



HAL
open science

Empowering well-being through conversational coaching for active and healthy ageing

Michael Mctear, Kristiina Jokinen, Gérard Chollet, Christophe Lohr, Jérôme Boudy, Mohnish Dubey, Sonja Dana Roelen, Wanja Mössing, Rainer Wieching

► To cite this version:

Michael Mctear, Kristiina Jokinen, Gérard Chollet, Christophe Lohr, Jérôme Boudy, et al.. Empowering well-being through conversational coaching for active and healthy ageing. International Conference on Smart Homes and Health Telematics (ICOST), Jun 2022, PARIS, France. 10.1007/978-3-031-09593-1_21 . hal-03955899

HAL Id: hal-03955899

<https://hal.science/hal-03955899>

Submitted on 25 Jan 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Empowering well-being through conversational coaching for active and healthy ageing

Michael McTear¹, Kristiina Jokinen², Mohnish Dubey³, Gérard Chollet⁴, Jérôme Boudy⁴, Christophe Lohr⁴, Sonja Dana Roelen⁵, Wanja Mössing⁵, Rainer Wieching⁶

¹Ulster University, UK, ²Artificial Intelligence Research Center, National Institute of Advanced Industrial Science and Technology, (AIRC/AIST), Japan, ³Institute für Angewandte Informatik, Germany (INFAI), Germany, ⁴Institut Mines-Télécom (IMT), France, ⁵Institut für Experimentelle Psychophysiologie GmbH (IXP), Germany, ⁶Universität Siegen (USI)
{mf.mctear@ulster.ac.uk, kristiina.jokinen@aist.go.jp, dubey@infai.org, gerard.chollet, jerome.boudy@telecom-sudparis.eu, christophe.lohr@imt-atlantique.fr, s.roelen, w.moessing@ixp-duesseldorf.de, rainer.wieching@uni-siegen.de}

Abstract. With life expectancy growing rapidly over the past century, societies are being increasingly faced with a need to find smart living solutions for elderly care and active ageing. The e-VITA project, which is a joint European (H2020) and Japanese (MIC) funded project, is based on an innovative approach to virtual coaching that addresses the crucial domains of active and healthy ageing. In this paper we describe the role of spoken dialogue technology in the project. Requirements for the virtual coach were elicited through a process of participatory design in workshops, focus groups, and living labs, and a number of use cases were identified for development using the open-source RASA framework. Knowledge Graphs are used as a shared representation within the system, enabling an integration of multimodal data, context, and domain knowledge.

Keywords: active and healthy ageing; dialogue system; knowledge graphs; participatory design

1 Introduction

With life expectancy growing rapidly over the past century, societies are being increasingly faced with new challenges related to an ageing population and the need to find smart living solutions for elderly care and active ageing [1]. This paper presents an innovative approach involving virtual coaching that is currently being developed in a 3 year joint European (H2020) and Japanese (MIC) funded research project^{1 2}. The virtual coach addresses the crucial domains of active and healthy ageing in cognition, physical activity, mobility, mood, social interaction, leisure, and spirituality, with the

¹ <https://www.e-vita.coach/>

² The e-VITA project has received funding from the European Union H2020 Programme under grant agreement no. 101016453. The Japanese consortium received funding from the Japanese Ministry of Internal Affairs and Communication (MIC), Grant no. JPJ000595.

aim of empowering older adults to better manage their own health and daily activities, resulting in improved well-being and improved stakeholder collaboration. The virtual coach provides individualized profiling and personalized recommendations based on big data analytics and social-emotional computing, detecting risks in the user's daily living environment by collecting data from external sources and non-intrusive sensors and providing support through natural interactions with 3D-holograms, emotional objects, or robotic technologies using multimodal and spoken dialogue technology, advanced knowledge graph representations, and data fusion. [2] provides a general overview of the project.

The remainder of the paper is structured as follows. Section 2 reviews related work in the area of active and healthy ageing, with particular reference to the use of dialogue systems to provide a virtual coaching application. Section 3 presents the technologies being employed in the e-VITA virtual coach, looking at multimodal data fusion from sensors, emotion detection, knowledge graphs, and dialogue technology. Section 4 describes the development of an initial prototype system, including a process of participatory design resulting in use cases and content development, some of which has been implemented in the Wave 1 prototype. Section 5 concludes by outlining the next steps in the project.

2 Related Work

There has been increasing interest in recent years in exploring how new developments in conversational AI and socially assistive robots can be applied to support active and healthy ageing. [3] presents a comprehensive state-of-the-art review in which several future research directions are identified, including: the need for unconstrained natural language processing and conversational strategies to enable robust and meaningful two-way conversations; the ability to interpret affective modalities in order to enhance user engagement and trust; and challenges in deployment of ensuring user adherence and data privacy. Looking at some individual relevant projects, [4] reports a month-long study of a virtual agent and robot that could interact with older adults in their homes through dialogue and gestures with the aim of providing companionship and reducing isolation. The importance of dialogue capability to enable social robot agents to provide natural interaction is also emphasized in [5], while [6] describes how information about the older adult's emotional status was extracted from an analysis of their verbal and non-verbal communication. [7] is an example of an ongoing project in the Netherlands with the aim of improving the lives of older adults through the use of voice technology. The e-VITA project also focusses on the issues addressed in these examples of related work but in addition extends the research to examine how data from sensors and emotion analysis can be used within dialogues with the older users, and how knowledge graphs can be used to store data relevant to active healthy ageing as well as personal data about the user to enable the virtual coach to offer personalized information and recommendations.

3 Technologies and Architecture of the Virtual Coach

The complete e-VITA platform is based on the Digital Enabler platform provided by the project partner Engineering (Italy)³. The platform is designed to include multiple devices and software components from across Europe and Japan that are based on different technologies and standards. The platform supports communication and integration among different smart devices such as sensors and robots, as well as the collection and management of data to provide coaching functionalities.

3.1 Multimodal Data Fusion from Sensors

In a smart home, sensors can be used to capture and monitor data in order to provide the user with assisted, safe and comfortable living. In the e-VITA project three types of sensors are used: holter-based sensors worn by the user (or possibly his/her smartphone) that sense physiological and actimetric parameters; environmental sensors that measure physical data to assess the level of comfort and the quality of the indoor environment; and home-based sensors to monitor user behaviour and activities. Data from these sensors is combined (or fused) to make inferences about users in their environment, for instance, postures, localisation in the home, and the users' physiological states. Contextual information extracted from these sensors can be stored in knowledge graphs and exploited by the interactive voice-based coaching system.

Various challenges being addressed include interoperability between the sensors involving a combination of multiple different data that may operate at different sampling frequencies and with different representation formats or scales for the targeted e-VITA applications, e.g., physical exercises, activities of daily living (ADL), or fall prevention [8]. Another goal is to perform Data Minimization for transmission of information into the Cloud.

To date software modules have been implemented that deliver data fusion to assess the user's situation (location, activity, vital state) and environmental conditions (temperature, humidity). The delivered data fusion software uses the Digital Enabler framework, aiming to provide semantical situation and environment labels that can be used by the knowledge graphs and the dialogue system. The Data Fusion algorithms constitute a first step for processing heterogeneous sensor data and signals in a multimodal way. Future work will address other modalities in order to obtain a complete set of data about the user's situation and environment.

3.2 Emotion Detection System

A range of emotional cues can be detected in acoustic parameters in the user's voice, including parameters in the Frequency domain (e.g., fundamental frequency, jitter, or pitch), in the Time domain (e.g., speech rate, speech pauses and syllable rates), in the Amplitude domain (e.g., intensity or energy), and in the Spectral Energy domain (e.g., relative energy in different frequency bands) [9]. The Emotion Detection System (EDS)

³ <https://www.eng.it/en/our-platforms-solutions/digital-enabler>

that is employed in the e-VITA platform enables the detection and classification of frequently used basic emotions such as fear, anger, disgust, happiness, surprise, and sadness during interactions between the coaching system and older adults [10]. The detected emotions can then be used by the dialogue system to provide appropriate interventions [11].

The exact classes detected by the e-VITA EDS may vary between the target languages in the e-VITA consortium depending on the availability of training data. In the current release, the German variant of the EDS can detect anger, disgust, fear, joy, neutral, and sadness. The Japanese variant can detect anger, contempt, disgust, fear, joy, neutral, sadness, surprise, and trust.

Within the e-VITA project, the EDS may need to be executed in a range of environments depending on upload bandwidth, data minimization, privacy concerns, and prototype environments. Thus the EDS may be implemented as a component within the Digital Enabler cloud environment, but could alternatively also be implemented on a local edge-computing device with yet-to-be-defined hard-/software specifications.

3.3 Knowledge Graphs

Knowledge Graphs are used to represent the knowledge that is used by intelligent AI systems [12]. A Knowledge Graph (KG) stores data about real-world things, events, and concepts. For example, an entity in a KG contains a semantic description and many characteristics that relate it to other entities, concepts, or events. Entities in a KG are represented as nodes and the properties of the entities are represented as edges in the graph that relate the nodes, resulting in a network of enhanced knowledge for a specific genre or topic.

In the current project the KG stores data about the user and the environment that is relevant to active healthy ageing and that enables the virtual coach to offer personalized information and recommendations. A localized KG stores personal information and is connected to the central KG. The central KG stores information required for various functions that are to be used across the platform. The central KG is connected to a database that stores procedural details that cannot be harnessed through KG triples. The local KG is on-premise and keeps the personal data safe, while being able to offer assistance and a personalized user experience to an individual user. For instance, the fact that a user is male is stored in the local KG enabling the entire system to recommend activities suitable for male users. Fig. 1 presents an excerpt of the graph that describes the prevention domain and the medical examinations that a male or female can take.

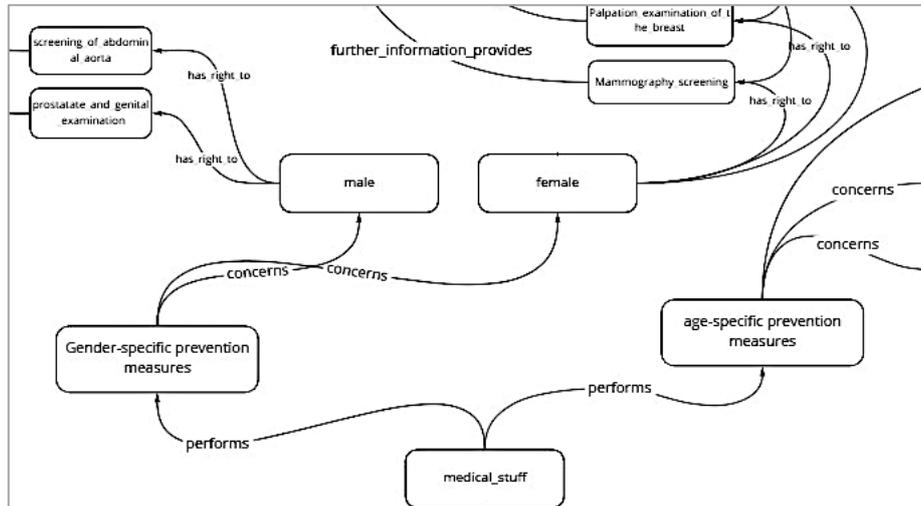


Fig. 1. Excerpt from the knowledge graph for the medical prevention domain

3.4 Dialogue System

The Dialogue System is the key component that enables interactions with the user. Its functions include understanding the user’s input messages and providing an appropriate responses in the given context (including the previous dialogue context and environmental state) [13]. In the e-VITA project the Dialogue Manager is implemented using the RASA Open Source Framework that will be integrated with speech technology components to enable spoken interaction. RASA supports a machine learning-based Natural Language Understanding pipeline for the interpretation of the user’s inputs and a combination of rule-based and machine learning-based dialogue policies to determine the system’s actions [14]. Knowledge Graph technology is integrated via specific knowledge-based actions supported by RASA to supply content for the system responses by querying the domain knowledge base [15].

4 Prototype Virtual Coach

A prototype version of the virtual coach was delivered at month 15 of the project. This section outlines how the requirements for the prototype were gathered and analysed, followed by a discussion of the use cases addressed in the Wave 1 prototype and an example of an interaction with the virtual coach.

4.1 Requirements Gathering and Analysis

Requirements for the virtual coach were gathered in interviews with older community-dwelling adults in Germany, Italy, and Japan, all of whom described themselves as regular users of smartphones and personal computers. In the first phase users were invited to living labs, while in the second phase studies were conducted in the homes of users. One of the studies conducted in Germany involved interactions with a nutrition chatbot and with the Nao robot in which a range of scenarios was explored including reminder functions, news, stories, and jokes, and general companionship. The end-user studies in Japan included interactions with real devices in Living Labs. One of these studies investigated users interacting with the Nao robot and a RASA dialogue system that included a general conversation of 10 minutes with the coach about daily living and a further 10 minute conversation about food and general question-answering about the news. The users were instructed to engage in several scenarios, for example:

- *Tell that you feel sad* – the user expresses sad emotions and the coach aims to provide empathic and consoling companionship.
- *Tell that you want some exercise* – the user tells the coach that he/she wants some exercise. The coach records the user’s exercise preferences and provides options that can help improve the user’s health and physical condition.

Results from the studies included comments that the users would like to be able to engage in longer conversations and not just receive short answers to their questions. In some cases, the system received user responses that were not included in the original design and that would have to be included in the machine-learning models of the next version of the system. Also required was the implementation of information about the user’s daily routines, events in the user’s calendar, and user preferences about topics such as music, and reminders.

4.2 Interactions with the Prototype Dialogue System

Based on the interviews with users the following use cases were identified for the prototype system to be developed and evaluated in Wave 1 of the project:

- Daily support;
- Health activity support;
- Environmental monitoring support;
- Question-Answering over Wikipedia, News;
- Social Activity support.

The following interaction is based on the RASA story represented in Table 1 which is activated following the recognition of the user’s utterance “what preventive examination should I have performed by my medical doctor” as the intent **ask_examination**. The system consults the knowledge graph shown in Fig. 1 and as the system does not know the user’s gender in this use case, it asks the user to which the user replies “male”. The system follows the **has_right-to** link from the **male** node in the knowledge graph, retrieves the required information, and outputs it to the user. A screenshot of the dialogue produced by this story is shown in Fig. 2.

Table 1. RASA story: preventive examinations.

- story: preventive examinations, the male user would like to know what preventive measures he can receive
steps:
- intent: ask_examinations
- slot_was_set: - user_gender: null
- action: utter_ask_gender
- intent: say_gender
- action: action_examination_KG #KG data
- action: utter_suggest_examination

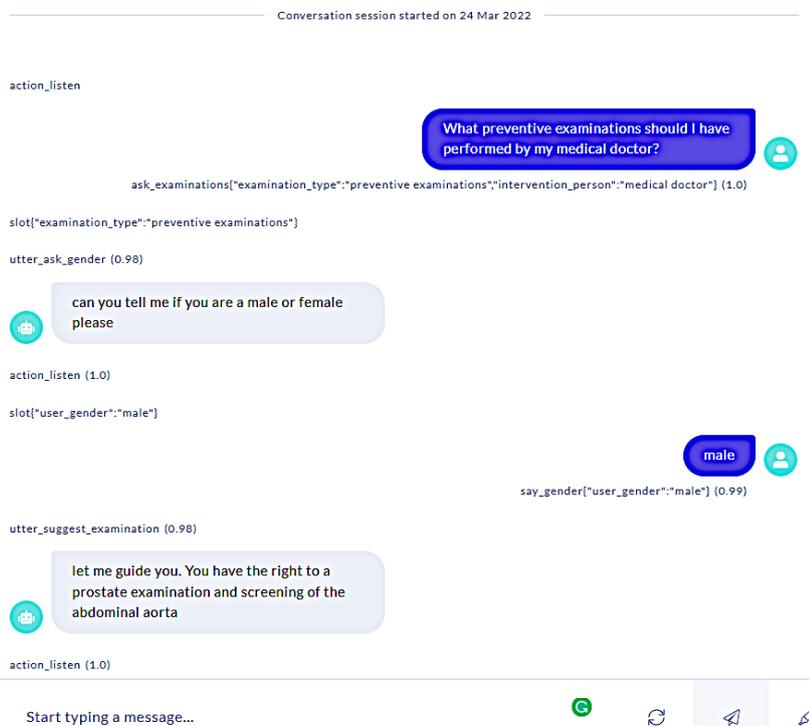


Fig. 2. Screenshot of dialogue in the medical prevention domain

5 Next Steps

In the next phase of the project the dialogue system will be tested and evaluated in the living labs and in the homes of the older adults. For the home-based trials the devices will be prepared by the study centres and installed in the homes of the users. A diary and end-user manuals will be provided and the system will be explained and

demonstrated. A human coach from the community will also visit once a week during the trials which will last for one month.

Future technical work will involve more complex knowledge modelling and knowledge graph creation and integration of the dialogue system with new knowledge graphs, sensor data, and data from the emotion detection system.

References

1. Spasova, S., Baeten, R., Coster, S., Ghailani, D., Peña-Casas, R., Vanhercke, B. Challenges in long-term care in Europe A study of national policies 2018, <https://www.age-platform.eu/publications/challenges-long-term-care-europe-study-national-policies-2018> (2018)
2. Jokinen, K., Homma, K., Matsumoto, Y., Fukuda, K. Integration and Interaction of Trustworthy AI in a Virtual Coach – An Overview of EU-Japan Collaboration on Eldercare. In: Artificial Intelligence. JSAI 2021. Advances in Intelligent Systems and Computing, vol 1423. Springer, Cham. (2022).
3. Lima, M.R., Wairagkar, M., Gupta, M., Rodriguez y Baena, F., Barnaghi, P., Sharp, D.J., Vaidyanathan, R. Conversational Affective Social Robots for Ageing and Dementia Support. *IEEE Trans. Cogn. Dev. Syst.*, (99):1-1 (2021)
4. Sidner, C.L., Bickmore, T., Nooraie, B., Rich, C., Ring, L., Shayganfar, M., Vardoulakis, L. Creating new technologies for companionable agents to support isolated older adults. *ACM Trans. Interact. Intell. Syst.* 1, 1, Article 1 (2017)
5. Jokinen, K. (2018). Dialogue Models for Socially Intelligent Robots. In: Social Robotics. ICSR 2018. Lecture Notes in Computer Science, vol. 11357. Springer, Cham., 127-138 (2018)
6. Justo, R., Ben Letaifa, L., Palmero, C., Gonzalez-Fraile, E., Torp Johansen, A., Vázquez, A., Cordasco, G., Schlögl, S., Fernandez-Ruanova, B., Silva, M., Escalera, S., deVelasco, M., Tenorio-Laranga, J., Esposito, A., Korsnes, M., Torres, M.I. Analysis of the interaction between elderly people and a simulated virtual coach. *J Ambient Intell Humaniz Comput*, 11:6125–6140 (2020)
7. Project Zilver: Research and expertise to develop voice technology for older adults. <https://projectzilver.com/en/>
8. Hussain, G, Jabbar, MS, Cho J-D, Bae S. Indoor Positioning System: A New Approach Based on LSTM and Two Stage Activity Classification. *Electronics* 8(4):375 (2019)
9. Patel, S., Scherer, K. R. Vocal Behavior. In: Hall, J. A., Knapp, M. L. (eds.), *Nonverbal Communication*, pp. 167–204. De Gruyter Mouton (2013)
10. Ekman, P. An argument for basic emotions. *Cogn. Emot.* 6(3-4), 169–200 (1992)
11. Picard, R. *Affective Computing*. MIT Press, Cambridge, Mass. (2000)
12. Ji, S., Pan, S., Cambria, E., Marttinen, P., Philip, S. Y. A survey on knowledge graphs: Representation, acquisition, and applications. *IEEE Trans Neural Netw Learn Syst* (2021)
13. McTear, M. *Conversational AI: Dialogue Systems, Conversational Agents, and Chatbots*. Morgan & Claypool Publishers, Kentfield, CA. (2020)
14. Bocklisch, T., Faulkner, J., Pawlowski, N., & Nichol, A. Rasa: Open source language understanding and dialogue management. arXiv preprint arXiv:1712.05181 (2017)
15. Wilcock, G., Jokinen, K. Conversational AI and Knowledge Graphs for Social Robot Interaction. Late-Breaking Reports. The 17th ACM/IEEE International Conference on Human-Robot Interaction (HRI-2022) (2022)