EACare: Embodied agent to support elderly mental wellbeing

Final report

Hedvig Kjellström, Jonas Beskow, Joakim Gustafson, KTH
Miia Kivipelto, Krister Håkansson, KI

Summary

The main goal of the multidisciplinary project EACare is to develop an embodied agent – a robot head with communicative skills – capable of interacting with especially elderly people at a clinic or in their home, analyzing their mental and psychological status via powerful audiovisual sensing and assessing their mental abilities to identify subjects in high risk or possibly at the first stages of cognitive decline, with a special focus on Alzheimer's disease. The interaction is performed according to the procedures developed for memory evaluation sessions, the key part of the diagnostic process for detecting cognitive decline.
This new diagnostic system will be one of the means by which medical doctors evaluate people for cognitive decline, in parallel to the existing methods such as memory evaluation sessions with a (human) clinician, MRI scans, blood tests, etc. Different parts of the framework can also be used for other purposes, such as to develop tools for dementia preventive training and for decision support during clinical memory evaluation sessions.

1. Background, objectives and organization of the project

1.1 Background, motivation and long-term vision

With an increasing number of elderly in the population, dementia is expected to increase. Dementia is an umbrella term for a range of diseases that affect the brain and corresponding cognitive functions in old age. The most common type of dementia is Alzheimer's disease (AD), carried by around 60% of all persons with dementia. Many old persons also have vascular lesions in the brain which can lead to vascular dementia (VaD), alone or in combination with other types of dementia. A typical early symptom especially of Alzheimer's disease is gradually increasing difficulties to remember things that happened recently. Alzheimer's disease is a progressive neuropathological disorder with progressing impairment of cognitive functioning; gradually increasing inability to recall past knowledge and memories, the loss of the ability to retain new information, implying an expanding erosion of specific neural network systems involved in collecting information, storing them as memories, and keeping them available for recollection. In the end, the pre-existing numerous colorful recollections from past events, which defines each individual, fades away, or may be totally lost.

Although many other changes occur during the aging process, such as limitations in the range of motion, speed of gait, muscular volume and strength, acuity of taste, smell, vision and hearing, none of these factors are as central to the core of a person as is the loss of their collective experiences. Dementia disorders and related illnesses such as depression in the aging population are devastating for the quality of life of the afflicted individuals, their families, and caregivers. Moreover, the increase in such disorders imply enormous costs for the Swedish health sector. There is still no cure available for dementia.

However, population studies with around 40 years follow-up show that it might be possible to reduce the risk for dementia in later life through a number of preventative measures (Exalto et al., 2014; Livingston et al., 2020). There is now additional experimental support for increased efficacy of interventions that combine several of the life-style factors that previous observational studies have suggested (World Health Organization, 2021; Kivipelto et al. 2020). Early diagnosis and prevention measures are thus key to counteract dementia. The proposed system will have a significant impact along both these lines, thus potentially providing huge improvements in the lives of the hundreds of thousands individuals suffering from dementia at different stages, as well as for their families and caregivers.

Bringing together expert medical and technical partners, the EACare project aimed to develop an embodied agent with a particularly active role in the pre-diagnosis of cognitive disorders, screening and monitoring of the elderly’s psychological status and cognitive abilities and at the same time systematically investigate alternative therapeutic interventions for the treatment of mild cognitive impairment supported by the proposed system’s capacities.

1.2 Concrete goals and objectives

We formulated this vision in a more concrete way by setting up five milestones.

The original milestones of the project were as follows:

M1 (month 12): A Wizard-of-Oz prototype of a game-like system for therapeutic training against early dementia. The first milestone is a Wizard-of-Oz prototype of a system, simulating functionality for providing daily therapeutic training against early dementia. In a Wizard-of-Oz prototype the system’s automatic analysis of the elderly test users verbal communication is simulated by a human operator (not visible to the test user).

M2 (month 36): A system for diagnosis of early dementia, based on verbal analysis. The second milestone is a system that, based on the verbal semantic interaction pattern between the embodied agent and the human user, detects early signs of dementia, depression or other cognitive disorders.

M3 (month 36): A game-like system for therapeutic training against early dementia, based on verbal anal- ysis. The third milestone concerns the other functionality of the proposed system, functionality for providing daily therapeutic training against early dementia.

M4 (month 48): A system for diagnosis of early dementia, based on verbal and non-verbal analysis. The fourth milestone is the integration of non-verbal signals, such as gaze patterns, facial expressions or head and upper body motion, into the diagnostics system.

M5 (month 60): An integrated system for diagnostics and training. The final milestone in the project is the full integrated system using an embodied agent that interacts with the user, providing daily therapeutic training against early dementia while at the same time monitoring the user and diagnosing early signs of dementia, depression or other cognitive disorders. When appropriate, the system gives feedback to the user with advice on further medical evaluations, and can also provide support to a medical specialist in the form of recorded patient behavior history.

Milestone M1 was carried out according to schedule, based on a simplified version of the procedures for memory evaluation used at the Memory Clinic of the Karolinska Hospital (KS), where all the KI researchers in EACare were active.

However, during the course of the project, it became apparent that the full procedure for interaction between a clinician and a patient during memory evaluations and therapeutic training are varying in rich manners, and also vary greatly with different clinicians and different patients. (See Sections 5.1 and 5.2 for a discussion on the multi-disciplinary collaboration process.)

We found that it is not feasible to develop protocols that describe this procedure in a deterministic way so that it could be implemented in a dialogue system, as proposed in the application. We therefore saw the need for a large data collection where we would record all aspects of the interaction between patients and clinicians, in real memory evaluations in the regular clinical practice at KS**. This data collection is a significant scientific achievement in itself, and we would therefore like to consider it a milestone. This is an additional contribution of EACare, and to accommodate it within the time and resource budget of the project, we refocused the project on memory evaluations only, thus removing the task of developing a therapeutic training tool as well.**
Moreover, the development of the robotic system depends on this data, and Milestone M2 thus needed to be moved forward in time.
Third, we also saw a benefit of developing the diagnostics part of the system independently of the embodied interaction part of the system.

These three changes lead us to, halfway through the project in the beginning of 2019, define the following updated milestones:

Mnew1 (month 12): A Wizard-of-Oz prototype of a memory evaluation system. The first milestone is a Wizard-of-Oz prototype of a system, simulating functionality for carrying out a memory evaluation to detect early dementia. In an Wizard-of-Oz prototype the system's automatic analysis of the elderly test users verbal communication is simulated by a human operator (not visible to the test user).

Mnew2 (month 42): A collection of data from memory evaluation sessions with real clinicians and patients in the regular clinical practice in the Memory Clinic at KS. Speech is recorded as well as video of both patient and clinician as well as heat video, gaze, heartrate, writing and drawing patterns of the patient. Each data session is labeled by a team of doctors, speech therapists and psychologists as part of the regular diagnostic session for this patient.

Mnew3 (month 54): An embodied dialogue system that carries out memory evaluation sessions in the same (but stylized) manner as the sessions recorded in Milestone *Mnew2*. The dialogue system is trained with the recorded verbal and non-verbal data from Milestone *Mnew2.*

Mnew4 (month 54): A diagnostics system that detects signs of cognitive decline from the patient's spoken communication and non-verbal behavior (such as gaze patterns, facial expressions or upper-body motion). The diagnostics system is trained with the recorded verbal and non-verbal data from Milestone *Mnew2*.

Mnew5 (month 60): An embodied dialogue and diagnostics system that carries out memory evaluation sessions and detects signs of cognitive decline from the patient's spoken communication and non-verbal behavior. The final milestone in the project is the full integrated system, the combination of the systems developed in Milestones *Mnew3* and *Mnew4*.

1.3 Organization

EACare was a collaboration between three research environments, which provided complementary expertise in different aspects of the project (see Appendix A1-A3). The project was coordinated by Prof Hedvig Kjellström, Robotics, Perception and Learning, KTH, and PIs were also Prof Joakim Gustafson and Prof Jonas Beskow, Speech, Music and Hearing, KTH, and Prof Miia Kivipelto, Neurobiology, Care Sciences and Society, KI.

Robotics, Perception and Learning, KTH

The research group of Hedvig Kjellström at Robotics, Perception and Learning (RPL), KTH is directed towards the development of methods for enabling artificial agents to behave and reason in ways interpretable to humans, and also to interpret human behavior and reasoning. The group consists currently of 5 PhD students and a varying number of MSc students and research engineers. There is tight interaction with the other research groups at RPL, a vibrant internationally competitive research environment in Machine Learning, Computer Vision and Robotics, with 13 faculty and more than 50 PhD students, funded by numerous current and previous ERC, EU FP7 and H2020, SSF, KAW and VR grants. RPL arranges three different reading groups, several PhD courses and frequent seminars with internal or external speakers.

During the course of the EACare project, two ongoing projects in Hedvig Kjellström's group that relate to the research of EACare. There were common aspects in how non-verbal behavior was observed and recognized in order to deduce underlying aspects, and with respect to the Machine Learning methods used to learn latent representations of the observed behavior.

Causal Healthcare (funded by SeRC, the Swedish e-Science Research Centre). In this project, we develop Machine Learning methods that discover causal structures from medical data, for automatic decision support to medical doctors in their work to diagnose different types of injuries and illnesses. The purpose is to determine the underlying causes of observed symptoms and measurements, making it possible to semi-automatically reason about the potential effects of different actions, and propose suitable treatment.

The project funds one PhD student and one Research Engineer, and is a collaboration with researchers at KI and Microsoft Research Cambridge, UK.

EquineML: Machine Learning methods for recognition of the pain expressions of horses (funded by VR, the Swedish Research Council). Recognition of pain in horses and other animals is important, because pain is a manifestation of disease and decreases animal welfare. The aim of this project was to develop a method for automatic recognition of pain in horses. The method employed video of the horse, and detected and recognized behavioral patterns related to pain, in an automated manner when the horse perceives itself as being alone. The project funded one PhD student and was a collaboration with researchers at the Swedish University for Agricultural Sciences.

Speech, Music and Hearing, KTH

The speech group at the Department of Speech, Music and Hearing (TMH), KTH has a strong background in the area of speech communication and technology. The Speech Group at KTH has been evaluated twice as part of the KTH International Research Assessment Exercise. In RAE2008 the evaluators commented the speech group with: This is an outstanding, world leading research group among the top and most respected (a national asset), and in RAE2012 its unit of assessment (Applied Computer Science) was selected as one of the most excellent at KTH, and the comments were: Research output is internationally excellent in all fields, with a substantial number of units reaching the level of world-leading quality.

During the course of the EACare project, TMH had several ongoing projects ranging from basic science in spoken communication and human-human interaction (e.g., VR and RJ-funded projects TillTal, MultiProm, ProGest, and SamBlick, that investigated fundamental functions of prosody, gesture and gaze in spoken interaction) to applied research in human robot interaction and speech technology. e.g. VR projects StyleBot, CorDial and RoboLearn that all dealt with various aspects of social human robot interaction involving expressive humanoid robots. Both of these classes of projects were relevant for the work in EACare - the former since they provided knowledge and analysis methods for human non-verbal and interactional behavior (on which we base diagnostic work in EACare), and the latter since expressive social robots were an integral part of the envisioned end-use case scenario in the project.

TMH is also heavily involved in infrastructure for speech research and speech technology. We are currently part of the VR-funded infrastructure Nationella Språkbanken. The speech technology tools and speech corpora developed in the proposed project will be made available to the research community through http://sprakbanken.speech.kth.se. The data collection activities in EACare are backed by this infrastructure, and the parts of the data that can be anonymized can also be made available through its channels.

Neurobiology, Care Sciences and Society, KI

The department of Neurobiology, Care Sciences and Society (NVS) is one of the largest departments at KI, consisting of eleven divisions with shared departmental management. Four centres are linked to the department; two of them of relevance to EACare: The Aging Research Center and the Alzheimers Centre. The mission of the Aging Research Center is to understand the biomedical and psychological aspects of the aging process in relation to social and physical contexts across the entire lifespan. The ultimate goal is to improve the health and well-being of older individuals by meeting the challenges and embracing the opportunities presented by the aging population.

During the course of the EACare project, there were several other projects in the research group at NVS, two of them of relevance to EACare: The Nordic Brain Network (NBN), led by Miia Kivipelto, and the Center for Alzheimer's Research (CAR). NBN is affiliated with various centers of excellence at KI, University of Eastern Finland, National Institute for Health and Welfare in Helsinki, and NBN is involved in several international research projects with over 50 research institutions across the world. Members of NBN have identified several modifiable and treatable risk factors for dementia and Alzheimer's disease, including interactions between genetic and environmental factors. They also conduct multidomain lifestyle interventions for the prevention of dementia, including exercise, diet, social stimulation, cognitive training and vascular risk management. Of these studies, the FINGER study has been described as a break-through in finding ways to counteract cognitive deterioration in elderly persons at risk of dementia (Kivipelto and Håkansson, 2018). NBN is actively involved in several phase I-III randomized controlled trials with potential disease-modifying drugs. NBN contributes additionally to understanding biological mech-anisms underlying the onset and progression of Alzheimer's disease, and to developing biomarkers that can aid early diagnosis, prognosis and treatment monitoring in Alzheimer's disease. More recently, they are developing and testing novel eHealth tools for dementia risk detection, dementia prevention, and for clinical decision support.

1.4 Changes made during the project

The project is highly multidisciplinary, and as such, it has been hard to foresee some factors that have had quite a large impact on the outcome. The first major change in the project was the data collection effort in the clinical practice at the Memory Clinic at Karolinska Hospital in Solna, which we realized was necessary only after starting the development of the interaction protocols (see Milestone *Mnew1* above). We made the decision to redefine the milestones in the beginning of 2019.

The work in the project was at the same time correspondingly reorganized to meet the new milestones as described in Section 2.1 below. A further reason for the replanning of work packages was that after about one year, we lost two key persons, one at KTH and one at KI. This was naturally a quite large setback in the development work, and meant that the work of the remaining three persons in the project, all PhD students, was affected. To counteract this and make it possible for the remaining three PhD students to carry out their projects, we replanned the work packages slightly.

The process of starting data collections was challenging and affected by several external factors. One such has been the new fast-track process at the new Memory Clinic at KS, which was designed during the first year of EACare, with explicit focus on integrating research and clinical practice. **It enabled us to perform data collections in the regular clinical practice – something that is quite unique internationally.** We decided to wait with the data collections until it was in place, as they would have been much more detached from the clinical reality in the old Memory Clinic, where there was less communication between research and clinical practice.

Furthermore, the application for ethical approval of data collection took 14 months. These factors meant that we were only able to start the data collections in August 2019. The data collections proceeded until March 2020 when they were halted by the covid-19 pandemic, and then recommenced in October of 2021. As a result of these delays, we were forced to change the project further.

Firstly, as described in Section 2.1 we decided to make a demo system with basic functionality in WP 2, which is modular in order to allow gradual incorporation of research results from WPs 3-5.

Secondly, in order to not delay the three PhD projects in WPs 3-5, these were replanned to make use of publicly available benchmark datasets rather than relying on the data collected in the project. This caused a very unfortunate separation between these three projects and the data collection and system building activities in WPs 1-2. In order to mitigate this and employ the data for analysis and training of machine learning models, a fourth PhD student was hired in 2020. He has a unique background as both computer scientist and psychologist, and provides an important bridge between the KTH and KI groups even after the finish of the project.

Finally, as described in Section 2.1 and 3.7, a new exploitation ("nyttiggörande") task has been added to the project, the development of software for a web-based version of the system. This online version of the diagnostics system uses only the speech signal from the user, rather than the multiple modalities envisioned for the full physical version. Moreover, the system does not have a physical Furhat agent, only an iPad interface. This task is carried out during the spring of 2022 and the work is led by the new PhD student.

2. The research of the project

2.1 Scientific accomplishments and results

These five and a half years have been incredibly interesting and productive in building knowledge in the intersection between Geriatrics and dementia research, and Machine Learning and Social Robotics (see Sections 5.1 and 5.2). Below, we describe the work package organization and its evolution over the course of the project, as well as the research within each work package. In the following, references to our own papers are on the form (A) and refer to the list in Appendix A4.

Original work package organization

The work packages in the original project plan were:

WP1 (month 1-12): Interaction Protocol Design. The goal of this WP is the design of clinically-informed user-agent interaction protocols, similar to those protocols that are currently used in clinical practice.

WP2 (month 1-60): Verbal and Non-Verbal Human Communicative Behavior Analysis. The goal of this WP is the design and development of an advanced human-computer interface capable of recording, analyz- ing and interpreting verbal and non-verbal behavior of the user by maintaining multimodal interaction with him/her as required by the protocols outlined in WP1.

WP3 (month 1-60): Learning Methods for Human Communicative Behavior Analysis. The goal of this WP is the design and implementation of a novel set of machine learning algorithms that will enable the computer framework in WP2 to assess the users emotional and cognitive state, based on the users verbal and non-verbal communication.

WP4 (month 13-60): System Integration and Interaction Design. The goal of this WP is integration of all the above into a series of accessible embodied healthcare system prototypes.

WP5 (month 13-60): System Evaluation. The goal of this WP is systematic evaluation of the embodied healthcare system prototypes developed in WPs1-4.

In order to make the plan more robust to delays, we proposed in 2018 an altered plan where the tasks were separated so that WP 1-2 comprised more application-specific (high-TLR) aspects and were headed by the two post-doctoral researchers, and that WP 3-5 comprised more general basic-scientific (low-TLR) aspects and were headed by the three PhD students. The PhD students each would carry out their project, where they developed basic methodologies which they would apply in the demonstrator systems of the project in collaboration with the post-doctoral researchers. We also proposed to accommodate the data collection efforts described in Section 1.4 above into WP 1-2. The new work packages are defined in the following subsections, each with a description of the research carried out within them.

WPnew1 (month 1-60): Interaction Protocol Design, Data Collection, and Consequence Analysis

Led by *KRISTER HÅKANSSON*
Papers: (H), (K), (N), (O), (X)

The purpose of this WP is to enable the flow of knowledge, data and information between the Computer Science research at KTH and the Geriatrics research at KI and clinical practice at the Memory Clinic at KS, and to make sure that all method development in the technical work packages is grounded in the needs of the actual clinical reality and does not detract in any way from the quality of the ordinary cognitive evaluation. To exemplify, several adaptations in cognitive testing procedures were made to enable behavior registration of relevance for this project, mainly by transforming paper-and-pen testing into performing the corresponding tests on an iPad. Ensuring that these translations were equivalent to traditional testing, was a major priority in the development of the design.

The writing of the ethical application played an important role in developing and specifying the design; detailing measurements and procedures helped to identify potential barriers and develop procedures how to overcome them in the final design that resulted and that was subsequently approved by the ethical committee. Important considerations were a) the protection of personal integrity as far as possible in a project that by necessity had important challenges in this respect (such as recording and analyzing patient behavior and speech), b) taking advantage of rapidly emerging technological advances in measurement equipment and the opportunities for increased precision and efficiency they offered, and c) thirdly alignment to the new fast track clinical process at the Memory Clinic at KS. Through this fast track clinical process, evaluations now take 1-2 weeks instead of the more common 2-3 months at other clinics. Besides being more comfortable for the evaluated patients, this fast-track process is also internationally quite unique in its tight interaction between research and practice.

Task 1.1: Interaction protocol design (month 1-36). The goal of this task was the design of clinically-informed user-agent interaction protocols, similar to protocols currently used in clinical practice, that will be carried out in Task 1.2 to collect a rich set of observations, both verbal, in the form of specific replies to questions, and non-verbal, e.g., natural or aberrant body movements and facial expressions, of a natural or a deranged condition.

Task 1.2: Data collection (month 31-still ongoing). The goal of this task was multi-modal recordings of patient behavior and patient-clinician interactions during clinical memory evaluations. The recordings took place and were aligned with the general clinical practice in the Memory Clinic at KS. Multi-modal behavior data from both patient and clinician are recorded using a range of sensors such as voice and speech from microphones, facial and body motion from regular RGB video, gaze direction and movements from eye trackers, changes in galvanic skin conductance, and heart rate. Moreover, video of the entire scene was captured, to enable analysis and labeling of the behavior of both persons. This labeling/analysis was done once a week in a session for the first 20 patients where clinicians and KTH researchers went through the recordings done earlier in the week. The recorded multi-modal behavior data will be used in WPnew 2-5. Moreover, the scene recordings and the knowledge and experience from the labeling/analysis sessions, will be of interest in training of clinicians and development of evaluation processes.
In more detail, the data collection was performed as follows:
1. When any of the three geriatric medical doctors involved in this project (Marie Rydén, Alexandre Bonnard or Sara Garcia at the Solna cognitive clinic of the Karolinska University Hospital) received a referral for examination of possible cognitive disorder from an external care-giver, they decided if the patient fulfilled inclusion criteria and did not meet any of the exclusion criteria.
2. Patients who were selected at this stage by the clinician were sent an invitation letter to participate in the project. This information contained information about the purpose of the project, what it means to participate, the right to abstain, and contact information.
3. Within a week after the invitation was sent out, the patient received a telephone call by a nurse to investigate whether the person accepts the invitation or not. The person was also invited to ask any further questions they might have about the project and their participation.
4. When the patient who had accepted the invitation arrived at the clinic, they met a nurse who asked them to fill in the consent form, if they still wanted to participate. They were also invited to ask any further questions they might have about their participation.
5. If they consented to participate, they received a bracelet (a smart watch) to record pulse changes, galvanic skin resistance and movements.

6. The doctor then accompanied the person to the room where the clinical investigation and meeting took place.
7. The doctor started the simultaneous and synchronized recording of the different devices. Beside the smart watch (with accelerometers, gyroscopes, and pulse sensor), which began recording already in step 5, these additional devices were:
(a) A video camera to record the face of the patient during the clinical interview and test.
(b) A video camera that recorded the face of the clinician during the clinical interview and test.
(c) A video camera that recorded an overview of the entire scene during the clinical interview and cognitive testing.
(d) A depth-sensing camera that recorded the 3D face position and the facial expression (so-called face geometry ”blend shapes”) of the patient during the clinical interview and testing.
(e)An IR camera that recorded the ambient temperature and the skin temperature of the patient during the clinical interview and testing.
(f) An iPad used for the the cookie test, the paragraph-reading test, and several parts of the MoCA test, and recorded the patient’s pen strokes during the trail-making, cube, and clock tests therein.
(g) A gaze-tracker that tracked the patient’s gaze when using the iPad during the cookie test, the paragraph reading test, and the MoCA test.
(h) An array of microphones that recorded audio from the clinical interview and test.
10. At the end of the clinical interview and testing the doctor terminated the parallel recordings.
11. The patient was thanked for having participated in the project and was again reminded that they have the right to abstain from participation also after this meeting, and that in this case the recording will not be used for further analysis and will in addition be erased.
12. A group of experts meets once a week to go through each recording that has been made since the previous expert session. At each of these meetings the responsible clinician participates along with at least one speech therapist, one psychologist, one nurse, and two persons who are experts in artificial intelligence analysis of verbal and non-verbal behavior. Each expert category is duplicated in the pool of experts, altogether twelve persons, so that at least one person from each expertise area can participate in each of these sessions.
13. The relevant part of the interview is approximately 20 minutes. We have developed a list of 42 behavior categories that will be used to mark different types of behavioral events with potential clinical relevance, including pauses, confusion, insight, repetitions and irrelevant digressions from the topic. A special focus will be on behaviors in conjunction with critical events that the clinician marked during the clinical examination, including the time point when the clinician arrived at a preliminary clinical judgment (that the patient is either healthy or suffering from an underlying disease). For all recordings we will in addition especially focus on the behavior during the first minutes of normal conversation and when three different tests are performed (the clock test (the person is asked to draw a clock that shows a certain time of the day), the Boston cookie theft description task (”the cookie test”), and during reading a progressively more complex text). During the expert group analysis session, audio-visual data as well as visualizations of sensor data will be used.
14. Each time any of the experts notices a behavior that potentially could indicate either a cognitive problem or that the person does not suffer from such a problem, that behavior is marked during the analysis, including the time stamp for subsequent automatic analysis.

15. After the analysis, each of the marked-up behaviors is, through support of artificial intelligence, related to events in other channels (temperature and pulse changes, motor activity data and changes in eye gaze).
16. Based on patterns of behavior signals that potentially indicate either cognitive health or a cognitive disorder, algorithms are developed for automatic analysis.
17. The resulting algorithms are finally tested through automatic analysis on a subset of 20 patients with either a known neurocognitive disorder or who are believed to be cognitively healthy, as judged by normal clinical examinations, including supporting biomarker data and results of cognitive testing.
18. If needed, the algorithms are fine-tuned and retested based on the outcome of the first test run.

When we restarted the recordings in October 2021, the expert sessions were removed from the protocol due to lack of personnel resources, only keeping the actual data recordings.

Task 1.3: Consequence Analysis (month 43-60). In this task, we will analyze how 1) the collected and annotated data, 2) the developed diagnostic method, 3) the embodied agent and dialogue system most efficiently can be used in the clinical practice, and potentially also in the education of future clinicians.
This task has been omitted from the project due to the covid-related delays.

WPnew2 (month 1-60): System Integration and Evaluation

Led by GUSTAV EJE HENTER
Papers: (R), (S), (U), (Ω)

This WP deals with the technical development of the embodied healthcare system, based on the Furhat robot and dialogue system. All three PhD students (see WPnew3-5) take active part in WPnew2, applying their methodology in different parts of the system.

Completed task: Wizard-of-Oz system conducting a memory test (month 1-12). As a concept demonstrator, we have implemented a Wizard-of-Oz prototype using the Furhat robot head and dialogue system, carrying out a range of established cognitive tests. The work in Task 2.1 will directly build on this prototype, adding more elaborate test procedures with a rich variability (see WPnew1) and employing methodology developed in WPnew3-5.

Completed task: Data recording setup design (month 28-30). The goal of this task has been to assemble the data recording setup. This has included building a table with stands for different sensors, but foremost, a piece of software for capturing and storing these heterogeneous, multi-device recordings in a robust and secure manner.

Task 2.1: System design (month 37-60). The goal of this task is the integration of results from WPnew1,3-5 into a sequence of accessible embodied healthcare system prototypes. The system will directly build on the prototype described above, adding automated dialogue design and diagnostic processes. The embodied dialogue system will work according to system design developed in WPnew3 and WPnew5, execute interaction protocols learned from data extracted in WPnew1 and, based on the information it has gathered about the user (see WPnew4), present detected symptoms to the human clinician overseeing the diagnostic process, and make suggested diagnoses.

Task 2.2: System evaluation (month 49-60). The goal of this task is systematic evaluation of the embodied healthcare system prototypes developed in Task 2.1. The system will be evaluated on real patients, as part of the clinical practice at the Memory Clinic at KS, according to the same procedures as the data collection described in WPnew1.
During 2019, we engaged the company Prototyp AB to develop a fully automated system prototype using the speech recognition and dialogue design features of the Furhat dialogue system API. The system performs a cognitive evaluation according to the protocol developed in Task 1.1, with spoken language modeled after one of the recorded cognitive evaluations. It is important to note that this prototype does not make a full diagnosis (which is the subject of research in WPnew4) but rather just follows the examination protocol and provides human-like dialogue feedback signals, such as head-nods and back-channel response. The system was designed in a modular fashion, to allow adding more sophisticated modules for back-channeling and user adaptation (WPnew3), diagnostics (WPnew4), and a richer and more varied non-verbal robot communication (WPnew5).
The system prototype has been shown publicly (see Appendix A5).

WPnew3 (month 1-60): Adaptive Spoken Human-Robot Interaction

Led by *PATRIK JONELL*
Papers: (A), (F), (M), (Q), (V), (Z)

This work package is a PhD project about method development for design of automatic dialogue systems.

This work presents several methods, tools, and experiments that contribute to the development of interlocutor-aware Embodied Conversational Agents (ECAs). Interlocutor-aware ECAs take the interlocutor's behavior into consideration when generating their own non-verbal behaviors. This thesis targets the development of such adaptive ECAs by identifying and contributing to three important and related topics:

1) Data collection methods are presented, both for large scale crowdsourced data collection and in-lab data collection with a large number of sensors in a clinical setting. Experiments show that experts deemed dialog data collected using a crowdsourcing method to be better for dialog generation purposes than dialog data from other commonly used sources.

2) Methods for behavior modeling are presented, where machine learning models are used to generate facial gestures for ECAs. Both methods for single speaker and interlocutor-aware generation are presented.

3) Evaluation methods are explored and both third-party evaluation of generated gestures and interaction experiments of interlocutor-aware gestures generation are being discussed. For example, an experiment is carried out investigating the social influence of a mimicking social robot. Furthermore, a method for more efficient perceptual experiments is presented. This method is validated by replicating a previously conducted perceptual experiment on virtual agents, and shows that the results obtained using this new method provide similar

insights (in fact, it provided more insights) into the data, simultaneously being more efficient in terms of time evaluators needed to spend participating in the experiment. A second study compared the difference between performing subjective evaluations of generated gestures in the lab vs. using crowdsourcing, and showed no difference between the two settings. A special focus in this thesis is given to using scalable methods, which allows for being able to efficiently and rapidly collect interaction data from a broad range of people and efficiently evaluate results produced by the machine learning methods. This in turn allows for fast iteration when developing interlocutor-aware ECAs behaviors.

WPnew4 (month 1-60): Automatic Diagnostics of Cognitive Function from Human Behavior

Led by *OLGA MIKHEEVA*
Papers: (B), (Y)

This work package is a PhD project about method development for contextual inference of cognitive state from communicative behavior.

In the current application, diagnostics of cognitive decline from behavior, the sequential data are measurements such as gaze, facial motion, body and head motion, heart rate, etc, and the underlying variable to be inferred is the cognitive status of the subject. Data set size constraints, combined with need for prediction uncertainty and interpretability, makes a probabilistic framework approach the most suitable.

Firstly we have carried out work on modeling time-series of biological signals using Gaussian Processes (B). This is of relevance for modeling of the time signals recorded during memory evaluations, as described in WPnew1.

Secondly, we have studied the problem of representing facial expressions so as to capture the factors that correspond to how humans perceive differences between different facial expressions (Y). The work on representation of facial expressions is presently continued within another project, and the focus is now disentanglement of facial motion due to speech from facial motion reflecting underlying emotional and physiological factors such as stress, engagement, focus etc.

A student project was carried out on developing a method for detecting early signs of Alzheimer's from speech only, and will be developed further in the exploitation/”nyttiggörande” project, see Section 3.7. We here use an American database of recorded speech during the Cookie Theft Test, which also is part of the memory evaluations at KS. This means that this work will be relevant to apply to the data recorded in WPnew1, when ready.

WPnew5 (month 1-60): Automatic Perception and Production of Human-Like Non-Verbal Behavior

Led by *TARAS KUCHERENKO*
Papers: (C), (D), (E), (G), (I), (J), (L), (P), (T)

This work package is a PhD project about method development for robot perception and production of non-verbal communicative behavior.

A large part of our communication is non-verbal: humans use non-verbal behaviors to express various aspects of our state or intent. Embodied artificial agents, such as virtual avatars or robots, should also use non-verbal behavior for efficient and pleasant interaction. A core part of non-verbal communication is gesticulation: gestures communicate a large share of non-verbal content. For example, around 90% of spoken utterances in descriptive discourse are accompanied by gestures. Since gestures are important, generating co-speech gestures has been an essential task in the Human-Agent Interaction (HAI) and Computer Graphics communities for several decades. Evaluating the gesture-generating methods has been an equally important and equally challenging part of field development. Consequently, this thesis contributes to both the development and evaluation of gesture-generation models.

This thesis proposes three deep-learning-based gesture-generation models. The first model is deterministic and uses only audio and generates only beat gestures. The second model is deterministic and uses both audio and text, aiming to generate meaningful gestures. A final model uses both audio and text and is probabilistic to learn the stochastic character of human gesticulation. The methods have applications to both virtual agents and social robots. Individual research efforts in the field of gesture generation are difficult to compare, as there are no established benchmarks. To address this situation, my colleagues and I launched the first-ever gesture-generation challenge, which we called the GENEA Challenge. We have also investigated if online participants are as attentive as offline participants and found that they are both equally attentive provided that they are well paid. Finally, we developed a system that integrates co-speech gesture-generation models into a real-time interactive embodied conversational agent. This system is intended to facilitate the evaluation of modern gesture generation models in interaction.

To further advance the development of capable gesture-generation methods, we need to advance their evaluation, and the research in the thesis supports an interpretation that evaluation is the main bottleneck that limits the field. There are currently no comprehensive co-speech gesture datasets, which should be large, high-quality, and diverse. In addition, no strong objective metrics are yet available. Creating speech-gesture datasets and developing objective metrics are highlighted as essential next steps for further field development.

2.2 Participating researchers

The participating researchers are listed in Appendix A3. The resource distribution has not been changed compared to the budget in the application.

Being in the Computer Science area, the project has a comparatively good gender balance; 3 (including the coordinator and a PI) out of 10 researchers are women, and 6 out of 10 experts are women.

2.3 Selected publications

The publications from the project are listed in Appendix A4. Section 2.1 indicates from which work package they were produced.

2.4 Important activities

Public presentations, demos, and news coverage from the project are listed in Appendix A5.

3. Strategic relevance

EACare is of large strategic relevance in the establishment of the fast-track diagnostic process at the new Memory Clinic at KS, as one of the core research projects that were first to be integrated with the clinical practice. It is frequently used as a show-case of multi-disciplinary collaboration at both KI and KTH.

3.1 Utilization of research results

Unfortunately we did not reach the stage where technology from the project is mature enough to be used clinically. We indeed hope for this to come true in a possible follow-up project, see Section 6.2.

3.2 Collaboration with healthcare

Collaboration with healthcare is at the core of this project, see Section 2.1.

3.3 Societal relevance

The societal relevance of the project is described in Section 1.1.

3.4 Intellectual property

The exploitation (“nyttiggörande”) part of the project we are currently developing software for a web-based version of the system, see Section 3.7. This will be of possible commercial / clinical interest in the future.

3.5 Implementation in healthcare

Unfortunately we did not reach the stage where technology from the project is mature enough to be used clinically. We indeed hope for this to come true in a possible follow-up project, see Section 6.2.

3.6 Public outreach

Public presentations, demos, and news coverage from the project are listed in Appendix A5.

3.7 Exploitation (“nyttiggörande”)

We have seen in the study (H) that different language and voice based features are clearly correlated to the biomarkers used in the present clinical methods for diagnosing Alzheimer’s disease. We will in the exploitation (“nyttiggörande”) project therefore develop a smartphone app that enables cognitive and memory tests based on voice control and voice analysis, in combination with different clinically validated cognitive tests.

While the physical system in the main project is intended for use at the memory clinic where patients are referred from primary care, the app will enable preclinical screening before seeking care, which can detect very early signs of dementia, before the person seeks medical help (which is a step that many hesitate to take). The app is intended to be used regularly and function as a quick cognitive health check that can be done e.g. on a weekly basis, and enables users to monitor their cognitive status over time, and observe the effect of various lifestyle changes.

In the first stage, we will invite former patients from the Memory Clinic at KS in Solna, as well as the interested public. Data from users will be collected (after ethical review and the individual's consent) and contribute to increased knowledge of changes over time, and where appropriate linked to previous data from examinations at memory clinics (e.g. biomarkers from spinal cord samples and MRI).

In healthcare, the app can also function as a first screening before you can see the general practitioner in primary care. It means that the basic examination, previously done by a general practitioner, can now be done by a computer system, which frees up doctor time.
In addition, the system can be used to enable early examinations and routine examinations at times when there are obstacles to physical meetings, such as during the covid-19 pandemic.

The ultimate goal of this software is to be used in healthcare, both clinically (as an aid for dementia diagnosis at health centers and memory clinics) and pre-clinically (in the form of an app where people can test themselves for advice on whether to apply for care for their symptoms).

For this to happen, the implementation of the method must first be fully developed, then the app further developed so that it is user-friendly and robust, and then validated clinically. This exploitation ("nyttiggörande") project will include the first and parts of the second stage, while the third stage, validation, is further ahead. We believe that a commercialization of the app is suitable for completing the second step and conducting clinical validation. There is a great need for this system in healthcare and we believe that the potential for commercialization is very good.

The exploitation ("nyttiggörande") project runs during the spring of 2022. Of the funding, 400 000 SEK will be used for salary to Birger Moëll, doctoral student at TMH, KTH, and main responsible for the development of the software. Moreover, 260 000 SEK will be used to engage experts at the Memory Clinic, KS, primarily psychologist Göran Hagman.

4. The graduate training of the project

4.1 Graduate training

No graduate courses were developed explicitly within the project. However, two graduate training activities should be noted.

In the summer of 2017, PhD students in the project participated in the eNTERFACE workshop, a 4 week project-focused summer school for collaborative research and development in the area of multi-modal human machine interaction where they designed a multi sensory data collection system for recording human social interaction. The software framework developed forms the basis for multi sensor synchronization in the data collection in the project.

In 2018, a team of PhD students including Patrik Jonell from the EACare project were selected as one of 8 teams worldwide to participate in the Amazon Alexa Prize, a global competition for university students dedicated to accelerating the field of conversational AI, where they used an innovative approach based on crowdsourcing to build a spoken dialogue system capable of social dialogue in unrestricted domains (V).

4.2 Development of independent group leaders

Gustav Eje Henter, postdoc in EACare between 2018 and 2020, was in 2020 accepted to a generously funded Assistant Professor position at TMH, KTH. His present research program is described on his homepage https://people.kth.se/~ghe.

4.3 Graduated PhD students

A list of graduated PhD students is given in Appendix A9, and a list of students about to graduate is given in Appendix A11.

5. Collaborations

5.1 Scientific collaboration within the project between groups

Considerable effort has been put down to ensure a lively multi-disciplinary interaction. Each site has a core group of researchers and PhD students. The combined group has met once every second week since the start of the project and all developments have been discussed and decided in this combined group. EACare also has a steering group of five persons (see Appendix A1), which discuss and decide on more strategic issues, such as budget planning, publication strategy, and main issues that have emerged and need to be dealt with.

The research environment at KS and KI is unique in the sense that the memory clinic at KS and the research unit at KI are located in neighboring corridors. The clinicians and the researchers work closely together and meet daily. This means that clinical practice and research are more integrated than perhaps anywhere else at KS, both in terms of geographical proximity and personal connections. The staff in the two corridors, i.e. researchers and clinicians, also have joint meetings to get to know each other better and to

discuss ongoing projects that involve both corridors. The integration is also facilitated by the fact that some of the researchers also work as clinicians in the clinical corridor, including one psychologist, one nurse, and one speech therapist. In addition, both medical doctors who are involved in the project, Marie Ryden, Sara Garcia and Alexandre Bonnard, who performed the clinical evaluations of all patients at the memory clinic, have their offices in the ”research” corridor.
Besides the two core groups and the steering committee, a wider network of around 20 researchers are involved in implementing the EACare project. In order to keep everybody informed and also to receive their input, we meet in this **larger network once a year during a two-day EACare workshop**. To develop new ideas and to get qualified feedback, external experts have participated as catalysts during these annual workshops.

5.2 Scientific collaboration between different disciplines and departments

See Section 5.1 for a discussion of this, as the different groups also span different disciplines and departments.

5.3 International collaboration, including participation in EU projects

Hedvig Kjellström’s group at KTH has several international collaborations in AI and Machine Learning as well as as medical and veterinary science applications, including Microsoft Research Cambridge, UK, Carnegie Mellon University, USA, University of California Davis and San Diego, USA, and Technical University of Munich, Germany. She is moreover an Affiliate Professor at the Max Planck Institute for Intelligent Systems in Tübingen, Germany, where she spends time regularly and conducts student exchange. Earlier collaborations through the EU projects PACO-PLUS and TOMSY included collaboration with Karlsruhe Institute of Technology, Germany, University of Edinburgh, UK, King's College London, UK, University of Stuttgart, Germany, and University of Göttingen, Germany. She has conducted several extended research visits, including MIT, Brown University and Xerox PARC, USA, and University of Surrey, UK.

The speech group at TMH have participated in a large number of national multi-site projects, and in 13 EU projects, e.g. NICE, MonAMI, Hearing@Home, IURO, GetHomeSafe, and BabyRobot. One of the PI:s spent 7 years as a senior speech and dialogue researcher at Telia Research, which gave him a large international network, both in academia and industry. World leading researchers in speech synthesis and prosodic modeling that have been guest researchers at KTH include Paul Taylor (at the time at CSTR), Julia Hirschberg, (Columbia University), Bernd Möbius (Uni Saarland), and currently Petra Wagner (Uni Bielefeld) is spending 12 month at KTH during 2018-2019 fully funded by RJ/Humboldt scholarship. Through EU projects he has collaborated with a large number of international speech and dialogue researchers, to name a few, Isabel Trancoso (Inesc) Jan Kleindeinst (IBM Research), Alexandros Potamianos (ICCS) and Stefan Kopp (Bielefeldt). The project team has collaborations with for instance David Traum, Stefan Scherer, Shri Narayanan (USC), Alex Rudnicky, Justin Cassell and Alan Black (CMU), Oliver Watts and Simon King (CSTR), Julie Carson-Berndsen (Univ College Dublin), Christer Gobl (Trinity Dublin), Bastiaan Kleijn (Victoria Univ), Junichi Yamagishi (NII) and Keiichi Tokuda (Nitech). The group also has informal collaborations with world-leading companies in speech synthesis (Acapela, Cereproc and Nuance).

The EACare research group at KI is part of the Nordic Brain Network (NBN), led by Miia Kivipelto. This is a group of approximately 50 researchers and PhD students in Sweden and Finland. Around 20 of them have their offices in the same corridor as the EACare group, next to the memory clinic at KS. NBN has a large number of ongoing research projects in the field of cognitive health and dementia together with partners in many different countries. We presently have a leading role in seven ongoing international projects:

- The Finnish Geriatric Intervention Study to Prevent Cognitive Impairment and Disability (FINGER),
- now extended to World Wide FINGERS
- The Multimodal Preventive Trial for Alzheimer's Disease (MIND-AD)
- Healthy Aging Through Internet Counseling in the Elderly (HATICE)
- The European Prevention of Alzheimer’s Dementia (EPAD)
- Impact of Nutritional Lipids on Neuronal and Cognitive Performance in Aging, Alzheimer's disease and Vascular Dementia (Lipididiet)
- Cardiovascular Risk Factors, Aging and Dementia (CAIDE)

Our key partners include University of Eastern Finland in Kuopio (all projects except EIT Health), National Institute of Health and Welfare in Helsinki, Finland (all projects except EIT Health), and Imperial College in London, UK (EIT Health, MIND-AD).

We also collaborate with a number of other partners in the various projects, including University of Oxford, UK (EPAD and EIT Health), University of Cambridge, UK (EPAD and HATICE), University of Leicester, UK (EPAD), Cardiff University, UK (EPAD), Trinity College Dublin, Ireland (EIT health), Delft University of Technology, Netherlands (EIT health), University of Amsterdam, Netherlands (MIND-AD), Saarland Uni- versity, Germany (MIND-AD), Universit Grenoble Alpes (UGA), France (EIT health), Barcelona Institute of Science and technology (BIST), Spain (EIT health), Eindhoven University of Technology, Netherlands (EIT health), Friedrich-Alexander-Universitat Erlangen-Nrnberg, Germany (EIT health), ETH Zurich, Switzerland (EIT health), Erasmus University Medical Centre Rotterdam, Netherlands (EIT health), Heidelberg University, Germany (EIT health and Lipididiet), LEITAT Technological Centerm Spain (EIT health), KU Leuven, Netherlands (EIT health), University of Copenhagen (UCPH), Denmark (EIT health), Tel Aviv University, Is- rael (Lipididiet), The University of Edinburgh, Scotland (EPAD), University of Tolouse, France (EIT health, HATICE, and MIND-AD), The University of New South Wales, Australia (World Wide FINGERS), National University Health System, Singapore (World Wide FINGERS), University of Southern California, USA (World Wide FINGERS), Wake Forest School of Medicine, USA (World Wide FINGERS), Kaiser Permanente North- ern California Division of Research, USA (World Wide FINGERS), Shandong University, China (World Wide FINGERS).

5.4 Collaboration with industry and other parts of society

The NBN group works closely with industry mainly in two projects in order to find partners that can help with implementation and give a broader access to citizens of products that result from our research. The European Prevention of Alzheimer’s Dementia (EPAD) is a unique and ground-breaking European initiative to streamline the testing and development of preventative treatments for Alzheimer's disease. This is the largest ever public-private partnership in Alzheimer's disease research. EPAD combines knowledge and expertise from 39 organizations across multiple sectors, both from academia and industry. The idea behind this project is to test treatments, both pharmacological and non-pharmacological, by combining them in a flexible way to develop optimal solutions for different groups of individuals. Industry partners are several leading pharmacological companies, including Pfizer, Lilly, Amgen, Roche, Janssen and Lundbeck.

EIT Health is a collaborative project to facilitate innovation and develop new technology such as Inter- net support and apps to improve the health of European citizens. EIT Health leverages the expertise of more than 140 leading organizations spanning key areas of healthcare, such as pharma, medtech, payers, research institutions and universities. Of these 140 partners, 66 are from industry, including Achmea Insurance Company, GE Healthcare, MEC - Interuniversity Microelectronics Centre, Janssen-Cilag GmbH, RISE Research Institutes of Sweden, Philips Electronics Nederland B.V., Medtronic, Siemens Healthineers, Roche Diagnostics GmbH, Thermo Fisher Scientific - Phadia AB, and UCB Biosciences GmbH.

The EACare project is designed as a bridge between academia and the health care sector, both in how the projects is carried out, by including clinicians in the research process, and in how the research results can be directly translated into clinical practice, thereby contributing to increased quality in the diagnostics and treatment of persons with early cognitive disorder. The benefits for the health sector are expected to be two-fold: a) the development of AI assistance to complement traditional clinical evaluations and b) potentially new discoveries of behaviors that can be used to develop clinical routines. In the first case there is a potential that robot assistance can become part of cognitive evaluation in the future if we succeed in developing algorithms that can be used for automatic evaluations in real time in parallel with the clinical interview. Automatic analysis has the potential to speed up the evaluation considerably, thereby saving time and money for a sector in society that is crippled by financial limitations and lack of personnel. In the second case, there is a chance that we will discover new signs of clinical relevance that can be translated into clinical practice; in one phase of the EACare project we will carefully analyze the recorded behaviors of patients during the clinical interview. These recordings will be analyzed by a group of experts from different areas of relevance, including geriatricians, psychologists, and speech therapists. If this detailed observation and analysis will lead to new discoveries of behavior signals that are relevant, observation and measurement of these signals could be included and enrich the diagnostic procedure in the future to make diagnoses more accurate and reliable.

5.5 How has the project affected the researchers?

The EACare project has significantly strengthened the Memory Clinic at KS as a research clinic. The close cooperation between clinicians and researchers has also contributed to significant mutual learning. One example is the observation sessions where recordings from “EACare patients” were jointly analyzed by an interdisciplinary group that represented diverse areas of clinical and scientific expertise: geriatricians, psychologists, neurologists, experienced nurses, and experts in behavior and speech analysis. In these sessions, observations were shared that were sometimes unique from a particular perspective, but not obvious to members from other areas. One learning point of particular importance was that careful observations of the recorded behavior often revealed events of clinical relevance that went unnoticed by the geriatrician during the actual interview, and that limitations in human information processing makes it virtually impossible to grasp the richness of behavior information during a clinical interview while it happens. Another learning point was that careful observation and classification of behaviors that are perceptible to eyes and ears (i.e. leaving out behavior signals that were additionally captured by our multi-sensor system) to a high degree predicted whether a person in the end of the diagnostic state-of-art evaluation would get a diagnosis of cognitive impairment or not.

At KTH, the EACare project has significantly strengthened the collaboration between the RPL and TMH division, by inspiring several other collaborative projects between the divisions. When the project started in 2016, contacts and collaborations between the divisions were quite sporadic and restricted to very few individuals. Now in 2022, there are a half-dozen large collaborative projects. The collaboration with KI/KS in EACare has also been a highly valuable learning experience for the KTH researchers, especially the opportunity to spend time and learn from the experts at the Memory Clinic at KS, and to perform recordings directly in the clinical practice, a unique opportunity enabled by the tight interaction between research and practice in the Clinic, see Section 2.1.

5.6 Awards

A list of awards is given in Appendix A14. Two conference papers have received prizes.

6. Continued work after the project is finished

6.1 Lessons learned

The most important scientific lesson learned is the following: A good way of employing Machine Learning methods for medical diagnostics is to train and combine detectors of different *digital biomarkers*, i.e. quite specific and well defined symptoms (that are more or less directly observable from the sensor signal), rather than aiming directly at a Machine Learning method that autonomously takes in a more composite behavior or signal and learns to recognize a more high-level output (such as a full diagnosis). There are three reasons for this:

1. Such a large “black-box” method would require very large amounts of data. In most medical applications, data is a very scarce resource.
2. Such a method would not be interpretable to the clinician making use of the method output and combining it with evidence from other examinations in the diagnostic process. The clinician needs to know *how* the method reasoned and came to the present conclusion, not only be notified about the conclusion itself.
3. A set of smaller methods for recognition of individual digital biomarkers can be combined with the current computer methods employed for fusing information from different types of other biomarkers, such as blood pressure, mri readings etc.

We have implemented this lesson in the last part of the project.

Regarding other aspects, the most important lesson learned is: Multidisciplinary work takes time.

1. The initial time needed to understand each other’s terminology and research challenges should not be underestimated.
2. In order to make full use of a collaboration with researchers from many different disciplines, the researchers must have the opportunity to spend extensive amounts of

time together in order to understand the details of each others' areas to a certain resolution level. This takes time during all phases of the project.

However, it should be emphasized that this extra time requirement is outweighed by the benefits of multidisciplinary work - if successful, the crosspollination between the disciplines can lead to truly novel results.

6.2 What will happen to the project?

Several different follow-up projects are in the pipeline. The aim is to integrate the EACare activities even further into the other research activity in the KI group, which is oriented towards early lifestyle interventions to counteract dementia. In order to set in early interventions, it is necessary to detect symptoms early, which brings a need for the type of Machine Learning-based methods proposed in EACare. The EACare consortium presently takes part in a large EU application that would fund such research.

The exploitation ("nyttiggörande") part of the project is ongoing during the spring of 2022 as described in Section 3.7. Moreover, EACare has also inspired a project that focuses on the ability to read, understand and remember texts of varying complexity. This project is now in the phase of submitting ethical applications.

In the meantime, we also continue our recordings in the clinical practice at the Memory Clinic. The collected data will be an asset to all three participating groups for smaller studies.

6.3 Long-term perspective on the field of the project

Alzheimer's disease and other neurocognitive disorders with a neuropathological origin develop gradually over many years before existing criteria of a clinical diagnosis are fulfilled. Making a correct diagnosis is a challenging task, especially in early stages of these diseases; it has been estimated that more than 50% of cases of dementia are undetected, and that the diagnostic accuracy is only between 70 and 90%, compared to what is revealed in post-mortem neuropathology.

The diagnostic uncertainty in neurocognitive disorders incurs great human and monetary costs to patients and society. For the patient, a false diagnosis inflicts unnecessary trauma with devastating consequences on quality of life, in addition to medication with likely negative side-effects. For society, large cost savings are possible if only persons with a high probability of neuropathology are referred to more detailed examinations.

In addition, if an underlying pathology can be correctly identified at an earlier stage, this will probably improve the efficacy of pharmacological as well as non-pharmacological counteractive measures. It is therefore of high priority to develop diagnostic tools for these diseases that are more sensitive, less invasive, more cost-effective, and easier to administer. The potential of the system we have successfully developed is that automatic capture and analysis of behavioral signals of potential clinical relevance could significantly add to diagnostic accuracy, primarily for two reasons: a) by reducing the risk that such signals are missed by the clinician during the medical examination and b) by adding new and complementary information beyond what is humanly perceptible, and beyond what is normally collected in the diagnostic process. Our preliminary results indicate that further development and implementation of such a system in clinical practice could make a significant difference, especially for the detection of cognitive disorders in early stages, and thereby for earlier intervention with better chances of success.

7. Costs

	2016	2017	2018	2019	2020	2021	Sum
PI salary	124 998	125 658	133 419	70 385	184 778	346 815	986 053
Seniors' salary	127 561	256 295	1 063 720	1 119 573	1 448 691	337 599	4 014 781
Postdocs' salary			551 250	742 930			1 294 180
PhDs' salary	31 192	1 675 851	1 653 588	1 162 179	1 742 184	775 418	6 476 012
Other salary			27 779	305 008	144 426	1 488 175	1 965 388
Equipment			99 710		10 224	15 887	125 820
Material	14 254	127 989	27 716	100 610	48 666	13 180	332 415
Travel	4 401	103 407	183 724	213 402	84 424	62 814	652 171
Administration							
Information							
Consultants				286 606			286 606
Other costs		9 271	111	155 627	181 849	12 730	359 587
Sum costs	302 406	2 298 471	3 741 017	4 156 320	3 845 242	3 052 618	16 493 014
Overhead	105 842	804 465	1 309 355	1 323 133	1 315 503	829 989	5 513 669
% OH	34,99	35,00	35,95	34,19	34,30	27,33	
Sum incl OH	408 248	3 102 936	5 050 372	5 479 453	5 160 745	3 882 607	22 006 683

8. External information and other activities

The EACare project has attracted media attention from the start, and the four principal investigators regularly promoted the project in media and in presentations. A list of coverage in news and presentations is given in Appendix A5.

The project also has its own homepage at http://www.eacare.se.

9. SWOT analysis

We looked back at the SWOT analysis in the mid-term report and made the following analyses of the outcome at the end of the project for a relevant selection.

Strengths

The data collection is directly integrated in the fast-track diagnostic process at the Memory Clinic at KS.
This has indeed proven important when the recordings started again after the pandemic. Thanks to the trust from the personnel in the Clinic, we have been allowed to continue, despite the strenuous situation.

Weaknesses

There was a lengthy process to set up the data collection and get the necessary ethical approvals. Part of the reason for that was that the KTH team initially underestimated the level of detail required in the ethical application, and that the KI team initially overestimated the capabilities of present Machine Learning and Robotics tools.
This is indeed true, and it forced a replanning of the project mid-way. See also Section 6.1 for a further discussion.

Opportunities

The tight integration of research and clinical practice at KI opens up for a real impact in how dementia diagnostics is done, via integration of AI support in cognitive clinical practice.
Unfortunately we did not reach the stage where technology from the project is mature enough to be used clinically. We indeed hope for this to come true in a possible follow-up project, see Sections 3.1, 3.5 and 6.2.

Threats

A threat is that postdoc contracts usually are two-year at KTH – i.e., shorter than the project time-span – and that the most fruitful path for the postdocs then is to move to another institution.
This threat actually did not materialize in our case. A postdoc hired in 2018, Gustav Eje Henter indeed finished his postdoc period after two years, but only to move to an Assistant Professor position at TMH, see also Section 4.2.

References

L. G. Exalto, C. P. Quesenberry, D. Barnes, M. Kivipelto, G. J. Biessels, and R. A. Whitmer. “Midlife risk score for the prediction of dementia four decades later”. In: *Alzheimer’s & Dementia* 10.5 (2014), pp. 562–570.

Kivipelto, M. & Håkansson, K. A Rare Success against Alzheimer’s. *Sci Am Mind* 27, 106–111, 2018.

Kivipelto, M. et al. World-Wide FINGERS Network: A global approach to risk reduction and prevention of dementia. *Alzheimers & Dementia* 1, PE424-17, 2020.

Livingston, G. et al. Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. *Lancet* 396, 413–446, 2020.

World Health Organization. Global status report on the public health response to dementia. https://www.who.int/publications/i/item/9789240033245, 2021.

Appendices

A1. Steering board

The project was led by a steering board consisting of the coordinator Hedvig Kjellström (2016-2021), the PIs Joakim Gustafson (2016-2021), Jonas Beskow (2016-2021) and Miia Kivipelto (2016-2021), and a senior researcher at KI, Dr Krister Håkansson (2018-2021).

A2. Activity of the steering board

The steering board convened about once every second month during the course of the project.

A3. Participating researchers

Senior researchers

HEDVIG KJELLSTRÖM, Professor, female. Role: Project coordinator, principal investigator at RPL, KTH, supervisor of Olga Mikheeva and Taras Kucherenko. Since project start.

MIIA KIVIPELTO, MD and Professor, female. Role: Principal investigator at NVS, KI and the Memory Clinic, KS. Main responsible for the medical part of EACare, development of study design, development of observational categories, and geriatric expertise during selected expert sessions. Since project start.

JOAKIM GUSTAFSON, Professor, male. Role: Principal investigator at TMH, KTH, co-supervisor of Patrik Jonell, Birger Moëll and Olga Mikheeva. Since project start.

JONAS BESKOW, Professor, male. Role: Principal investigator at TMH, KTH, supervisor of Patrik Jonell and Birger Moëll, co-supervisor of Taras Kucherenko. Since project start.

KRISTER HÅKANSSON, PhD, male. Role: Project coordinator at the Memory Clinic, KS, leader of WPnew1. Since 2017.

GUSTAV EJE HENTER, PhD, male. Role: Project coordinator at TMH, KTH, leader of WPnew2. Postdoc 2018-2020, Assistant Professor since 2020.

PhD students

THOMAS STORSKOG, MD, male. Role: PhD student, leader of WP1. Until 2017.

JOSEPH MENDELSON, MSc, male. Role: PhD student, leader of WP2. Until 2017.

PATRIK JAMES JONELL, PhD, male. Role: PhD student, leader of WPnew3. Since 2017.

OLGA MIKHEEVA, MSc, female. Role: PhD student, leader of WPnew4. Since 2017.

TARAS-SVITOZAR KUCHERENKO, PhD, male. Role: PhD student, leader of WPnew5. Since 2017.

BIRGER MOËLL, MSc, male. Role: PhD student, leader of exploitation ("nyttiggörande") project. Since 2020.

JASPER HOLLEMAN, MSc, male. Role: PhD student, participant in WPnew1 (monitoring of patient recordings). Since 2020.

Domain experts

MARIE RYDÉN, MD, female. Role: Selection of patients based on selection criteria, conducting interviews with patients that are recorded, development of observational categories, and geriatric expertise during expert sessions. Since project start.

ALEXANDRE BONNARD, MD, male. Role: Selection of patients based on selection criteria, conducting interviews with patients that are recorded, development of observational categories, and geriatric expertise during expert sessions. Since project start.

GÖRAN HAGMAN, Psychologist, male. Role: Experienced in therapy and psychological testing of elderly persons with possible cognitive impairment. Contributed with development of study design, psychological expertise during expert sessions. Since project start.

VESNA JELIC, MD, female. Role: Geriatric expertise during selected expert sessions. Until 2018.

MARIA SUNDELL, MD, female. Role: Geriatric expertise during selected expert sessions. Since 2019.

SARA GARCIA, MD, female. Role: Geriatric expertise during selected expert sessions. Since 2019.

PÄR ÖSTBERG, PhD, male. Role: Experienced speech therapist. Contributed in expert sessions, especially in observations of speech characteristics. Since project start.

SARA STORMOEN, PhD, female. Role: Speech therapist with special expertise of speech characteristics in persons with dementia. Development of texts that will be used during the interview, benchmarking with other researchers in the same field, regular participation in expert sessions, especially in observations of relevant speech characteristics. Since project start.

ULRIKA AKENINE, Geriatric nurse, female. Role: Contributed with development of study design, development of ethical applications, observations during expert sessions. Since project start.

A4. Selected publications

(A) Patrik Jonell. *Scalable Methods for Developing Interlocutor-aware Embodied Conversational Agents: Data Collection, Behavior Modeling, and Evaluation Methods*. PhD Thesis, KTH Royal Institute of Technology, 2022.

(B) Olga Mikheeva, Ieva Kazlauskaite, Adam Hartshorne, Hedvig Kjellström, Carl-Henrik Ek, and Neill DF Campbell. Aligned multi-task Gaussian process. In *International Conference on Artificial Intelligence and Statistics*, 2022.

(C) Taras Kucherenko, Rajmund Nagy, Michael Neff, Hedvig Kjellström, and Gustav Eje Henter. Multimodal analysis of the predictability of hand-gesture properties. In *International Conference on Autonomous Agents and Multi-Agent Systems*, 2022.

(D) Taras Kucherenko. *Developing and evaluating co-speech gesture-synthesis models for embodied conversational agents*. PhD Thesis, KTH Royal Institute of Technology, 2021.

(E) Taras Kucherenko, Rajmund Nagy, Patrik Jonell, Michael Neff, Hedvig Kjellström, and Gustav Eje Henter. Speech2Properties2Gestures: Gesture-property prediction as a tool for generating representational gestures from speech. In *ACM International Conference on Intelligent Virtual Agents*, 2021.

(F) Patrik Jonell, Youngwoo Yoon, Pieter Wolfert, Taras Kucherenko, and Gustav Eje Henter. HEMVIP: Human evaluation of multiple videos in parallel. In *International Conference on Multimodal Interaction*, 2021.

(G) R Nagy*, T Kucherenko*, B Moëll, A Pereira, H Kjellström, and U Bernardet. A framework for integrating gesture generation models into interactive conversational agents. In *International Conference on Autonomous Agents and Multiagent Systems*, demo track, 2021. (*Joint first authors)

(H) P Jonell*, B Moëll*, K Håkansson*, G Eje Henter, T Kucherenko, O Mikheeva, G Hagman, J Holleman, M Kivipelto, H Kjellström, J Gustafson, and J Beskow. Multimodal capture of patient behaviour for improved detection of early dementia: Clinical feasibility and preliminary results. *Frontiers in Computer Science*, doi: 10.3389/fcomp.2021.642633, 2021.

(I) T Kucherenko*, P Jonell*, Y Yoon*, P Wolfert, and G E Henter. A large, crowdsourced evaluation of gesture generation systems on common data: The GENEA Challenge 2020. In *International Conference on Intelligent User Interfaces*, 2021. (*Joint first authors)

(J) T Kucherenko, D Hasegawa, N Kaneko, G Eje Henter, and H Kjellström. Moving fast and slow: Analysis of representations and post-processing in speech-driven automatic gesture generation. *International Journal of Human-Computer Interaction*, doi: 10.1080/10447318.2021.1883883, 2021.

(K) K Håkansson, J Beskow, H Kjellström, J Gustafsson, A Bonnard, M Rydén, S Stormoen, G Hagman, U Akenine, K Morales Peres, G Henter, M Sundström, and M Kivipelto. Robot-assisted detection of subclinical dementia: progress report and preliminary findings. *Alzheimer's & Dementia* 16(S6), e043311, 2020.

(L) P Jonell, T Kucherenko, G Eje Henter, and J Beskow. Let's Face It: Probabilistic Multi-Modal Interlocutor-Aware Generation of Facial Gestures in Dyadic Settings. In *ACM International Conference on Intelligent Virtual Agents*, 2020.

(M) P Jonell*, T Kucherenko*, I Torre, and J Beskow. Can we trust online crowdworkers? Comparing online and offline participants in a preference test of virtual agents. In *ACM International Conference on Intelligent Virtual Agents*, 2020. (*Joint first authors)

(N) Krister Håkansson, Jonas Beskow, Hedvig Kjellström, Joakim Gustafsson, Alexandre Bonnard, Marie Rydén, Sara Stormoen, Göran Hagman, Ulrika Akenine, Kristal Morales Pérez, Gustav Henter, Maria Sundell, and Miia Kivipelto. Kan en robot bidra till tidigare och bättre diagnoser vid misstänkt kognitiv störning? *Svensk Geriatrik*. Mars 2020.

(O) K Håkansson and M Kivipelto. Början till en global livsstilsrevolution? *Äldre i Centrum* vol 2020 1-2, 2020.

(P) T Kucherenko, P Jonell, S van Waveren, G Henter, S Alexanderson, I Leite, and H Kjellström. Gesticulator: A framework for semantically-aware speech-driven gesture generation, In *ACM International Conference on Multimodal Interaction*, 2020.

(Q) Patrik Jonell, Taras Kucherenko, Erik Ekstedt, and Jonas Beskow. Learning non-verbal behavior for a social robot from YouTube videos. In *ICDL-EpiRob Workshop on Naturalistic Non-Verbal and Affective Human-Robot Interactions*, 2019.

(R) Éva Székely, Gustav Eje Henter, Jonas Beskow, and Joakim Gustafson. Spontaneous conversational speech synthesis from found data. In *Interspeech*, 2019.

(S) Éva Székely, Gustav Eje Henter, Jonas Beskow, and Joakim Gustafson. Off the cuff: Exploring extemporaneous speech delivery with TTS. In *Interspeech*, 2019.

(T) Taras Kucherenko, Dai Hasegawa, Gustav Eje Henter, Naoshi Kaneko, and Hedvig Kjellström. Analyzing input and output representations for speech-driven gesture generation. In *ACM Intelligent Virtual Agents Conference*, 2019.

(U) Éva Székely, Gustav Eje Henter, and Joakim Gustafson. Casting to corpus: Segmenting and selecting spontaneous dialogue for TTS with a CNN-LSTM speaker-dependent breath detector. In *ICASSP*, 2019.

(V) Patrik Jonell, Mattias Bystedt, Fethiye Irmak Dogan, Per Fallgren, Jonas Ivarsson, Marketa Slukova, Ulme Wennberg, Jose Lopes, Johan Boye and Gabriel Skantze. Fantom: A Crowd- sourced Social Chatbot using an Evolving Dialog Graph. 2nd *Proceedings of Alexa Prize*, 2018.

(X) Håkansson K, Ngandu T, Kivipelto M. The Patient with Cognitive Impairment. In: Hachinsky V, editor. *Treatable and Potentially Preventable Dementias*. 1st ed. Cambridge University Press; 2018. pp. 5280.

(Y) Olga Mikheeva, Carl Henrik Ek, and Hedvig Kjellström. Perceptual facial expression representation. In *IEEE International Conference on Automatic Face and Gesture Recognition*, 2018.

(Z) Patrik Jonell, Catharine Oertel, Dimosthenis Kontogiorgos, Jonas Beskow, and Joakim Gustafson. Crowd-powered design of virtual attentive listeners. In *International Conference on Intelligent Virtual Agents*, 2017.

(Ω) Patrik Jonell, Joseph Mendelson, Thomas Storskog, Göran Hagman, Per Östberg, Iolanda Leite, Taras Kucherenko, Olga Mikheeva, Ulrika Akenine, Vesna Jelic, Alina Solomon, Jonas Beskow, Joakim Gustafson, Miia Kivipelto, and Hedvig Kjellström. *Machine Learning and Social Robotics for Detecting Early Signs of Dementia*, http://arxiv.org/abs/1709.01613, 2017.

A5. Important activities

News coverage

Robot kan upptäcka demens tidigt, Framtidens Forskning, June, 2020.

Sverige behöver satsa på AI, IVA Aktuellt, 2:18-23, April, 2018.

The system has also been presented to members of the Swedish Royal family during the inauguration of the new Memory Clinic at KS in 2018, which rendered us news coverage in Damernas Värld, quite unique for a research project.

Social robot ska kunna upptäcka demens tidigare, Dagens Nyheter, July, 2016.

Public demo

Public demo of the dialogue system prototype (Task 2.2 in Section 2.1) at the conference *Digitalize in Stockholm*, https://www.cwtnordicevents.com/digitalizeinsthlm2019, at Kistamässan, November, 2019.

Presentations

Miia Kivipelto: Online live lecture Future Healthcare Istanbul “Brain health and Alzheimer’s disease: The future healthcare with multidomain interventions, digital biomarkers and precision prevention”, October 20, 2021.

Krister Håkansson: Vad är tidigt stöd när minnet sviktar, Stockholm Demensförenings Samtalsserie Del 1, December, 2020.

Hedvig Kjellström: AI och sociala robotar - en ny väg framåt i demensdiagnostik. Presentation vid seminarium för KTH-donatorer, Stockholm, 23 maj, 2019.

Hedvig Kjellström: Perception and cognition of non-verbal behavior: Two very different clinical applications. StratNeuro conference, Stockholm, October 2, 2019. Annual conference gathering researchers in Neuroscience and adjacent areas in Sweden.

Joakim Gustafson: Toward User-Oriented Agents: Research Directions and Challenges. NSF Workshop at CMU, Pittsburgh, USA, October 24-25, 2019. Arranged by CMU to inform funding agencies of the directions that research should take in the future according to experts in the field - with observers from NSF and DARPA.

Krister Håkansson: EACare: Robotassisterad diagnostik vid misstänkt tidig neurokognitiv störning. Öppen föreläsning, ABF-Huset, Stockholm, 10 december 2019. Anordnad av Stockholms Demensförening.

Joakim Gustafson: ”Towards situated interaction with social robots” Keynote at the CHILI Winter School, EPFL, Feb 2019

Krister Håkansson: EACare Fas 1: Kartläggning av kliniskt relevanta beteenden vid tidig kognitiv störning, En lägesrapport för ett samarbetsprojekt mellan KI, KS och KTH. Presented at Geriatriskt Forskarforum, 2019-02-06, Stockholm. (National conference arranged by Svensk geriatrisk förening, Sveriges Läkarförbund.)

Joakim Gustafson: Social robots for the elderly companionship with a purpose, Stockholm-Tokyo Workshop on Multi-disciplinary collaboration for sustainable development, Hongo Campus, Tokyo, 2018 (Keynote)

Hedvig Kjellström: ”EACare: Embodied Agent to support elderly mental wellbeing”, Keynote at AI4Healthcare, Stockholm, March 2018.

Joakim Gustafson: Human-robot interactions in the wild, Shonan Meeting on Multimodal Agents for Ageing, Tokyo, 2018 (Keynote)

A9. PhD exams

Two PhD students have defended their thesis in the project:

PATRIK JAMES JONELL, PhD in March 2022, male. TMH, KTH, 100% funded from EACare.

TARAS-SVITOZAR KUCHERENKO, PhD in December 2021, male. RPL, KTH, 100% funded from EACare.

A11. Future exams

Two PhD students are on their way to defend their thesis:

OLGA MIKHEEVA, PhD estimated in November 2022, female. RPL, KTH, 90% funded from EACare.

BIRGER MOËLL, PhD estimated in November 2024, male. TMH, KTH, 40% funded from EACare.

A13. Awards

Paper (P) won the **ICMI 2020 Best Paper Award**
https://icmi.acm.org/2020/index.php?id=award

Paper (S) won the **“Interspeech Best Show and Tell Paper 2019” Award**
https://www.isca-speech.org/iscaweb/index.php/honors/awards

B1. What would you do differently if the project started today?

If we would get the opportunity to start over again, we would firstly draw conclusions from the changes made in the present project, see Section 1.4, and organize the project as it ended up being organized after a quite time-consuming iterative process. We would also take great care to make contingency plans towards situations similar to the covid-19 pandemic.

Secondly, we would draw inspiration from the lessons learned, see Section 6.1, and we would, already from the start of the project, aim at designing interpretable and modular diagnostic methods based on structured combinations of recognition of digital biomarkers. Moreover, we would assign even more time for spending extended periods of time in each others' environments.

B2. What will ultimately be the main impact of the project on society and academia?

Alzheimer's disease and other forms of degenerative cognitive disorders are highly age-dependent, and life expectancy is steadily increasing in most parts of the world. Approximately 50 million persons today carry some form of neurocognitive disorder, a number that is expected to grow to around 150 million in 2050. Due to high-intensive need of care in later phases, these diseases put a high burden on limited care resources and societal economies. Combating these disorders has been declared a priority by the World Health Organization.

The potential of the system we have built is that it could significantly contribute to improving diagnostic accuracy for early detection of cognitive disorders, thereby enabling earlier interventions, both pharmacological and non-pharmacological. There is a good chance that earlier, well-tailored interventions will be more efficient to counteract cognitive disorders. For similar reasons, the system could also contribute to improving research on neurocognitive disorders, especially in early phases, where accurate classification of participants is crucial to test the efficacy of both pharmacological and non-pharmacological treatments. Evidently, if persons in these studies are misclassified, results will be misleading. The potential of the system we have built to increase detection and accuracy in early stages of a neurocognitive disorder could therefore also be used for research purposes.

B3. What do you expect will happen to the activities when the funding has expired?

See Section 6.2 for a discussion of this topic.

B4. What were the problems of the project?

Section 1.4 gives an account of the problems that occurred during the project, and how we dealt with them. As discussed in Appendix B1, there are many things we would do differently if starting over again.

However, the greatest frustration was something we did not anticipate in the SWOT analysis in the mid-term report: the covid-19 pandemic. A learning from this is to make even more careful contingency plans.

B5. What was the most fun with the project?

All project participants agree that the most fun part was the highly multidisciplinary environment. The EACare project has created a sustainable bond between the three research groups.

B6. Your feedback to the Foundation

We are very grateful for the support of SSF in allowing extensions and reorganization of the project, and for the pragmatic and efficient administration around the project.

A point of criticism is that the interaction between projects in the program has been non-existent, except for a meeting in the beginning. It would have been good with at least one program conference where we could have shared experiences.