



HAAL

AAL-2020-7-229-CP

Responsible Innovation Report
[Deliverable D1.7]

Document Information	
Workpackage	WP1
Leading partner	Vilans
Due Date	M12
Type	Report
Status	V1.0
Security	PU - Public
Project URL	https://www.haal-aal.com

Versioning and contribution history

Version	Description	Comments
1.0	Dirk Lukkien + check by Henk Herman Nap	First version

Deliverable process schedule

No	Process step	Responsible	Timing	Comments
1.0	First draft of deliverable	Dirk Lukkien	15-07-2022	First version

Table of Contents

Executive summary.....	4
1. Introduction	5
1.1 Purpose of the document	5
1.2 Document structure	5
1.3 Contribution of partners	5
1.4 Relations to other activities in the project	6
2. About HAAL.....	7
2.1 HAAL project	7
2.2 Focus on development of a dashboard	8
3. About Responsible Innovation.....	10
3.1 Relevance of Responsible Innovation in HAAL	10
3.2 Background on Responsible Innovation	11
4. Plan of action.....	13
4.1 Overall research approach	13
4.2 Chronological steps in the RI research	15
5. Provisional results	18
References.....	20
Appendixes.....	22
Appendix 1: Internal HAAL survey on Responsible Innovation	22
Appendix 2: Protocol for RI interviews with primary end-users (i.e. care professionals).....	29

Executive summary

Health(care) technologies will not simply be used on a largescale base by and for older citizens and their caregivers just because they address relevant health(care) related issues and fit certain needs. It is important that technologies fit within the dynamic context they are implemented in and that they do not interfere with what a large group of stakeholders, older people included, find important when it comes to the implementation of digital technologies in everyday life and practices. Therefore, in the HAAL project, research is being conducted to explore both the positive (and desirable) and the negative (unintended) impacts of the HAAL platform (in particular the AI-based dashboard) and what can be done to achieve a more meaningful and responsible innovation.

Our methodology thus far included workshops during consortium meetings, a survey amongst all project partners, and interviews with end-users. Our goal is to explore what factors can support the socially desirable, ethically acceptable and sustainable design and use of the AI-based HAAL dashboard as a decision aid of professional caregivers in the homecare for people with dementia. In other words, we explore how to innovate responsibly with AI within the context of the HAAL project and beyond.

In the next (and already running) phases of the HAAL project, we will more extensively engage with the perspectives of end-users on the positive and potential negative impacts of using the HAAL dashboard. The insights from this phase will be fed back to the consortium to discuss (again in workshops) what can be ultimately done within the context of the HAAL project (in the design, development, implementation and marketing of the HAAL dashboard) to address RI and account to what end-users and others stakeholders find important to the implementation of supportive technologies in everyday life and practices.

1. Introduction

1.1 Purpose of the document

Within the HAAL project, responsible innovation (RI) research is being conducted to learn how to achieve socially desirable, ethically acceptable and sustainable (implementation of) innovation in HAAL. The research explores the perspectives of different stakeholders, and at the same time guides these stakeholders (both project partners and end-users) to reflect on the different values at stake (e.g. privacy, autonomy, quality of life) and to facilitate a meaningful and responsible development and implementation of technology in the HAAL project and beyond. The central question to the RI research in HAAL is: *What factors can support the socially desirable, ethically acceptable and sustainable design and use of the AI-based HAAL dashboard as a decision aid of professional caregivers in the homecare for people with dementia?* This document reports on the research and its results. The document firstly serves as a (iterative) guideline and plan of action for research related to responsible innovation in the context of the HAAL project. Secondly, it is a running document where research results are collected.

1.2 Document structure

The present deliverable is divided in four parts (chapters):

Chapter 2 provides an introduction **about the HAAL project** and its main objectives.

Chapter 3 provides an introduction **about responsible innovation** and why it is important in the context of the HAAL project.

Chapter 4 presents the **plan of action**, including the overall approach taken to explore responsible innovation in the HAAL project and the various research steps taken to feed this exploration.

Chapter 5 reports **provisional results** that were gathered throughout the steps taken to explore responsible innovation.

This document is considered public and all the information reported is open.

1.3 Contribution of partners

All partners involved in the HAAL consortium are actively contributing to the RI activities during the entire project duration. In particular, the following are the tasks assigned to the partners:

Vilans, who is responsible for the current deliverable, is in the lead of the RI research, which amongst others includes studying the perspectives of both HAAL consortium partners and end-users from each country. Vilans takes the lead in gathering insights and ideas about RI from the perspective of HAAL consortium partners. All consortium partners contribute to this line of research by their active involvement in the internal research activities (e.g. workshops, surveys, etc.) organized by Vilans. Furthermore, Vilans (The Netherlands), The Municipality

of Aarhus (Denmark), INRCA (Italy), YZU, and NCKU (Taiwan) are contributing to the data collection with end-users in each country. Vilans prepares the materials (i.e. a project description, research protocols, brochures etc.) for the RI research and guides the partners involved from other countries. All consortium partners will be involved, to a greater or lesser extent, in the analysis and discussion of insights. Vilans will communicate the insights on RI in the HAAL context via the current Responsible Innovation Report, scientific publications and its own communication platforms and those of HAAL. All partners will also take care of the dissemination of the RI results and activities through already existing social networks of partners and individuals in the consortium (e.g.. LinkedIn and Twitter).

1.4 Relations to other activities in the project

An important stream of (research) activities in the HAAL project is focusing on co-design with end-users. The co-design activities (meaningful try-outs, focus group sessions and workshops, interviews, etc.) with end-users that are organized in each country, are focused on answering the question: *What can be the role of an (AI-driven) dashboard as a decision support of professional caregivers in the homecare for PwD, and could such a dashboard look like?* While the co-design focuses on the type of dashboard, including the functionalities and design, that is needed. The RI research (as described in this document) on the other hand focuses on *how to responsibly* design and use such an dashboard, i.e. how to ensure that the design and use of the dashboard is not only socially desirable, but also ethically acceptable and sustainable. In principle, the co-design (and related technical development) and RI research are two sides of the same coin. The RI research as described in this document is set up in parallel to the co-design activities. Yet, it will be continuously discussed amongst consortium partners involved in the co-design and RI research how these streams and results can be optimally integrated; what the one stream can learn from the other and how both streams can be most efficiently organized without doing things unnecessarily double. One way in which the streams are already integrated is by (Vilans) adding some RI-questions in the co-design protocols in all four countries.

2. About HAAL

This chapter firstly serves to remind us of the core ideas behind the HAAL project and the objectives we are bound to. It therefore firstly discusses the general scope of the HAAL project, and secondly discuss how the focus (for the RI research) is on the (responsible development of the) HAAL dashboard.

2.1 HAAL project

The world is facing similar challenges in respect to the ageing population and shortages of care personnel, with dementia as the largest and one of the most expensive challenges. Across Europe and beyond, various dementia friendly eco-systems are in place within municipalities with AAL services to empower clients, informal carers and care professionals. With the HAAL project, The Netherlands, Taiwan and Italy combine their strength, co-create, evaluate and share their experiences in supporting dementia care by state-of-the-art AAL bundles. HAAL project is a 30-month, international project, partially funded by the AAL Programme. Within HAAL the goal is to support people with dementia (PwD), their loved ones and care professionals¹ during the different stages of dementia, by using different types of care and well-being technologies researched in previous AAL projects. The functionalities and data of these HAAL devices are integrated and combined to create valuable insights. It is highly relevant to involve – iteratively – frail seniors, informal and formal carers in all design steps and continuously evaluate and reflect on the HAAL goals, and platform characteristics. We want to support formal and informal carers in their work, reduce the care burden and support the independence of PwD at home.

The HAAL platform includes eight digital care and wellbeing devices (HAAL devices)² developed by technical partners. The technologies are either on the market or in the final stage of development, and have yield positive results. The products were selected based on a diagram of product suitability throughout dementia stages.³ The initial assumptions on the HAAL devices that are suitable per dementia stage are shown in Figure 1.

¹ Care professionals or care professionals include all relevant care and support workers, such as nurses, district nurses, care assistants, case managers dementia, data analysts in care organizations, and alarm centralists.

² When referring to the HAAL devices we refer to separate AAL devices (including the developed functionalities separate from the HAAL project): Sensara, Whizpad, CogvisAI, Kompy Pico, Tessa, Tipr, Whiztoys, Medido and Compaan. When referring to the HAAL platform we refer to the whole HAAL ecosystem including the data from the different HAAL devices and the corresponding HAAL dashboard.

³ Van der Leeuw, J., Cornelisse C.C., Suikerbuijk, S. & Nap, H. H. (n.d.). Infographic: Technology for People with Dementia. <https://www.vilans.org/item/infographic-technology-for-people-with-dementia/>. Retrieved August 17, 2022.

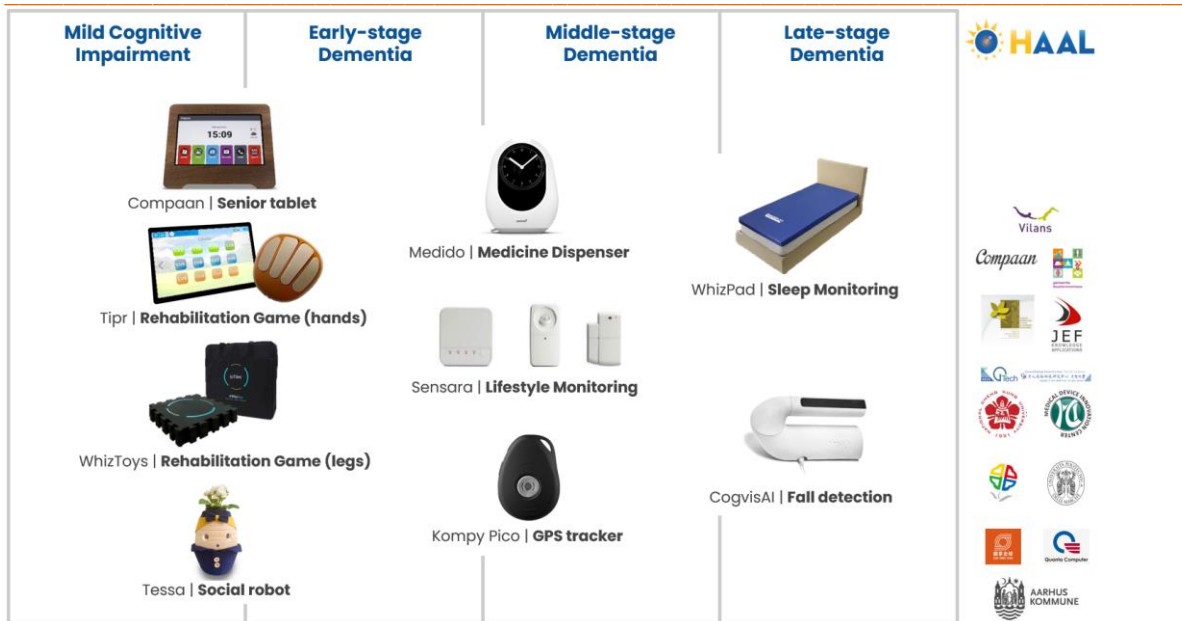


Figure 1. An initial diagram of HAAL technologies throughout the dementia stages

To maximize success and market fit, HAAL will be developed using an iterative methodology with four streams – i.e. co-creation, business modelling, responsible innovation and AI – machine learning – and with special attention for their coherence. All stakeholders – i.e. homecare nurses, informal caregivers, frail seniors and care organizations, but also the developers of the technological solution – are involved in sessions focused on these streams and in pilots in all three participating countries.

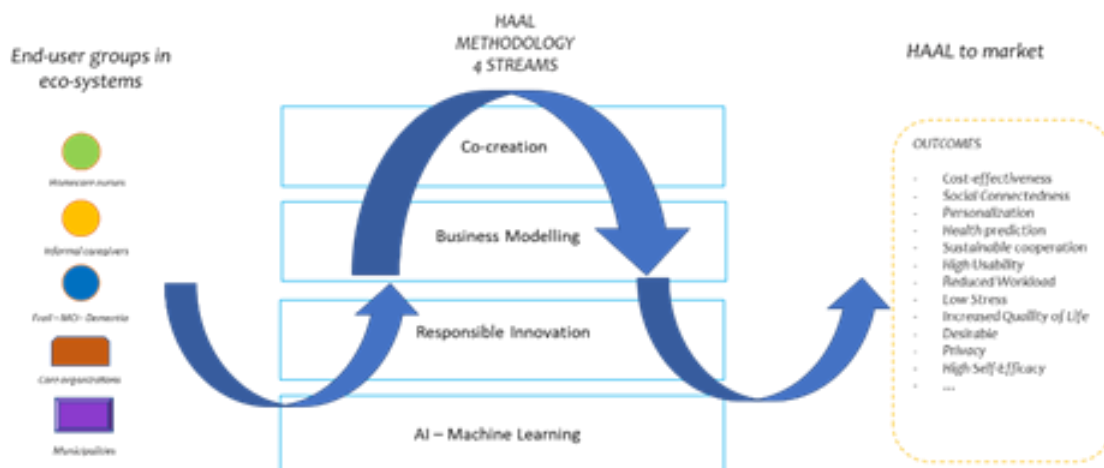


Figure 1: HAAL methodology & streams

2.2 Focus on development of a dashboard

In the research on responsible innovation in HAAL, we focus on the role of the overarching HAAL dashboard (that is to be developed in co-creation with end-users). The different technologies included in the HAAL ecosystem all generate data that directly or indirectly tell

something about the health, well-being or safety of community-dwelling people with dementia. Such data can, for example, help caregivers assess and determine the care and support needs of their clients. Within the HAAL project we are developing a dashboard in which we make the data from the various technologies available in one place. The dashboard is in fact a tool for linking the data back to caregivers (and any other stakeholders) in a manageable way so that they can do something with it.

The dashboard is intended to make various data available to healthcare providers in one place, so that they do not have to look in separate overviews and apps. The dashboard could provide a picture of the current situation in the client's home based on data about e.g. exercise, nutrition, medication intake and social contact. Eventually, the dashboard could help to automatically translate the data into insights (e.g. warnings or risk scores are automatically displayed about possible emergency situations such as a fall) and recommendations (e.g. possible follow-up actions such as stimulating a client's physical activity if the data shows that they have been relatively inactive in the past few days) for caregivers and other stakeholders.

3. About Responsible Innovation

This chapter explains what Responsible Innovation (RI) entails, firstly by discussing its relevance in the context of the HAAL project, and secondly by providing more (theoretical) background information about the concept and how it can be studied.

3.1 Relevance of Responsible Innovation in HAAL

The increasing advancements of AI-technologies such as the HAAL dashboard come with benefits such as better, faster and more efficient data-analysis and decision-making, but also with challenges from a social and ethical perspective, e.g. related to privacy, transparency, human control and trust. In this line, it is broadly acknowledged that the proper embedding of healthcare technologies driven by AI requires innovators and other stakeholders to actively anticipate and reflect on, and be responsive to promises and risks and to societal values, needs and expectations (Brundage, 2016; European Commission, 2019; Morley et al., 2019; Tsamados et al., 2021). In the HAAL project, we therefore explore what decisions and actions can be taken in the design of the HAAL platform and its implementation and use in different countries (and contexts) in order to account for the needs and values of end-users, and to achieve meaningful and responsible innovation (RI).

Responsible innovation is concerned with the actions, behaviour and activities that innovators, users and other stakeholders undertake to align both the processes and outcomes of innovation with societal needs and concerns (Fraaije & Flipse, 2019; Owen et al., 2013; Von Schomberg, 2013). Health(care) technologies will not simply be used on a largescale base by and for older citizens and their caregivers just because they address relevant health(care) related issues and fit certain needs. It is important that technologies fit within the dynamic context they are implemented in and that they do not interfere with what a large group of stakeholders, older people included, find important when it comes to the implementation of digital technologies in everyday life and practices.

RI requires collaboration between diverse stakeholders such as technology developers, care providers, users and other people who are ultimately affected by the technology (Owen et al., 2013; Stahl et al., 2019). Moreover, the impact of AI technologies does not depend only on decisions in technology design, but for a large part also on how these systems are introduced in practice and used in everyday situations (Dignum, 2022). This stresses the importance of a user-centred approach when addressing RI. In this study, we will therefore engage with the views and perspectives of both end-users and project partners to gain insight into the expected⁴ positive and negative impacts of using HAAL and what can be done to address RI. More about our approach and the specific research steps taken to explore these perspectives can be found in chapter 4.

⁴ A limitation of this study is that we will primarily explore the expectations about the (positive and negative) impact of using HAAL and only to a limited extent on real-world experiences. Thereby, the study is to some extent still speculative about what matters and what can be done to foster RI .

3.2 Background on Responsible Innovation

The term Responsible Innovation (RI) or Responsible Research and Innovation (RRI) has become increasingly fashionable in recent years and notably in European policy circles (Owen et al., 2012, 2013; Von Schomberg, 2013). Within this discourse, hereafter stated as RI, it is generally accepted that “modern innovation hardly ever have a single “author” who can be held responsible for the use (by others)” (Von Schomberg, 2013, p.13) and that innovators and societal actors have co-responsibility to build and deploy innovations in a responsible way. Within the discourse on RI, it is generally agreed that both the *process* and *outcomes (product)* of innovation should be responsible. When exploring (and addressing) RI, it is useful to distinguish between process and product dimension, which are naturally interrelated. Therefore, we briefly elaborate on the process-oriented (i.e. procedural) and outcome-oriented (i.e. substantive) approach to RI.

The process-approach or procedural approach frames RI as an ongoing and iterative process that follows a set of normative procedural criteria and is applied to design and develop technologies in a responsible way (Owen et al., 2012; Stilgoe et al., 2013). Four process-oriented dimensions of RI are distinguished:

- **anticipation** of the risks and opportunities of innovation, by systematically exploring possible and plausible futures and their associated societal and ethical impacts and dimensions.⁵
- **reflection** towards the value systems and social practices in which innovations are embedded, which is like “*holding a mirror up to one’s own activities, commitments and assumptions, being aware of the limits of knowledge and being mindful that a particular framing of an issue may not be universally held*” (Stilgoe et al., 2013, p.1571).
- **inclusion** of the views of professional and lay users and the broader public in the innovation process.⁶
- **responsiveness** of the innovation trajectories to emerging and changing views, norms and knowledge and to unforeseen consequences that may occur along innovation trajectories and across innovation ecosystems.

The product-approach and largely substantial approach to RI by Von Schomberg (2011, 2013) suggests that a set of shared values and norms has to first be developed and then applied in order to define the characteristics a given technology should possess, such as being, ethically acceptable, sustainable, socially desirable.

In section 4.1, we discuss how six principles for responsible AI innovation from the WHO guidelines on Ethics and governance of AI provide substantive starting points for HAAL from a product-perspective. Importantly, we acknowledge – in a critical sense – that such principles can merely provide starting points. As suggested by Boenink & Kudina (2020), values and principles are highly context-dependent and flexible. values are not ready-made entities, at once knowable and available for deliberation. Rather, as they state, “*values result from and as well as guide human action, and are continuously reshaped in and by the interactions of humans with their material and social environment*”. Their existence and meaning can be

⁵ Note that anticipatory activities should thus not only support the identification of potential risks but also the desirable outcomes.

⁶ Based on this philosophy, for example, the importance of the co-creation process within HAAL can be clearly underlined.

found inside human practices and they are constantly being (re-)experienced and worked out within these practices. From this, we take both the product- and process-oriented approach to RI only provide meaningful directions for us to explore and address RI in the context of HAAL, and while doing so, the work of making values visible and retrieving them from practice should not be underestimated (Boenink & Kudina, forthcoming).

4. Plan of action

This chapter discusses the plan of action for our research on responsible innovation in HAAL, including the overall approach taken to explore RI and the specific research steps taken.

4.1 Overall research approach

We conduct explorative research on responsible innovation in the context of HAAL from three perspectives:

1. HAAL project partners⁷
2. End-users, including care professionals (such as nurses, district nurses, care assistants, case managers dementia, data analysts in care organizations, and alarm centralists), informal carers and older people themselves.
3. Experts (outside the consortium) in the area of technology, AI, aged healthcare and ethics.

Thus, as part of the project, we study both innovators' and users' views and perspectives on the opportunities and risks of using such a platform to explore what decisions and actions can be taken in the design and deployment of the platform to achieve meaningful and responsible innovation. As we explain in the following, we have specific attention for the (increasing) role of AI in the HAAL dashboard and we use both scenarios and framework with high-level principles on responsible (AI) innovation to explore RI. Furthermore, throughout the research, we gather and compare between the views of Dutch, Danish, Italian and Taiwanese innovators and users, thereby exploring how to be sensitive to both the wider applicability and the local embedding of technology.

Focus on the (increasing) role of AI

In HAAL, artificial intelligence (AI) is driving some of the AAL (point) solutions such as ADL monitoring, but it is also being explored how AI can drive the overarching HAAL dashboard. Artificial intelligence (AI) is widely positioned and predicted to become a key element of intelligent technologies that are used in the long-term care for older people. Human intelligence is what people use to answer the question: what should I do now to reach a certain goal, given everything I know, everything I perceive, and how the environment is changing? Artificial intelligence is the attempt to mimic parts of that intelligence in machines. A specific definition of AI in a recommendation of the Council on Artificial Intelligence of the OECD states that an AI-based system is a “machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing the real or virtual environments they interact with” (Yeung, 2020). AI systems perform functions such as image, speech and pattern recognition, and natural language processing which are normally associated with the human brain (McCarthy et al., 2006). They can be driven by both pre-programmed, rule-based algorithms that capture and embody the

⁷ In this sense, our research can be perceived as participatory action research (Baum et al., 2006), which is aimed at directly improving the social setting under investigation. We namely explore ‘from within’ how RI can be addressed throughout the project activities and the practices of the various organisations and individuals involved.

knowledge of human experts in a specialized domain (Lucas & Van Der Gaag, 1991; Miller et al., 1982) and self-learning, case-based algorithms that independently learn to execute tasks and improve on the basis of machine learning on historical, exemplary data (Aamodt & Plaza, 1994; Jordan & Mitchell, 2015; LeCun et al., 2015; Samuel, 1959). Accordingly, AI technologies are designed to operate with varying levels of autonomy (Yeung, 2020).

Instead of seeing AI as technology that replaces the role of humans, we adopt a more productive approach to AI (or IA; intelligence augmentations). We aim to create synergy between the irreplaceable expertise of people and the increasing possibilities of computers, and thus commit to the development of AI-based tools that ‘augment’ or ‘amplify’ human intelligence (Asaro, 2013; Skagestad, 1993).

Using scenarios with differing roles for AI

Throughout the research, we will use scenarios (as a source of inspiration for research participants) about the HAAL dashboard that illustrate how both the level of autonomy of AI-systems and their ability to learn are matters of degree. Accordingly, each scenario comes with different opportunities and risks. The two scenarios will be used as a source of inspiration and research tool to explore the views and perspectives of HAAL project partners, end-users and other stakeholders on the (expected) ethical and social impact of the HAAL dashboard.

Framework with principles

To investigate how AI can be used responsibly, it will be explored how a set of values and principles for responsible AI innovation can be promoted in the design and implementation of the HAAL dashboard. Values and principles can be seen as conceptions of what is good, proper, and desirable (Graeber, 2001) and as beliefs or moral compasses that could guide responsible innovation. The WHO guidelines for Ethics & Governance of AI for Health provide a set of principles that will be used as a starting point in this study. The following six principles are central in these guidelines:

- protect autonomy
- promote human well-being, safety
- ensure transparency, explainability and intelligibility
- foster responsibility and accountability
- ensure inclusiveness and equity
- promote AI that is responsive and sustainable

Many organisations have already produced statements of the values or principles that should guide the development and deployment of AI in society. Far less attention has been paid to the implementation and impact of such principles in the actual design and implementation of AI in practice (Hagendorff, 2020; Jobin et al., 2019; Lukkien et al., 2021). In the HAAL project, and in particular in the RI research, we aim to move the dialogue forward, from values and principles to concrete opinions about how they are important in the context of the HAAL dashboard - and if possible even into recommendations about how to apply these principles in the HAAL project. We will explore the importance of these values and principles and how they (can) take shape or even be routinized in the design and implementation of the HAAL

dashboard. We are not so much going to test the WHO principles, but rather see them as a starting point for discussing and thinking about what is important for the responsible application of AI. The principles will be used as sensitizing concepts that provide “a *general sense of reference and guidance*” and “*directions along which to look*” during both the data collection and analysis (Blumer, 1954, p.7). We will ask different stakeholders to react on the principles and operationalize them into their own context (deductive reasoning). Also, we will offer them the opportunity to talk about other factors that are important (inductive reasoning), but which (at least in their view) fall outside these predefined (WHO) categories. Thus we combine an open, bottom-up approach with a top-down approach based on the WHO guidelines, or in other words, we combine inductive and deductive reasoning about

Exploring the role of (cultural) contexts

An important point of departure for this study is that a variety of countries are involved in HAAL and that RI thus needs to be addressed in (and across) different ‘context-of-use’. In this line, the HAAL project will need to be sensitive to both the wider applicability and the local embedding of decisions related to RI. In other words, it should on the one hand be explored and articulated how - through what decisions in design and/or implementation – it can be fostered that the HAAL platform is robust, socially desirable and ethically acceptable across the different countries’ cultures and contexts-of-use. On the other hand, it should be explored what customization is needed to deal with local differences between end-users’ views and perspectives about the social and ethical impact of HAAL. Therefore, this RI study will compare the views and perspectives on RI from end-users of different countries (The Netherlands, Italy, Taiwan and eventually Denmark). Also, we will explore the ideas of various stakeholders, including project partners and end-users, about how these contextual differences can and should be taken into account in the design and implementation of the HAAL platform.

A comparison between the views and perspectives from end-users from the different European countries may already be interesting in itself, but we expect that especially the comparison between the European and Taiwanese context could be relevant, as significant differences in needs and values could be expected here. As ÓhÉigearthaigh et al (2020) for instance state, Western and Eastern cultures are founded on divergent philosophical traditions, leading to significant and perceivably irresolvable value differences on key issues such as data privacy. As ÓhÉigearthaigh et al. (2020) explain, the Western “independent” and Eastern “interdependent” perspective on ageing have different ways of valuing privacy, parent–child relationships, connectivity, and outsourcing health and safety monitoring (in full) to technology. Besides, Whittlestone et al. (2019) discuss that everyone might agree in principle that ‘fairness’ is important, but that there nevertheless exist deep political disagreements about what exactly constitutes fairness. Groups may also vary in how much weight they put on one value relative to others in situations of conflict: more individualist cultures may put more weight on personal privacy than more collectivist cultures, for example (Whittlestone et al., 2019).

4.2 Chronological steps in the RI research

In this section, we discuss what steps are, and will be taken in chronological order to explore responsible innovation in HAAL from the perspectives of project partners, end-users and (external) experts.

Step 1 (2021): Workshop during consortium Kick-off meeting

Workshop with a high-level introduction about RI and its relevance for HAAL to raise general awareness. Also we gathered insights (through Mural) into the initial ideas of project partners about RI, what it means and how it is relevant in context of HAAL. Also intended to ensure that all project partners have some basic knowledge about RI.

Step 2 (2021): Workshop during hybrid consortium meeting in Utrecht

We presented a framework (Guidance ethics approach) through which RI can be explored. Through the Mural, we gathered results about opportunities and risks / expected positive and negative effects of HAAL + the values that are affected by its use.

Step 3 (2022): Responsible AI innovation survey amongst project partners

First we give a presentation during the consortium meeting. Then we set out the survey (see appendix 1)

Step 4 (2022): Workshop with consortium during hybrid meeting in Ancona

Through a FG (during the consortium meeting in Ancona), we deepened the insights gained through the survey and zoomed in on some of the results (and differences between countries) that emerge.

Step 5 (2022-2023): Interviews with end-users from different countries

First, we start with interviews with primary end-users, i.e. different types of care professionals, including nurses, district nurses, care assistants, case managers dementia, data analysts in care organizations, and alarm centralists), informal carers and older people themselves. Aim: $N=20$ in The Netherlands, $N=10$ in Denmark, $N=10$ in Italy, $N=10$ in Taiwan. For the interview protocol, see appendix 2.

Second, we conduct interviews with secondary end-users, i.e. dyads of older people and their informal carer. Aim = $N=5-7$ in The Netherlands, $N=3-5$ in Denmark, $N=3-5$ in Italy, $N=3-5$ in Taiwan. Interview protocol will be developed later on, amongst others based on initial insights from the interviews with primary end-users.

Step 6 (2023): Interviews with experts outside the consortium

After interviewing end-users, we conduct interviews with approximately experts in the area of elderly care, technology, AI and/or ethics who are not part of the HAAL consortium. These interviews will be focusing on exploring if any factors that are important to the responsibly use of the HAAL dashboard (or AI decision support in general) have been overlooked in the interviews with end-users and to explain the results and dynamics. We focus on interviews, where possible physically, with native or English speaking experts that are located in the four countries involved. Aim = $N=5-7$ in The Netherlands, $N=3-5$ in Denmark, $N=3-5$ in Italy, $N=3-5$ in Taiwan. Interview protocol will be developed later on, amongst others based on the insights from step 3-5.

Step 7 and etc. (2023): Workshops with the consortium.

As a follow-up on the insights from interviews with end-users and experts (steps 5 and 6), we will organize multiple workshops with the HAAL consortium to study how the HAAL

consortium approaches and implements RI; how the project promotes the social desirability, ethical acceptability and sustainability of the dashboard, how Responsible AI principles are translated to this specific innovation project and how dealing with these principles is routinized. We will later develop methodology and protocols for these workshops. The workshops are focused at ultimately *doing things to address RI*, and at learning about the actual *responsiveness* of the HAAL project to insights about RI.

5. Provisional results

Step 1

No results reported

Step 2

Examples of some risks discussed by the consortium partners:

- Older adults could end up living like a technologically served guinea pig in their homes, especially when living alone
- Loose of human contact
- People constantly looking – privacy infringement
- Risk that data is used for other purposes (e.g. costs of insurance)
- Carers gets 'exposed' in his/her work as care processes are monitored
- Risk of contrasting decisions between AI and carer
- Anxiety for carers that continuously check/read data
- Examples of potential actions to address RI:
 - notify only when a valuable pattern has been identified
 - put sufficient work into the annotation of data & algorithms
 - hierarchical decision order – caregiver over AI

Step 3

Some highlights:

- Most project partners are excited for the possibilities of using AI, and to develop a more advanced (and proactive) dashboard, since it can greatly support (and reduce the workload of) professional carers.
- The main added value for both carers and clients is the automated detection of subtle risks which might not be manually recognized. The dashboard could for instance help carers to observe changes and deviations in the health and wellbeing of clients earlier.
- At the same time, multiple partners express that they feel anxious about the potential side effects of using advanced AI-based technology, such as the complexity and hidden logic of some methods to achieve AI (such as machine learning) and the potential lack of human judgement. It is generally acknowledged that the consortium needs to have (and think about) the right intentions to use AI to avoid any negative impacts from the start.
- A major concern lies in the training of the people that will use the output of AI applications: are they able to assess the reliability and usefulness of the data?

- Project partners do not yet have a unanimous judgement as to whether the consortium should aim for scenario 1 (relatively simple dashboard) or 2 (relatively advanced dashboard). Various arguments are put forward in favour of either scenario 1 or 2:
 - Preferring scenario 1: aged care is not ready yet a dashboard that prompts action; the dashboard could be more easily integrated in the current care process; clients may prefer a simple dashboard as they trust the clinical experts to judge their information.
 - Preferring scenario 2: could provide a much better solution to the current gap between care needs and care staff; adding more data in one place without elaborating on it doesn't really reduce the burden to the caregiver > some summarization is needed.
- Some action opportunities (for the design process) were identified, such as:
 - In order to be able to support people in decision-making, it is important that data and algorithms are transparent, so that you know the basis on which you make your choices. To maintain a good overview in the HAAL dashboard, more in-depth explanations and data analyses should be accessible in submenus.
 - There should be a feedback button where a client could assess if the highlighted info/recommendations were indeed efficient/needed action.
 - End-users should at least have training in the type of data that is gathered by the technologies, so they can judge the validity of this and take this into account while looking at the dashboard.
 - Caregivers should be taught and made conscious that they will always in some degree be influenced by the information on the dashboard, and even recommended to make their own judgement on the decision first, before getting the information from the machine.
 - There should be inclusion and exclusion criteria for applying the dashboard. For instance, if there are personal differences, such as physical or visual handicaps, in that case, does the dashboard give appropriate information? You shouldn't ask someone who has trouble walking to make more strolls around the block.
 - It should be investigated how the dashboard affects mental stress or workload of both carers and older people.

The results from step 3 will be more extensively reported in the next version of the RI Report. Based on the results from step 3 and 4, we can state that the consortium already gained a clearer picture of the HAAL platform and its potential impact. It was suggested that the scenarios provided good inspiration and a nice frame to discuss desired functionalities and potential unintended effects that should be avoided.

Step 4

Currently running.

References

- Aamodt, A., & Plaza, E. (1994). Case-based reasoning: Foundational issues, methodological variations, and system approaches. *AI Communications*, 7(1), 39–59. <https://doi.org/10.3233/AIC-1994-7104>
- Asaro, P. M. (2013). From Mechanisms of Adaptation to Intelligence Amplifiers: The Philosophy of W. Ross Ashby. *The Mechanical Mind in History*, 149–184. <https://doi.org/10.7551/MITPRESS/9780262083775.003.0007>
- Baum, F., MacDougall, C., & Smith, D. (2006). Participatory action research. *Journal of Epidemiology and Community Health*, 60(10), 854. <https://doi.org/10.1136/JECH.2004.028662>
- Blumer, H. (1954). What is Wrong with Social Theory? *American Sociological Review*, 19(1), 3. <https://doi.org/10.2307/2088165>
- Boenink, M., & Kudina, O. (2020). Values in responsible research and innovation: from entities to practices. *Journal of Responsible Innovation*, 1–21. <https://doi.org/10.1080/23299460.2020.1806451>
- Brundage, M. (2016). Artificial Intelligence and Responsible Innovation. In *Fundamental Issues of Artificial Intelligence* (pp. 543–554). Springer International Publishing. https://doi.org/10.1007/978-3-319-26485-1_32
- Dignum, V. (2022). *Relational Artificial Intelligence*. <https://arxiv.org/abs/2202.07446v1>
- European Commission. (2019). *Ethics Guidelines for Trustworthy AI*. High-Level Expert Group on Artificial Intelligence. <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>
- Fraaije, A., & Flipse, S. M. (2019). Synthesizing an implementation framework for responsible research and innovation. *Journal of Responsible Innovation*, 1–25. <https://doi.org/10.1080/23299460.2019.1676685>
- Graeber, D. (2001). *Toward an anthropological theory of value: The false coin of our own dreams*. Springer.
- Hagendorff, T. (2020). The ethics of AI ethics: An evaluation of guidelines. *Minds and Machines*, 30(1), 99–120. <https://doi.org/10.1007/s11023-020-09517-8>
- Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 1(9), 389–399. <https://doi.org/10.1038/s42256-019-0088-2>
- Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245), 255–260. <https://doi.org/10.1126/science.aaa8415>
- LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444. <https://doi.org/10.1038/nature14539>
- Lucas, P., & Van Der Gaag, L. (1991). *Principles of expert systems*. Addison-Wesley Wokingham.
- Lukkien, D. R. M., Nap, H. H., Buimer, H. P., Peine, A., Boon, W. P. C., Ket, J. C. F., Minkman, M. M. N., & Moors, E. H. M. (2021). Toward Responsible Artificial Intelligence in Long-Term Care: A Scoping Review on Practical Approaches. *The Gerontologist*, XX, 1–14. <https://doi.org/10.1093/geront/gnab180>
- McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (2006). A proposal for the

- Dartmouth summer research project on artificial intelligence, August 31, 1955. *AI Magazine*, 27(4), 12–12. <https://doi.org/10.1609/AIMAG.V27I4.1904>
- Miller, R. A., Pople Jr, H. E., & Myers, J. D. (1982). Internist-I, an experimental computer-based diagnostic consultant for general internal medicine. *New England Journal of Medicine*, 307(8), 468–476. <https://doi.org/10.1056/NEJM198208193070803>
- Morley, J., Machado, C., Burr, C., Cowls, J., Taddeo, M., & Floridi, L. (2019). The debate on the ethics of AI in health care: a reconstruction and critical review. *SSRN 3486518*. <https://doi.org/10.2139/ssrn.3486518>
- ÓhÉigearthaigh, S. S., Whittlestone, J., Liu, Y., Zeng, Y., & Liu, Z. (2020). Overcoming Barriers to Cross-cultural Cooperation in AI Ethics and Governance. *Philosophy & Technology*, 1–23.
- Owen, R., Macnaghten, P., & Stilgoe, J. (2012). Responsible research and innovation: From science in society to science for society, with society. *Science and Public Policy*, 39(6), 751–760.
- Owen, R., Stilgoe, J., Macnaghten, P., Gorman, M., Fisher, E., & Guston, D. (2013). A framework for responsible innovation. In *Responsible innovation: managing the responsible emergence of science and innovation in society* (pp. 27–50). Wiley.
- Samuel, A. L. (1959). Some studies in machine learning using the game of checkers. *IBM Journal of Research and Development*, 3(3), 210–229. <https://doi.org/10.1147/rd.33.0210>.
- Skagestad, P. (1993). Thinking with machines: Intelligence augmentation, evolutionary epistemology, and semiotic. *Journal of Social and Evolutionary Systems*, 16(2), 157–180. [https://doi.org/10.1016/1061-7361\(93\)90026-N](https://doi.org/10.1016/1061-7361(93)90026-N)
- Stahl, B. C., Borsella, E., Porcari, A., & Mantovani, E. (2019). Responsible innovation in ICT: challenges for industry. In *International Handbook on Responsible Innovation*. Edward Elgar Publishing.
- Stilgoe, J., Owen, R., & Macnaghten, P. (2013). Developing a framework for responsible innovation. *Research Policy*, 42(9), 1568–1580. <https://doi.org/10.1016/j.respol.2013.05.008>
- Tsamados, A., Aggarwal, N., Cowls, J., Morley, J., Roberts, H., Taddeo, M., & Floridi, L. (2021). The ethics of algorithms: Key problems and solutions. *AI & Society*. <https://doi.org/10.2139/ssrn.3662302>
- Von Schomberg, R. (2011). Towards responsible research and innovation in the information and communication technologies and security technologies fields. *Available at SSRN 2436399*.
- Von Schomberg, R. (2013). A vision of responsible research and innovation. In *Responsible innovation: Managing the responsible emergence of science and innovation in society* (pp. 51–74). Wiley Online Library.
- Whittlestone, J., Nyrup, R., Alexandrova, A., & Cave, S. (2019). *The Role and Limits of Principles in AI Ethics: Towards a Focus on Tensions*. <https://doi.org/10.1145/3306618.3314289>
- Yeung, K. (2020). Recommendation of the council on artificial intelligence (OECD). *International Legal Materials*, 59(1), 27–34. <https://doi.org/10.1017/ilm.2020.5>

Appendixes

Appendix 1: Internal HAAL survey on Responsible Innovation

Exploring the potential and desired impact of the AI-based HAAL dashboard

Section 1. Introduction In HAAL, we are developing a dashboard in which the data from the different AAL (point) solutions is combined, analysed and coherently presented or even transformed into recommendations to support caregivers in their decision-making. In this survey, we ask HAAL project partners about your views and expectations about the positive and negative impacts of the HAAL dashboard on end-users. We use two scenarios about the HAAL dashboard to explore and define the role of artificial intelligence (AI) and its impact, as well as directions for further development. Also, we ask you about principles for responsible AI innovation and how we can promote these principles in the design and implementation of the HAAL dashboard. Through this survey, we aim to learn about your perspectives on innovation that is meaningful and responsible for end-users, i.e. people with dementia and their formal and informal caregivers. Please note that we ask you to answer the survey from your own perspective, so from your role in the HAAL consortium. By completing this survey you give permission for the processing of the data by Vilans. Preferably you do not include person-specific information such as your name in your answers, except for the last section of the survey, where you can optionally leave person-specific information. Your personal information can help us to track your views on responsible innovation in HAAL over time, and to compare the views of different HAAL project partners. The results will be used (anonymised where possible) as input for further research and development within the HAAL project. The results may also be used (anonymously) for publications. More information about the processing of data by Vilans can be found in the privacy statement of Vilans. This survey contains 10 sections. This is the end of section 1.

Section 2. Artificial intelligence Importantly, the HAAL dashboard will be driven by artificial intelligence (AI). Artificial intelligence is the attempt to mimic parts of human intelligence in machines. A specific definition from the OECD states that AI technologies are “*machine-based systems that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing the real or virtual environments they interact with*”. Both pre-programmed (rule-based) algorithms and self-learning (data-based or machine learning) algorithms can be at the basis of AI technologies. The increasing advancements of AI-technologies such as the HAAL dashboard come with benefits such as better, faster and more efficient data analysis and decision-making, but also with challenges from a social and ethical perspective, e.g. related to privacy, transparency, human control and trust. Therefore, we are exploring **what decisions and actions can be taken in the design and implementation of the HAAL dashboard to achieve responsible innovation (RI)** that is socially desirable, ethically acceptable and sustainable.

Question: How do you feel about the increasing role of data and AI in healthcare and society (e.g. hopeful, suspicious, or otherwise)? Please explain your answer.

Section 3. Types of AI-based analyses Through an iterative user-centered design approach, the HAAL consortium is carefully determining what kind of overviews and AI-based analyses the dashboard should provide. Finally, the HAAL dashboard could provide descriptive, diagnostic, predictive and even prescriptive analyses. In the following, it is briefly explained what these types of analyses could entail in HAAL: • Descriptive analyses could mean that data from the various AAL products about the current situation in the client’s home is displayed and eventually translated into a coherent and comprehensible overview of insights. The data and insights can be about the client’s behaviour, health, well-being and

safety, i.e. exercising, nutrition, medication intake and social contact. • Diagnostic analyses could mean that data gathered by several HAAL products is combined to create new meanings. For example, when data shows that the client’s sleep quality during the night is declining while activity during the day also declines, then the system may automatically link and display these two pieces of information together, as a way to imply that the day-night rhythm of the senior may be disturbed. • Predictive analyses could mean that based on the data from HAAL products warnings or risk scores are given about possible future situations or emergencies such as a fall. • Prescriptive analyses could mean that recommendations are derived from the data that prompt caregivers to take a certain action, for instance to stimulate physical activity if a client is has been relatively inactive over the last days.

Section 4. Two scenarios about the dashboard In this survey, we ask you to reflect on two scenarios about the HAAL dashboard in order to learn more about what is feasible, desirable and (ethically) acceptable. As shown in the figure below, the scenarios focus on different types of AI-based analyses. Scenario A primarily focuses on descriptive analyses with relatively low complexity and a low level of automation, as relatively much human input is required to translate these analyses into real-world decisions and actions. Scenario B provides various types of analyses (i.e. descriptive, diagnostic, predictive and even some prescriptive analyses) that involve more complexity and a higher level of automation. We are very interested in your views and opinions about both scenarios. After showing each scenario, we will ask you questions about it.

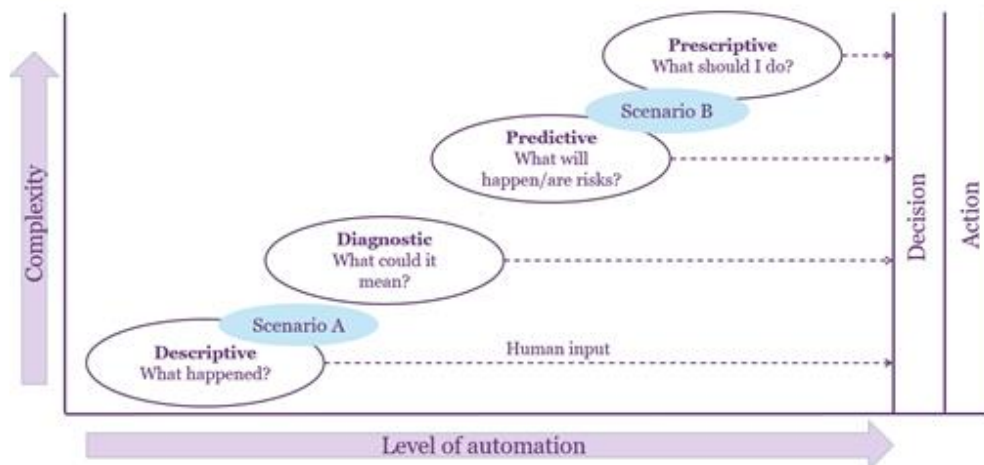


Figure: Types of AI-based analyses

Section 5. Scenario A: an integrated dashboard with a descriptive overview The HAAL dashboard is a graphic user interface that provides professional caregivers a central entrance point to descriptive analyses about the behaviour, health, well-being and safety of their clients. The analyses are based on the data that is gathered through various AAL products that are deployed in the individual senior’s home setting. In the dashboard, caregivers can assess current situation and how this has changed over time. **Examples** For instance, prior to a home visit to a client a caregiver looks into the dashboard. In a graph, he sees that for the last few days, the amount of physical activity of the client has been lower than usual. The caregiver interprets this as a health risk and assumes it will be good to stimulate physical activity. The caregiver arranges that the client is taken out for a walk the next day. For another client the caregiver sees a notification in the dashboard that the client is not getting out of bed, while being awake for a few hours. This has been happening for multiple days in a row now. Also, the data shows that the client is lying in the same position during most of the time in bed. The caregiver wonders if something might be going on and decides to ask the client about it during the next visit, and to assess the risk for bed sores. **Supportive system** Overall, the insights from the HAAL dashboard can support caregivers when making decisions about

the care and support for a specific client, or at least provide starting points for further investigation. The dashboard is a low-complex system that can be easily understood and interpreted by caregivers. **Pre-programmed system** The HAAL dashboard does not simply show all the data that is gathered through the technologies. Only the information that is deemed to be most relevant at that moment is shown. In order to distil relevant insights from across the datasets, the dashboard uses algorithms that are defined by a team of human programmers and professional caregivers. In addition, caregivers that use the platform have some options for manually tweaking the rules, in order to better tailor the dashboard to the context of the care organisation or a specific client. For instance, they could decide to let the dashboard show more detailed information about physical activity, or to prolong the period over which data is shown. **HAAL bundle adjustments** In the dashboard, caregivers can also see what other technologies are available and could be connected to the platform. Through the dashboard, caregivers can manually initiate or stop the use of a particular technology. The purchase, installation, user training and other necessary steps are then automatically initiated and coordinated.

Questions:

1. What do you think about the added value of this dashboard for professional caregivers and clients?
2. What do you think could be the role of the informal caregiver in the use of the dashboard?
3. What are the main strengths of this scenario in your opinion?
4. What weaknesses do you (fore)see for caregivers in this scenario?
5. What weaknesses do you (fore)see for clients in this scenario?
6. What suggestions do you have about how to deal with these weaknesses?

Section 6. Scenario B: a forward-looking decision aid The HAAL dashboard is a personal assistant that provides professional caregivers with real-time insights and predictions about the well-being, safety and health of clients. Also the dashboard gives recommendations for next steps in the care and support of these clients. The personal assistant provides these insights and recommendations both proactively and on demand of the caregiver, and can be interacted with through voice. **Examples** For instance, in the morning, the HAAL dashboard notifies to a caregiver that five of her clients might need closer attention based on erratic behaviour patterns in the past week. After looking briefly at graphs about these clients' behavioural patterns, the caregiver quickly decides that three of the five clients need an urgent home visit. Through the HAAL dashboard, her working time on the screening process is shortened. Besides, more time is devoted to personalised care for clients with high risks. During, and sometimes before home visits, the case manager looks into the HAAL dashboard to assess the situation of specific clients. For one client, the platform displays that the client takes more time to get out of bed in the past 2 months. There is an orange flag next to the graph about the time for 'rising-getting out'. The system comments that this indicates a high risk for a fall, and suggests a leg rehabilitation program to reduce the risk. During further talks with the client at home, the case manager decides to refer to a physiotherapist for a fall prevention program. **Supportive system** Overall, the HAAL dashboard relocates the workload for caregivers from intensive rehabilitation to prevention, for instance by helping to avert the risk of health complications that occur after a fall. **Adaptive system** Over time more and more

data are collected about the individual client, whereby the system's performances and insights verifiably become more and more reliable. Based on a combination of pre-programmed and self-learning algorithms, the dashboard is automatically interpreting all data, learning and reasoning about the client's individual situation, and providing suggestions about the care situation and the best courses of action. Based on newly gathered data, the system is able to

Haal Survey DEF - 2930 Dirk Lukkien / Nathalie Stolkwijk Page 10

update its internal decision-making logic over time. In other words, the system is continuously improving the rules that are used to process raw data into relevant insights. At the same time, the rules used by the system become more difficult to explain over time, as it is not transparent how the system processes input to certain output. **HAAL bundle adjustments** The AI-based dashboard proactively supports caregivers in adapting the eco-system of AAL technologies to the current needs of the older adult and his/her care network. Based on the data from the various HAAL products, the dashboard proactively provides suggestions to caregivers about care technologies that cannot or need not be used anymore, and other technologies and services that could be introduced to better support the situation at that stage. Through the dashboard, caregivers only need to approve or reject suggestions.

Questions:

1. What do you think about the added value of this dashboard for professional caregivers and clients?
2. What do you think could be the role of the informal caregiver in the use of the dashboard?
3. What are the main strengths of this scenario in your opinion?
4. What weaknesses do you (fore)see for caregivers in this scenario?
5. What weaknesses do you (fore)see for clients in this scenario?
6. What suggestions do you have about how to deal with these weaknesses?

Section 8. Principles for responsible AI innovation Next, we will explore how a set of values and principles for responsible AI innovation can be promoted in the design and implementation of the HAAL dashboard. Values and principles can be seen as conceptions of what is good, proper, and desirable and as beliefs or moral compasses that could guide responsible innovation. Many organisations have already produced statements of the values or principles that should guide the development and deployment of AI in society. In HAAL, we strive to move the dialogue forward, from values and principles, to concrete recommendations about how to apply these principles in the project. In the following, we therefore present a set of principles for responsible AI innovation (adopted from Guidance on Ethics and governance of artificial intelligence for health from the World Health Organization) and we ask you to reflect on these principles in the light of the two scenarios.

When answering the questions, it could help to consider if certain elements from scenario A or B could either promote or undermine this principle. This may provide directions along which to think about measures that could be taken to ensure that this principle is promoted. At the same time, all suggestions are welcome, also apart from the scenarios.

Principle 1: Autonomy Adoption of AI can lead to situations in which decision-making could be or is in fact transferred to machines. The principle of autonomy requires that any extension of machine autonomy not undermines human autonomy. In the context of healthcare, this means that humans should remain in full control of healthcare systems and medical or care-related decisions. AI systems should be designed to assist humans in making informed decisions, whether they be care providers or clients.

Question: What are your suggestions for the design and/or implementation of the HAAL dashboard to promote the autonomy of clients and their (formal and informal) caregivers?

Principle 2: Well-being & safety AI technologies should not harm people. They should satisfy regulatory requirements for safety, accuracy and efficacy before deployment, and measures should be in place to ensure quality control and quality improvements.

Question: What are your suggestions for the design and/or implementation of the HAAL dashboard to promote the well-being and safety of clients and their (formal and informal) caregivers?

Principle 3: Transparency & explainability AI should be intelligible or understandable to developers, users and regulators. AI technologies should be explainable to the extent possible and according to the capacity of those to whom the explanation is directed. Many AI technologies are complex, and the complexity might frustrate both the explainer and the person receiving the explanation. There is a possible trade-off between full explainability of an algorithm (at the cost of accuracy) and improved accuracy (at the cost of explainability).

Question: What are your suggestions for the design and/or implementation of the HAAL dashboard to promote the transparency and explainability of insights and outcomes shown by the HAAL-dashboard?

Although AI technologies perform specific tasks, it is the responsibility of human stakeholders to ensure that they can perform those tasks and that they are used under appropriate conditions. And when something does go wrong in application of an AI technology, there should be accountability.

Question: What are your suggestions for the design and/or implementation of the HAAL dashboard to promote human responsibility and accountability?

Principle 5: Inclusiveness and equity Inclusiveness requires that AI used in health care is designed to encourage the widest possible appropriate, equitable use and access, irrespective of age, gender, income, ability or other characteristics. AI technologies should be designed by and evaluated with the active participation of those who are required to use the system or will be affected by it, including providers and patients, and such participants should be sufficiently diverse. This means, for instance, that AI technologies should not be biased. Bias is a threat to inclusiveness and equity because it represents a departure, often arbitrary, from equal treatment.

Question: What are your suggestions for the design and/or implementation of the HAAL dashboard to promote inclusiveness and equity?

Principle 6: Responsiveness Responsiveness requires that designers, developers and users continuously, systematically and transparently examine an AI technology to determine whether it is responding adequately, appropriately and according to communicated expectations and requirements in the context in which it is used. For instance, the identification of a health need requires a response to that need. Or when an AI technology is ineffective or engenders dissatisfaction, there is a duty to be responsive by resolving the problem or even terminating the use of the technology. Also, AI technologies should be introduced only if they can be fully integrated and sustained in the health-care system.

Question: What are your suggestions for the design and/or implementation of the HAAL dashboard to promote responsiveness of both the HAAL project and its outcomes (i.e. the AI-based HAAL dashboard)?

Section 9. Prioritizing the principles Below we ask to rank the principles you saw in section 8 three times. Make sure you drag the answers up so the answer is green outlined.

- 1. Which of these principles is in your view most important to pay attention to in HAAL? Please rank them from most important to least important. Can you explain your prioritization?**
- 2. Which of the principles do you expect is hardest to achieve in HAAL? Please rank them from hardest to easiest to achieve. Can you explain your prioritization?**
- 3. In relation to which principles do you see the largest risks if no explicit action is taken by the HAAL consortium to actively promote these principles? Please rank them from highest risk to lowest risk. Can you explain your prioritization?**

Section 10: Final questions

- 1. What is your name?** *(optional)* This is relevant for us to know, as we may ask you more questions later in the project, and this allows for a comparison over time. However, you can always decide to stay anonymous.
- 2. For which partner organisation of the HAAL project do you work?** *(optional)*
- 3. Can you briefly describe your function?** *(obligatory question)*
- 4. What is your age?** *(optional)*
- 5. What HAAL work package(s) do you mostly work on?** *(optional) Multiple answers possible*

Appendix 2: Protocol for RI interviews with primary end-users (i.e. care professionals).

Explanation for the interviewer:

- The text in black is an explanation that may be read literally to the respondent.
- **The green-printed (numbered) questions can all be asked in chronological order. Please make sure that you have the time to ask all these questions. Especially the questions in part 2 are important.**
- The blue-printed questions (a, b, c, etc) are additions to the numbered questions. These questions are optional and can be asked at the interviewer's discretion, for instance when respondents have a lot to tell about the respective theme, or as 'probing questions' to help the respondent get started with an answer when they initially find it difficult to answer. Please note that it is not necessary to ask the blue-printed questions.
- Please be vigilant during the interview that you as an interviewer are not too steering, and that you allow the respondent to share as much as possible about what they believe is important.
- During the interview, you can use PowerPoint slides to show the images.

Part 1

Introduction

Thank you for joining the interview. This is very valuable for the HAAL project because we like to match the technology as closely as possible with practice and those who will eventually use it.

Can we record the interview just for ourselves to process the data?

The interview will last between 1 and 1.5 hours. There are quite many questions we want to ask you. Therefore, we will run through some parts of the interview pretty quickly. The first part of the interview is meant to get to know you, and to introduce you to the project. The second part delves deeper into the ethical and social impact of technology.

<turn on recording & transcript>

Let us introduce ourselves first. And then there are some questions about you and your work situation.

<researchers introduce themselves>

- 1. What is your job title? And can you shortly explain what this job entails?**

2. *What is your age?*

3. *How long have you been working in care for people with dementia?*

4. *On average, how often do you work with people with dementia?
Daily/weekly/monthly/yearly/barely*

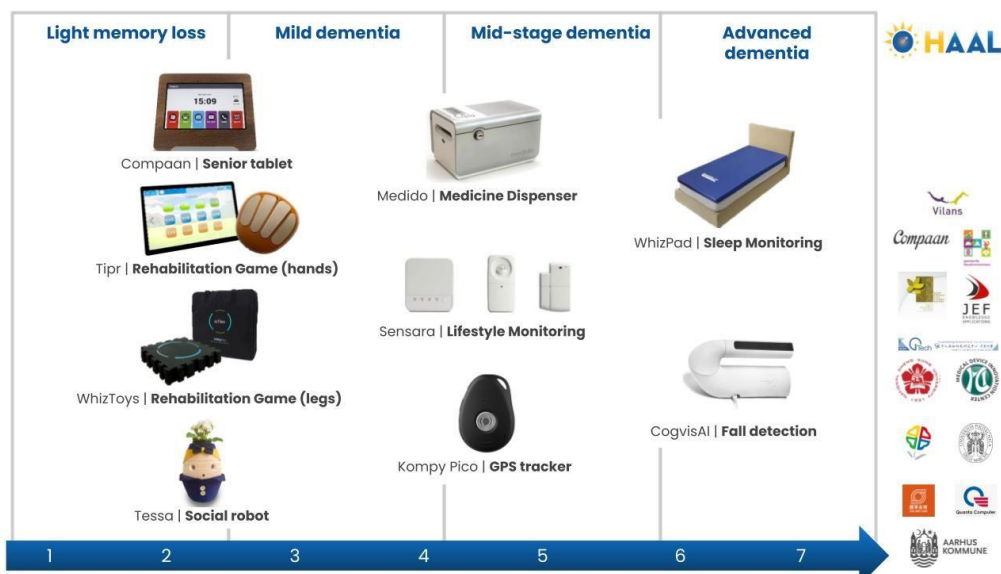
5. *Can you briefly tell us more about the role you play in their care?*

6. *In which phase of dementia are the clients you mainly care for?*

Overview of technologies

In short, within the HAAL project we are investigating how a bundle of care and welfare technologies can be used for people living at home with dementia. In the image below you can see the different technologies that are part of the HAAL project. I will briefly explain them.

<show flowchart>



1. de Compaan, a seniors tablet with which you can, for example, view social online contact of the client.
2. Medido, a medicine dispenser, with which you can see whether the client is taking his medication.
3. Whizpad, a smart mattress, with which you can see, among other things, how well the client sleeps, time in bed, how often he leaves his bed, and whether there is a risk of bedsores.
4. Tipr, a hand-drag rehabilitation game that allows us to see cognitive decline.
5. Sensara, sensors for the home, with which you can see in which room the client is located.
6. Whiztoys, also a game that allows us to see cognitive decline.
7. Cogvis AI, a fall detector, which measures, among other things, the danger of falling.

8. Tessa, a social robot who provides daily structure, can ask questions.
9. GPS tracker, which knows where the client is and can measure activity.

Familiarity with the technologies

7. *Do you know one or more of these technologies, or similar products? Do you or your clients already use them themselves?*
 - a. *Check: Do you have any questions about any of the technologies or the data generated by the technologies?*

HAAL dashboard

The different technologies all generate data that directly or indirectly tell something about the health, well-being or safety of community-dwelling people with dementia. Such data can, for example, help caregivers assess and determine the care and support needs of their clients. Within the HAAL project we are developing a dashboard in which we make the data from the various technologies available in one place. The dashboard is in fact a tool for linking the data back to caregivers (and any other stakeholders) in a manageable way so that they can do something with it. To investigate how we can properly design this dashboard, and how it can be applied in practice, we conduct these interviews.

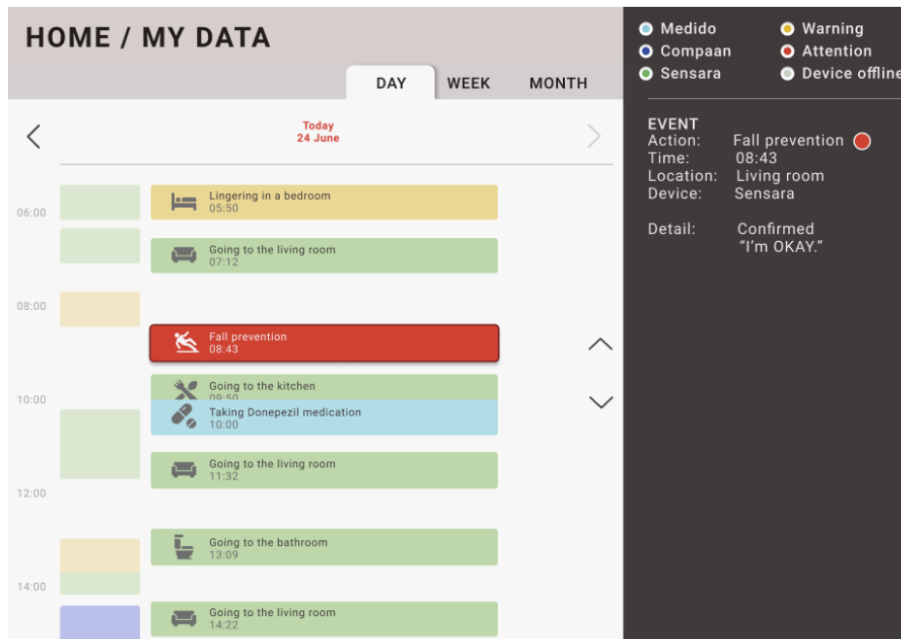
8. *Now that you have heard this information, what information (data) would you find interesting to see in the dashboard?*
(Note: We are not interested specifically in the function of the technological tools, but rather in the data they generate.)
 - a. *Why do you consider this information to be relevant?*
 - b. *How would this information affect your work? For example, which interventions would you deploy based on this data?*
 - c. *Do you have any other information that you find relevant?*
9. *For whom else then yourselves do you think this dashboard could be relevant? Are there any types of colleagues in your organization, or people outside your organization such as the client himself, for whom such a dashboard can be useful? Why?*

To make it more tangible and understandable what the HAAL dashboard could look like, we have made up two scenarios. Both scenarios are fictional and intended to inspire you about what the dashboard might look like.

Scenario 1

Scenario 1 is about a low-complex, descriptive dashboard. Below you see an example that illustrates the scenario. *[show illustration in powerpoint]*

The dashboard provides a picture of the current situation in the client's home based on data about e.g. exercise, nutrition, medication intake and social contact. The dashboard is mainly intended to make various data available to healthcare providers in one place, so that they do not have to look in separate overviews and apps. Important about this scenario is that no major interpretation of the data has yet been made in the dashboard, except that colours are used to indicate, among other things, whether there is potentially a risky situation (red), a situation that requires attention (orange) or if there is no action appears to be required (green). of the descriptive dashboard.



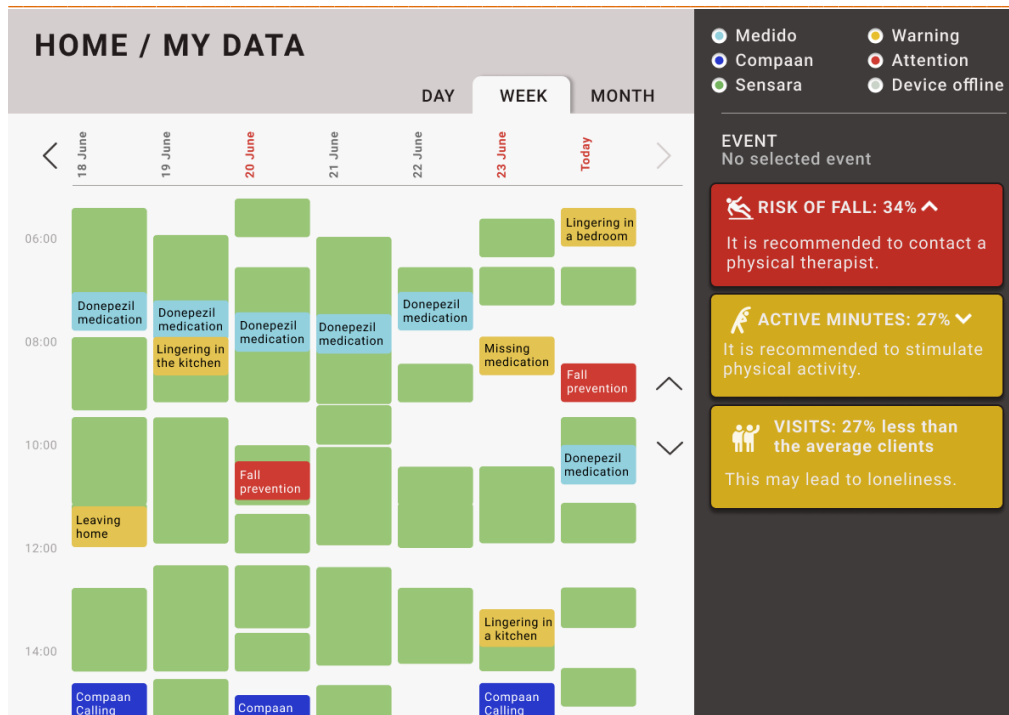
10. To what extent do you think such a descriptive dashboard could be of added value in home care for older people with dementia?

- a. *What do you think of the way in which data and information is fed back to the user, i.e. of the interpretation that the dashboard makes on the data?*

Scenario 2

Here you see an illustration for scenario 2. *[show illustration in powerpoint]*

Scenario 2 is an advanced dashboard that automatically translates the data into insights and recommendations for healthcare providers. In this collaborative dashboard, for example, based on data, warnings or risk scores are automatically displayed about possible emergency situations such as a fall. And the dashboard makes recommendations on certain possible follow-up actions, for example to stimulate a client's physical activity if the data shows that they have been relatively inactive in the past few days.



11. To what extent do you think such a proactive and collaborative dashboard could be of added value in the home care for older people with dementia?

a. *What do you think of the way in which data and information is fed back to the user, i.e. of the interpretation that the dashboard makes on the data?*

Role of AI

Regardless of whether the actual development of the HAAL dashboard will move more towards scenario 1 or 2, AI will play a central role. The term AI stands for artificial intelligence and refers to the attempt to mimic parts of human intelligence in machines. With AI you can think about computers that make predictions, recommendations or decisions based on data and decision rules that are programmed into the systems. An algorithm tells the computer which specific steps must be performed in a specific order to achieve a certain end result.

12. Do you understand what AI is? And if so, what examples of AI come to mind from everyday life? These don't need to be examples in healthcare.

<Depending on their answers, you can give examples, such as:

- voice assistant Siri from Apple (including speech-to-text)
- Google search engine (including both the search results and the word suggestions you get when you type a few letters in Google)
- navigation on Google Maps
- personalized timeline and suggestions on Facebook, Spotify, Netflix
- spam filters in your mailbox

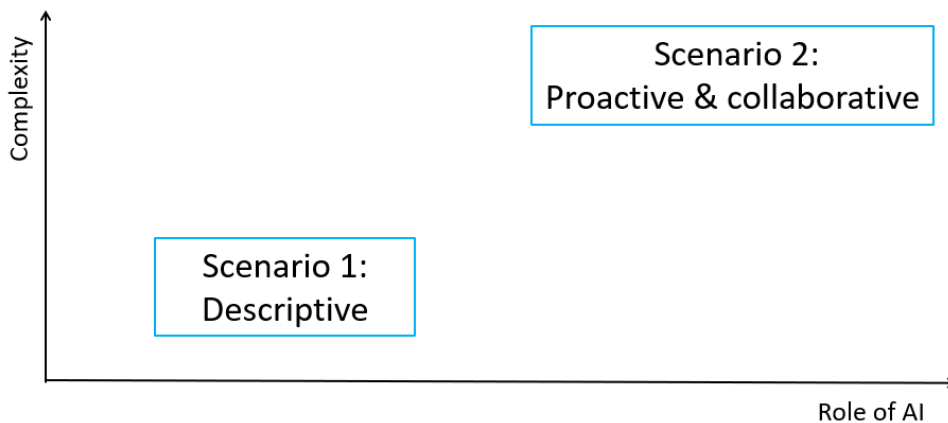
- *drones that monitor crop growth or count livestock in agriculture*
- *image interpreting cameras that can determine if someone has fallen*>

13. What role do you hope AI will have in your work?

Comparison scenarios

The figure below shows a simple representation of how the two scenarios we just discussed differ from each other in terms of complexity and the extent to which AI plays a role. In scenario 1, AI plays a minor role. Algorithms determine here, for example, with which colour certain information is marked, but apart from that, the interpretation by the computer and the complexity of the analyses is low. Therefore, it may take a little more thinking to extract meaning from the data.

In scenario 2, AI plays a major role. Based on data, the dashboard automatically makes predictions and recommendations about follow-up actions. The collaborative dashboard has a higher complexity and based on predictions and recommendations, this dashboard takes some of the cognitive workload over from the caregiver.



14. If you had to choose between scenario 1 and 2, which would you prefer?

- Why?*
- <repeat the reason and ask again why it is important>*

Part 2

Risks and downsides

During the HAAL project, we investigate how we can use AI in a meaningful and responsible way as decision support for caregivers and other stakeholders. We also pay attention to the ethical aspects of AI, because the use of AI in healthcare offers many opportunities, but also entails certain risks.

15. Can you think of any risks or disadvantages that come with using a dashboard like in scenario 1 or 2?

- a. *For whom are these risks / disadvantages?*
- b. *Why? (for example, privacy can be a risk for several reasons, so it's interesting to ask what's behind it)*

16. Do you see any differences between scenario 1 and 2 in terms of risks or disadvantages? If so, which one?

17. What are your first thoughts about how to deal with the discussed risks or disadvantages? Do you have any thoughts about how the developers of the dashboard, or users themselves, can deal with this?

Responsible Innovation themes

We first asked you openly about the risks and points of attention that you see for the responsible use of the HAAL dashboard. We will now also ask you about a few themes that are important to take into account when deploying AI. For each theme, we first briefly tell you something about the theme, and then ask you what risks you see in this area. Please tell as much as possible about what comes to mind when we discuss the themes. We ask you about ‘risks’ related to each theme. If something does not really feel like a big risk to you, but rather like a ‘point of attention’ – i.e. something to be aware of and take into account – then it is just as relevant to share with us.

1 Inclusion and equality

One theme to consider when deploying AI is inclusion and equality. AI systems must be designed and deployed in such a way that they can support the broadest possible target group (i.e. people of various ages, backgrounds and personal characteristics) and that the datasets they learn from, for example, are not biased and therefore disadvantage certain individuals or groups of people. If, for example, the HAAL dashboard learns from data of mainly older clients, there is a risk that the system makes unreliable predictions and recommendations about younger clients (e.g. people diagnosed with young onset dementia).

7. Can you think of any other factors to take into account to make the HAAL dashboard representative and fair?

8. Do you see any risks with respect to inclusion and equality when developing and using the HAAL dashboard?

- a. *Are these risks different for scenario 1 than for scenario 2? If so, how?*
- b. *What are your thoughts about how to deal with the discussed risks?*

2 Human autonomy

Another important theme in AI is human autonomy. The use of AI systems can lead to situations in which certain decision-making, or steps in decision-making, are partly or even completely transferred to technology.

9. Do you think that the use of the HAAL dashboard can have an impact on the autonomy of users? Do you see any risks in this area?

<If they doubt or answer no, you can explain: even if technology only makes recommendations, it consciously or unconsciously quickly influences the decisions you make as a person. Moreover, AI is often something that mainly plays a role in the background. As a user, you often do not see (or only to a limited extent) how certain input is converted into results by algorithms. Think for example of the search engine of Google: as a user you do not see how the system determines which results you get for a certain search.>

- a. *So what do you think of this in the context of the HAAL dashboard? Could it affect your autonomy as a caregiver?*
- a. *And do you also see a possible influence on the autonomy of the client?*
- b. *Do you see any differences between scenario 1 and 2 in terms of risks regarding people's autonomy? If so, which ones?*
- c. *What are your thoughts about how to deal with the discussed risks or disadvantages?*

3 Transparency and explainability

We are also interested in your view on the transparency and explainability of the outcomes of AI. Many AI systems are complex, which means that it is not always entirely clear how their outcomes come about. The performance of AI systems often improves as they can learn from larger amounts of data - because more data often means more evidence on which outcomes are based. But this is often at the expense of transparency and explainability of AI – it is then not clear, or only to a limited extent, how AI processes certain inputs into outcomes.

10. What would be more important to you?; that the dashboard shows you the most accurate recommendations OR that you can fully understand the recommendations provided by the dashboard and interpret how the system came to these insights? Can you elaborate why?

11. Do you see any risks with respect to transparency and explainability in the HAAL dashboard?

- a. *Are these risks different for scenario 1 than for scenario 2? If so, how?*
- b. *What are your thoughts about how to deal with the discussed risks or disadvantages?*

4 Responsibility and liability

While AI systems perform specific tasks, for example by making predictions or recommendations, the question is who can take responsibility for the consequences of those outcomes of AI.

12. Of course, it is possible that technology like the HAAL dashboard sometimes points users in the wrong direction. For instance, imagine a scenario where based on the currently provided information and recommendations from the HAAL dashboard, a caregiver concludes that no additional measures are needed to reduce the risk of falling for this client. The next week this turns out to be an adverse decision, because the client fell anyway

Who do you think should be liable if a decision made based on the dashboard turns out to be an adverse decision about the care for a client? In other words, who do you think should be responsible if something goes wrong when using the HAAL dashboard?

13. Do you see any risks with respect to responsibility and liability?

- a. Are these risks different for scenario 1 than for scenario 2? If so, how?*
- b. What are your thoughts about how to deal with the discussed risks?*

5 Well-being and safety

AI systems should of course not have negative consequences for people. They must be safe, accurate and effective before being deployed.

14. Do you see any risks in this area?

- a. Are these risks different for scenario 1 than for scenario 2? If so, how?*
- b. What are your thoughts about how to deal with the discussed risks or disadvantages?*

6 Align with, and adapt to circumstances

Furthermore, AI systems must of course match the context of and circumstances in practice, even if these change over time. In fact, they should be developed and introduced only if they can be adequately and fully integrated in the care process.

15. Do you have any ideas about what in this sense would be needed to ensure that AI systems match the needs and context of care practice?

- a. Do you see any risks?*
- a. Are these risks different for scenario 1 than for scenario 2? If so, how?*
- b. And what are your thoughts about how to deal with the discussed risks?*

Closing the interview

We are at the end of this interview.

16. Do you have any other suggestions or things you would like to discuss about the social and ethical impact of the HAAL dashboard?

We use the results of the interviews to reflect, together with the other organizations we collaborate with in this project, on what is important for the meaningful and ethical application of the HAAL dashboard. We try to incorporate the insights from users directly into the development process. We will also write publications based on the results, so that other innovators and users of AI applications in healthcare can also learn from our lessons about responsible innovation.

17. Do you know of any other healthcare professionals that we could interview? These can be different types of professionals for which such a dashboard could be relevant, for example district nurses, geriatric specialists, but also data experts, for instance.

18. May we approach you more often in later research to explore this topic in more depth?

<Thank you and goodbye>