requirements derived from multiple and mixed research methods, for all three primary end-user groups: older adults with age-related impairments, chronic pain patients and diabetes type 2 patients (D2.3).
- A framework of higher-level requirements, aimed to ensure responsible research and to prioritize the research areas. This framework was developed and further optimized during deliberation processes among the project partners (D2.3).
- Elaborate patient journeys per patient group that describe how the Council of Coaches technology can support patients with a chronic disease in multiple stages in their daily life and the roles of virtual coaches during these stages (D2.3).
- User evaluation of the first functional prototype that serves as input for the design cycles (D2.4).
- Categorization of positive and negative aspects of the characters, background stories and roles of the virtual coaches, based on users' thoughts and opinions which are derived from focus groups (D2.4).
- Additional user feedback on alternative functional prototypes (Low-fidelity demonstrator and a 2D/Web demonstrator) for improving and supporting the innovation process (D2.4).

6.3.3.4 Progress (M19-M36)

The second half of T2.3 ran until August 2019. The remaining two deliverables for this task were completed and submitted, namely:

- D2.5: Evaluation results of the second functional prototype and updated requirements (M18)
- D2.6: Evaluation results of the third functional prototype and updated requirements (M24)

As stated in the M1-M18 progress, deliverable D2.5 was completed with a slight delay. It was submitted just before the midterm review in April 2019 (M20) and reported on an end-user evaluation with users from three countries (the Netherlands, Denmark and Scotland) who were part of the project's target population. Following the points of attention and feedback from the midterm review, deliverables D2.3, D2.4, D2.5, and D2.6 were restructured to clarify the types of participants and studies that were

performed. A 'corrections' section has been added to each of these deliverables, detailing the adjustments. The main changes for each deliverable involved the addition of a structured abstract for every study reported and the addition of overview tables for the reported studies and their participants.

While the first version of the functional prototype for the third evaluation cycle was released in May 2019 (M21), it was updated with functionalities such as an activity book, the coach selection screen, Fitbit connection, and the account creation procedure before being evaluated with end users. Deliverable D2.6 was completed in November 2019 (M27) and reported on the outcomes of the third user evaluation with a focus on the newly added elements, the new design, new coaching content, and task performance of commonly performed actions, which participants were asked to perform. The deliverable was submitted later than scheduled, but this delay did allow us to test the prototype's fully redesigned user interface (following user's feedback from the second evaluation round) with the target population before the final evaluation would start.

Overall, the continuous evaluations of the functional prototypes have contributed a lot to the design process of the Council of Coaches application. With additional evaluation methods included as the technology readiness level of the prototype increased (e.g. the inclusion of the TAM and task metrics in the third evaluation; D2.6), user involvement provided valuable insights and feedback for the developers. The original requirements from D2.3 and updated requirements from all evaluation cycles (D2.4, D2.5, D2.6) have been continuously used to adjust, update and extend the project's prototypes and a summarizing review of all requirements resulting from end user involvement is reported in the deliverable on final coaching actions and content (D3.4).

6.3.3.5 Achievements (M19-M36)

The following achievements resulted from this task in the second half of the project:

- User evaluation of the second functional prototype that serves as input for design cycles (D2.5).
- Categorization of usability issues encountered during the evaluation of the second functional prototype, classified in four categories (Navigation & Structure, Content & Information, Design & Presentation, and Other). Recommendations for improvement for each of these categories were written, and served as input for the updated requirements (D2.5).
- Using the same prototype as the user evaluation, an online experiment was conducted that investigated the tailoring of motivation strategies to users' type of motivation (D2.5). A paper about this experiment has been submitted to the Computers in Human Behavior journal and is currently under review.
- User evaluation of the third functional prototype that serves as input for design cycles (D2.6).
- User feedback on newly added system elements (activity book, coach selection screen, Fitbit connection, account creation process) and redesigned user interface. Feedback was gathered through usability issues extracted from interviews as well as using evaluations scales (SUS and TAM) (D2.6).
- Insight in target population user performance of common tasks in health applications (e.g. account creation, Fitbit connection) using five different task metrics (D2.6).
- Both the second and third usability evaluation did not only provide insight into negative usability issues, but also provided insight into the positive usability 'issues'; thus clarifying which elements of the prototype were appreciated by end users in our target population (D2.5, D.6).
- Both evaluations were conducted with end users from the Netherlands, Scotland and Denmark. These end users were all part of the project's target population. (D2.5, D2.6)

6.3.3.6 T2.3 – Deviations from DoA

D2.3 describes multiple studies besides the diary study for the elicitation of user requirements. Also, in deliverable D2.4 additional demonstrators were developed besides the official functional prototype. These extra efforts help to accelerate the technical innovation process and to improve the quality of its outcome. Unforeseen issues in medical ethical permission and recruitment caused some delay in finishing D2.3 according. These issues have been solved.

In the second period, we had some additional delays for deliverables D2.5 and D2.6. Overall the planning from demonstrator releases to evaluation deliverable release has been to tight in the Document of

Action. In order to ensure the quality and maximize the usefulness of these evaluations we have opted to delay the documents.

6.3.4 T2.4: Socio-Technical Integration

6.3.4.1 Objectives

The purpose of this task is to guide and document the development of the COUCH RRI Vision and its implementation in the project. Drawing on the concept and methodology of socio-technical integration research (STIR), this task supports the integration of ethical and societal perspectives in the research processes conducted in the project. Social scientists from the consortium interacts directly with the scientists and engineers in the labs with the purpose of stirring reflection as well as complementing and sustaining the RRI-oriented activities in T2.2 and T2.3. The interactions are guided by semi-structured interaction protocols specifically designed to enhance reflection upon research decisions in the light of broader considerations, to support anticipation and responsiveness, and to bring in additional perspectives and stakeholder views. The aim is also to challenge the COUCH scientists and engineers to actively reflect upon potential outcomes of their work while it is being conducted in the labs, and to adjust research practices and research directions accordingly. The task includes the development of a socio-technical integration approach tailored specifically to the needs and the conditions of COUCH. This entails the development of the semi-structured interaction protocol, preparation and training of the social scientists, continuous formative evaluation of the process, and systematic documentation. For practical and logistic reasons, the STIR method adapted in this task follows a logic of reflexive monitoring in the form of mini interventions. The task thus follows the same rhythm as the overall coordination of the project and is taken up in consortia meetings as well as in ongoing online coordination meetings when relevant and possible. By 'reflexive monitoring' is implied that the RRI team keeps monitoring progress and supporting ongoing reflexion about the meaning of the RRI vision as laid down in deliverable D2.1. By 'mini-interventions' is implied that the RRI team facilitates individual talks, mini-workshops, and distribute further analysis at strategic points in the project's development. This intensifies typically just before and just after the production of prototypes.

6.3.4.2 Progress (M1-M18)

Reflexive monitoring:

- Introduction of the RRI Vision to the consortium and taking on-board feedback.
- A series of online negotiations about how the RRI issues should be prioritised at the given moment in time.

Mini interventions:

- STIR interventions tied to the RRI workshops as reported under D2.1

6.3.4.3 Achievements (M1-M18)

The STIR activities are documented in the internal reports from the RRI workshops as reported under D2.1.

6.3.4.4 Progress (M19-M36)

- The STIR process was finalized and D2.7 submitted.

6.3.4.5 Achievements (M19-M36)

- D2.7 was submitted

6.3.4.6 Deviations from DoA

D2.7 was submitted later than planned in the DoA, as approved by the PO.

7 WP3: Coaching Strategies and Knowledge Base

7.1 Planning

WP / Task	Title	Participants	Months
WP3	Coaching Strategies and Knowledge Base	RRD	M1-M27
T3.1	Definition of Tailored Coaching Strategies	RRD, CMC, DBT, UDun	M1-M12
T3.2	Definition of Coaching Actions and Content	UDun, CMC, RRD	M4-M27
T3.3	Development of Shared Knowledge Base	RRD, CMC, UDun, UPV	M6-M27

7.2 Main Achievements

The following main achievements result from this work package:

7.2.1 First Period (M1-M18)

The following main achievements result from this work package:

- Provide an initial set of definitions for key terminology used in WP3 and the project in general.
- Overview of relevant theories of behaviour changes and coaching methods, among which contextualization and framing.
- Identified key knowledge components (parameters) required for successful and tailored coaching.
- Set the required theoretical framework for the automated coaching engine: *Coaching Goals – Coaching Strategies – Coaching Actions – Dialogue Actions*.
- A set of example scenarios and coaching dialogues.
- An initial set of concrete coaching actions that aim to bridge the gap between current and desired states in the area of physical, cognitive, mental or social status.
- Provide an initial set of requirements, an overview of potential technologies and first design and technical framework of the Knowledge Base component.
- A shared knowledge base has been built and is being updated and extended with each prototype in which static and dynamic knowledge gathered through interactions with the user is used to tailor the suggested strategies and content of dialogues to the user.

7.2.2 Second Period (M19-M36)

The following main achievements result from this work package:

- Coaching content has been developed for seven coaches and the assistant agent. This process has taken into account the feedback from the user evaluations and updated requirements;
- The WOOL Dialogue Platform was developed to facilitate content authoring by domain experts;
- The use of WOOL dialogues has been integrated in both the functional and technical demonstrator (allowing for both scripted dialogues as well as dialogue games from the Dialogue and Argumentation Framework to be used);
- Research on tailoring coaching strategies to users' motivation to live healthy has been performed in coordination with WP2;
- Research on the effect of tailored conversational topic suggestions on users' engagement with the system (the micro-randomized trial) has been performed in coordination with WP7;
- A shared knowledge base has been further developed and integrated with both the functional and technical demonstrator;
- A proof of concept topic selection engine has been developed and integrated with both the functional and technical demonstrator.

7.3 Individual Tasks

7.3.1 T3.1: Definition of Tailored Coaching Strategies

7.3.1.1 Objectives

The objective of this task is to define the theoretical framework that will be used in the project that individual coaches will employ to plan their coaching actions. The coaching strategies will be based on established theories of behaviour change. This task will first deliver a functional requirement specification of the Shared Knowledge Base, and finally a literature-based model of behaviour change covering the domains of physical-, cognitive-, mental-, and social wellbeing as well as condition specific coaching strategies for Diabetes Type 2 and Chronic Pain.

7.3.1.2 Progress (M1-M18)

This task has run from the start of the project (M1) throughout the first year, having finished in August of 2018 (M12). This task has delivered *D3.1: Initial knowledge base design and coaching strategies* in M6 and *D3.3: Definition of tailored coaching strategies* in M12 without delays.

7.3.1.3 Achievements (M1-M18)

The following major actions have been achieved by the task:

- Provide an initial set of definitions for key terminology used in WP3 and the project in general (D3.1);
- Provide an in-depth overview of 13 different relevant theories of behaviour change (D3.1);
- Based on theories of behaviour change – provide an overview of 8 different relevant coaching methods (D3.1);
- Provide a literature-based overview of contextualization and framing methods used in coaching (D3.1);
- Identified key knowledge components (parameters) required for successful and tailored coaching (D3.1);
- Provide an initial set of requirements, an overview of potential technologies and first design of the Knowledge Base component (D3.1);
- Set the required theoretical framework for the automated coaching engine: *Coaching Goals – Coaching Strategies – Coaching Actions – Dialogue Actions* (D3.3);
- Provide an initial technical framework for determining tailored coaching strategies and delivering coaching actions (D3.3);

7.3.1.4 Progress (M19-M36)

Even though the deliverables reporting on T3.1 had been submitted on time in August of 2018 (M12), the definition of tailored coaching strategies was still investigated during the second half of the project as this process was largely interwoven with the definition of coaching actions and content. As tailoring involves the adjustment of something to someone, we investigated further how coaching strategies could be tailored to the project's target population. This was done in coordination with the end user studies and evaluations that took place in WP2 and WP7.

In the deliverable detailing the second prototype evaluation (WP2, D2.5), an online experiment has been described. In this experiment, we investigated if higher level motivational strategies could be tailored to users' motivation to live healthy. The principal result was that tailoring of higher level strategies was indeed possible. Based on these results and contemplating their connection to the content we were creating, we concluded that each of these strategies involved specific topics that were relevant to discuss between a coach and a user. For example, when following the *health education* strategy, topics that should be discussed focussed on providing information on the what, why and how of healthy behaviour. On the other hand, the *implementation intentions* strategy would require the emphasis of discussed topics to shift more towards goal-setting and planning.

Connecting the insight that topics might need to be tailored to the conceptual framework described in D3.3, in practice this meant that the execution of a coaching action was implemented as the discussion of a certain topic; e.g. the coaching action 'inform on what healthy behaviour is' (for a domain) would

appear to the user as a conversation with a coach on 'what is healthy behaviour for [e.g. physical activity]'. A card-sort study with cards representing coaching actions and groups representing conversational topics has been conducted with the aim of identifying conversational topics and a topic structure for health coaching. A paper describing this study is currently being written and will be published in a Journal as part of PhD research in the project.

Based on the notion that tailoring conversational topics could benefit users, a topic selection engine was implemented as a proof of concept for tailoring conversational topics, and integrated in both the functional and technical prototypes. To evaluate whether tailoring the selection of topics would improve users' interaction with the Council of Coaches system, a micro-randomized trial study was performed as part of the final evaluation. This study compared users' interactions that started with the coach suggesting a tailored topic ('Shall we discuss [topic]?') to interactions that started with a more conventional statement ('How can I help you?'). The deliverable reporting on the final evaluation (D7.7) includes the details of this study and its preliminary results.

7.3.1.5 Achievements (M19-M36)

In the second half of the project, the following major actions have been achieved by this task:

- A study was conducted on tailoring motivational strategies to users' type of motivation to live healthy (see deliverable D2.5). A paper about this study has been submitted to the Computers in Human Behavior journal and is currently under review;
- A study on determining relevant topics and a topic structure for coaching conversations (that can be used in tailoring coaching strategies) has been conducted. A paper about this study is currently being written;
- A proof of concept topic selection engine was implemented and integrated in the functional and technical demonstrators;
- A study was conducted on the effect on interaction for the proof of concept topic selection engine (see micro-randomized trial in deliverable D7.7). A paper about this study is currently being written.

7.3.1.6 Deviations from DoA

In the first period this task has set out and worked towards achieving the objectives as described in the Document of Action and has achieved these in large parts. As concluded in D3.3, the planned objective of having a "final" definition of coaching strategies has proven too ambitious for completion in the time period set for the first task (the first 12 months of the project). The process of defining strategies has been largely interwoven with the actual definition of coaching actions and content as they were being defined and implemented in T3.2 and T3.3 – as such, activities towards "finalizing" these coaching strategies were an effort that continued after the task had officially ended, but has finally managed to fulfil the task objectives.

7.3.2 T3.2: Definition of Coaching Actions and Content

7.3.2.1 Objectives

The objective of this task is to specify an initial set of coaching actions and content which will be used in the development of the first functional prototype. This will be based on the relevant parameters defined in T3.1, and the collection of real interactions between coaches and users.

7.3.2.2 Progress (M1-M18)

This task has run from the start of the project (M1) until the end of May 2018 (M9). This task has delivered D3.2: Initial coaching actions and content without delays.

7.3.2.3 Achievements (M1-M18)

The following major actions have been achieved by the task:

- A set of example scenarios that encapsulate ways in which patients might interact with the system (D3.2);

- A set of example coaching dialogues between patients and their health councillors (to show how insights from behaviour change literature can be incorporated in coaching dialogues and how behaviour change techniques can concretely be incorporated) (D3.2);
- An initial set of concrete coaching actions that aim to bridge the gap between current and desired states in the area of physical, cognitive, mental or social status for all target populations and specific health outcomes for the diabetes and chronic pain groups (D3.2);
- A set of guidelines for editing the living document D3.2 in order to add further coaching actions (D3.2).

7.3.2.4 Progress (M19-M36)

While started in December 2017 (M4), this task ran until November 2020 (M27). The remaining and final deliverable (D3.4: Final coaching actions and content) was submitted without delays. In addition to presenting the final coaching actions and content this deliverable also elaborates on the following:

- Design decisions for the domains that are addressed by the Council of Coaches.
- Presentation of the developed coaches including their role, expertise, and personalities.
- Motivation and description of the coaching actions and content available for each coach.
- Design and description of the WOOL Dialogue Platform.
- A review of the projects requirements that were elicited from users throughout the project.

Based on the coaching literature described in deliverable D3.1, the initial coaching actions and content from D3.2, the conceptual framework described in D3.3 that defined the relation between *Coaching Goals*, *Coaching Strategies*, *Coaching Actions* and *Dialogue Actions* and results from user studies on coaching strategies (also see the previous description in T3.1), a methodology was defined during the project for structuring coaching content that is also reflected in the projects technical and functional demonstrators. That is, coaching actions were related to dialogues about a certain topic of conversation and a hierarchy was defined for each coach that allowed us to structure dialogues.

For the functional demonstrator the defined coaching actions, content and structures meant that every WOOL script would contain one dialogue on a certain topic and that WOOL scripts could then be linked to each other following the defined structure. For the technical demonstrator this meant that there were two options for including content from the functional demonstrator: either by directly using the created WOOL script on a certain topic or by starting a dialogue game on that topic using the Dialogue and Argumentation Framework (DAF; WP5). Both have been realized. One major adjustment was made to the Council of Coaches' content after the deliverable had been submitted, namely: a five-week sleep coaching programme was added to the Chronic Pain coach.

The development of the WOOL Dialogue Platform allowed us to involve experts in the authoring process for coaching content. Not only did this ensure that the content provided to our participants was suitable according to the domain experts, but the 'preview dialogue' functionality also made the conceptual gap between the developed application and expert's intended dialogue experience smaller. This meant that experts could also test the flow of dialogues themselves, which helped the iterative process of dialogue authoring. Furthermore, WOOL provided tools to tackle translations, which allowed us to translate dialogues to Danish, Dutch and Scottish.

7.3.2.5 Achievements (M19-M36)

The following major actions have been achieved by the task:

- A motivation for the domains that are addressed by the Council of Coaches has been provided (D3.4);
- Coaches have been designed based on the user evaluations and studies performed in the project; in this task their content and roles have been defined (D3.4);
- Coaching content has been developed for seven coaches and the council's assistant (D3.4);
- Each of the seven coach's content is described by providing an overview of the coaching actions and content that are available (D3.4);
- For each of the seven coaches' a definition of their domain, background from the literature (featuring the relevant guidelines), and additional resources that can be referenced are also

included with the aim of making the content creation process insightful for those looking to create content for their own coaching applications featuring conversational agents (D3.4);

- The open source WOOL Dialogue Platform is described and a paper describing it is being written (D3.4);
- A review of the project's requirements from the user involvement process and how these were addressed (D3.4).

7.3.2.6 Deviations from DoA

The primary objective of this task was to specify a set of coaching actions and content, to be used in the Council of Coaches application. This content has been developed in an iterative process in which the requirements from WP2 were included. Deliverable D3.4 therefore not only describes the final coaching actions and content, but also provides insight into domain choices, coach design, authoring tools and reviews the requirements that resulted from the user involvement process. T3.3: Development of Shared Knowledge Base

7.3.3 T3.3: Development of Shared Knowledge Base

7.3.3.1 Objectives

The objective of this task is to construct a model of a shared knowledge base based on the relevant identified parameters from T3.1 and as well as the input from the user requirements gathering process in T2.3. The knowledge base will contain all the information regarding the users (User Model), their previous interactions with the Council of Coaches or individual coaches (Interaction Model), relevant identified contextual information (Context Model) and health specific parameters (Domain Models) for Diabetes Type 2 and Chronic Pain. Both user- and context models are partly static (e.g. gender, language) and partly dynamic (e.g. self-efficacy, stage of change), dependent on interaction topic, issues under discussion and context. For the technical implementation of the shared knowledge base, the project will explore the possibility of extending the FIWARE Orion Context Broker.

7.3.3.2 Progress (M1-M18)

This task started in M6, and updates to the shared knowledge base have been made in M9 and M15, coinciding with two functional prototype releases. Further updates will follow at M21 and M27, with the remaining functional prototype release and the release of the technical prototype. In addition to the updates, the deliverable corresponding to this task (*D3.5: Shared Knowledge Base component*) is scheduled to be completed with the last update and will be a demonstrator.

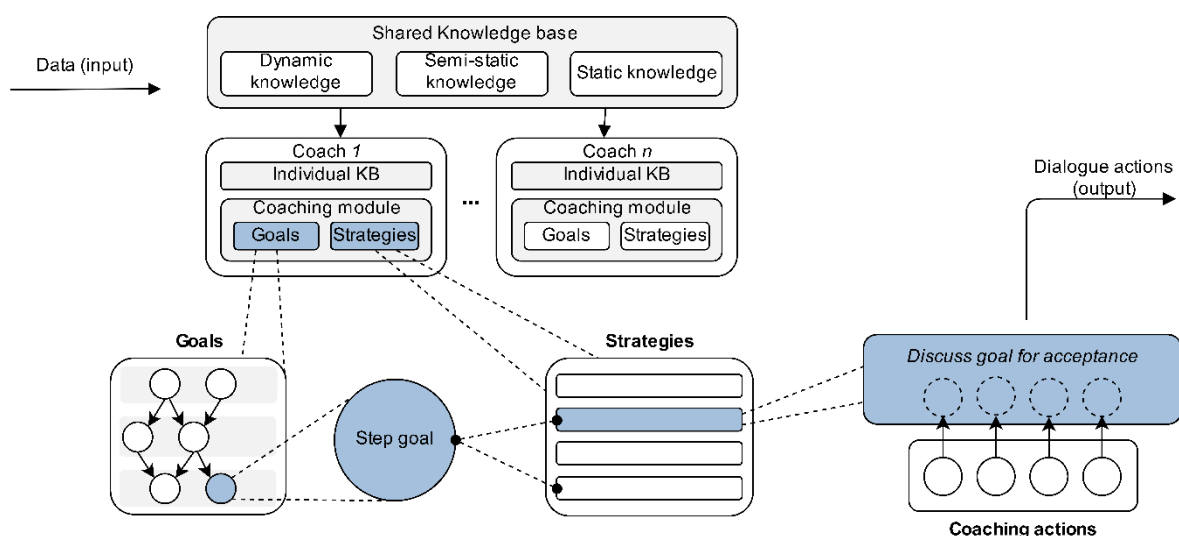


Figure 7: Initial design of the knowledge base.

7.3.3.3 Achievements (M1-M18)

The following major actions have been achieved by the task:

- A shared knowledge base has been built and is being updated and extended with each prototype.
- Knowledge gathered through interactions with the user is stored in the shared knowledge base and is used to deduce dynamic constructs that were found relevant in T3.1 (e.g. user's motivation type).
- Static and dynamic knowledge is used to tailor the suggested strategies and content of dialogues to the user.
- The possibility of extending the FIWARE Orion Context Broker was explored. It was decided to develop FIWARE-compatible data models. Further integration may be studied.

7.3.3.4 Progress (M19-M36)

The shared knowledge base has been further developed and delivered as deliverable D3.5. The component was the main part of the functional demonstrator and was integrated with the technical demonstrator. It stores user data, stores the scripted WOOL dialogues, provides functionality for execution of these dialogues and uses a variable store to drive the dialogue execution. It also provides user management in both the functional and technical demonstrators.

A detailed record of all variables collected and used in the Council of Coaches has been kept. Variables use a structured naming strategy and their source is described, thus allowing authors of new dialogues to include knowledge about the user in new dialogues. In order to facilitate 3rd party developers to use the Knowledge Base software as part of the Agents United Platform – a stand-alone version of the SKB has been developed as part of the open source WOOL Platform (called WOOL Web Service).

7.3.3.5 Achievements (M19-M36)

- Knowledge gathered through the connected sensors (Fitbit in functional demonstrator, HBAF in technical demonstrator) is stored in the Shared Knowledge Base and is available for tailoring dialogues between coaches and users.
- The shared knowledge base has been integrated in the functional and technical demonstrators; it is the main backend component for the functional prototype and provides data storage and user management for the technical prototype.
- A stand-alone version of the Shared Knowledge Base has been developed to facilitate uptake as a component of the Agents United and WOOL platforms.

7.3.3.6 Deviations from DoA

There have been no significant deviations from the DoA.

8 WP4: User Behaviour Sensing, Modelling and Analysis

8.1 Planning

WP / Task	Title	Participants	Months
WP4	User Behaviour Sensing, Modelling and Analysis	CMC	M1-M36
T4.1	Inference of Short-Term Behaviours from Sensor Data	<u>SU</u> , CMC, RRD, UPV	M1-M15
T4.2	Identification of Long-Term Behaviours from Short-Term Information	<u>CMC</u> , UPV	M9-M21
T4.3	Automatic Detection of Behaviour Changes	<u>CMC</u> , RRD	M18-M27
T4.4	Evaluation of Holistic Behaviour Analysis Framework	<u>RRD</u> , CMC	M25-M36

8.2 Main Achievements

The following main achievements result from this work package:

8.2.1 First Period (M1-M18)

- Set of definitions for key terminology used in WP4 and the project in general.
- In-depth review of sensing technologies to measure physical, emotional, cognitive and social behaviours.
- Initial set of requirements, use case specifications and first design of the Holistic Behaviour Analysis Framework component.
- Short-term physical, social, emotional and cognitive behaviour detection models.
- Preliminary evaluation of the short-term behaviour detection models.
- Software components implementing the developed short-term behaviour detection models
- Definitions for long-term behaviours or routines out of short-term behaviours.
- Knowledge-driven long-term behaviour detection model based on short-term physical and social behaviours.
- Data-driven long-term behaviour detection model based on short-term physical and social behaviours.

8.2.2 Second Period (M19-M36)

- Refinement, update and elaboration of new definitions for key terminology used in WP4 and the project in general.
- Long-term physical, social, emotional and cognitive behaviour detection models.
- Preliminary evaluation of the long-term behaviour detection models.
- Software components implementing the developed long-term behaviour detection models.
- Couch-as-a-sensor concept and related narratives/questionnaires.
- Behaviour change detection models.
- Preliminary evaluation of the behaviour change detection models.
- Software components implementing the developed behaviour change detection models.
- Novel holistic human behaviour dataset (6-week study, 20 participants).
- Overall evaluation of the holistic behaviour analysis framework (HBAF) methodology

8.3 Individual Tasks

8.3.1 T4.1: Inference of Short-Term Behaviours from Sensor Data

8.3.1.1 Objectives

The objective of this task is to investigate and develop the necessary methods and tools for translating multimodal sensory data into short-term behaviours or primitives. The measurement of behaviour will be approached through both physical sensors (e.g., accelerometers for registering body motion, video for capturing facial expressions or audio for recording voice tone) and virtual sensors (i.e., coaches or virtual agents for sampling user experiences and feelings). The implementation of the sensor data collection will be conducted in this task. Machine learning and natural language processing techniques will be used for the definition of the short-term behaviour recognisers. This task will particularly contribute with the development of new enablers for recognising each domain of behaviour, namely physical, emotional, cognitive and social.

8.3.1.2 Progress (M1-M18)

This task has run from the start of the project (M1) throughout the first year, having finished in November of 2018 (M15). This task has delivered *D4.1: State-of-the-art, requirement analysis and initial specification of the Holistic Behaviour Analysis Framework* in M6, *D4.2: Methods for inferring short-term behaviour from multimodal sensor data* in M12 and *D4.3: Short-term behaviour analysis prototype* in M15 without delays.

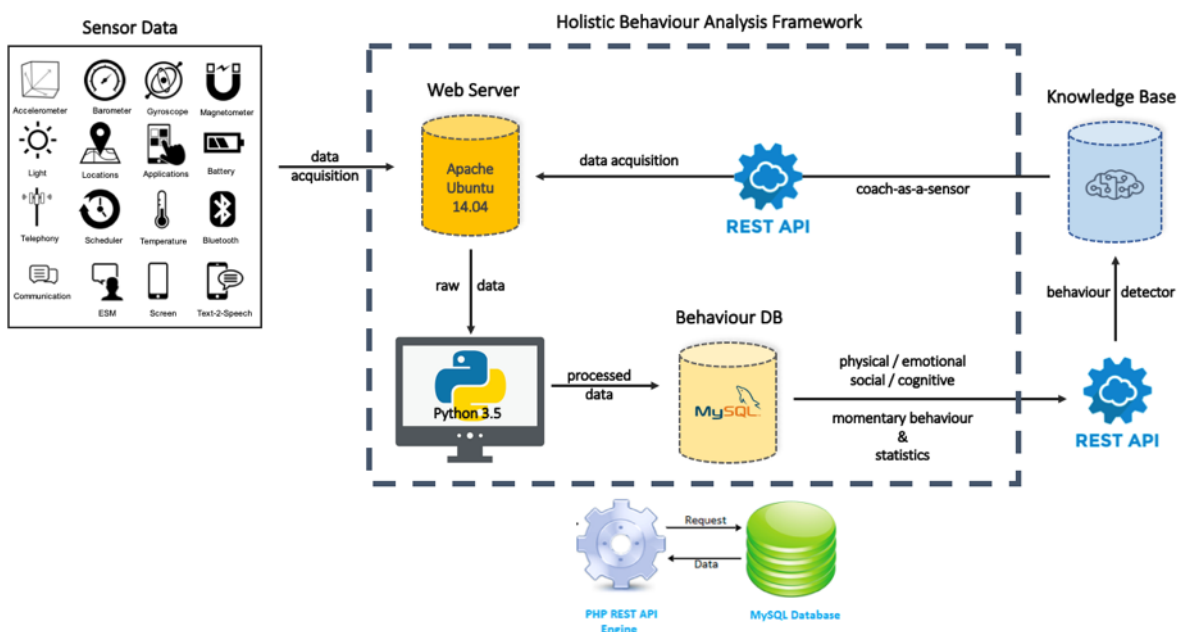


Figure 8: Operation flow of the HBAF.

8.3.1.3 Achievements (M1-M18)

- Provide a set of definitions for key terminology used in WP4 and the project in general (D4.1);
- Provide an in-depth review of on-body and off-body sensing technologies to measure physical, emotional, cognitive and social behaviours (D4.1);
- Provide an initial set of requirements, use case specifications and first design of the Holistic Behaviour Analysis Framework component (D4.1);
- Identify relevant sensor data for the detection of short-term (a.k.a. momentary) behaviours (D4.2);
- Develop a momentary physical behaviour detection model based on accelerometer and GPS data to detect steps and some commonplace activities (walking, commuting, biking, tilting or remaining still - standing/sitting) (D4.2);

- Develop a momentary social behaviour detection model building on Bluetooth, call and SMS logs, and ambient noise data to detect social activeness / inactiveness (D4.2);
- Develop a momentary emotional behaviour detection model based on audio and video data to detect arousal and valence (D4.2);
- Develop a momentary cognitive behaviour detection model based on audio and video data from natural dyadic interactions to detect levels of user-coach engagement (D4.2);
- Perform a preliminary evaluation of all developed short-term behaviour detection models (D4.2);
- Build a software component that implements the developed momentary physical and social behaviour models (D4.3);
- Build a software component that implements the developed momentary emotional and cognitive behaviour models (D4.3);
- Developed three video demonstrators to showcase the operation of each developed software component (D4.3);

8.3.1.4 Progress (M19-M36)

This task was completed during period 1.

8.3.1.5 Achievements (M19-M36).

This task was completed during period 1.

8.3.1.6 Deviations from DoA

This task has set out and worked towards achieving the objectives as described in the Document of Action and has achieved these in large parts. The original plan was to have real time speech recognition during the coaching sessions. During one of the consortium meetings, the researchers present, all agreed on the state of the art of real time language recognition was not good enough to be used within the project. The freedom of speech, where an end user goes on a long rant about different topics, would lead to an extremely branched out tree of coaching options, that would probably decrease the quality of the specific advice and coaching strategies the coaches can provide. Therefore, we chose to limit the input option for the user to specific answers, that can be provided through on the screen. This way the coaching advice will be more to the point and therefore more valuable to the end user.

8.3.2 T4.2: Identification of Long-Term Behaviours from Short-Term Information

8.3.2.1 Objectives

The objective of this task is to research and develop the techniques for intelligently combining the behaviour primitives generated in T4.1 into more descriptive representations of behaviour. The combination or fusion process will be carried out in a temporal and conceptual level. The temporal fusion will allow to merge sequences of granular actions or primitives (e.g., gestures) into more general behaviour definitions (e.g., activities). This horizontal fusion will be performed over primitives of the same domain of behaviour (e.g., physical). The conceptual fusion will allow to unite cross-domain behaviours (e.g., activities and emotions) into more abstract representations of behaviour (e.g., routines or lifestyles). Topic models, probabilistic models and semantic models will be primarily used to fuse the behaviour information at both temporal and conceptual levels.

8.3.2.2 Progress (M1-M18)

This task started in May of 2018 (M9) and will finish in May of 2019 (M21). This task has delivered *D4.4: Methods for inferring long-term behaviours from short-term behaviour information* in M18 without delays and it is expected to deliver *D4.5: Long-term behaviours analysis prototype* in M21.

8.3.2.3 Achievements (M1-M18)

- Provide a set of knowledge-driven definitions for long-term behaviours or routines out of short-term behaviours (D4.4).
- Select relevant statistical and heuristic features to detect long-term behaviours (D4.4).
- Provide a brief overview of regression models relevant to identify long-term behaviours (D4.4).

- Develop a knowledge-driven long-term behaviour detection model based on short-term physical and social behaviours.
- Develop a data-driven long-term behaviour detection model based on short-term physical and social behaviours.
- Deployment of sensor data collection platform.
- Development of REST API to facilitate data exchange with other COUCH components.
- Interface with UniversAAL.

8.3.2.4 Progress (M19-M36)

CMC has worked on this task that started in May of 2018 (M9) and finished in May of 2019 (M21). This task has delivered *D4.4: Methods for inferring long-term behaviours from short-term behaviour information* in M18 and *D4.5: Long-term behaviours analysis prototype* in M21 without delays. The developed models and software components have been preliminary tested in this task and will be further evaluated during the longitudinal evaluation of the Holistic Behaviour Analysis Framework in T4.4. UPV has been working on the integration with the rest of the system.

8.3.2.5 Achievements (M19-M36)

- Define the concept and types of long-term behaviours (D4.4);
- Identify the most relevant statistics and machine learning techniques to transform short-term behaviours into long-term behaviours (D4.4);
- Define a set of descriptive features to characterize physical, social, emotional and cognitive short-term behaviours in order to calculate long-term behaviour data series (D4.4);
- Develop a long-term behaviour machine learning detection model based on the time series of descriptive features (D4.4);
- Build a software component that implements the developed long-term behaviour detection model (D4.5);
- Evaluate the developed long-term behaviour detection models (D4.5);
- Developed a video demonstrator to showcase the operation of the software component (D4.5);

8.3.2.6 Deviations from DoA

There have been no deviations from the DoA.

8.3.3 T4.3: Automatic Detection of Behaviour Changes

8.3.3.1 Objectives

The objective of this task is to investigate and develop the methods for automatically discovering relevant behaviour changes for a given user or group of users from the analysis of the long-term behaviours generated in T4.2. The analysis will consist of detecting changes between time periods (change detection), determining the significance of the detected changes (change assessment) and analysing the nature of the change (change explanation). Statistical multivariate analysis and regression techniques will be primarily used for the detection, assessment and explanation of changes.

8.3.3.2 Progress (M1-M18)

This task has not started yet.

8.3.3.3 Achievements (M1-M18)

This task has not started yet.

8.3.3.4 Progress (M19-M36)

CMC has worked on this task that started in February of 2019 (M18) and finished in November of 2019 (M27). This task has delivered *D4.6: Methods for detecting behaviour changes from short-term behaviour information* with a 3-months delay (as approved by the PO) and *D4.7: Behaviour change detection analysis prototype* in M27 without delay. The developed models and software components have been

preliminary tested in this task and will be further evaluated during the longitudinal evaluation of the Holistic Behaviour Analysis Framework in T4.4. UPV has been working on the integration with the rest of the system. RRD has been working on the integration with the knowledge base.

8.3.3.5 Achievements (M19-M36)

- Define the couch-as-a-sensor concept (D4.6);
- Develop behaviour-specific questionnaires for the realization of the couch-as-a-sensor (D4.6);
- Define the concept of behaviour change (D4.6);
- Develop a behaviour change detection model based on change point detection statistical techniques to identify tipping points in short-term behaviour time series (D4.6);
- Create a synthetic dataset of short-term behaviour time series to generate controlled changes (D4.6);
- Evaluate the performance of various change point detection techniques on the synthetic dataset and also a realistic dataset (D4.6);
- Build a software component that implements the developed behaviour change detection model (D4.7);

8.3.3.6 Deviations from DoA

This task was originally planned to determine the nature of the change (change explanation). While this is a concept intrinsically connected to the detection of behaviour changes, it has been deemed during the course of the project to be much more approachable and objective when realized via the user-council conversations. Hence, this task has rather focused its efforts on developing the methodology for the automatic detection of variations in the behaviour data time series, which in turn can trigger the possible intervention of the council to scrutinize the reason(s) for such change(s).

8.3.4 T4.4: Evaluation of Holistic Behaviour Analysis Framework

8.3.4.1 Objectives

The objective of this task is to assess the functional operation of the Holistic Behaviour Analysis Framework. The evaluation will be performed in real-world scenarios and it will particularly assess the reliability and resilience of each of the three core components of the framework (T4.1, T4.2 and T4.3).

8.3.4.2 Progress (M1-M18)

This task has not started yet.

8.3.4.3 Achievements (M1-M18)

This task has not started yet.

8.3.4.4 Progress (M19-M36)

RRD and CMC have worked on this task that started in September of 2019 (M25) and finished in August of 2020 (M36). This task has delivered *D4.8: Evaluation of the Holistic Behaviour Analysis Framework* without delays.

8.3.4.5 Achievements (M19-M36)

- Develop a research protocol to perform a data collection study for the assessment of the holistic behaviour analysis framework (D4.8);
- Obtain the approval of the University of Twente Ethics Committee for the realization of the study (D4.8);
- Elaborate the call for participation and related documentation (D4.8);
- Recruit the participants of the study via social media channels (D4.8);
- Conduct a 6-week data collection study involving 20 participants (D4.8);
- Curate the collected data prior its use for evaluation (D4.8);
- Evaluate the performance of the HBAF models on the collected dataset (D4.8);

8.3.4.6 Deviations from DoA

The evaluation of the HBAF framework has been conducted via a study that does not include people from the relevant target groups (diabetes, chronic pain, elderly). The HBAF methodology is disease-independent, that is the reason why it was not stated in the DoA in the first place. In either case, the COVID-19 pandemic has made it difficult to reach out participants of those groups. However, the lockdown and de-escalation has presented an unparalleled opportunity for us to evaluate relevant changes in people's behaviour. This does not represent a deviation per se but rather an opportunity for this methodology to be tested in a rather dynamic and potentially reproducible context, which certainly affects the project's target population and the society as a whole.

9 WP5: Dialogue and Argumentation Framework (UDun)

9.1 Planning

WP / Task	Title	Participants	Months
WP5	Dialogue and Argumentation Framework	UDun	M1-M36
T5.1	Theoretical Model	UDun, CMC, RRD	M1-M12
T5.2	Abstract Model	UDun, RRD, SU	M10-M21
T5.3	Computational Model	UDun, CMC, SU, UPV	M18-M30
T5.4	Evaluation of Framework	RRD, CMC, UDun	M25-M36

9.2 Main Achievements

The following main achievements result from this work package:

9.2.1 First Period (M1-M18)

- A corpus of analysed real-life interactions between a patient and at least two medical professionals.
- The design of formal dialogue games, complete with locution, structural, commitment, termination, and outcome rules, based on the analysed interactions. Provide a design for the Dialogue and Argumentation Framework component that will fit into the overall Council of Coaches architecture.
- Integration of the Dialogue Game Execution Platform with ASAP and Flipper.
- Initial coaching dialogue game implementation, accepted for publication and presentation at the Computational Models of Argument (COMMA) 2018 conference.

9.2.2 Second Period (M19-M36)

- Final coaching dialogue implementation as a result of testing
- Implementation of a web-based testing interface for the Dialogue and Argumentation Framework
- Development of the “utterance generator” for finding content for dialogue moves based on dynamic argument models that are updated as a dialogue progresses
- Manuscript accepted for publication in the journal *Argument & Computation*, on ethical issues for using dialogue and argumentation in health care
- Paper on selection of dialogue move types and content accepted for presentation and publication at the *Computational Models of Argument (COMMA) 2020* conference
- Submission of revised version of *Language Resources and Evaluation* manuscript on the *Patient Consultation Corpus*
- Final integration of the Dialogue and Argumentation Framework with the rest of the technical demonstrator modules

9.3 Individual Tasks

9.3.1 T5.1: Theoretical Model

9.3.1.1 Objectives

The objective of this task is to develop theoretical foundations and a formal dialogue system. Using a corpus of real-life interactions, we will research and develop a flowchart of dialogue moves and from this a formal dialogue game. This will necessitate identifying locution rules, structural rules, commitment rules, termination rules and outcome rules.

9.3.1.2 Progress (M1-M18)

This task has run from the start of the project (M1) throughout the first year, having finished in August of 2018 (M12). This task has delivered *D5.1: Dialogue and Argumentation Framework Design* in M6 and *D5.2: Design of Dialogue Game* in M12 without delays.

9.3.1.3 Achievements (M1-M18)

The following major actions have been achieved by the task:

- Provide a design for the Dialogue and Argumentation Framework component that will fit into the overall Council of Coaches architecture (D5.1).
- A corpus of analysed real-life interactions between a patient and at least two medical professionals (D5.2).
- The design of formal dialogue games, complete with locution, structural, commitment, termination, and outcome rules, based on the analysed interactions (D5.2).
- Integration of the Dialogue Game Execution Platform with ASAP and Flipper.

9.3.1.4 Progress (M19-M36)

This task was fully completed in Period 1.

9.3.1.5 Achievements (M19-M36)

This task was fully completed in Period 1.

9.3.1.6 Deviations from DoA

This task has set out and worked towards achieving the objectives as described in the Document of Action and has achieved these in large parts. During the analysis of real-life interactions (D5.2) it was found that pathos, logos and ethos were not found in measurable quantities such that they could be incorporated into the design of the dialogue game. This was something that could not have been foreseen prior to the gathering and analysis of the interaction data. This data has however allowed for the specification of dialogue games that accurately reflect the processes involved in a coaching session.

9.3.2 T5.2: Abstract Model

9.3.2.1 Objectives

The objective of this task is to investigate moving from a formal dialogue system to a set of directed graphs, over which we can reason. We expect to use the Argument Interchange Format standard in the execution of our dialogue game to create argument maps which represent a collective knowledge base as a directed graph. Using established theories of argumentation, we will explore the construction of a shared theory that is accepted by the virtual coaches and the user at a given stage of the game.

9.3.2.2 Progress (M1-M18)

This task has run from June 2018 (M10) and is currently ongoing until the end of May 2019 (M21). The first deliverable from this task, *D5.3: Methods and tools for moving from dialogue game to updating individual and shared knowledge bases*, will be delivered on-time at the end of February 2019 (M18).

9.3.2.3 Achievements (M1-M18)

- Initial coaching dialogue game implementation, accepted for publication and presentation at the Computational Models of Argument (COMMA) 2018 conference.

9.3.2.4 Progress (M19-M36)

This task has run from June 2018 (M10) and was completed at the end of May 2019 (M21). Following submission of D5.3 (in Period 1), the remainder of the task focused on testing the dialogue descriptions reported in that deliverable.

9.3.2.5 Achievements (M19-M36)

- Final coaching dialogue implementation based on the results of testing

9.3.2.6 Deviations from DoA

During development of the Dialogue and Argumentation Framework (DAF), it was identified that translating from the Argument Interchange Format (AIF) to ASPIC+ was only deemed necessary because the current version of the Dialogue Game Execution Platform (DGEP) generated AIF as an output. Since the DAF required a new and updated version of DGEP to be implemented, it was decided to bypass AIF and instead focus on ASPIC+ directly.

9.3.3 T5.3: Computational Model

9.3.3.1 Objectives

We will embody both theoretical and abstract model computationally, in conjunction with WP5. This will involve, for example, specifying and executing the protocols developed in T4.1, using tools such as the Dialogue Game Description Language (DGDL) and the Dialogue Game Execution Platform (DGEP) (Bex et al., 2014). We will explore inducing abstract frameworks from the AIF structures created as a side-effect of executing the DGDL specification, over which we could then compute acceptability semantics to create the shared knowledge base.

9.3.3.2 Progress (M1-M18)

This task has not started yet.

9.3.3.3 Achievements (M1-M18)

This task has not started yet.

9.3.3.4 Progress (M19-M36)

This task ran from March 2019 (M19) to February 2020 (M20). The first deliverable from this task, *D5.4: Design of computational model based on theoretical and abstract model* was written by UDun, RRD and CMC, and has been delivered on-time at the end of February 2019 (M30). The second deliverable from this task, *D5.5: Methods for evaluating the Dialogue and Argumentation Framework* was written by UDun and RRD, and has been delivered on-time at the end of February 2019 (M30).

9.3.3.5 Achievements (M19-M36)

- Web-based demonstrator for the Dialogue and Argumentation Framework
- Implementation of the “utterance generator” for finding dialogue move content
- Collaboration with DBT leading to submission of a manuscript to the *Journal of Argument and Computation* on ethical issues for the use of dialogue and argumentation in a health care context (final decision reported in T5.4).

9.3.3.6 Deviations from DoA

There have been no deviations from the DoA.

9.3.4 T5.4: Evaluation of Framework

9.3.4.1 Objectives

We will develop methods for evaluating our resulting framework, focusing in particular on correspondence to real life coaching scenarios, logical viability and resulting knowledge bases. This will be a semi-automatic evaluation, with the aim of analysing the outputs of dialogues entirely between autonomous agents representing the individual coaches. Evaluating the dialogue framework as a standalone module will allow for fine-tuning before it is integrated with the user interface for full-scale simulations in WP7.

9.3.4.2 Progress (M1-M18)

This task has not started yet.

9.3.4.3 Achievements (M1-M18)

This task has not started yet.

9.3.4.4 Progress (M19-M36)

This task ran from September 2019 (M25) to August 2020 (M36). The deliverable from this task, *D5.5: Evaluation of the Dialogue and Argumentation Framework* was written by UDun, and has been delivered on time at the end of August 2020 (M36).

9.3.4.5 Achievements (M19-M36)

- Paper on selection of dialogue move content accepted for presentation at the Computational Models of Argument (COMMA) 2020 conference.
- Acceptance of *Argument and Computation* submission (see achievements T5.3)
- Submission of revised version of *Language Resources and Evaluation* manuscript on the *Patient Consultation Corpus*
- Final integration of components of the Dialogue and Argumentation Framework into the Technical Demonstrator

9.3.4.6 Deviations from DoA

The evaluation of the Dialogue and Argumentation Framework (DAF) was adjusted to account for new and updated requirements as the project advanced. Instead of focusing on outcomes of autonomous dialogues between agents, the evaluation is now based on being able to replicate scripted outcomes using dynamic protocols, and create new outcomes using those same protocols.

10 WP6: Human-Computer Interfaces

10.1 Planning

WP / Task	Title	Participants	Months
WP6	Human-Computer Interfaces	SU	M1-M30
T6.1	User Interface Framework for Home and Mobile	CMC, RRD, SU, UPV	M1-M30
T6.2	Design and Creation of Virtual Agents	SU, CMC	M1-M20
T6.3	Modelling Group Interaction	SU, CMC, UDun	M6-M24
T6.4	Engagement Model	SU, CMC	M1-M30
T6.5	Simulation of Council of Coaching Sessions	SU, CMC	M16-M30

10.2 Main Achievements

The following main achievements result from this work package:

10.2.1 First Period (M1-M18)

- Definition of platforms and key terminologies used in WP6.
- Development of a Unity scene for the main user interface of Home UI.
- Integration of ASAP realiser and Greta agents to interact in the same scene.
- Integration of the Dialogue Game Execution Platform with ASAP and Flipper.
- Extension of Greta platform to support multiple agents.
- Development of six virtual characters and mapping their behaviours and traits to create a baseline.
- In-depth review of the existing multiparty models that are capable of handling turns using multiple virtual agents.
- Support of gaze behaviours for multiple agents in the Greta platform.
- Implementation of an engagement model to detect user's engagement.
- Understanding user's perceived level of persuasion about the virtual coaches, accepted for publication and presentation at the Intelligent Virtual Agents (IVA) 2018 conference.

10.2.2 Second Period (M19-M36)

- Development and evaluation of mobile interfaces for technical demonstrator (whatsapp/chatbox/skype).
- Development and evaluation of virtual reality interfaces for technical demonstrator (various environments and interaction modalities).
- Development of speech-based interface in functional demonstrator.
- Development of content for modelling conflicts between coaches based on theories from psychology and sociology. Conducting a user study to understand the impact on persuasiveness, group cohesion, interaction experience, perceived quality of coaching, and perception of the coaching team.
- Development of conflict mediation content for a peer-coach based on conflict mediation theories to mediate the conflicts from the content from the aforementioned study. Conducting a user study to understand the impact on persuasiveness, group cohesion, interaction experience, perceived quality of coaching, and perception of the coaching team, and compare it to the previous group discussion involving conflict study on the same measures.
- Improvement of the gaze behaviour for the Greta agents.

- Development and evaluation of a gesture prediction model based on prosodic features.
- Annotation of a data corpus for cohesive values and identifying the most prominent non-verbal cues associated with group cohesion.
- Development and evaluation of a group behaviour model to generate cohesive group behaviour for group of virtual agents.

10.3 Individual Tasks

10.3.1 T6.1: User Interface Framework for Home and Mobile

10.3.1.1 Objectives

Objective 1: *Home UI* – The focus in this task is the development of a “meeting room” in which the virtual coaches can be plugged in. The main User Interface for the Council of Coaches is the Home UI in which the primary user is able to interact with a group of virtual coaches. For the Home UI, the GRETA/VIB platform will form the basis of the technology platform, extending it with the capabilities of mixed-initiative multi-party dialogues between user and a group of 2 to 5 virtual coaches that can participate at the same time.

Objective 2: *Mobile UI* – The second objective of this task is to design and develop an easy-to-use, non-obtrusive mobile application that enables one-to-one user-to-coach interaction. This Mobile UI will be developed in such a way that it integrates with his/her own customary use of the mobile device. To reach this aim we will port the Greta/VIB platform onto Android mobile. The Greta/VIB platform is written in java and is integrated within Unity3D which is compatible with Android. In T6.1 we will develop the Server-Client connection between Greta/VIB (server) and Unity (Client).

10.3.1.2 Progress (M1-M18)

A Unity scene is developed that functions as the main User Interface of the Home UI where users can interact with the council of coaches. Within the scene three agents controlled by ASAP realiser (Kolkmeier, Bruijnes, Reidsma, & Heylen, 2017) and an agent controlled by Greta realiser (Ravenet, Cafaro, Biancardi, Ochs, & Pelachaud, 2015) exist. Both realisers have their own strong points and for that reason we would like to have access to both realisers and have the agents controlled by these realisers cooperate.

The ASAP realiser was turned into a multi-agent platform so it can control multiple coaches and the behaviours can be synchronised between the agents. Adaptations were made to BML to specify behaviour for a certain agent and to synchronise behaviour between such as gaze behaviour and turn-taking. First steps to integrate and synchronise behaviours between ASAP and GRETA were made.

The dialogue engine called Flipper 2.0 (van Waterschoot, et al., 2018) is used to integrate different components of the COUCH system. Flipper and DGEP can communicate to create a new DGEP powered dialogues where ASAP and GRETA powered agents can join the dialogue. Possible moves in the dialogue are generated by DGEP and translated into BML by Flipper. Flipper templates are used to generate UI buttons in the Unity scene where end users can select preselected user input. User input is fed back into the system. Different user evaluation tests were executed as described in deliverable D6.3.

For the second objective we investigated different options for the mobile UI (deliverable D6.2), but we postponed the development of the mobile app. We first focused on the integration of different components of the system to facilitate the development of the mobile application.

10.3.1.3 Achievements (M1-M18)

The following achievements have been reached:

- Unity scene with ASAP and Greta controlled agents
- ASAP multi agent platform.
- BML adaptations for synchronizing behaviour between agents
- Integration of DGEP dialogue with ASAP and GRETA.

10.3.1.4 Progress (M19-M36)

Additional user interfaces have been developed to further complement the main Unity scene in the technical demonstrator. Users can interact with the council of coaches through several different types of interfaces and modalities. We now have UIs that run on a mobile phone, based on recognizable styling inspired by WhatsApp and traditional chat boxes (see Figure 9). Furthermore, we have created a Unity scene with multiple cameras that resembles a video conference between the user and the coaches (see Figure 10). Finally, we have created several scenes that can be displayed on a virtual reality headset, allowing the user to interact with the council in an immersive 3D environment. We have implemented a living room, forest and beach house setting.

We implemented state of the art speech recognition in the functional demonstrator, accompanied with a text to speech system. This allows users to initiate in spoken dialogues with the council of coaches. A user study is planned where we compare the usability of the spoken dialogue approach with the traditional text-based interaction.

We have developed a generation model to predict when to produce a communicative gesture (gesture timing) and which gesture type. Gesture types, be beat gesture (i.e., marking speech rhythm, not related to the content of the speech) or ideational (i.e., related to the content of the speech) affect the impressions the agents give in terms of warmth and competence. The gesture class is computed based on the speech prosody. The model learns their relationship using a recurrent neural network with an attention mechanism. The model was evaluated objectively through ablation tests and subjectively asking participants about gesture naturalness and speech-relation.

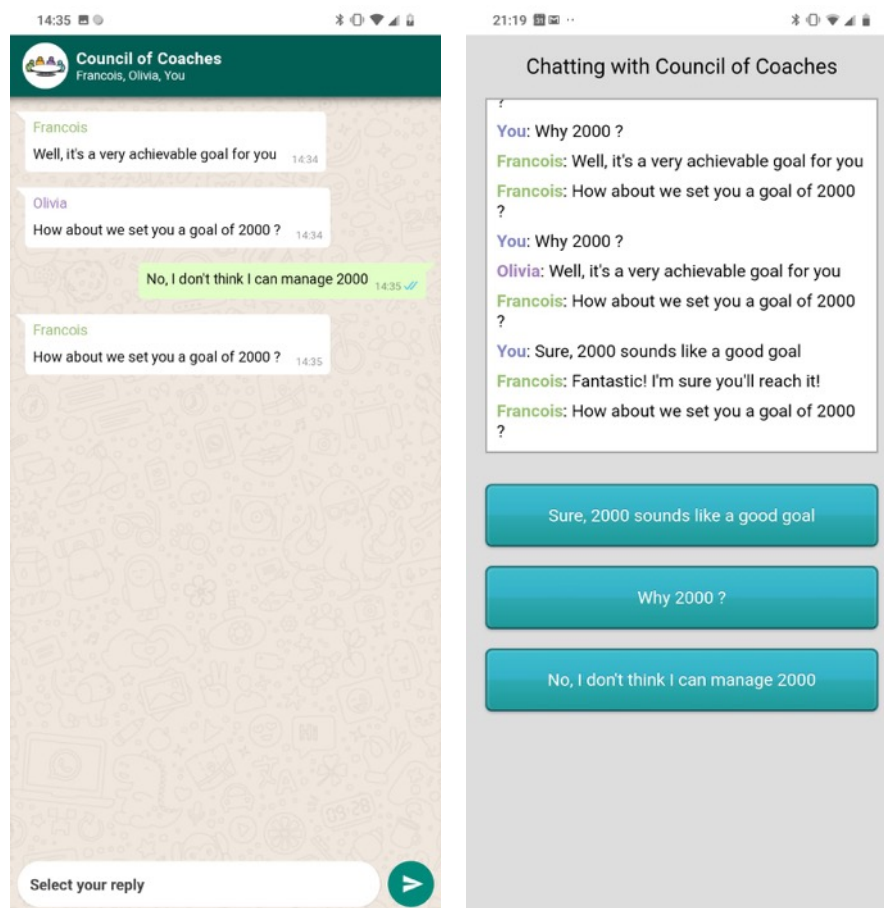


Figure 9: Mobile UIs inspired by WhatsApp and chatbox styling.

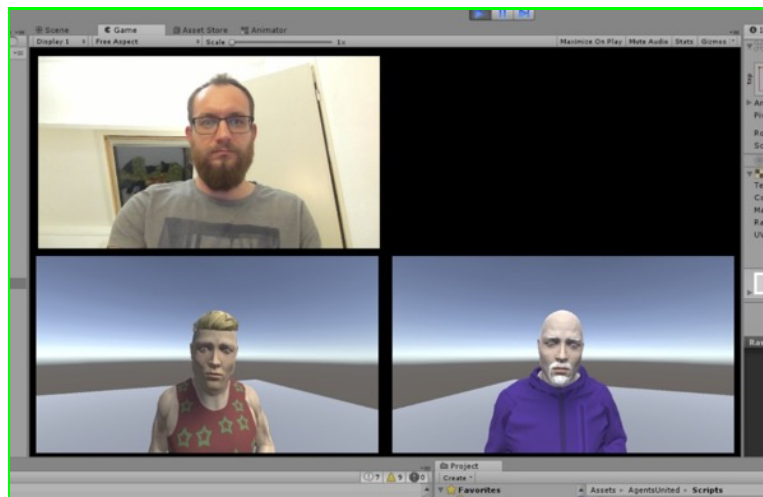


Figure 10: A Unity scene with multiple cameras, resembling video conferencing systems like Skype or Microsoft Teams.

10.3.1.5 Achievements (M19-M36)

- Development of several additional user interfaces and interaction modalities
- A machine learning model to predict gesture timing has been developed
- The gesture predictive model has been evaluated objectively by using a sequence comparison algorithm
- An ablation study was also performed to find out which prosody features are more pertinent
- The gesture prediction model was extended to find out the efficacy of including eyebrow movements to predict beat gesture timing

10.3.1.6 Deviations from DoA

We combined ASAP and GRETA in the COUCH system. ASAP was not mentioned in the DoA. Since both realisers have their own strong points, we decided to have access to both realisers and have the agents controlled by these realisers cooperate.

Mobile application development has seen some changes from the original vision. From the perspective of the end-user and the final demonstrations in WP7, we did not consider the effort required to build a high quality “mobile companion” app to be worth the gain – and it even would distract from the evaluation’s main focus: the conversations with multiple coaches at home on the Council of Coaches web application. Instead, mobile application development has found its focus as being part of the Agents United platform. As demonstrated at the project’s midway point, the Agents United platform supports seamless interaction between virtual agents on mixed desktop-mobile settings. This multi-platform support is now a core feature of the platform, allowing for creative multi-device apps (e.g. using UI on mobile, agents on desktop, or having mobile-to-desktop conversations between agents).

Some delays happened at the end of the project due to the COVID situation. The final technical development week had to happen remotely. Evaluation studies had to be redesigned to be conducted offline.

10.3.2 T6.2: Design and Creation of Virtual Agents

10.3.2.1 Objectives

The objective of this task is to develop a tool to create virtual coaches with their own communicative and emotional characteristics. A virtual coach is defined by its sensibilities, behaviour expressiveness and behaviour lexicon i.e., (i) *Behaviours*: their types and their expressivity; (ii) *Emotional sensibility*: Agents will be defined by their tendency to feel certain emotions and thus display certain multimodal behaviours; (iii) *Attitude*: Social attitudes that evolve dynamically is displayed through the variation of the verbal and nonverbal behaviours of the agent; (iv) *Interactional sensibility*: Coaches will be defined

by the type of feedbacks they have the tendency to provide (acknowledgment feedback or more emotional feedbacks) and by their degree of empathy toward user's emotions.

10.3.2.2 Progress (M1-M18)

This task has run from the start of the project (M1) throughout the first year, and will finish in April 2019 (M20). SU and CMC have delivered partially *D6.2: Initial user interface design for Home UI and Mobile UI* in M9 where initial virtual coach design was presented. The progress so far has been reported in a deliverable *D6.4: First Virtual Coach Design and Model* due in M18.

The consortium has developed a first “cast” of six characters for the virtual coaches: Helen Flores (Cognitive Coach), François Dubois (Diet Coach), Hank (Social Coach), Owen (Life Coach), Melissa (Mental Health Coach) and Alexa (Physical Activity Coach).



Figure 11: Character designs, from left to right Helen, Hank, Owen, Alexa, Francois, and Melissa.

The initial set of coaches are provided with a background story, physical description (appearance details e.g., build, height, skin, hair colour, clothes), mannerisms, strengths and weaknesses. These details have been reported in deliverable *D6.2*.

Further we started working on the behaviour characteristics for the agents to map their traits and attitudes to a baseline following Mancini's model of distinctive agents. So far, we have worked on two characters: Hank and Alexa which will be reported in *D6.4*.

10.3.2.3 Achievements (M1-M18)

Achievements so far are:

- *D6.2: Initial user interface design for Home UI and Mobile UI*
- *D6.4: First Virtual Coach Design and Model* due in M18.
- Six basic characters for virtual coaches
- Two more elaborated characters with behaviour characteristics.

10.3.2.4 Progress (M19-M36)

This task finished at M20, so only a small part of the task falls within the second reporting period. We finalized the creation of virtual coaches in term of behaviour expressivity and interaction sensibility. The new version of *D6.4* contains the detailed mapping for two virtual coaches.

A first evaluation study (see Figure 12) was performed to measure first impressions of and preference for static virtual coach images differing in gender, age and role. This study was published in the



Figure 12: The agents subjected to testing, differing in gender, age and role.

10.3.2.5 Achievements (M19-M36)

- D6.4, version 2: detailed description of two virtual coaches where the mapping of the agents' personalities is used to create a baseline for behaviour expressivity and interaction sensibility.

10.3.2.6 Deviations from DoA

There have been no significant deviations from the DoA.

10.3.3 T6.3: Modelling Group Interaction

10.3.3.1 Objectives

The objectives of this task is to model multiparty conversation behaviours. In particular, we will focus on turn-taking management and an associated dynamic group phenomenon i.e., cohesion. While indication of what the virtual coaches would say to whom and when will be provided by the dialogue framework (WP5), the turn-taking model will instantiate which behaviours the agents will display.

10.3.3.2 Progress (M1-M18)

This task has run from the first year of the project (M6), and will finish in August 2019 (M24). The progress so far will be reported in a deliverable D6.4: First Virtual Coach Design and Model due in M18.

We have conducted a literature review on the existing multiparty models and group cohesion models. Also, we have designed the annotation scheme for annotating the non-verbal behaviours that will be incorporated to facilitate turn-taking. The turn taking will also consider a higher-level group phenomenon which is cohesion. This is further categorised as task cohesion i.e., the performance of the group during task solving or decision making and social cohesion i.e., the behaviours of the group during social interactions. This will be reported in D6.4: First Virtual Coach Design and Model.

Previously the Greta platform supported a single agent only. An improvement was done to include multiple agents in the same environment. This will facilitate in developing a model to simulate group interactions. The work has been reported in D6.3: *First Prototype description and evaluations of the virtual coaches platform* (includes software documentation).

10.3.3.3 Achievements (M1-M18)

In this task the achievements have been:

- D6.3: First Prototype description and evaluations of the virtual coaches platform.
- A literature study has been performed.
- Annotation scheme for non-verbal behaviours.
- GRETA as a multi agent platform.

10.3.3.4 Progress (M19-M36)

To model the group interaction in human-agent interaction, we incorporate a group phenomenon named cohesion. In order to develop a cohesive model, our first step was to recognize the behaviours

associated to group cohesion. We conducted an analysis on annotations of (social and task) cohesion, dialog acts and non-verbal behaviours. Regarding the later, we considered: head movement, smile, eyebrow movement, gaze, laughter, turn-taking and interruption. Results are reported in D6.5.

We constructed a neural network using Keras toolkit with TensorFlow to model non-verbal behaviour associated with cohesion. In particular, we focus on the gaze, smile and turn-taking behaviours. This will be reported in *D6.6: Final virtual coach design and model* in M36.

The intent planner module has been extended with a component which is responsible for orchestrating the floor management between multiple agents. This component integrates with Greta's Group Behaviour Module, to ensure proper turn selection, turn taking and floor battling while generating appropriate nonverbal behaviours.

10.3.3.5 Achievements (M19-M36)

- Annotation of the Patient interview corpus for cohesion values.
- Statistical analysis and reporting on the prominent non-verbal behaviour cues associated with group cohesion and their prediction using machine learning methods.
- Constructing a model to generate cohesive listener behaviours for the group of virtual agents.

10.3.3.6 Deviations from DoA

There have been no significant deviations from the DoA.

10.3.4 T6.4: Engagement Model

10.3.4.1 Objectives

The objective of this task is to develop strategies to ensure engagement of the elderly in his interaction with the virtual coaches. One of the aims of the virtual coach is to gather data on a user's state. This will be done during the interaction between the coaches and the elderly. Therefore, in this task we will endow the virtual agent with different strategies to ensure engagement for the interaction to last as long as possible. This will be done by the agents building rapport by providing feedbacks. The feedbacks will indicate that the agent is interested in the user's speech and is paying attention to what user says. Further, the agent will also be capable of providing emotional feedbacks depending on their degree of empathy.

10.3.4.2 Progress (M1-M18)

This task has run from the start of the project (M1) throughout the first year, and will finish in February 2020 (M30). The progress so far will be reported in a deliverable *D6.6: Final Virtual Coach Design and Model* due in M30.

To model the engagement level of the user in human-agent interaction, we have constructed a neural network using Keras toolkit with TensorFlow. The model was designed to use the facial actions units, the head rotation and the conversational state of the interaction to predict the engagement, arousal and valence level of the user. This has been reported in *D4.2: Methods for inferring short-term behaviour from multimodal sensor data* in M12 and *D4.3: Short-term behaviour analysis prototype* in M15.

To maintain engagement, we are integrating a backchannel model based on Bevacqua's backchannel model. This model includes a set of rules that indicate when a backchannel should be triggered. We will develop on this model to decide the type of backchannel depending on the agent's state and their specific characteristics (that are developed in T6.2).

10.3.4.3 Achievements (M1-M18)

The achievements within T 6.4 are:

- D4.2: Methods for inferring short-term behaviour from multimodal sensor data.
- D4.3: Short-term behaviour analysis prototype.
- Model for engagement level of the user in human-agent interaction.

10.3.4.4 Progress (M19-M36)

We have explored the contribution of different multimodal features, namely gaze, head and action units, to the engagement prediction. We have also investigated the importance of considering one interlocutor's behavior for predicting the engagement of another interlocutor. We have designed a neural network for the purpose of predicting the engagement level of an human participant. The model has three layers: a LSTM hidden layer to extract features from the input layer, followed by an output layer (dense) for predicting the engagement level.

10.3.4.5 Achievements (M19-M36)

The achievements within T6.4 are:

- Model for engagement level of the user taking into account the different interlocutors' behaviours.

10.3.4.6 Deviations from DoA

There have been no significant deviations from the DoA.

10.3.5 T6.5 Simulation of Council of Coaching Sessions

10.3.5.1 Objectives

The objective of this task is to develop and evaluate a non-interactive presentation of our dialogue framework. This implements the notion of "dialogue as performance" where a user can passively watch agents argue without interacting themselves. The main idea of this task is to provide simulations of coaching sessions to the elderly. The simulation will go from scripted scenarios at the beginning of the project to fully interactive scenarios. Through this task we aim: (i) To test user interface and presentation styles at early stages of the development of the virtual coach dialogue framework; (ii) To test the functional design of the argumentation framework developed in WP4 and the interfaces with the GRETA/VIB platform for behaviour generation; (iii) To demonstrate the concept of COUCH as a 3D virtual environment where users can watch using a VR headset.

10.3.5.2 Progress (M1-M18)

This task was planned to start from the second year of the project (M16) and will run throughout the second year and finish in February 2020 (M30). However, this task has started early and an initial evaluation study was conducted which has been reported in D6.2: *Initial user interface design for Home UI and Mobile UI* in M9. The progress will be reported in various deliverable throughout the project incrementally.

This task started earlier than planned (M16). An evaluation study was conducted to understand the effects of status of the agent and persuasion techniques used by the agents on user's persuasion. We conducted this initial study where agents employed persuasive arguments about a very neutral and popular topic of discussion: movies.

From this study, we confirmed that using multiple agents to provide persuasive arguments is better than using single agents and that authoritative agents are perceived to be more credible and persuasive. Further we also observed a tendency where vicarious persuasion (where the aim is to persuade the audience rather than the person with whom a proponent is directly engaged in discussion) was effective.

10.3.5.3 Achievements (M1-M18)

Achievements that have been reached are:

- Evaluation study into agent status and persuasion techniques.

10.3.5.4 Progress (M19-M36)

We performed two evaluation studies of the initial technical and of the initial functional prototypes developed for the Council of Coaches project which were reported in D6.5: Final prototype description

and evaluations of the virtual coaches platform. We performed several user evaluations which are reported in *Deliverable 6.6: Final virtual coach design and model*:

- Conducted an online user study with around 240 participants, in which they interacted with a coaching team consisting of four 2D agents in group discussions to have different models on coach behaviour in group discussions evaluated. We evaluated the impact on persuasiveness, group cohesion, interaction experience, perceived quality of coaching, and perception of the coaching team.
- Conducted an online user study with around 240 participants, in which they interacted with a coaching team consisting of five 2D agents in group discussions based on the same model. However, the fifth coach that was added played the role of a peer mediator, and in both models would mediate conflicts using mediation approaches found in mediation literature. Here we evaluated the impact this had in both models on persuasiveness, group cohesion, interaction experience, perceived quality of coaching, and perception of the coaching team, as compared to the previous study.
- Conducted an online user study with around 30 participants, in which they interacted with a coaching team consisting of four 3D agents in group discussions, using the dialogue developed previously. We had two conditions where one generated group behaviours randomly and the other generated cohesive group behaviour based on our model. We measured the perceived level of cohesion, trust and persuasiveness of the agents by the user.
- Conducted an online user study with around 30 participants, in which they viewed videos of virtual agent exhibiting communicative behaviours. We measure the impact of our computational model of communicative gestures.

10.3.5.5 Achievements (M19-M36)

- Evaluation studies of the functional and technical prototypes
- Evaluation studies related to user interface usability evaluation, multi-device interaction, impact of verbal conflict presentation styles in group discussion, peer mediation in conflict in group discussion, gesture generation, and cohesive group evaluation

10.3.5.6 Deviations from DoA

Due to the COVID-19 situation there were some delays in conducting the evaluation studies. They had to be redesigned to be run offline.

11 WP 7: Continuous Integration and Demonstration

11.1 Planning

WP / Task	Title	Participants	Months
WP7	Continuous Integration and Demonstration	UPV	M4-M36
T7.1	Technical Requirements, Specifications and Maintenance of APIs	UPV, CMC, RRD, SU, UDun, iSPRINT	M4-M30
T7.2	Development and Integration of the Functional Prototype	UPV, CMC, RRD, SU, UDun	M6-M21
T7.3	Development and Integration of the Technical Prototype	UPV, CMC, SU, UDun	M22-M36
T7.4	Final Demonstration	RRD, CMC, DBT, SU, UDun, UPV	M25-M36

11.2 Main Achievements

The following main achievements result from this work package:

11.2.1 First Period (M1-M18)

- Performed study on FIWARE and universAAL platforms looking for modules of interest. Created a Webinar on the topic and published it on the project's YouTube channel.
- Studied and agreed on a methodology for architecture formal definition: ARCADE. Trained partners on the ARCADE framework for our architecture definition.
- Defined an initial version of the architecture, including requirements, data models and interfaces (D7.1). This architecture has been continuously updated with any changes and additions to software modules developed in the technical WPs, including new requirements from user studies. These updates made are available as updates to the architecture repository.
- Set up the project repositories in GitLab for later publication as open source. Populated and initialized tools: requirements, issue boards, milestones, members, assignees...
- Set up an issue management process for the creation, tracking and resolution of issues, bugs, and requirement implementations based on the tools available in GitLab.
- Set up a development management procedure with periodic conference calls to report on progress, based on the development boards available in GitLab.

11.2.2 Second Period (M19-M36)

- As part of the Agents United Open Agent Platform, we transferred our code and development tools from GitLab to GitHub and made it public, with the LGPLv3 open source license.
- Updated, tested, and evaluated the Functional Demonstrator.
- Held multiple technical integration meetings for coordinating the development of the Technical Demonstrator.
- Updated the Technical Prototype with the remaining features to fulfil the requirements and recommendations.
- Development, packaging and release of the Third Functional Prototype and Final Council of Coaches Technical Prototype.
- Elaborated the demonstration protocol and obtained the approval of the ethical committee for the final demonstration.
- Performed the evaluation of the final demonstration and analysed the results.

11.3 Individual Tasks

11.3.1 T7.1: Technical Requirements, Specifications and Maintenance of APIs

11.3.1.1 Objectives

This task defines technical requirements for the project's integrated demonstrator(s). The architecture reported in this task will adopt a loosely integrated modular approach, with a focus on clearly defined APIs. Another focus of this task is to select the relevant modules within the FIWARE and universAAL frameworks that will facilitate development and deployment. Finally, this task sets up the process of continuous integration and technical support.

11.3.1.2 Progress (M1-M18)

This task started in M4 and will finish in M36, but the bulk of the effort was concentrated in the first period, until the release of the deliverable D7.1: *System architecture and design of APIs* in M9. The rest of the task is spread out until its finalization in M36, focusing on support and updates to the architecture and requirements.

11.3.1.3 Achievements (M1-M18)

The following major actions have been achieved by the task:

- Performed study on FIWARE and universAAL platforms looking for modules of interest. Hosted a Webinar on the topic and published it on the project's YouTube channel.
- Studied and agreed on a methodology for architecture formal definition: ARCADE. Trained partners on the ARCADE framework for our architecture definition.
- Defined an initial version of the architecture, including requirements, data models and interfaces (D7.1). This architecture has been continuously updated with any changes and additions to software modules developed in the technical WPs, including new requirements from user studies. These updates made are available as updates to the architecture repository.
- Set up the project repositories in GitLab for later publication as open source. Populated and initialized tools: requirements, issue boards, milestones, members, assignees...
- Set up an issue management process for the creation, tracking and resolution of issues, bugs, and requirement implementations based on the tools available in GitLab.
- Set up a development management procedure with periodic conference calls to report on progress, based on the development boards available in GitLab.

11.3.1.4 Progress (M19-M36)

The remainder of this task was initially focused on maintaining the architecture up to date and oversee the development efforts to keep them in line with requirements. An additional effort was added to manage the transition to public open source as part of the Open Agent Platform (Agents United).

11.3.1.5 Achievements (M19-M36)

- Established a new communication structure, with periodic teleconference calls for development management: 1) "Technical demonstrator to open agent model" and 2) "Functional demonstrator to evaluation".
- As part of the Agents United Open Agent Platform, we decided to transfer our source code from GitLab to GitHub, and make it public. This implied transferring the code repositories, developer accounts and roles, readmes, documentation, issue tracking...
- As part of the transfer to GitHub, the structure of the code repositories was reorganized, with two main purposes: 1) Use a cleaner and more independent separation of modules and 2) make a clear distinction between modules created in COUCH hosted by the Agents United Alliance, and existing modules still hosted in external repositories.
- Established Continuous Integration processes, and procedures for nightly builds.
- This architecture has been continuously updated with any changes and additions to software modules developed in the technical WPs, including new requirements from user studies.

11.3.1.6 Deviations from DoA

Some of the models of the architecture definition, as required if following the ARCADE methodology, are not yet completed. This is because it was difficult at the moment of publication of D7.1 to have a clear idea of the more complete and advanced models, such as deployment, which require implementation details that are not yet available in the development. As details are settled and implementation advances to more complete stages of the prototype, these models will be filled and published.

The intention of incorporating FIWARE and universAAL platforms has resulted in two different approaches: The data models used in Council of Coaches (in particular the Knowledge Base) are designed so that they are interoperable with FIWARE data models. Regarding universAAL, once the final system is complete, it will be made into a universAAL-compliant application by implementing the appropriate connectors and ontologies.

Finally, since the only official deliverable of this task, D7.1, was released in M9, some of the progress made afterwards in this task is tangentially reported in other deliverables in this WP (for instance, the definition of the issue management process, which is included in D7.3).

In the second period, some of the models of the architecture definition based on ARCADE will not be completed. Those missing ones are mostly deployment-focused models, which are not relevant anymore given the multi-platform technologies used in the system (e.g. Docker) and multiple deployment options. The existing models represent the architecture sufficiently enough.

The change in the engine of the Knowledge Base and its migration from a document-based to a variable-based storage system means that it is no longer compatible with FIWARE as it was in its first version. Compatibility with FIWARE, if needed, will be left to the future maintenance of the Open Agent Platform in the Agents United Alliance.

11.3.2 T7.2: Development and Integration of the Functional Prototype

11.3.2.1 Objectives

This task develops and maintains the functional prototype, which at first serves as a driver and testbed for the platform specifications as defined in T7.1. The functional prototype evolves across the planned milestones, starting with a working Wizard-of-Oz type prototype in M9. Then the functional prototype is constantly updated and expanded with features developed in WP3, WP4, WP5 and WP6, as well as input from user-engagement efforts in WP2. The functional prototype provides a realistic view of the capabilities of the Council of Coaches applications. Its main purpose is to demonstrate the system's external behaviour, functionalities, and visual aspects.

11.3.2.2 Progress (M1-M18)

The task started in M6 and will finish in M21. It is divided in prototype iterations, after which a deliverable reports the release of each prototype as a packaged software. The first two iterations have been completed and reported: D7.2 *Initial functional prototype* and D7.3 *Second functional prototype*. The third and final functional prototype before the completed system is ongoing.

11.3.2.3 Achievements (M1-M18)

The following major actions have been achieved by the task:

- Uploaded initial codebase to GitLab repositories.
- Set up Technical Integration Taskforce, in charge of coordinating work of developers across technical WPs in order to integrate their progress into a cohesive prototype.
- Split development efforts between Functional Prototype and Technical Prototype. The Technical Prototype holds the current overall integrated system, which in some aspects may still be rough to be evaluated by end users in WP2. The Functional Prototype wraps the more stable inner modules with mock-up interfaces (starting with the so-called Wizard of Oz prototype) of evolving complexity, representing the idealized interfaces that users can evaluate.
- Held face-to-face workshops to coordinate efforts of the Technical Integration Taskforce and speed up development in a coordinated way.

- Development, packaging and release of the First Functional Prototype (D7.2).
 - Modified DGEP to support message-oriented architecture.
 - Send BML from Flipper and Behavior Planner, via ActiveMQ, into ASAP & Greta.
 - Full integration between DGEP, Flipper, ASAP and Greta.
 - Set of sample dialogues to be used in Initial Functional Prototype.
 - Enhanced representation and movement of coaches according to personality.
- Development, packaging and release of the Second Functional Prototype (D7.3).
 - Analyse results of evaluation of First Functional Prototype and identify new requirements.
 - Determine enhanced scenarios for Second Functional Prototype: concrete age-related scenario.
 - Inclusion of Greta agent into Unity scene along with embedded browser.
 - Enhanced sensor data gathering and native integration with FIWARE platform.
 - Design of data models and interfaces for the Knowledge Base
 - Conversion of Yarn to DGD L
 - Enhanced gaze behaviour

11.3.2.4 Progress (M19-M36)

The task finished shortly after the first period, producing the deliverable D7.4 *Third functional prototype*. However, due to the dual approach to development explained in the first report (Functional Prototype + Technical Prototype) the development structure set up in T7.2 continues essentially unaltered in T7.3 to produce the Final version of the prototype(s).

11.3.2.5 Achievements (M19-M36)

- Migrated the Shared Knowledge Base engine to R2D2 storage system (RRD tool) using a variable storage approach.
- Online test and evaluation of the Functional Demonstrator beta.
- Technical integration meeting in Twente for preparing the Third Prototype, 25th-29th March 2019.
- Development, packaging and release of the Third Functional Prototype (D7.4).
 - Improved user experience of the Functional Demonstrator, including resolution scaling, login mechanism, privacy section and tutorials
 - Integrated Holistic Behaviour Analysis Framework based on AWARE, including Short Term Behaviour Analysis and universAAL Bluetooth devices, and linked it to the Knowledge Base for storing sensor data.
 - Enabled DGEP to manage conversations between agents, and sub-dialogues.
 - Introduced the Topic Selection Engine module, along with initial content for coaching topics and goal-setting.
 - Enabled use of WOOL scripts into the Dialog and Argumentation Framework.
 - ASAP and Greta agents in the same Unity scene.
 - Greta can now receive Function Mark-up Language and send feedback about time markers, automatically generate gestures, and incorporates the new Gaze model towards the user.

11.3.2.6 Deviations from DoA

The initial plan was to have a Functional Prototype that would eventually become a finalized Technical Prototype. However, because the development is performed in an integral way, these two efforts were set up in parallel: The Technical Prototype holds the current overall integrated system, which initially lacked some modules that were under development. The Functional Prototype wraps the more stable inner modules at each milestone with mock-up interfaces (starting with the so-called Wizard of Oz prototype) of evolving complexity, representing the idealized interfaces that users can evaluate. The development of the functional and technical demonstrators have been extensively described in the first periodic report.

11.3.3 T7.3: Development and Integration of the Technical Prototype

11.3.3.1 Objectives

The output of this task is the final system prototype that will be used for the project's demonstration in T7.4. The final demonstration requires a fully working client-server architecture in order to facilitate real-time sensing and data collection and the connection between home and mobile based user interfaces. This task deals with the provisioning of the infrastructure that enables the running of the various modules defined and developed in WP3-WP6, developing and supporting the prototype's deployment.

11.3.3.2 Progress (M1-M18)

This task has not started yet.

11.3.3.3 Achievements (M1-M18)

This task has not started yet.

11.3.3.4 Progress (M19-M36)

This task started in M22 by continuing the work of T7.2, but with the objective of delivering the final version of the prototype (actually, both prototypes, Functional and Technical, following the dual approach mentioned earlier). The task focused on finalizing all remaining development tasks to fulfill all requirements and recommendations, while setting the stage for the future code base of the « COUCH as a Product » and « Open Agent Platform ».

11.3.3.5 Achievements (M19-M36)

- As part of the Agents United Open Platform initiative, adopted LGPLv3 License as the open source license for the Agents United code. Applied the license to the code, and analysed, cleaned and refactored dependencies and libraries of all modules to get rid of incompatible licenses.
- Created demo videos of independent features and the full demonstrator
- Added possibility for the user to input text through the Unity interface, to add data that can later be used in dialogues.
- Single sign-on login mechanism in the Unity user interface, integrated with account management through AWARE, HBAF, SKB.
- Created the dialogues for the final demonstration including different coaching content, agent personalities and delivery style.
- Additional developments outside the scope of the Technical Demonstrator: Virtual Reality demo, packaging HBAF and sensor framework into Docker container, basic speech recognition experiment, scripted dialog processing is now possible in a companion mobile app, integrating the entire COUCH system as a universAAL application.
- Technical integration meeting in Twente for preparing the Final Prototype, 14th-18th October 2019.
- Remote technical integration meeting for managing the remaining development tasks, 20th-24th April 2019.
- Development, packaging and release of the Final Council of Coaches Technical Prototype (D7.5).
 - Visual overhaul of Functional Demonstrator, with coaches that are more realistic, custom selection of coaches, and more data for account creation.
 - Created different start scripts for different demos of the Technical Demonstrator.
 - HBAF and sensor integration updated with Long-Term Behaviour Analysis, link to Fitbit data, and the possibility to include sensor data in dialogues.
 - Added coaching conversations based on sensor data with actual sensor content.
 - Introduced Coach-as-a-Sensor (CaaS) feature, including questionnaires for psychosocial assessment, sending results back to SKB.

- Updated Shared Knowledge Base: definition of variable types, backup feature, initial demo data (including Fitbit activity), data feedback from other modules.
- Integrated new Utterance Generator module into the architecture.
- Multiple ASAP and multiple Greta agents in the same Unity scene, with non-verbal group social behaviour, floor management, gaze between the agents, and automatic gesture generation.

11.3.3.6 Deviations from DoA

There have been no significant deviations from the DoA.

11.3.4 T7.4: Final Demonstration

11.3.4.1 Objectives

This task designs the demonstration trial, considering medical ethical research issues and official approvals from committees. Then the full Council of Coaches system will be evaluated in a real-life setting, focusing on user experience and potential effect, with a non-experimental pre-test/post-test. The system will be deployed among 50 participants in the Netherlands and United Kingdom. Participants will be included based on pre-defined inclusion and exclusion criteria. User experience will be assessed qualitatively and quantitatively, and potential effect will be assessed by health-related factors. Participants will be asked to complete surveys before, during, and after interacting with the system. Systematic data from this activity will be used in large clinical-grade, regulatory-driven demonstrators for generating the evidence necessary for large scale uptake by health systems. To assess the effectiveness of the system, a micro-randomized trial (MRT) will be executed, which enables researchers to study proximal effects of specific intervention components, changes over time, and the factors that moderate those effects.

11.3.4.2 Progress (M1-M18)

This task has not started yet.

11.3.4.3 Achievements (M1-M18)

This task has not started yet.

11.3.4.4 Progress (M19-M36)

This task ran from December 2019 (M28) until August 2020 (M36). Two deliverables resulted from this task, namely:

- D7.6: Demonstration protocol, Ethical approval (M30)
- D7.7: Final demonstration results (M36)

Deliverable D7.6 was submitted in April 2020 (M32) to the commission. This demonstration protocol was also published in the journal of medical internet research in April 2020 (Hurmuz, Jansen-Kosterink, op den Akker, & Hermens, 2020). Deliverable D7.7 was finished on schedule in August 2020 (M36). The results of this task will be published in several journal/conference papers.

The final demonstration of the functional demonstrator was conducted in two rounds in two countries: The Netherlands and United Kingdom. For this, a pre-/post-test longitudinal study was set up at RRD and UDun. In December 2019 (M28), we started recruiting participants by advertisements in local newspapers and on University of Dundee's website and their weekly newsletter, messages on social media, an interview on a local TV station in Scotland, and snowball sampling. The aim was to include 50 participants in total, but during the recruitment process, we aimed at including 25 participants in each round in each country. In the end, we almost doubled the minimally required sample size. We reached a total of 94 older adults who participated in this study. These participants were enrolled in this study for 5-9 weeks: baseline week (week 1), implementation phase (week 2-5), and facultative phase (week 6-9).

11.3.4.5 Achievements (M19-M36)

The following achievements resulted from this task:

- Several publications in journals/conferences:
 1. Protocol of the final demonstration: already published (Hurmuz, Jansen-Kosterink, op den Akker, & Hermens, 2020)
 2. Results about design of virtual agents: submitted
 3. Results about user experience, potential health effects, and applicability of the virtual coaches: not published yet
 4. Results about the Micro Randomized Trial (MRT): not published yet
- Improvements of the functional demonstrator due to feedback of the older adults who participated in the study (Overview in D7.7).

11.3.4.6 Deviations from DoA

There have been no significant deviations from the DoA.

12 WP 8: Dissemination and Exploitation

12.1 Planning

WP / Task	Title	Participants	Months
WP8	Dissemination and Exploitation	iSPRINT	M1-M36
T8.1	Dissemination, Communication and Pre-Marketing Activities	iSPRINT, CMC, RRD, DBT, SU, UDun, UPV	M1-M36
T8.2	Exploitation and Business Planning	iSPRINT, CMC, RRD, DBT, SU, UDun, UPV	M25-M36
T8.3	Ecosystem Building of the Open Agent Platform	UPV, CMC, RRD, SU, UDun	M25-M36
T8.4	Standardization Activities	CMC, RRD, UPV	M7-M30
T8.5	Training activities	RRD, CMC, DBT	M25-M36

12.2 Main Achievements

The following main achievements result from this work package:

12.2.1 First Period (M1-M18)

- Development of communication strategy.
- Early launch of project website and continuous update of content.
- Setup of social media channels and content creation, in line with the project's strategy.
- Creation of dissemination material.
- Submission and publication of scientific papers.
- Plans for standardization and exploitation.
- Iterative and systematic approach of the COUCH business model definition.
- Inventory of standards used in the project.
- Inventory of the possible exploitable outcomes of the project.
- Workshop on regulatory challenges and Opportunities organised.

12.2.2 Second Period (M19-M36)

The following main achievements result from this work package:

- Communication strategy for the two exploitable outcomes of the project.
- Blog posts with Technical Episode Series.
- Submission and publication of scientific papers.
- Broad usage of social media channels.
- Organization of Virtual workshop that attracted 65 participants.
- Following standardization and using relative standards in the project.
- Exploitation plans for the two project outcomes.

12.3 Individual Tasks

12.3.1 T8.1: Dissemination, Communication and Pre-Marketing Activities

12.3.1.1 Objectives

Task T8.1 is related to the dissemination and communication activities, including publications in journals, presentations in conferences and exhibitions, establishing a strong presence in social and electronic media, participating in prominent events and more. A strategic plan will be developed from the start of the project, in order to detail the dissemination and communication activities in terms of specific dates, events, forums etc. Prior to this, T8.1 the task will deliver the project website to foster collaboration between consortium members and raise awareness about project activities.

12.3.1.2 Progress (M1-M18)

The activities of T8.2 are progressing according to the plan with fruitful outcomes that support the awareness creation of the project. More specifically, the project follows a 3-step communication approach that focuses on the communication aspects, as a pre-planned and ongoing activity and continuously and gradually evolving process, far beyond the project consortium: (1) awareness raising, (2) dissemination, (3) exploitation. In the first reporting period the project has covered the 1st phase and is currently executing activities for of the 2nd period.

So far, the dissemination activities contained:

- 4 press releases
- 9 blog posts on the website
- 2 campaigns on the website
- 1 YouTube webinar
- 1 Facebook live interview
- 70 Twitter posts (142 followers)
- 35 Facebook posts (>3k users reached)
- 12 Google+ posts
- 19 LinkedIn posts
- 9 YouTube videos with >500 views
- 12 oral presentations
- 4 poster presentations
- 3 demonstrations/demo booth
- 4 workshops for external participants
- 1 mention in a radio show
- 1 mention in a newspaper
- 12 scientific publications
- 1 training activity
- 2 mentions during educational activities

Overall, the project has both excellent scientific and popular dissemination activities, and we estimate the total audience reached to be over 10.000 individuals.

12.3.1.3 Achievements (M1-M18)

The achievement of T8.1 are listed below:

- On-time submission of deliverables D8.1, D8.2, D8.3.
- Development of communication strategy.
- Early launch of project website and continuous update of content.
- Setup of social media channels and content creation, in line with the project's strategy.
- Creation of dissemination material.
- Submission and publication of scientific papers.

12.3.1.4 Progress (M19-M36)

In the second reporting period all major stakeholders and target groups have been identified and their needs analyzed, many communication tools were developed and key activities were prepared. We capitalized on the achievements of the Council of Coaches project, securing the maintenance of stakeholders' interest and engagement, and encouraging participation. Furthermore, all partners were very active in conferences in total 9 participations with several publications until the Covid19 pandemic started affecting the usual activities of partners. We moved some events online and continued to use all online tools like Social Media, website and Newsletter to continue dissemination activities. We ended the phase with the Council of Coaches Workshop that attracted a lot of people from the virtual coaching research community. The final numbers on the dissemination activities are reported on in D8.5.

12.3.1.5 Achievements Period 2

The achievement of T8.1 are listed below:

- On-time submission of deliverables D8.4, D8.5, D8.10.
- Applied the communication strategy defined in the reporting period 1.
- Continuous content creation and social media updates, in line with the project's strategy.
- Submission and publication of scientific papers.
- Organization of the COUCH workshop with 65 attendees.

12.3.1.6 Deviations from DoA

There have been no significant deviations from the DoA.

12.3.2 T8.2: Exploitation and Business Planning

12.3.2.1 Objectives

This task has the goal to align the exploitation plans of the partners towards an overall exploitation plan of the project. The execution of the overall exploitation plan of the project will be managed in this task ensuring exploitation of COUCH results as early as possible. As part of T8.2, the project's exploitable assets are specified, along with the exploitation modalities that each partner will employ in order to benefit from these assets. An exploitation agreement regulating IPR shares and the partners' relationships in the scope of the foundation will be also prepared. As part of the exploitation plans, the partners will also specify business plans for specific exploitable assets, including market analysis and financial analysis.

12.3.2.2 Progress (M1-M18)

Even in the first year of the project, COUCH has identified key exploitable outcomes, as well as a strategy towards standardization and exploitation.

The project has organized a **workshop on "Innovation uptake in eHealth with patient-centeredness and gamification. Regulatory challenges and opportunities"**. The workshop was held on 7th November 2018 with the scope to brainstorm the status of the regulatory and ethical scenario in the context of eHealth and expose the state of the art of COUCH, jointly with two other EC projects. Experts from the industry and the regulatory environment discussed the process-to-market implications especially at the light of the Regulatory framework in the clinical research and uptake in the healthcare scenario.

12.3.2.3 Achievements (M1-M18)

The achievement of T8.2 are listed below:

- Submission of deliverables D8.6 and D8.7
- Inventory of the possible exploitable outcomes of the project has been performed.
- Iterative and systematic approach of the COUCH business model definition
- Workshop on regulatory challenges and Opportunities organised.

12.3.2.4 Progress (M19-M36)

The Council of Coaches project has defined an ambitious exploitation strategy with main and secondary outcomes. In the 2nd reporting period, the initial plans have been complemented with outcomes from the Innovation Managements workshops and were included in the exploitation deliverable D8.8. The project has decided to go for a 2-fold exploitation strategy, one focusing on the Open Agent Platform (Agents United Platform) and one on the COUCH product, targeting DTx services.

12.3.2.5 Achievements (M19-M36)

The achievement of T8.2 are listed below:

- Finalized the inventory of the primary and secondary exploitable outcomes of the project.
- Defined exploitation strategy for the two paths that the project identified.
- Submission of deliverable D8.8.

12.3.2.6 Deviations from DoA

Deliverable D8.6 was submitted on M15, in order to include tangible outcomes of the exploitation and standardization efforts. The late submission has been agreed with the EC, in order to include more information that will be the basis for exploitation in the next months of the project.

12.3.3 T8.3: Ecosystem Building of the Open Agent Platform

12.3.3.1 Objectives

Linked to the dissemination and marketing activities of the previous tasks, this task will focus on building a network of interested parties for the Open Agent Platform. This network will include developers, vendors, service providers and other stakeholders.

12.3.3.2 Progress (M1-M18)

The task has not started yet.

12.3.3.3 Achievements (M1-M18)

This task has not started yet.

12.3.3.4 Progress (M19-M36)

The Open Agent platform has been relabelled as Agents United Platform. It has been made publicly available on GitHub. The platform will be adopted by the Personalized eHealth Technology programme at the UT. An MoU has been signed between all partners who want to take part in the Agents United Alliance. A Postdoc will be hired from Sept 2020 on, to continue work on the ecosystem, the technical development and the organisation of the platform. A website has been setup, a newsletter is in the works and the platform has been announced through the COUCH newsletter. A press release bringing attention to the platform will be published in September 2020 and so far, we have received our first request from an external party to join the Agents United Alliance already.

12.3.3.5 Achievements (M19-M36)

- Agents United Platform has been made publicly available on github.
- The Agents United Alliance has been setup and a MoU has been signed.
- The platform has been adopted to be continued after the project has ended and staff has been hired.
- Communication measures have been set up to extend the community after the end of the COUCH project.
- The first external researcher has contacted us that they wanted to join the Alliance and work with the platform.

12.3.3.6 Deviations from DoA

There have been no significant deviations from the DoA.

12.3.4 T8.4: Standardization Activities

12.3.4.1 Objectives

This task will focus on the project's contribution to standards. The goal of this task is to monitor standardization/pre-standardization activities that are linked to the activities of the project and ensure the compliance of the project results with existing standards or drafts. The task will work on innovation management through standardisation, by monitoring closely related standardisation activities and identifying gaps and missing links in current standards/drafts (especially related with eHealth/eCare solutions) and provide feedback to WP3, WP4, WP5 and WP6 for pursuing innovative research to address these shortages.

12.3.4.2 Progress (M1-M18)

The progress in this task has been reported on in D8.6 and D8.7, the initial and intermediate plans for standardization. D8.6 has been delayed since the initial plans for exploitation that are described in the project were not completely developed at such an early stage in the project. Therefore, we chose to submit the deliverable in a later stage, when ideas have become more tangible.

For the standardization activities, an initial overview has been made that describes the different standards used in the project. Possible contributions to these standards have been listed in the initial plan. An update is provided in D8.7, however, since D8.6 was significantly delayed, this will not be a very significant update.

12.3.4.3 Achievements (M1-M18)

- An inventory set up of the different standards used in the project.

12.3.4.4 Progress (M19-M36)

The progress in this task has been reported on in D8.8, the final report on standardisation. Throughout the project we have made extensive use of standards and have made contributions and extensions to the standards where applicable. This ensures that our work is in a good position for further integration and interoperability with existing platforms and tools. In addition to using and contributing to existing standards, the WOOL standard was developed. It consists of language definitions, tools, documentation and tutorials for specifying and running virtual agent conversations. WOOL is described in more detail in D8.8 and D8.9. For the universal standard, we have developed a connector that turns the Technical Demonstrator into a universAAL-compliant app: provides service to start a coaching session through a specifically developed (basic) coaching ontology. Further contributions to standards such as the AWARE platform, ASAP, SAIBA, BML, FML and DGEP are described in D8.8

12.3.4.5 Achievements (M19-M36)

- Developed a completely new standard: the WOOL standard.
- Finalised the inventory of the different standards used and contributions to these standards made in the project.
- Developed connector to make the functional demonstrator a universal compliant app.

12.3.4.6 Deviations from DoA

There have been no significant deviations from the DoA.

12.3.5 T8.5: Training activities

12.3.5.1 Objectives

This task is devoted to raising awareness of the Open Agent Platform, including virtual coaching training in general and in the projects' solutions in particular. Relevant training materials will be developed, which will be used in the scope of both on-line training and face-to-face training of different stakeholders. As part of the task the project will develop online training resources (e.g., webinars), while at the same planning face-to-face training workshops.

12.3.5.2 Progress (M1-M18)

This task has not started yet.

12.3.5.3 Achievements (M1-M18)

This task has not started yet.

12.3.5.4 Progress (M19-M36)

In the last year of the project, many of the software platforms have reached advanced levels of maturity. Since one of the core objectives of Council of Coaches is to contribute to Open Innovation by making results of research as well as tools openly available, this task has focused on making those open results more accessible to the public. A lot of efforts in the final year of the project have been put into setting the foundations for the Agents United platform to be successful, and the core tasks of providing great documentation will be ongoing after the project's end, supported by the Agents United Alliance.

12.3.5.5 Achievements (M19-M36)

- Created detailed documentation and guides for the various GitHub Repositories of the Agents United Platform.
- Created detailed language documentation and first set of beginner tutorials for the WOOL Platform, populating the dedicated website for this platform component at www.woolplatform.eu.
- Created course material for training students in HCI (Interaction Technology) in using the Council of Coaches platforms in their own course assignments.
- Created an inventory and overview of all educational activities (student supervision) for the Council of Coaches website.
- Documented all activities in D8.9: Training Material (submitted on time).

12.3.5.6 Deviations from DoA

Focus has shifted slightly towards online training material (as opposed to organizing face-to-face sessions). There is also a greater focus on technical training (as opposed to training of healthcare professionals), which is simply put much more relevant given the outcomes of the project.

13 WP9: Ethics

This automatically injected work package has been fully executed during the first project period, and is included here for completeness sake.

13.1 Planning

WP / Task	Title	Participants	Months
WP9	Ethics	CMC	M1-M12
T9.1	Human	CMC, RRD, DBT, SU, UDun, UPV, iSPRINT	M1-M12
T9.2	Protection of Personal Data	CMC, RRD, DBT, SU, UDun, UPV, iSPRINT	M1-M12

13.2 Main Achievements

The following main achievements result from this work package:

- Creation of awareness of ethical issues in the council of Coaches project.
- Council of Coaches guidelines on how to handle personal data have been created
- Privacy Impact Analyses has been performed
- Templates for informed consent forms have been created
- Design of the informed consent procedure.

13.3 Individual Tasks

13.3.1 T9.1: H- Requirements No. 1

13.3.1.1 Objectives

The objectives as described in the DoA are:

- Details on the procedures and criteria that will be used to identify/recruit research participants must be provided.
- Detailed information must be provided on the informed consent procedures that will be implemented for the participation of humans.
- Templates of the informed consent forms and information sheet must be submitted on request.
- The applicant must clarify whether vulnerable individuals/groups will be involved. Details must be provided about the measures taken to prevent the risk of enhancing vulnerability/stigmatisation of individuals/groups.
- Copies of ethics approvals for the research with humans must be submitted.

13.3.1.2 Progress (M1-M18)

In some projects, the required ethics deliverables are treated as just another checkbox on a list. Once the content of the deliverable has been created, one can write it down and continue as before. However, in the Council of Coaches project, one of the aims of the project management has been to create awareness within the project about the ethical aspects of the project. Privacy and the vulnerability of our end users should be respected and taken into account when designing and testing our prototypes.

Incorporating these topics in the project has been a bit of a balancing act, especially because in WP 2, an RRI vision has already been created. Fortunately, as the project and the RRI vision came to ripen, it became clear that the deliverables as described in WP 9 were a nice start to several RRI discussions. In return, the awareness created through the RRI vision workshops was helpful in collecting the data necessary for the deliverables.

For the deliverable about studies involving human subjects, we asked everyone in the project to describe their test setups during the bi-weekly telco's. This way the entire team could provide feedback and, more importantly, learn from each other in how to set up a study involving humans the correct way. For the deliverable, each team that has set up studies were asked to submit their information leaflets and consent forms. These were checked by the project management before being used in the testing.

If necessary, the partner that conducted the tests was asked to get ethical approval from their local ethical boards. In some cases, this has been granted, but in others it was not deemed necessary by the ethical boards themselves.

13.3.1.3 Achievements (M1-M18)

Awareness has been created among the researchers in Council of Coaches on how to handle ethical issues in the project. An overview of all studies involving human subjects has been provided, including information letters, consent forms, and, if possible, ethical approvals.

13.3.1.4 Progress (M19-M36).

This task was completed during period 1.

13.3.1.5 Achievements (M19-M36)

This task was completed during period 1.

13.3.1.6 Deviations from DoA

There have been no significant deviations from the DoA.

13.3.2 T9.2: POPD- Requirements No. 2

13.3.2.1 Objectives

The objectives as described in the DoA are:

- Copies of opinion or confirmation by the competent Institutional Data Protection Officer and/or authorization or notification by the National Data Protection Authority must be submitted (which ever applies according to the Data Protection Directive (EC Directive 95/46, currently under revision, and the national law).
- If the position of a Data Protection Officer is established, their opinion/confirmation that all data collection and processing will be carried according to EU and national legislation, should be submitted.
- Detailed information must be provided on the procedures that will be implemented for data collection, storage, protection, retention and destruction and confirmation that they comply with national and EU legislation.

13.3.2.2 Progress (M1-M18)

The first year of the project has been interrupted by the introduction of the GDPR on May 25th, 2018. It was our conscious decision to follow the rules and regulations as stated by the GDPR even before this law was enforced. This way we would create a uniform approach to personal data over time as well as over partners, since with the introduction of the law, all partners in the project would be submitted to it anyway.

However, the GDPR is no easy feat. Especially, with the collection of medical data as we are foreseeing in this project, we need to be really careful on how to handle these data. As project management we therefore translated the GDPR into guidelines within the project on how to handle personal data as well as possible. A privacy impact analyses has been conducted to see which aspects could be improved even further.

A "cheat sheet personal data" has been created, for everyone to hang above their desks. It is meant as a quick reference guide that researchers can easily use in their day to day life. This way, we have tried to give the GDPR as much attention as possible in the day to day practice. The deliverable in which the guidelines were discussed in more detail has been sent to all researchers to implement during their work. During the consortium meetings, several presentations have been held specifying the GDPR in

general and the specific project guidelines in particular. This way we could discuss specific topics, take away worries and educated everyone on how to handle personal data properly.

Together with the RRI researchers, through the RRI discussion in which privacy was and is a big issue, we worked on incorporating as much of the privacy aspects in the prototypes as possible. This has resulted in a clear idea on how to inform the end users about their rights as data subjects. This approach has been described in section 2.3.3 about the consent dialogue.

Each partner has had their approach to personal data (the general Council of Coaches approach) validated by their DPO, if these were appointed.

13.3.2.3 Achievements (M1-M18)

We have created awareness on GDPR, personal data handling and privacy. Dedicated Council of Coaches Guidelines have been laid out in the deliverable and been printed in a "Personal Data Cheat Sheet". DPO approvals have been received. A solution towards the informed consent procedure has been developed.

13.3.2.4 Progress (M19-M36)

This task was completed during period 1.

13.3.2.5 Achievements (M19-M36)

This task was completed during period 1.

13.3.2.6 Deviations from DoA

We have gotten in contact with the PO since some of the descriptions in the DoA were not quite clear to us. As discussed with the PO, point 4.1 as specified in the objectives could be disregarded. *All partners should seek approval from their local DPO's. In cases where that was not possible, expert advice should be obtained.*

Some of the partners did not have a DPO set up by the time we required the approvals. However, since all partners submitted the same set of information (description of the guidelines in the deliverables and the cheat sheet), we assumed getting extra expert advice would not be necessary, if the other DPO's had already given their approval.

14 Bibliography

- Isbister, K. (2006). *Better Game Characters by Design: A Psychological Approach (The Morgan Kaufmann Series in Interactive 3D Technology)*. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc.
- Sloan, R. J. (2015). *Virtual Character Design for Games and Interactive Media*. London, UK: CRC Press Taylor & Francis Group.
- Kolkmeier, J., Bruijnes, M., Reidsma, D., & Heylen, D. (2017). An ASAP realizer-Unity3D bridge for virtual and mixed reality applications. *International Conference on Intelligent Virtual Agents*. (pp. 227-230). Springer.
- Ravenet, B., Cafaro, A., Biancardi, B., Ochs, M., & Pelachaud, C. (2015). Conversational behavior reflecting interpersonal attitudes in small group interactions. *International Conference on Intelligent Virtual Agents*. (pp. 375-388). Springer.
- van Waterschoot, J., Bruijnes, M., Flokstra, J., Reidsma, D., Davison, D., Theune, M., & Heylen, D. (2018). Flipper 2.0: A Pragmatic Dialogue Engine for Embodied Conversational Agents. *International Conference on Intelligent Virtual Agents* (pp. 43-50). ACM Press.
- Klaassen, R., Hendrix, J., Reidsma, D., op de Akker, H. J., van Dijk, E. M., & op den Akker, H. (2013). Elckerlyc Goes Mobile Enabling Natural Interaction in Mobile User Interfaces. *International Journal on Advances in Telecommunications*, 6(1-2), 45-56.
- Huber, M., van Vliet, M., Giezenberg, M., Winkens, B., Heerkens, Y., Dagnelie, P., & Knottnerus, J. (2016). Towards a "patient-centred" operationalisation of the new dynamic concept of health: a mixed methods study. *BMJ Open*, 6(1), e010091.
- Hurmuz, M., Jansen-Kosterink, S., op den Akker, H., & Hermens, H. (2020). User Experience and Potential Health Effects of a Conversational Agent-Based Electronic Health Intervention: Protocol for an Observational Cohort Study. *JMIR Research Protocols*, 9(4), e16641.
- ter Stal, S., Tabak, M., op den Akker, H., Beinema, T., & Hermens, H. (2020). Who Do You Prefer? The Effect of Age, Gender and Role on Users' First Impressions of Embodied Conversational Agents in eHealth. *International Journal of Human-Computer Interaction*, 36(9), 881-892.

Acknowledgements



The Council of Coaches project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement #769553. This result only reflects the author's view and the EU is not responsible for any use that may be made of the information it contains.

Headings and titles in this document, as well as the Council of Coaches logo use the Comfortaa font, designed by Johan Aakerlund and Cyreal and licensed under the Open Font License¹.

Additional text in this document uses the Roboto font, designed by Christian Robertson and licensed under the Apache License, Version 2.0².

The Council of Coaches logo and Blobmen graphics were *drawn freely* in Inkscape, licensed under the GNU General Public License³.

¹ Open Font License: http://scripts.sil.org/cms/scripts/page.php?site_id=nrsi&id=OFL_web

² Apache License, Version 2.0: <http://www.apache.org/licenses/LICENSE-2.0>

³ Inkscape License Information: <https://inkscape.org/about/license/>