Establishing the Foundations of Emotional Intelligence in Care Companion Robots to Mitigate Agitation among High-Risk Dementia Patients: Protocol for Empathetic Patient-Robot Interactions

Adeline Nyamathi, Nikil Dutt, Jung-Ah Lee, Amir Rahmani, Mahkameh Rasouli, Donna Krogh, Erik Krogh, David Sultzer, Humayun Rashid, Hamza Liaqat, Riyam Jawad, Farhan Azhar, Ali Ahmad, Bilal Qamar, Taha Yasin Bhatti, Chet Khay, Jocelyn Ludlow, Lisa Gibbs, Julie Rousseau, Mahyar Abbasian, Sabine Brunswicker

Submitted to: JMIR Research Protocols
on: December 22, 2023

Table of Contents

Original Manuscript **5**
Supplementary Files **36**
Figures **37**
Figure 1 .... **38**

Establishing the Foundations of Emotional Intelligence in Care Companion Robots to Mitigate Agitation among High-Risk Dementia Patients: Protocol for Empathetic Patient-Robot Interactions

Adeline Nyamathi[1]; Nikil Dutt[2] PhD; Jung-Ah Lee[1] RN, PhD, FAAN; Amir Rahmani[1] PhD; Mahkameh Rasouli[1] BSN, MSN; Donna Krogh[3] PhD; Erik Krogh[3] PhD; David Sultzer[4] MD; Humayun Rashid[5] DBA; Hamza Liaqat[5] BSc; Riyam Jawad[5] MS; Farhan Azhar[5] BSc; Ali Ahmad[5] MS; Bilal Qamar[5] BSc; Taha Yasin Bhatti[5] BSc; Chet Khay[6] MSN, FNP; Jocelyn Ludlow[1] PhD, RN, CNE, CHSE, CMSRN; Lisa Gibbs[7] MD; Julie Rousseau[7] PhD; Mahyar Abbasian[2] MSc; Sabine Brunswicker[8] PhD

[1]Sue & Bill Gross School of Nursing, University of California Irvine Irvine US
[2]Donald Bren School of Information and Computer Sciences, University of California Irvine Irvine US
[3]Smart Forward Rancho Palos Verdes US
[4]Department of Psychiatry and Human Behavior, and Institute for Memory Impairments and Neurological Disorders, University of California Irvine Irvine US
[5]NaviGAIT Irvine US
[6]Amore Senior Living Laguna Niguel US
[7]School of Medicine, University of California Irvine Irvine US
[8]Purdue University West Lafayette US

Corresponding Author:
Adeline Nyamathi
Sue & Bill Gross School of Nursing, University of California Irvine
854 Health Sciences Quad
Room 434
Irvine
US

Abstract

Background: There are an estimated 6.7 million persons living with dementia in the U.S.; expected to double by 2060. Those at highest risk, persons experiencing moderate to severe dementia (P-MSD), are 4-5 times more likely to fall than those without dementia, as they often experience unpredictable agitation, leading to unsteady gait. Socially assistive robots fail to address the dynamically changing emotional states associated with agitation, and there is a lack of understanding how emotional states change, how they impact agitation and gait over time, and how social robots can best respond by showing empathy.

Objective: Design and validate a foundational model of emotional intelligence for empathic patient-robot interaction that mitigates agitation among those at highest risk, P-MSD.

Methods: A design science approach will be used to: 1) collect and store granular, personal, chronological data (Personicle), using real-time visual, audio and physiological sensing technologies in a simulation lab and Board & Care facilities; 2) develop statistical models to understand and forecast a P-MSD's emotional state, agitation level and gait in real-time using ML/AI and the Personicle; 3) design and test an empathy-focused conversation model, focused on storytelling; and 4) test and evaluate the empathy-focused conversation model for the Care Companion Robot (CCR) in the community.

Results: The architecture development for the Personicle has been initiated with existing open source data and a non-Human Subject approval obtained in November 2023. A Community Advisory Board was formed and met in December 2023, and an ethical board was established with international colleagues. Full IRB was submitted in December 2023.

Conclusions: This innovative caregiving approach is designed to recognize signs of agitation, and upon recognition, intervene with empathic verbal communication. This proposal thus has the potential to have a significant impact on an emerging field of computational dementia science by reducing unnecessary agitation and falls of P-MSD, while improving caregiver burden.

(JMIR Preprints 22/12/2023:55761)

DOI: https://doi.org/10.2196/preprints.55761

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?
✓ **Please make my preprint PDF available to anyone at any time (recommended).**
Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.
Only make the preprint title and abstract visible.
No, I do not wish to publish my submitted manuscript as a preprint.
2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?
✓ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**
Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v
Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

JMIR Research Protocols

Establishing the Foundations of Emotional Intelligence in Care Companion Robots to Mitigate Agitation among High-Risk Dementia Patients: Protocol for Empathetic Patient-Robot Interactions

Adeline Nyamathi[1], ANP, PhD., FAAN., Nikil Dutt[2], PhD., Jung-Ah Lee[1], PhD., FAAN; FGSA; Amir Rahmani[1], PhD., Mahkameh Rasouli[1], BSc; Donna Krogh[3], PhD., Erik Krogh[3], Ph.D., David Sultzer[4], MD; Humayun Rashid[5], DBA., Hamza Liaqat[5] BSc; Riyam Jawad[5], MS; Farhan Azhar[5], BSc; Ali Ahmad[5], MS; Bilal Qamar[5]; BSc; Taha Yasin Bhatti[5], BSc; Chet Khay[6], MSN, FNP; Jocelyn Ludlow[1], PhD., RN, CNE, CHSE, CMSRN; Lisa Gibbs[7], MD., Julie Rousseau[7], PhD., Mahyar Abbasian[2] MSc; Yutong Song[2],BS., Cheonkam Jeong[1,8]PhD.,Sabine Brunswicker[8], PhD.

[1] Sue & Bill Gross School of Nursing, University of California Irvine, Irvine, CA, United States
[2] Donald Bren School of Information and Computer Sciences, University of California Irvine, Irvine, CA, United States
[3] Smart Forward, 4101 Sea Horse Lane, Rancho Palos Verdes, CA 90275 United States
[4] Department of Psychiatry and Human Behavior, and Institute for Memory Impairments and Neurological Disorders, University of California Irvine
[5]NaviGAIT, 2211 Michelson Drive, Suite 900, Irvine, CA 92612, USA
[6] Amore Senior Living, 25011 Monte Verde Dr., Laguna Niguel CA 92677 United States
[7] School of Medicine, University of California Irvine, Irvine, CA, United States
[8]Purdue University, West Lafayette, Indiana, United States

Corresponding Author:
Adeline Nyamathi, ANP, PhD., FAAN
Distinguished Professor
Sue & Bill Gross School of Nursing
854 Health Sciences Quad, Office 4234
Irvine, CA, 92697
United States
Phone: 949 824-8932
Email: anyamath@hs.uci.edu

Abstract

Background: There are an estimated 6.7 million persons living with dementia in the U.S.; expected to double by 2060. Those at highest risk, persons experiencing moderate to severe dementia (P-MSD), are 4-5 times more likely to fall than those without dementia, as they often

experience unpredictable agitation, leading to unsteady gait. Socially assistive robots fail to address the dynamically changing emotional states associated with agitation, and there is a lack of understanding how emotional states change, how they impact agitation and gait over time, and how social robots can best respond by showing empathy.
Objective: Design and validate a foundational model of emotional intelligence for empathic patient-robot interaction that mitigates agitation among those at highest risk, P-MSD.
Methods: A design science approach will be used to: 1) collect and store granular, personal, chronological data (Personicle), using real-time visual, audio and physiological sensing technologies in a simulation lab and Board & Care facilities; 2) develop statistical models to understand and forecast a P-MSD's emotional state, agitation level and gait in real-time using ML/AI and the Personicle; 3) design and test an empathy-focused conversation model, focused on storytelling; and 4) test and evaluate the empathy-focused conversation model for the Care Companion Robot (CCR) in the community.
Results: The study was funded in October 2023. For Aim 1, the architecture development for the Personicle was initiated with search for existing open-source data in January 2024 with non-Human Subject approval. A Community Advisory Board was formed and met in December 2023 to provide feedback on the use of a CCR and begin gathering personal stories. Full IRB was approved in March 2024 to prepare for placement of the cameras and CCR in Board & Care Facilities. Beginning in March 2024, the development of Atomic Markers was underway. For Aim 2, two main tasks were completed as a foundation for the CCR's empathic conversational interface. First, work began on establishing an emotional classifier after a careful review of open-source data of P-MSD patients. Data labeling started in April 2024 and was completed in June 2024. The labeled data is currently undergoing validation. Second, in parallel, the team established a baseline multimodal model trained and validated on healthy datasets using transformer architecture in a semi-supervised manner. This model has since been retrained on the labeled P-MSD dataset, and a robust model is under construction to handle imbalanced data, forming the basis for the CCR's empathic conversational interface. In April 2024, the team also initiated research on empathy alignment of large language models (LLMs) using prompt engineering and reinforcement learning, as well as furthering the use of Aim 1 Atomic Markers.
Conclusions: This innovative caregiving approach is designed to recognize signs of agitation, and upon recognition, intervene with empathic verbal communication. This proposal thus has the potential to have a significant impact on an emerging field of computational dementia science by reducing unnecessary agitation and falls of P-MSD, while improving caregiver burden.

Keywords: Persons with dementia; Empathy-based Care Companion Robot; Agitation; Fall Risk

Introduction

The Impact of Dementia Among Elderly

Alzheimer's disease and related dementias (ADRD, hereafter dementia) has been increasing among the elderly. In the US, an estimated 6.7 million currently have dementia; a number expected to increase to 13.8 million by 2060 [1]. The majority of persons living with dementia suffer from neuropsychiatric symptom of dementia [2], that are noncognitive, such as agitation, defined as inappropriate verbal, vocal, or motor activity [3] Agitation is problematic, as not only can it impact a person's morbidity and mortality, but as well impair motor functions during events of agitation and increase the risk of falling and sustaining severe injuries, and subsequent hospitalization [4,5]. Fall is a major cause of morbidity and mortality of older adults. Indeed, persons experiencing moderate to severe dementia (P-MSD) are 4-5 times more likely to fall than older people without cognitive impairments [4,6,7]. Despite close observation and monitoring by caregivers, the onset of agitation is often unpredictable [8]. In fact, research has showed that agitation in persons with dementia in long-term care settings is associated with higher medication use and an increased likelihood of experiencing falls, fractures, infections, and other neuropsychiatric symptoms compared to older adults without agitation (OR=1.58 for fall, 95% CI 1.41-1.77, P-value <0.001). Thus, it is important to develop an approach to effective management of agitation symptoms in this population to prevent falls [9].

Impact of Dementia on Caregivers

While much of the literature focuses on the fact that caring for P-MSD creates a heavy burden on caregivers, leading to caregiver burnout [10], a few papers have highlighted ways that caregivers of P-MSD can bring joy and happiness to their loved ones [11,12]. Further, evidence has demonstrated that P-MSD can continue to express joy for minutes beyond remembering an event that was joyful [13]. However, more typically, the literature documents how informal caregivers (e.g., spouses or adult children) for P-MSD experience chronic stress, depression, sleep disorders, poor health, low quality of life, and early mortality [14]. Family caregivers who live together with P-MSD show severe sleep disturbance compared to caregivers for persons with other health conditions such as stroke or cancer [15]. Moreover, family caregivers for persons with dementia are twice as likely to suffer from depression as caregivers for someone without dementia [14,16]. It is not uncommon for early mortality of spousal caregivers of persons with dementia [17].

As a result, P-MSD are often removed from their homes due to worsening dementia, and the heavy burden on caregivers [14,18]. Further, despite evidence that person-centered intervention strategies may improve a person's state [19], such strategies are not widely adopted, and factors impacting its success are not fully understood [10,20]. Anecdotes, personal experiences from caregivers, and recent literature suggest that the success of person-centered care may depend upon a caregiver's ability to relate to the patient and show empathy, or the ability to imagine what they may be feeling or thinking [21]. However, the successful implementation of such care requires intense training, particularly when abusive behavior is present [10,20].

Current Status of Use of Socially-Assistive Robots as a Healthcare Approach

Socially-assistive robots are innovative solutions to deliver quality care for dementia patients, and to improve agitation via verbal and non-verbal communication [22-24]. Social robots are typically equipped with onboard sensing technologies, including cameras and motion sensors to sense the patient's state, and can support verbal and non-verbal communication with the person. Some robots come in the form of "animals" (i.e., a dog), and have been shown to help persons with dementia build social relationships [23]. However, human-like robots (humanoids) that can demonstrate social patient-robot relationships can also have therapeutic effects as measured in positive emotional responses [25]. Such social relationships are established through the social "robot" presence and sensorimotor interaction with the robot using haptics and vision (e.g. touching and seeing the robot) and through use of sound (e.g. a laughter). Currently, it is not clear if and how robots may lower the risk of agitation through verbal communication. Further, existing work on social robots fails to address one important issue: That agitation is associated with dynamically changing emotional disturbances [3].

Moreover, there is not sufficient research on how to design a care-companion robot (CCR) that is equipped with such a conversational and empathic intelligence that combines sound in the form of speech features with language-based interaction. Currently, there is a lack of understanding in how to best design emotionally aware robots that use empathic conversations to trigger positive emotional change. Thus, research is needed that integrates both streams of research to close the gap on how emotional intelligence and empathic conversations with caregiving robots can have therapeutic effects.

The specific research gaps we identified include three points:1) the forms and the time-variant nature of emotions expressed by P-MSD are not sufficiently understood; 2) exiting technical solutions for emotion detection using data-driven machine learning algorithms, trained and fine-tuned on large amount of data collected and labeled for health individuals, have not been sufficiently evaluated in the context of dementia given the scarcity of data emotional states for P-MSD patients. While there is an active research community on detecting dementia based on language, speech signals, facial expression and body movements, emotion detection for dementia patients is not well studied. 3) existing social robots lack the emotional intelligence to respond to the person-specific emotional state through language-based communication that is empathic by displaying understanding for the persons' feelings and emotions through words and speech features (e.g. warm tone). Existing conversational robots use standard implementations of large language models (LLMs) without considering the challenges of fine-tuning such language models for greater empathy. Research on empathy alignment in general, and in the particular context of healthcare and dementia patients, is scarce, and requires significant research. The few available efforts to align LLMs either use simple prompt engineering or rely on datasets from social media platforms that have been labelled without a deep consideration of the language-specific properties of empathy [26-28]. 4) Finally, beyond the technical research gaps related to ML/AL and LLMs, the correlation relationship between different forms of emotions with motor impairments expressed in agitation and subsequent falls is unclear. While new research papers continue to be published on the impact of socially assistive robots on P-MSD [25], the research is still in a nascent stage. We believe that our study will provide significant contributions by building on recent research and addressing the points identified above. Table 1 articulates a lack of emotional intelligence and empathy-based conversations by existing robots, vs our proposed CCR which overcomes these limitations.

Proposed Care Companion Robot

A CCR is an advanced technological solution designed to provide assistance, support, and companionship for elderly individuals, while addressing specific challenges associated with aging and cognitive decline. Equipped with autonomous indoor navigation, human-safe movement capabilities, and robust open robotics framework (ROS)-based middleware, the robot seamlessly navigates indoor environments while ensuring safety and reliability. Its multimodal interaction capabilities enable intuitive communication with users through speech, touch, and physical gestures using head/neck movement, facilitated by an Android tablet interface. With a suite of environment monitoring sensors and customization options, the robot can be used to detect anomalies and adapt to various use cases, including the early detection of behavior changes indicative of conditions like dementia. Powered by an onboard NVIDIA Jetson Orin Device, it processes sensor data and executes tasks with high performance and efficiency, making it a valuable companion and caregiver support tool for the elderly.

Table 1. Scenario

Limitations of existing social robot capabilities	Future capabilities of our social robot – the CCR
Paul, an 82-year-old resident with P-MSD, wakes up most mornings, forgetting how to get out of bed, using his walker. He gets upset and angry, trying to get up. His social robot in the room notices a movement and approaches. It starts talking using a standard dialogue flow with a normal voice pitch: *Hello Paul, how can I help you?* This makes Paul even more upset, worsening his emotional state. He gets very angry, and agitated but the robot keeps asking: *How can I help you?* Paul's agitation escalates. He shouts at the robot and yells: *You can't help me stupid robot. I do not need your help. Paul* tries to get up without the walker and almost falls. Luckily the caregiver observing Paul comes into the room, and prevents the fall by calming him down using soft language, and talks about exciting events in Paul's past where his memory is still strong.	Paul, an 82-year old resident with P-MSD, wakes up most mornings, forgetting how to get out of bed, using his walker. His CCR in the room is sensing a potential emotional shift ahead of time as Paul's facial expression, body movement and utterance indicate that he is getting angry. When Paul starts trying to get out of bed, the CCR starts talking to Paul using a soft language: *Good morning Paul! Shall we go to breakfast?... (CCR pauses to ensure the comment was understood) Let's use our walker this morning. (CCR shows an image of the walker which is close by)… But before we go, let us talk about your colleague at Boeing that you worked with.* Paul's response: Y*ou mean Alfred? He was such a great guy.* Paul continues his personal story, and the CCR responds, adding facts that Paul has forgotten. In the meantime, the caregiver, who had been informed about a potential onset of agitation, moves the walker closer to Paul so he can get out of bed safely. Paul gets up. While walking to the breakfast room, the CCR and Paul are talking about Alfred. Paul has a happy start to his day.

While the theoretical mechanism for story telling in soothing agitation is yet to be articulated, Rios and colleagues [29] reported that storytelling has helped to support memory, reminiscence, identity and self-confidence; potential mechanisms under which story telling may reduce agitation. Further, other authors found that quality of life, depression and quality of the carer-patient relationship improved at post intervention with storytelling [30]; perhaps helping to build a therapeutic relationship, bond, rapport, comfort, and trust between the elder and the caregiver [31], possibly leading to reduced agitation and fall risk. By providing mental stimulation, such as by storytelling, cognitive functions may be improved, including attention, executive function, memory as well as agitation, depression and anxiety [32,33].

Objectives

This study describes the design and implementation of computational patient-robot interaction research designed to predict agitation and mitigate fall risk, and tests the strategies in both the simulation lab with up to 50 undergraduate and graduate students and 12 P-MSD. The study aims to: 1) collect and store new datasets used to train a CCR; 2) design new computational models focused on P-MSD; 3) develop new systems of empathy-based conversational models, using LLM and theories of linguistics; and 4) pilot an empathy-focused intervention model for the CCR using a quasi-experimental intervention design, and evaluate the conversational intervention longitudinally using mixed method approaches.

Methods

Study Design

The proposed 2-year study builds on the expertise of a multidisciplinary research team (i.e., nursing, computer sciences, machine learning (ML), artificial intelligence (AI), engineering, linguistics, and medicine). We will use computational ML/AI methods to first collect non-invasive data and develop models which can forecast a person's emotional state, agitation level and gait in real-time, using ML/AI and Personicles, and then test and evaluate the empathy-focused conversation model for the CCR in the community, using quasi-experimental and mixed method approaches (see Figure 1). We will use dynamic event analysis of empathy-focused CCR and P-MSD conversations to examine the impact of conversational events during heightened emotional state on agitation level and gait pattern.

The aims of the study are as follows:

Aim 1: Collect and store granular, Personicle, using real-time visual, audio and physiological sensing technologies in a simulation lab and Board & Care facilities;

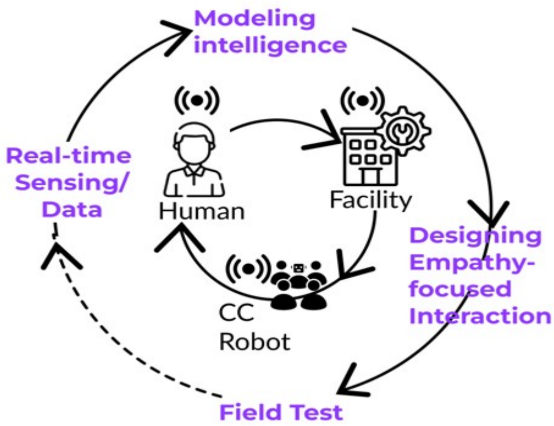
Aim 2: Develop statistical models to understand and forecast a P-MSD's emotional state, agitation level and gait in real-time using ML/AI and the Personicle;

Aim 3: Design and test an empathy-focused conversation model, focused on storytelling; and

Aim 4: Test and evaluate the empathy-focused conversation model for the CCR in the community.

Figure 1 Design Components

Prior to the onset of the study, a Community Advisory



Board (CAB) will be formed to ensure community feedback will be integrated into the logistics of the study, ensuring privacy and safety, and relevancy of empathy-focused conversational models delivered by the CCR. The team will meet three times in the first year of the study. The CAB is designed to help guide CCR observation and modeling; as well as over the 2-year study, provide ongoing evaluative feedback on usability and acceptability collected from family and professional caregivers. All members of the CAB will be compensated for their time.

Recruitment

Our study is being conducted in four Board & Care (B/C) facilities in Southern California, owned and operated by a certified family nurse practitioner. Each of the 4 facilities is licensed by the department of social services. Each B/C facility has an occupancy of 6 residents, who are cared for by a multidisciplinary team composed of nurse practitioners, medical assistants and periodic visitation by a physician and music and exercise therapist. The residents are typically 65 years and older who were admitted, the majority typically experience moderate to severe Alzheimer's disease; over 50% also experience co-morbidities including hypertension, congestive heart failure, or have experienced a cerebrovascular accident. These conditions require 24-hour supervision and assistance due to safety concerns. The research staff will work closely with the owner who will ensure all eligible residents (or their surrogates) will receive a flyer about the study.

Sample and Setting for Human Subject Engagement for Specific Aims 1, 3 and 4

For Aim 1 (Observation of P-MSD, Interviews with Family and Professional Caregivers), Aim 3 (Simulation Lab role plays with undergraduate and graduate students) and Aim 4 (Testing and Evaluation of the impact of CCR modeling), the following is relevant:

Aim 1: The sample will include P-MSD who are residents of four B/C facilities in Southern California, where a total of up to 24 patients reside, with turnover approximately every 6 months; about 25% stay for 2-5 years. Current demographics reveal that the residents are equally male/female, and predominantly White. The average age is 88, and at least one family caregiver is involved. In total, 14% of residents have mild dementia, 43% moderate dementia, and 24% severe dementia. Particularly with residents who experience agitation, risk of falls is a reality that is very concerning. We estimate, at minimum, that 6 residents will meet eligibility criteria and remain in the facility throughout the 2-year program, with new patients admitted every month.

Inclusion Criteria: 1) admitted by a MD as a P-MSD; 2) residing in a facility; 3) age $\geq$ 40; 3) consent provided by a legally authorized representative. Exclusions include: 1) severe medical conditions that would limit participation in the study, including terminal illness.

Aim 1 and Aim 4: Family Caregiver: The sample will include up to 14 family caregivers of P-MSD who will be asked for their interest in participating in informal interviews to discuss any experiences in dealing with agitation and fall risk of their family member, and about the joyful times their family member experienced (Aim 1). In Aim 4, up to 6 family caregivers whose family member with dementia is enrolled in the three-month intervention and provide consent will be administered two structured instruments at baseline and three-month follow-up and a formal interview about how the CCR functioned in mitigating agitation and fall risk. Inclusion Criteria include age 18 and older and English speakers, and self-reported family members of the P-MSD.

Aim 1 and Aim 4: Professional Caregiver: The sample will include up to 5 professional caregivers of P-MSD who work in the B/C facility and who will be asked if interested in participating in informal interviews to discuss any experiences in dealing with agitation and fall risk of P-MSD who they care for in the B/C facility. They will also be asked to share experiences of what makes the P-MSD joyful vs sad (Aim 1). In Aim 4, up to 5 professional caregivers who care for enrolled P-MSD in the three-month intervention, and provide consent, will be administered two structured instruments at baseline and three-month follow-up and a formal interview about how the CCR functioned in mitigating agitation and fall risk. Inclusion Criteria include age 18 and older, English speaker, and working in the B/C facility.

Undergraduate/Graduate Student Volunteer: In Aim 3, about 50 student volunteers will be consented to spend about an hour in the simulation lab, and will be trained by our investigative team to role-play a person with dementia who is experiencing emotions such as agitation. To capture authentic emotions, students may also be provoked to display emotions as well. Inclusion Criteria: age 18 or older, enrolled undergraduate/graduate student at UCI, and English speaker.

Ethical Considerations

Institutional Review Board

Two Institutional Review Board (IRB) requests have been submitted to the grant recipient institutional IRB. The first IRB (Non-Human Subjects Research [UCI IRB #3795]), approved in December 2023, related to accessing publicly available datasets containing multimodal and sensory input data for modeling emotion states, degree of agitation, and gait as well as training CCR for empathy-based conversational interventions for Aims 2 and 3. The second IRB was approved in March 2024 for human subject research in the minimal-risk biomedical track (UCI IRB #4150) for all aims. The study includes access to medical records and observation of clinical care, and a partial waiver of HIPAA authorization has been requested.

Informed Consent

Family Caregivers and Healthcare Professionals

At the beginning of the study, a study team member will provide family and professional caregivers with comprehensive information about the study and obtain their consent.

Participants with Dementia

As the dementia participants might be cognitively impaired, the procedure of informed consent will utilize a validated decision-making assessment tool [34] tailored to the cognitive abilities of participants with moderate to severe dementia. This tool will assess the participant's capacity to comprehend the study objectives and procedure and determine their capability of providing informed consent regarding participation in the study. If the assessment indicates the participant's lack of decision-making capacity to provide informed consent independently, a surrogate decision-maker who is a legally authorized representative or a family member will be identified and will be informed about the study by the study team and will provide comprehensive information about the study and obtain their consent on behalf of the participants.

Additionally, despite requiring surrogate consent, efforts will be made to obtain assent from the participant whenever possible. The assent process involves presenting information about the study that aligns with their cognitive abilities, seeking their agreement or disagreement (by words or non-verbal cues, including body language, facial expressions, gestures, or signs of agitation or discomfort when study-related discussions arise), and respecting their preferences to the best extent feasible.

However, as the study involves individuals with moderate to severe dementia, we anticipate fluctuations in decision-making capacity due to the natural progression of the disease. Consequently, participants' ability to provide or withdraw informed consent may vary throughout the study duration. Therefore, as regular assessments of decision-making capacity are essential, qualified professionals continue to evaluate participants' ability to comprehend study information and make informed decisions. If, during the study, individuals demonstrate a diminished capacity to provide consent, the protocol outlines procedures for involving surrogate decision-makers or legally authorized representatives to act in the participant's best interests. It is essential to mention that participation in the study is completely voluntary, and all participants can opt out at any time without penalty.

Research Description

AIM 1: Collect and store granular, *Personicle* using real-time visual and audio sensing technologies, (including the CCR), to gain a deeper understanding of a P-MSD's emotional state, agitation level, and gait over time.

Activity 1: Real-time data collection and sensing.

Milestone & Outcomes – Activity 1:

1a) A database storing a P-MSD's hourly activities and behavioral states in real-time;

1b) A database storing their stories as episodic knowledge graphs.

Activity 1.1.: Real-time sensing to create Personicles: In activity 1.1., we will capture a P-MSD's behavioral states and patterns through real-time "lifelogging" using motion sensing RGB cameras with video capture and voice recording capacity (Microsoft Kinect), optical motion sensors (LIDAR) and remote photoplethysmography (rPPG) [35]. rPPG is a non-invasive optical technique used to estimate physiological parameters (e.g. heart rate), remotely from a distance, without the need for direct contact with the skin. The cameras and sensors will be installed in two bedrooms and the living room in one facility at a time. We will begin observation of two patients at a time over at least a two-month period, starting in month 2, and continuing until month 15, to capture logs of 8 to 14 P-MSDs. At least 6-8 P-

MSD will remain for the full 2 years. In month 16, we will introduce the CCR, and also use it as a real-time sensing device (using its camera and lidar sensor onboard). The vision system of CCR is equipped with a human identification system. It can recognize more than one patient and track them. The CCR is also capable of identifying the caregivers as well. This allows the CCR to log gait, activities, agitation of the person under observation. This system would be configured during the startup/onboarding process.

These sensing technologies will capture lower-level “atomic”, person-centric data - the “logs” - with a timescale of a minute that can be enriched with higher-level, more meaningful behavioral states using the models developed in Activity 2. Specifically, we will model a P-MSD’s emotional state, agitation level, and gait and design a multi-model taxonomy and activity/states ontology, extending the existing “Personicle”, a person-centric temporal activity/states database developed at the UCI Institute for Future Health. For the purpose of recognizing agitation, our initial approach involves constructing models that leverage audio-visual sensing technologies. These models are designed to accurately recognize and categorize standard agitated behaviors as outlined in the Cohen-Mansfield Agitation Inventory (CMAI) [36,37]. This inventory includes a range of specific activities such as kicking, grabbing, pushing, throwing objects, and screaming. After successfully capturing and classifying these activities, the next phase will focus on utilizing these models to evaluate and quantify agitation levels in individuals with P-MSDs. We will store activities/states at the minute-level. This ontology will be used in Activity 2 to find the relationships among the hourly/daily states and activities and to create a robust temporal activity model.

The ultimate goal of the CCR system is to develop a dynamic model that accurately captures an individual's behavioral patterns. This model is intended to iteratively adapt and respond in a manner that positively influences and enhances these behaviors. In pursuit of this objective, a certain degree of tolerance for inaccuracy is considered acceptable, particularly in the precise identification of specific emotional states or the exact nature of agitated symptoms. This approach acknowledges that while exact details are beneficial, the primary focus is on the overall effectiveness of the model in improving behavioral outcomes.

Activity 1.2.: Collecting personal stories: We will collect personal stories from our sample as input in Activity 3. Our sensing devices will “lifelog” personal stories shared without probing when P-MSD converse with their caregivers in our facilities. In addition, we will also elicit personal stories through informal “interviews” of the caregivers at least twice a week. By monitoring P-MSD’s behavioral state in their rooms, we can also label such events with emotions and agitation. Data will undergo manual review and labeling by our team of qualified clinical experts, performed offline for optimal focus. In addition, 2-3 interviews will be conducted with at least one family caregiver, and the trained facility staff to collect information about their experiences about personal stories that have helped reduce agitation. We also ask family members to share archival data (e.g. videos of the P-MSD in the past, recordings of social, family events, etc.) that are representative of the P-MSD's past and positive/negative stories. The unstructured natural language data will be translated into a knowledge graph to represent events/stories, and additional attributes (e.g. people and emotions) over time as an episodic relationship.

AIM 2: Develop statistical models to understand and forecast emotional state, agitation level, and gait in real-time using ML/AI.

Activity 2: Modeling behavioral and cognitive states.

Milestone & Outcomes - Activity 2: Statistical models for modeling emotions, agitation, and gait.

Activity 2.1. Modeling emotional states: To model a P-MSD's emotional state, we will design multi-modal ML/AI models using four types of data: natural language content, facial expression, body movements, and vocal features. We will focus on modeling emotions from each modality (e.g., text as natural language content) and then integrate these models to create a comprehensive multi-modal model. The goal is to evaluate the emotional content of a conversation at a particular point in time. In establishing models, we consider both theoretical and technical aspects to ensure robustness. In establishing models, we consider both theoretical and technical aspects to ensure robustness. Specifically, in the case of the text modality, we will adapt existing theories (e.g., Parrott's theory [38]) and ML/AI techniques (e.g., pre-trained models, like RoBERTa [39], with fine-tuning) for inferring emotions from natural language to the context of dementia when modeling emotions using text extracted via speech-to-text techniques (for a comprehensive review, see [40]). This is especially pertinent to feature representation processes. In those processes, we will obtain word embeddings for emotions with the assumption that multiple emotional states can co-exist simultaneously (for further information, see [41].), as well as six distinct emotional states (i.e., sadness, fear, joy, love, surprise, and anger). In the modeling process, we will pre-train our models on open-source datasets widely used for emotion recognition tasks in the NLP community (e.g., IEMOCAP [42]) and then re-train them on datasets focused on conversations from P-MSD, using both open-source datasets (e.g., DementiaBank [43]) and our own (i.e., National Institute of Aging -NIA [44]).

Beyond the text modality, we will also train a deep learning model, using images and videos containing a P-MSD's facial expression, to infer emotions like sadness or anger from their smiles, frowns, brow movements, and lip configuration. We will train our models using large image datasets like FER [45], while also considering re-training on datasets specific to dementia patients [46, 47]. Additionally, videos and motion sensors will be utilized to infer emotions from a patient's body movements. Audio data will be employed to infer emotions based on markers such as speech rhythm, voice intensity, and other spectral features (e.g., mel-frequency cepstral coefficients (MFCCs). Empathy modeling, including the alignment of large language models (LLMs) using prompt engineering and reinforcement learning techniques, will also be conducted, considering the relationship between emotion and empathy.

Our research aims to make scientific contributions to ML/AI for emotional intelligence and empathetic conversation of CCRs in two main ways: 1) designing a new multi-modal model for empathy classification that enables CCRs to sense a patient's emotional state in real-time, and 2) aligning LLMs for empathy using various techniques, such as prompt engineering, reinforcement learning, or a combination of both. For the first point, we plan to build a semi-supervised ML model that uses a transformer architecture and pre-train it on healthy datasets. We will evaluate this baseline using dementia datasets carefully labeled by domain experts with respect to the emotional state of dementia patients. We will then improve this baseline model by enhancing its self-supervision based on features identified from four modalities (i.e., natural language content, facial expression, body movements, and vocal features) in data collected from dementia patients. We will examine different ML/AI models and evaluate their performance in both quantitative and qualitative ways.

Our major contribution will be the development of a semi-supervised emotion classifier that integrates multiple data modalities, including text, speech features, facial images, and body movements. We will use established evaluation protocols and metrics, such as accuracy, precision, F1-score, recall, and confusion matrices, to assess and compare various models to evaluate this baseline model. We will complement these evaluation protocols with a qualitative evaluation and inspection of our emotion classifier. Further, we will also examine the model's run-time given the need for real-time evaluation of emotions.

For the second point, we will first evaluate existing open LLMs (e.g. openlama and Neuro Lama) using our own benchmarking dataset and those published by others [28]. Departing from that, we will then employ prompt-engineering methods to enhance the empathetic capabilities of the LLMs' conversations. To assess the performance of these fine-tuned LLMs, we will employ well-established metrics and protocols in alignment research, including F-1 score, BLEU score, accuracy, precision, and recall. Beyond prompt engineering, we will also leverage reinforcement learning to optimize the LLMs for empathy. In addition to prompt engineering, we will leverage reinforcement learning to optimize the LLMs for empathy. Specifically, we plan to first use inverse reinforcement learning to develop a reward function for a highly empathetic caregiver.

Activity 2.2. Modeling degree of agitation: We will develop an ML/AI model for P-MSD's agitation that captures disturbances in verbal, vocal, or motor activity. Instead of relying on caregiver assessments, this model measures agitation in an objective way, using image, video, motion-sensor, and also non-intrusive physiological data collection using rPPG sensors (see above). We first aim to extend recent novel research for modeling agitation via inertial body sensors to our setting of non-intrusive rPPG sensing [48]. We plan to enrich such models using image analysis following recent efforts to infer agitation via facial grimacing rather than body sensing [49]. We also aim to integrate vocal/verbal data to infer states of agitation based on disturbances in verbal motor action. Using this multi-model ML/AL approach, we can enrich a P-MSD's Personicle with states of degree and type of agitation (e.g. motoric functions). We will then build predictive models that allow us to predict both emotions and agitational states.

Activity 2.3. Modeling Gait: To facilitate continuous yet non-intrusive monitoring of gait and mobility, we will use a vision-based method to conduct comprehensive evaluations of crucial gait factors. Through the application of state-of-the-art computer vision techniques, the method will assess gait speed, step length, step width, knee flexion angle, and the Time Up and Go (TUG) test. The foundation of the vision-based method lies in its utilization of dependable and adaptable body pose estimation algorithms, which can be tailored to monitor and assess transitional activities. Thus, a more comprehensive comprehension of the P-MSD's gait patterns will be achieved, leading to a more precise identification of potential fall risks [50]. Integrating time-variant, person-level emotional, agitation, and mobility states will allow us to build models for understanding their causal relationships and prediction of high-risk events.

The system has been designed to operate in a normal home setting. Standard gait parameters such as gait speed, step length, step width, cadence are calculated using combination of computer vision, statistical ML and trigonometry. TUG which is standard test has been estimated from the above parameters using a ML model [51].

AIM 3: Design an empathy-focused conversation model for successful human-robot intervention that considers a P-MSD's emotional state and associated events (using outcomes of Activity 1 & 2).

Activity 3: Designing empathy-focus CCR-patient interaction & simulated evaluation.

Milestone & Outcomes – Activity 3: NLP-based language models incorporating time-variant emotional states and events; an interface for empathy-based interaction (speech plus text display); and "simulated" validation of CCR usability as well as the impact of the conversational model on human emotional states, trust, and cognitive load.

In Activity 3, we focus on designing the empathy-focused intervention model for the CCR. The CCR is a social robot (3'9" & 26 kg) that is primarily focused on acting as a companion for communication with the P-MSD. In addition, it also acts as a sensing device for Activity 1. It is distinct from other social robots like PARO, AIRBO, or MINI that are "pets". The CCR is equipped with a 12-inch display and audio features to communicate with the P-MSD via speech and sound (e.g. music) but also text and visual content (e.g. pictures and movies displayed). The software/hardware platform architecture of the CCR is modular and uses a widely adopted and ROS. With state-of-the-art Simultaneous Localization and Mapping (SLAM) and human-aware navigation capabilities, our CCR can expertly navigate complex environments while prioritizing the safety and well-being of humans.

To ensure safety and well-being, the research will build upon the well-established stream of research using SLAM to ensure safe navigation. Indeed, SLAM algorithms are used in may robotics application in healthcare setting, not just for research but for day-to-day practice. Thus, the contribution to the field of SLAM research is not the priority of this research, but instead, the focus lies on emotional intelligence and empathic conversation. However, to ensure safety using SLAM, the team will advance existing SLAM models using vision for human recognition (in combination with other sensors such as LIDAR) to account for the dynamic environment in safety critical conditions like caregiving using visual SLAM (VSLAM) that integrates state-of-the-art research from fast advancing computer vision research. Similar to Fang et al. [52], we aim to improve the map construction of VSLAM to allow for dynamic scene construction.

Activity 3.1. Design of emotionally aware and empathy-based conversational model: To realize the CCR's emotional intelligence, we will leverage the results of Activity 2.1, to design an AI model for understanding and predicting a P-MSD's emotional state and agitation, using real-time sensing via the sensing technologies in the room. Using this AI model, the CCR will interact with the P-MSD when the likelihood of an emotional disturbance or agitation increases, using an emotionally aware and empathy-focused conversation model. To realize such a conversational model, we extend an open LLMs in a way that the CCR is able to engage in empathetic storytelling with the P-MSD. To do so, we will design a rule-based prompt engineering model that changes the degree of empathy of the CCR's language. This empathy prompting technique builds upon recent research of Dr. Brunswicker integrating AI with theories of empathy display in linguistics, reflecting the rules about how syntax and rhetoric in language display shape empathy in social conversations [53]. In this research, the effect of AI-enabled empathy display in conversations with an artificial agent has been validated in randomized behavioral experiments with healthy subjects, having different degrees of emotional disturbances (e.g. anger). We will advance these rule-based models in the context of dementia patients to investigate the linguistic dimension of empathy for lowering the risk of agitation. Further, we will integrate the LLM with an episodic knowledge graph that represents the P-MSD's personalized stories as connections of events, people, and activities [54]. Then the CCR is able to engage in personalized storytelling at the onset of an emotional arousal using empathetic language.

Activity 3.2. Design of user interface: Using the results of Activity 3.1., we will design a conversational interface that uses a dual-mode communication approach, incorporating both audio and text-based interactions. This approach will allow the CCR to display empathy through its voice and on the display screen. Through these two ways, personal stories are conveyed to the P-MSD. Our design will focus on usability, aiming to reduce the P-MSD's cognitive load. To achieve this, we will draw upon established usability heuristics for interactive applications and revise our existing interface through an iterative and heuristic-based usability design process. This process will involve developing Personas, which describe fictional or real individuals that represent end-user needs. These Personas will support the product design of the audio and text interface.

In evaluating the usability of this interface, we will base our tests on established literature, specifically focusing on two widely used methods in the practitioner community: Nielsen's [55, 56] ten general principles for interaction design and the heuristics for conversational agent design proposed by Langevin et. al. [57]. We specifically plan to utilize criteria by Langevin et al. with some modification as the basis for the design of our audio and speech interface. Evaluation heuristics can be broadly categorized into three main areas: design (e.g., aesthetic and minimalist design), content (e.g., designing content for dialogue), and security (e.g., data privacy). We will iteratively evaluate and update the interface to ensure optimal usability and performance accordingly.

Activity 3.3. Simulated Validation: Before evaluating our new empathy-focused CCR in the facilities with P-MSD, we will work in the UCI School of Nursing's simulation lab [58]. First, we will use the lab for "simulated", randomized experiments to examine whether and how our newly designed conversational model for the CCR responds to videos and audio-recordings of our P-MSD residents collected in Activity 1 and played on a large screen to perform out-of-sample validation experiments related to our P-MSD. We will evaluate the CCR's response time to approach the screen, its choice of a personal story, as well as the ability to align with the emotional state of the P-MSD using quantitative and qualitative techniques. Specifically, we will examine the content of the language, voice pitch, and the content logic of the stories. Afterwards, we will perform exploratory evaluation studies and also RCTs with healthy humans. We will carefully sample ~80 undergraduate/graduate students. Some will be provoked to be emotionally disturbed, using established scientific methods of emotion induction [59,60]. In the RCTs, we will use a simple 2x2 factorial treatment design (with CCR/without CCR and emotionally stable versus emotionally disturbed), lasting up to 1 hour. The goal is to evaluate usability, trustworthiness, helpfulness, cognitive load, and emotional states, and agitation before and after the interaction with the CCR (see activity 2) [59,60].

AIM 4: Pilot an empathy-focused intervention model for the CCR using a quasi-experimental intervention design and evaluate the conversational intervention longitudinally using mixed method approaches.

Milestone & Outcomes – Activity 4: Evaluate the impact of the CCR on the emotional state, agitation level, and gait pattern of 6 P-MSDs, and assess the acceptability/usability of caregivers. We will use relational event analysis of empathy-focused conversations of patients and a CCR, combined with P-MSD pathway analysis to examine dynamics over time.

Activity 4: Pilot implementation and field study.

Activity 4.1. Pilot implementation: Starting in month 16, 3 patients at a time will be interacting with 3 CCRs over a period of 3 months (one introductory month for patient and CCR to become familiar with each other, while assessing for and mitigating agitation caused by the CCR, as well as usability/acceptability periodically with family and professional staff every two weeks, during a two-month intervention period). Thereafter, an additional 3 P-MSD will be followed over the next 3 months, with observation ending in month 21, and final analysis and dissemination to follow. During the testing period, the CCR will be placed in the corner of the designated rooms and the family caregivers or staff do not need to program the CCR in any way. At the onset of potential emotional disturbance, the CCR will move to the P-MSD and observe closely and start talking to the P-MSD using empathy-focused conversations. If the P-MSD is expressing fear, the conversation will be different than when anger is expressed. The CCR will use storytelling with the P-MSD to prevent an onset of agitation, when signals are picked up by the CCR. If the CCR's empathy-based conversations to prevent agitation are not successful, the traditional de-escalation strategy will then be applied, such as facility caregiver support and behavioral intervention (always notified of any agitation events), and if necessary, the use of medication indicated and previously prescribed for the individual P-MSD.

Activity 4.2. Evaluation (Quantitative): We will measure the P-MSD's emotions, agitation level, and gait [49-54, 58-66] using models developed in Activity 2. This computational and non-intrusive granular measurement is very novel for dementia research and allows us to associate conversational events with a P-MSD's state. We will then perform a quantitative cross-sectional event-level analysis (comparing similar event periods within and without CCR interactions) and an intervention analysis using a dynamic event-state model for a few case settings, allowing us to understand how interactions with the CCR affect P-MSD outcomes over time (e.g. relational event modeling). We will also use longitudinal analysis for the family, assessing at baseline (beginning at month 1 of the 3-month intervention with CCR) and post intervention (end of 3rd month) using the Caregiver Self-Assessment Questionnaire [67], to assess perceived stress and health. For the facility caregivers, we use the Maslach Burnout Inventory [68] to assess emotional exhaustion and depersonalization.

Activity 4.3. Evaluation (Qualitative): We will also conduct interviews with both types of caregivers to assess Acceptability and Usability, with questions to assess why (or why not) the CCR impacted emotions, agitation, and falls. Thematic analysis will be conducted by qualitative investigators (AN, JAL), along with trained research staff on transcribed data from caregivers, using a rigorous process [69]. In vivo coding will be generated from the first and second cycle coding of the participants' responses [70]; these data will then be placed into an Excel file.

The rigor of the iterative data analysis will be ensured by the trustworthiness of data (i.e., credibility, confirmability, transferability, and dependability [71]. Credibility will be ensured by delivering the sessions in safe and confidential areas to ensure open communication, and audio-recording and transcribing all conversations to ensure the data is credible. Credibility will be further established by having two trained coders involved in the coding process and a senior coder overseeing the entire process. Confirmability will be ensured by ongoing discussions of caring for P-MSD with the caregivers and asking for feedback with data heard earlier. Finally, dependability will be ensured by a detailed method description.

Data Management/Data Security

All study data will be entered and managed using REDCap, a secure, web-based application designed to support data entry and storage for research data, entered using electronic tablets Layer (SSL) protocol, and protected via a Firewall service provided by OIT. Data encryption and network-based access/key management services will be used for data storage.

Privacy and Confidentiality

Participants' personal information and identifiers will be kept separately from the information, and a code will be assigned to each participant. The de-identified data will be utilized for analysis and training purposes. When capturing audio and visual data on participants with whom they or their family member or legal guardian has provided consent, we will use the ARX Data Anonymization Tool for anonymizing sensitive personal data. All the geographical locations, contexts, and environmental data will be abstracted, so that the collected data is not traceable to its original providers and will not be transmitted or stored in our cloud servers. In the cloud server, subjects' data will be processed using automated tools. We will make all relevant de-identified data and publications available following publication and will honor intellectual property.

The single electronic file containing participant identifiable information will be password-protected and stored on a laboratory computer in the Principal Investigator's locked laboratory, to which only the lead researcher, co-researchers, and study coordinators will have access. Research data will be stored electronically on a secure cloud-based data collection storage system maintained by the UCI Computer Science team led by Co-Investigator.

As the audio recordings serve as primary training data for the CCR, it's crucial to preserve the integrity of identifiable audio recordings without transcription to preserve Voice Characteristics such as tone, intonation, and nuances in speech patterns, which are crucial for capturing the full complexity of human speech in developing machine learning algorithms. Therefore, the audio recordings will not be transcribed to prevent a loss or distortion of training data quality. However, De-identification procedures commence immediately after the audio recordings are obtained, ensuring minimal delay in protecting participant confidentiality.

The de-identification process for audio recordings follows 1) utilizing specialized software such as Audacity/ Adobe Audition designed for de-identification purposes, 2) voice masking methods (including pitch shifting, time stretching, or applying filters to modify the frequency characteristics of the voice) within these tools to alter the pitch, tone, or timbre of the participants' voice to make it less recognizable, 3) quality checks to ensure all identifiable information has been effectively removed or altered, 4) speech intelligibility metrics such as Signal-to-Noise Ratio (SNR), Mean Opinion Score (MOS), and Perceptual Evaluation of Speech Quality (PESQ) to objectively measure the clarity of the modified audio, and 5) at last storing the de-identified audio recordings securely in a protected database to prevent unauthorized access. Similarly, preserving identifiable video recordings without transcription is crucial for comprehensive emotion and agitation detection as well as gait analysis, especially in training the CCR. These recordings contain critical contextual cues, facial expressions, body language, and environmental elements and De-identification risks the integrity of developing precise training models. So, the video recordings will not be de-identified.

Compensation

In aims 1 and 4 where interviewers will take place with family and professional caregivers, a compensation of $20 will be provided per interview. In total, we anticipate conducting interviews with 8 - 14 family and professional caregivers in Aim 1 and 6 - 12 family and professional caregivers in Aim 4. In addition, in Aim 3, we will compensate approximately 50 students in the simulation lab with $10.

Results

To provide cultural nuances for the development and training of the CCR, a CAB was formed in December 2023. The CAB is composed of six members, three of which were family caregivers of dementia patients residing in Board & Care facilities and three were directors of or professional caregivers working in these facilities. During the first of several meetings which will occur several times over the period of the grant, the members expressed gratitude for the work that is planned and highlighted the importance of the work in mitigating fall risk among PWD.

Activity 1: Architecture development was initiated for the Personicle, beginning January 2024, using an existing open-source data and foundations from the skeletal-based recognition module developed several years earlier [72]. To date, 16 actions (Atomic, Transitions, and Agitation Markers) have been designed using client-server architecture. Accuracy was shown to be 99% on the test set and 96% on the validation set to date. Currently the tracker is being developed for P-MSD and the caregivers with facial recognition.

Further, to set the stage for capturing audio and visual images of the PWD in their residential settings, camera stations were meticulously designed and customized according to the specific needs of the facilities we inspected. These internet-connected cameras are capable of detecting movement and start recording for 30 seconds upon sensing it. In addition, a workstation computer was purchased in June 2024, and we will be utilizing its graphical processing unit (GPU) for training models in Aims 1 and 2, and serving as substantial storage for all the collected data. In addition, an automated and secure pipeline was developed, allowing the cameras to compress and transmit recorded videos, including point-cloud, images, and audio data, to our high-performance workstation, where team members can securely access and analyze the data.

To test the functionality of our cameras, we conducted role-playing scenarios simulating interactions between dementia patients and caregivers, guided by dementia experts and facility observations. These simulations helped us identify potential issues, such as the high volume of TVs in the rooms interfering with the detection of conversations between patients and caregivers, thereby limiting our ability to detect signs of agitation or emotional changes from audio cues. We addressed these issues by finding solutions to improve the accuracy of our recordings under such conditions. Through these activities, we created a simulated dataset that mirrors real-life situations involving dementia patients. This dataset has been instrumental in training and testing our models to detect possible agitation or emotional changes, ensuring our system's reliability and effectiveness in real-world applications.

Activity 2: In Activity 2.1, we initiated the development of ML/AI models by identifying open-source datasets and pre-trained models relevant to primary of interests, such as emotion recognition and action recognition. These models serve as baselines for our project. This stage is particularly pertinent for problems related to classification or detection. Therefore, we focused on finding suitable open-source datasets and integrating them into our research objectives. The results of this stage are two-fold: dataset identification/construction and modeling.

Our initial objective was to establish an emotion classifier in various ways (e.g., semi-supervised way). To achieve this, we explored various datasets for emotion classification in the context of P-MSD and identified DementiaBank, which consists of both text and audio files from 78 dementia patients. As the dataset was not labeled, seven human labelers, including three domain experts, were involved. In this process, we pre-processed the dataset, including audio segmentation and utterance extraction from text. During a training session, we built a guideline for labeling based on discussions with the labelers. Thereafter, the labelers labeled the data accordingly. The labeling activity was completed, and the labeled data is currently in the validation process.

In parallel to the labeling activity, we established a baseline model for emotion classification using IEMOCAP and MELD [73], open-source datasets widely used in emotion classification tasks. IEMOCAP includes 151 dyadic dialogues, while MELD consists of 1,039 dyadic dialogues; both datasets contain text and audio data in the context of healthy datasets. We obtained features from pre-trained models, RoBERTa, for text and speech representation, respectively. Thereafter, multi-modal modeling for emotion was implemented using transformers with self-attention and cross-attention layers, followed by emotion prediction. Through both quantitative and qualitative analyses of the results, we found that enhancing feature representations of each modality contributes to model performance.

Activity 2.2. Concurrently, the team from NaviGAIT and computer scientists from UCI are working on agitation activity. We used the collected data and trained several models to detect different motions, such as sitting, standing, etc which are capable of annotating the video every 3 seconds. Currently, we are at the phase of training the model and improving its accuracy for agitation. In terms of audio, we are able to detect audio events by the Agitation Maker. In which we convert the audio into text and use semantic scoring from GPT to get an agitation score. This work is under progress.

Activity 3.2: In terms of Aim 3, the Design Use subgroup began developing personas by April 2024 and to date has developed 10 personas characterizing P-MSD that will be useful for training the models. The personas include the background of real persons residing in board & care facilities, with first names and pictures replaced, and includes brief description of the person, interests, personality, sleep quality, health conditions and medications prescribed, conversational ability, entertainment and communication devices used, and what makes the person happy or upset. Finally, more recently added have been behaviors to watch for that often leads to agitation, and caregiver instructions in dealing with these behaviors. A second document was also developed which outlines the typical daily activities of each person.

Next Steps

In Activity 1, using client-server architecture, work is ongoing in predicting an action/activity every 3 seconds. With multiple camera stations to be in place in the next month, continued functionality of the cameras will be tested using role playing. Further, we will continue to advance the dataset that mirrors real-life situations involving dementia patients, and will be used to train and test our models to detect possible agitation or emotional changes, ensuring our system's reliability and effectiveness in real-world applications. Ongoing design-use personas and workplans will be created to add rich pre-model training data.

Once an appropriate baseline model is identified, we will proceed to gather data. In scenarios where suitable datasets are unavailable, we will undertake the collection of in-house data, tailored to our specific requirements. In instances where an open-source dataset is available but inadequately labeled, our team will conduct a meticulous process of re-labeling the data to ensure alignment with our research needs. Following data acquisition and preparation, we will embark on the training and fine-tuning of our model. This process leverages the baseline model and is informed by the unique characteristics of our collected or curated dataset.

The next step involves re-training the baseline model on the labeled data of P-MSD (i.e., DementiaBank), after training and fine-tuning the baseline model with healthy datasets (i.e., IEMOCAP and MELD), followed by both quantitative and qualitative analyses. This process includes building and testing various architectures, such as enhancing modality representations (e.g., text representation), to improve model performance. Considering the nature of the data for emotion recognition (e.g., imbalance and data size), we will adopt and develop various ML techniques (e.g., attention weight) to mitigate potential issues that can arise. In parallel, we will establish a baseline model for empathy recognition using LLM alignments, including prompt engineering and reinforcement learning, considering the close relationship between emotion and empathy. Both quantitative and qualitative analyses of the results of all the interim processes will be conducted. After scrutinizing these results and gaining insights from the analyses, we will improve our baseline models to build a robust CCR empathic conversational model with data to be collected.

The subsequent phase involves rigorous testing of the trained model in various stages to ascertain its efficacy and reliability. It is imperative to note that the transition from a trained model to its application in real-world settings involves additional steps. Given our use of Jetson devices for edge computing, we convert the trained model into a TensorRT format. This conversion is a critical step to facilitate the deployment of the model on edge devices. Such deployment enables real-time inference while also adhering to privacy considerations, a crucial aspect of our research protocol.

Finally, Aim 4 will be the culmination of our evaluation, wherein the CCR will be stationed in the residences of the P-MSD enrolled in our study. Our goal is to then test and evaluate the empathy-focused conversation model for the CCR with the enrolled residents and gather qualitative data from both family and professional caregivers.

Plan for Pre-training the Robot

We plan first to develop a robust and efficient pre-training strategy for emotion detection in dementia patients using publicly available datasets. We first will follow an unsupervised pre-training plan using speech feature extraction (e.g., Mel-frequency cepstral coefficients, pitch, and energy), visual feature extraction (e.g., facial features related to emotion expressions, such as action units and facial landmarks), and physiological feature extraction (physiological signals like heart rate and heart rate variability). The second step involves utilizing supervised pre-training through finetuning the pre-trained model on smaller, labeled datasets. We finally further fine-tune the pre-trained model on the final emotion detection task specific to dementia patients.

Current literature review includes publicly available datasets, including Dem@Care, Dementia Bank, and University of California Irvine Alzheimers Disease Research Center (ADRC) dataset collections. These available data formats are text, voice, video, and sensor data (physiological signals). All datasets are in English, as the inclusion criteria were set and includes healthy and/or dementia participants. The datasets were collected in controlled settings such as research laboratories or uncontrolled settings such as home care or participants' homes.

To enhance the existing literature review seamlessly, we will delve deeper into datasets that provide a nuanced perspective on dementia related research. For example, one dataset that stands out is the Dem@Care dataset, which serves as a resource with its collection of video and audio recordings captured in both laboratory and home settings using the Kinect RGB D device. This dataset goes further by including information on sleep patterns and physiological aspects. For another example, the TIHM dataset may be promising for remote healthcare monitoring as it leverages sensor data like motion and sleep mat recordings. What makes this dataset particularly interesting is its inclusion of labels for agitation events in the 6 hours allowing for correlations with sleep and activity data. Another valuable resource may be the Prompt database from i2 labs, which adds diversity by featuring video, audio and facial expressions of participants in a home-based setting. This database provides a resource for training classification models.

In contrast there is the DEAP dataset that focuses more on classification but does not specifically revolve around dementia patients. For the dimension, we came across two datasets: Mobile Device Voice Recordings at King's College London (MDVR KCL) and The ADReSS Challenge. These datasets serve as voice banks for training audio-based models on individuals with dementia. However, both datasets require annotations to develop emotion-based audio models.

This careful choice of datasets each, with its characteristics not only expands the range of our exploration but also forms the basis for a thorough and detailed investigation into dementia, across different methods and situations. As we progress, it will be important to consider the validity of such modeling and training using datasets from people without dementia or mild dementia and may not be fully applicable to persons with more advanced dementia. We will also consider cultural adaptation in this companion robots including customs, languages, and cultural social norms in the designing phase. Engaging with communities to understand their values and preferences is key to ensuring these robots are viewed as helpful companions rather than intrusive tools.

Discussion

Worsening dementia places P-MSD at high risk for significant declines in cognition, emotion, and behavioral disturbances, including falls and serious injury due to unpredictable agitation. The findings of our study to date include for Aim 1 and its activities, identifying, generating and collecting public datasets for the training and testing of our models. We identified the key activities P-MSD demonstrate, including signs of agitation (like standing up, pushing and striking), as well as located datasets for some of them as well as generated sample datasets with staff in our simulation lab and NaviGAIT team internal recordings to train AI models to detect key activities and extract events from them. The result of this activity was a list of motions and audio events that we found to be correlated with agitation and collecting their respective datasets. Finally, we acquired a computer workstation and camera design for costume dataset creation in preparing for Aim 4 and testing real world situations along with capturing new data in our simlab.

In Aims 2, we established a baseline model with benchmark datasets (i.e., IEMOCAP and MELD) for the emotion classification task before incorporating the labeled data of P-MSD (i.e., DementiaBank), which was labeled by our team. We extracted features from each modality (i.e., text and audio) using pre-trained models (e.g., RoBERTa) and leveraged them for multimodal modeling using transformer-based architectures. Furthermore, our analyses indicate that text representation is more challenging than speech representation. This discrepancy may be due to the fact that emotion is not heavily dependent on text. Specifically, emotion is often implicit and not always conveyed through emotion-specific words like "mad." As a result, text-based emotion recognition can be more complex, especially as contextual information plays a crucial role in detecting emotions within text. This discrepancy in modality representation merits further discussion.

Given our findings and analyses of the results, our next steps include the following. First, we will navigate and develop embedding techniques (e.g., ontology-based approaches) to enhance feature representations, especially text representation. In parallel, we will apply our baseline model to the labeled data and conduct quantitative and qualitative analyses of the results to gain insights into the languages of emotion expression by P-MSD. Another step includes finding ways to mitigate potential problems caused by highly skewed data (e.g., a lot of “neutral” labels in the labeled data) since data quality is of paramount importance to model performance. Our interim goal is to improve text and speech modeling further. Following these steps, we will focus on empathy modeling and LLM alignments, as well as incorporating more modalities (e.g., from video). We expect that these interim processes will significantly contribute to achieve our overarching goal: developing a robust CCR empathic conversational model with our data to be collected, ultimately benefiting

P-MSD.

Finally, as part of Aim 3, we have begun to develop personas and workplans focused on real P-MSD to help in pre-train the models which will then set the stage for Aim 4.

Comparison to Prior Work. This proposed study is innovative in many ways. It is a caregiving approach, developed with advice from community stakeholders, and will be tested both in our state-of-the-art simulation space and in the community. The Care Companion robot will be designed to forecast and recognize signs of agitation and other emotional behaviors, and upon recognition, intervene with empathic verbal communication. Care Companion Robots will be assessed for their ability to objectively observe, record, analyze, and appropriately respond to P-MSD, relieving human caregivers of some of the time-consuming and stressful work they currently do. Robot caregivers can also be able to potentially reduce the escalation of agitation among persons with moderate to severe dementia, as robots, unlike human caregivers, will not be prone to displays of distressed emotions that can exacerbate the distress and behavioral disturbance of P-MSD.

By the end of our study, we hope to develop a comparable CCR with our partnering small facility-based settings where P-MSD reside. Currently under development is the notification system that the CCR will send to professional staff in the facilities when an issue arises that their support to the patient is needed.

The critical aspects we will be focusing on relate to whether we can train the CCR to predict agitation and be able to use storytelling to calm the person down while a support person is notified. We will be analyzing the communication and behaviors of individuals with dementia within their environments. This technique can then be applied to broader contexts, ensuring that the care strategies are versatile enough to benefit a wider range of individuals in various care settings (from homes to communities).

Strengths and Limitations

The strengths of our study are significant. Our team have been collecting data from multiple sources, and the use of the Personicle which will gather data through real-time “lifelogging” will be a novel approach. In terms of our emotion recognition model, by leveraging diverse benchmark datasets such as IEMOCAP and MELD, our model benefits from exposure to varied conversational contexts, enhancing its generalizability. Moreover, the multimodal approach, which extracts and integrates features from text and speech, allows for a comprehensive analysis of emotional information. Furthermore, utilizing transformer-based architectures, our model effectively processes both local and global emotional context, handling the short-term and long-term dependencies inherent in conversations. Additionally, our model’s performance benefits from the enhanced feature representation for each modality.

Despite these strengths, there are several prerequisites for the CCR to deliver on this vision. However, first is the need for sufficient data to train and test the algorithmic models. There is a dearth of public datasets of P-

MSD due in large part to the sensitive personal health information that would be revealed in an open format. Our team is overcoming these challenges by the combination of searching and identifying new datasets, utilizing the simulation lab to create datasets and in Aim 4, to test the CCR in the B&C facilities with P-MSD

Further, building models for emotion classification comes with its own set of challenges. First, the data distribution of currently available benchmark datasets is highly skewed, with many "neutral" labels. Moreover, the diverse data sources of the benchmark datasets add to the complexity of the task. Specifically, IEMOCAP includes dyadic conversations while MELD includes multi-party conversations. Additionally, IEMOCAP was created by actors in a laboratory setting, resulting in high-quality audio, whereas MELD was based on the TV show "Friends." Consequently, model performance is sensitive to ambient environments, such as noise.

Conclusion

As family caregivers of dementia patients often experience high levels of stress when dealing with a loved one's progressive dementia, family disruption occurs as these persons are moved to community facilities. Moreover, even at these facilities, there is an inability of the professional staff to constantly observe and manage signs of agitation and fall risk. Our study has been making great progress in gathering existing datasets, and collection of data via cameras, and subsequently, IoT devices. This granular, personal, chronological data in the Personicle will provide one of the first implementations of the framework envisioned by Jalali et. al [74]. The development of novel models that receive input from P-MSD, algorithmically assess their emotional state and determine the appropriate intervention strategy, and deliver an empathic conversational intervention is perhaps the most significant innovation in our proposal.

Dissemination Plan and Future Directions

While this study will be conducted in a facility-based study, agitation is an issue experienced by persons with dementia who reside in many types of facilities as well as at home. Thus, regardless of setting, the critical aspects we will be focusing on whether we can train the CCR to predict agitation and be able to use storytelling to calm the person down while a support person is notified. We will be analyzing the communication and behaviors of individuals with dementia within their environments. This technique can then be applied to broader contexts, ensuring that the care strategies are versatile enough to benefit a wider range of individuals in various care settings (from homes to communities).

Future research will focus on developing the baseline model by fine-tuning and enhancing feature representations. This includes assessing the model both quantitatively and qualitatively. Thereafter, we will retrain the model with the data from DementiaBank, which has been carefully labeled by our domain experts. We will explore and develop various model architectures (e.g., semi-supervised models) to analyze emotions in both healthy and dementia datasets and evaluate the model performance, after which features from another modality, video, will be incorporated into the model.

This proposal thus has the potential to have a significant impact on an emerging field of computational dementia science and on society, by reducing episodes of falls, injuries, and unnecessary hospitalizations of P-MSD, while helping to relieve caregiver burden.

Funding Statement
Funding for CCR was provided by NOYCE Institute

Conflict of Interest
All authors declare no conflict of interest.

Data Availability
Data sharing is not applicable to this article as the data collection phase is not started yet.

Author Contributions
A.N. wrote the grant, assisted in conducting the research and drafted the manuscript. N.D. supervised the methodology and reviewed the manuscript critically from technical perspective. J.A.L. reviewed the manuscript critically from the nursing perspective. D.K., E.K., C.K., M.A, A.M.R., M.R., and S.B. played a key role in assisting in designing the study, and revising the manuscript critically. H.R., H.L., R.J., F.A., A.A., B.Q., and T.Y.B contributed to software development. J.L., L.G., and J.R. reviewed the manuscript critically.

Abbreviations
ADRC: Alzheimers Disease Research Center
ADRD, dementia: Alzheimer's disease and related dementias
AI: Artificial intelligence
B/C: Board & Care
CAB: Community Advisory Board
CCR: Care Companion Robot
CMAI: Cohen-Mansfield Agitation Inventory
Humanoids: Human-like robots
LIDAR: Optical motion sensors
LLMs: Large Language Models
MDVR KCL: Mobile Device Voice Recordings at King's College London
MFCC: mel-frequency cepstral coefficient
ML: Machine Learning
NIA: National Institute of Aging
OIT: Office of Information Technology
P-MSD: Persons experiencing moderate to severe dementia
Personicle: Personal, chronological data
ROS: Widely adopted and robust open robotics framework
SLAM: Simultaneous localization and mapping
SSL: Secure Sockets Layer
TUG: Time Up and Go
VSLAM: Visual SLAM

References

1. Alzheimer’s Association, “2023 Alzheimer’s Disease Facts and Figures,” 2023. [Online]. Available: https://www.alz.org/media/documents/alzheimers-facts-and-figures.pdf
2. Tampi RR, Jeste DV. Dementia is more than memory loss: neuropsychiatric symptoms of dementia and their nonpharmacological and pharmacological management. American Journal of Psychiatry. 2022 Aug 1;179(8):528-43. doi: 10.1176/appi.ajp.20220508. PMID: 35921394.
3. Cerejeira J, Lagarto L, Mukaetova-Ladinska EB. Behavioral and psychological symptoms of dementia. Frontiers in neurology. 2012 May 7;3:73, Accessed: Jul. 21, 2023. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fneur.2012.00073
4. Fernando E, Fraser M, Hendriksen J, Kim CH, Muir-Hunter SW. Risk factors associated with falls in older adults with dementia: a systematic review. Physiotherapy Canada. 2017;69(2):161-70. doi: 10.3138/ptc.2016-14.
5. Fillit H, Aigbogun MS, Gagnon-Sanschagrin P, Cloutier M, Davidson M, Serra E, Guérin A, Baker RA, Houle CR, Grossberg G. Impact of agitation in long-term care residents with dementia in the United States. International Journal of Geriatric Psychiatry. 2021 Dec;36(12):1959-69. doi: 10.1002/gps.5604. Epub 2021 Aug 27. PMID: 34286877; PMCID: PMC9291552.
6. Beauchet O, Sekhon H, Schott AM, Rolland Y, Muir-Hunter S, Markle-Reid M, Gagne H, Allali G. Motoric cognitive risk syndrome and risk for falls, their recurrence, and postfall fractures: results from a prospective observational population-based cohort study. Journal of the American Medical Directors Association. 2019 Oct 1;20(10):1268-73. doi: 10.1016/j.jamda.2019.04.021.
7. Booth V, Logan P, Harwood R, Hood V. Falls prevention interventions in older adults with cognitive impairment: a systematic review of reviews. International Journal of Therapy and Rehabilitation. 2015 Jun 2;22(6):289-96. doi: 10.12968/ijtr.2015.22.6.289
8. Koenig AM, Arnold SE, Streim JE. Agitation and irritability in Alzheimer’s disease: evidenced-based treatments and the black-box warning. Current psychiatry reports. 2016 Jan;18:1-0. doi: 10.1007/s11920-015-0640-7.
9. Fillit H, Aigbogun MS, Gagnon-Sanschagrin P, Cloutier M, Davidson M, Serra E, Guérin A, Baker RA, Houle CR, Grossberg G. Impact of agitation in long-term care residents with dementia in the United States. International Journal of Geriatric Psychiatry. 2021 Dec;36(12):1959-69.
10. Narme P. Burnout in nursing staff caring for patients with dementia: role of empathy and impact of empathy-based training program. Geriatrie et psychologie neuropsychiatrie du vieillissement. 2018 Jun 1;16(2):215-22. doi: 10.1684/pnv.2018.0735.
11. Cohen-Mansfield J, Marx MS, Freedman LS, Murad H, Thein K, Dakheel-Ali M. What affects pleasure in persons with advanced stage dementia?. Journal of psychiatric research. 2012 Mar 1;46(3):402-6.
12. Person M, Hanssen I. Joy, happiness, and humor in dementia care: A qualitative study. Creative nursing. 2015 Feb;21(1):47-52.
13. Guzmán-Vélez E, Feinstein JS, Tranel D. Feelings without memory in Alzheimer disease. Cognitive and behavioral neurology. 2014 Sep 1;27(3):117-29.
14. Lee JA, Kim HJ, Ju E, Guo Y, Rousseau J, Gibbs L, Tran TM, Tom CE, Sabino-Laughlin E, Kehoe P. A culturally and linguistically appropriate telephone support intervention for diverse family caregivers of persons with dementia during the COVID-19 pandemic. Journal of Family Nursing. 2022 Aug;28(3):231-42. doi: 10.1177/10748407221106531.
15. Ju E, Burton C, Kim J, Guo Y, Park JI, Qu A, Gibbs L, Lee JA. Sleep disturbances and interrelationship between persons with dementia and family caregivers: The lived experiences of Korean American Dyads. Geriatric Nursing. 2024 Jan 1;55:144-51.

https://doi.org/10.1016/j.gerinurse.2023.10.028. (In press).
16. Le CT, Lee JA. Home visit based mindfulness intervention for vietnamese american dementia family caregivers: a pilot feasibility study. Asian/Pacific Island Nursing Journal. 2021;5(4):207. doi: 10.31372/20200504.1096. PMID: 33791408 PMCID: PMC7993887
17. Kim HJ, Kehoe P, Gibbs LM, Lee JA. Caregiving experience of dementia among Korean American family caregivers. Issues in mental health nursing. 2019 Feb 1;40(2):158-65. doi: 10.1080/01612840.2018.1534909. PMID: 30620625.
18. Leggett AN, Polenick CA, Maust DT, Kales HC. Falls and hospitalizations among persons with dementia and associated caregiver emotional difficulties. The Gerontologist. 2018 Mar 19;58(2):e78-86. doi: 10.1093/geront/gnx202.
19. Kales HC, Lyketsos CG, Miller EM, Ballard C. Management of behavioral and psychological symptoms in people with Alzheimer's disease: an international Delphi consensus. International psychogeriatrics. 2019 Jan;31(1):83-90. doi: 10.1017/S1041610218000534. Epub 2018 Aug 2. PMID: 30068400.
20. Teri L, Logsdon RG, Peskind E, Raskind M, Weiner MF, Tractenberg RE, Foster NL, Schneider LS, Sano M, Whitehouse P, Tariot P. Treatment of agitation in AD: a randomized, placebo-controlled clinical trial. Neurology. 2000 Nov 14;55(9):1271-8. doi: 10.1212/WNL.55.9.1271.
21. Brown EL, Agronin ME, Stein JR. Interventions to enhance empathy and person-centered care for individuals with dementia: a systematic review. Research in gerontological nursing. 2020 May 1;13(3):158-68. doi: 10.3928/19404921-20191028-01.
22. Huschilt J, Clune L. The use of socially assistive robots for dementia care. Journal of gerontological nursing. 2012 Oct 1;38(10):15-9. doi: 10.3928/00989134-20120911-02.
23. Yu R, Hui E, Lee J, Poon D, Ng A, Sit K, Ip K, Yeung F, Wong M, Shibata T, Woo J. Use of a therapeutic, socially assistive pet robot (PARO) in improving mood and stimulating social interaction and communication for people with dementia: Study protocol for a randomized controlled trial. JMIR research protocols. 2015 May 1;4(2):e4189. doi: 10.2196/resprot.4189.
24. Yu C, Sommerlad A, Sakure L, Livingston G. Socially assistive robots for people with dementia: Systematic review and meta-analysis of feasibility, acceptability and the effect on cognition, neuropsychiatric symptoms and quality of life. Ageing research reviews. 2022 Jun 1;78:101633. doi: 10.1016/j.arr.2022.101633.
25. Otaka E, Osawa A, Kato K, Obayashi Y, Uehara S, Kamiya M, Mizuno K, Hashide S, Kondo I. Positive Emotional Responses to Socially Assistive Robots in People With Dementia: Pilot Study. JMIR aging. 2024 Apr 11;7(1):e52443.
26. Sorin V, Brin D, Barash Y, Konen E, Charney A, Nadkarni G, Klang E. Large language models (llms) and empathy-a systematic review. medRxiv. 2023:2023-08.
27. Hasan MR, Hossain MZ, Gedeon T, Rahman S. LLM-GEm: Large language model-guided prediction of people’s empathy levels towards newspaper article. InFindings of the Association for Computational Linguistics: EACL 2024 2024 Mar (pp. 2215-2231).
28. Rashkin H, Smith EM, Li M, Boureau YL. Towards empathetic open-domain conversation models: A new benchmark and dataset. arXiv preprint arXiv:1811.00207. 2018 Nov 1.
29. Rios Rincon AM, Miguel Cruz A, Daum C, Neubauer N, Comeau A, Liu L. Digital storytelling in older adults with typical aging, and with mild cognitive impairment or dementia: a systematic literature review. Journal of Applied Gerontology. 2022 Mar;41(3):867-80.
30. Subramaniam P, Woods B. Digital life storybooks for people with dementia living in care homes: an evaluation. Clinical interventions in aging. 2016 Sep 16:1263-76.
31. Damianakis T, Crete-Nishihata M, Smith KL, Baecker RM, Marziali E. The psychosocial impacts of multimedia biographies on persons with cognitive impairments. The Gerontologist. 2010 Feb 1;50(1):23-35.

32. Strong J. Immersive virtual reality and persons with dementia: a literature review. Journal of gerontological social work. 2020 Apr 2;63(3):209-26.
33. Hirt J, Beer T. Use and impact of virtual reality simulation in dementia care education: A scoping review. Nurse education today. 2020 Jan 1;84:104207.
34. University of California, Irvine. “Use of Surrogate Consent in Research,” https://research.uci.edu/human-research-protections/subject-enrollment/informed-consent/use-of-surrogate-consent-in-research, Accessed: July 2024.
35. Botina-Monsalve D, Benezeth Y, Miteran J. Performance analysis of remote photoplethysmography deep filtering using long short-term memory neural network. BioMedical Engineering OnLine. 2022 Dec;21(1):1-27. doi: 10.1186/s12938-022-01037-z.
36. Cohen-Mansfield J, Marx MS, Rosenthal AS. A description of agitation in a nursing home. Journal of gerontology. 1989 May 1;44(3):M77-84.
37. Finkel SI, Lyons JS, Anderson RL. Reliability and validity of the Cohen–Mansfield agitation inventory in institutionalized elderly. International Journal of Geriatric Psychiatry. 1992 Jul;7(7):487-90.
38. Parrott, W. G. (Ed.). (2001). *Emotions in social psychology: Essential readings*. psychology press.
39. Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., ... & Stoyanov, V. (2019). Roberta: A robustly optimized bert pretraining approach. *arXiv preprint arXiv:1907.11692*.
40. Zad S, Heidari M, James Jr H, Uzuner O. Emotion detection of textual data: An interdisciplinary survey. In2021 IEEE World AI IoT Congress (AIIoT) 2021 May 10 (pp. 0255-0261). IEEE. doi: 10.1109/AIIoT52608.2021.9454192.
41. MohammadiBaghmolaei R, Ahmadi A. TET: Text emotion transfer. Knowledge-Based Systems. 2023 Feb 28;262:110236. doi: 10.1016/j.knosys.2022.110236.
42. Busso C, Bulut M, Lee CC, Kazemzadeh A, Mower E, Kim S, Chang JN, Lee S, Narayanan SS. IEMOCAP: Interactive emotional dyadic motion capture database. Language resources and evaluation. 2008 Dec;42:335-59.
43. Becker, J. T., Boller, F., Lopez, O. L., Saxton, J., & McGonigle, K. L. (1994). The natural history of Alzheimer's disease: description of study cohort and accuracy of diagnosis. *Archives of Neurology, 51*(6), 585-594.
44. National Institute on Aging. “Data Resources for Behavioral and Social Research on Aging,” https://www.nia.nih.gov/research/dbsr/data-resources-behavioral-and-social-research-aging, Accessed: July 2024.
45. Goodfellow, I. J., Erhan, D., Carrier, P. L., Courville, A., Mirza, M., Hamner, B., ... & Bengio, Y. (2013). Challenges in representation learning: A report on three machine learning contests. In *Neural information processing: 20th international conference, ICONIP 2013, daegu, korea, november 3-7, 2013. Proceedings, Part III 20* (pp. 117-124). Springer berlin heidelberg.
46. Khaireddin Y, Chen Z. Facial emotion recognition: State of the art performance on FER2013. arXiv preprint arXiv:2105.03588. 2021 May 8. doi: 10.48550/arXiv.2105.03588.
47. Minaee S, Minaei M, Abdolrashidi A. Deep-emotion: Facial expression recognition using attentional convolutional network. Sensors. 2021 Apr 27;21(9):3046. doi: 10.3390/s21093046.
48. Bankole A, Anderson M, Knight A, Oh K, Smith-Jackson T, Hanson MA, Barth AT, Lach J. Continuous, non-invasive assessment of agitation in dementia using inertial body sensors. InProceedings of the 2nd Conference on Wireless Health 2011 Oct 10 (pp. 1-9). doi: 10.1145/2077546.2077548.
49. Becouze P, Hann CE, Chase JG, Shaw GM. Measuring facial grimacing for quantifying patient agitation in critical care. Computer methods and programs in biomedicine. 2007 Aug 1;87(2):138-47.

50. James K, Schwartz AW, Orkaby AR. Mobility assessment in older adults. The New England journal of medicine. 2021 Nov 11;385(21):e76. doi: 10.1056/NEJMvcm2009406.
51. Ma J. Predicting TUG score from gait characteristics with video analysis and machine learning. InChinese Intelligent Automation Conference 2023 Sep 23 (pp. 1-12). Singapore: Springer Nature Singapore.
52. Fang B, Mei G, Yuan X, Wang L, Wang Z, Wang J. Visual SLAM for robot navigation in healthcare facility. Pattern recognition. 2021 May 1;113:107822.
53. S. Brunswicker, C. Rashidian, M. Attal, Z. Yifan, and D. Linna, "The Impact of Empathy in Conversational AI," in *9th International Conference for Computational Social Science,* Copenhagen, Denmark, Jul. 2023.
54. Pan S, Luo L, Wang Y, Chen C, Wang J, Wu X. Unifying Large Language Models and Knowledge Graphs: A Roadmap. arXiv preprint arXiv:2306.08302. 2023 Jun 14. doi: 10.48550/ARXIV.2306.08302.
55. Nielsen, J. (1995). "10 Usability Heuristics for User Interface Design". Nielsen Norman Group.
56. Nielsen, J. (1994, April). Usability inspection methods. In Conference companion on Human factors in computing systems (pp. 413-414).
57. Langevin, R., Lordon, R. J., Avrahami, T., Cowan, B. R., Hirsch, T., & Hsieh, G. (2021, May). Heuristic evaluation of conversational agents. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (pp. 1-15).
58. Quiñones D, Rusu C. How to develop usability heuristics: A systematic literature review. Computer standards & interfaces. 2017 Aug 1;53:89-122. doi: 10.1016/j.csi.2017.03.009.
59. "UCI nursing school simulation laboratory." https://nursing.uci.edu/2020/05/nursing-simulations/
60. Lobbestael J, Van Vreeswijk MF, Arntz A. An empirical test of schema mode conceptualizations in personality disorders. Behaviour research and therapy. 2008 Jul 1;46(7):854-60. doi: 10.1016/j.brat.2008.03.006.
61. Siedlecka E, Denson TF. Experimental methods for inducing basic emotions: A qualitative review. Emotion Review. 2019 Jan;11(1):87-97. doi: 10.1177/1754073917749016.
62. Borowicz A, Zasadzka E, Gaczkowska A, Gawłowska O, Pawlaczyk M. Assessing gait and balance impairment in elderly residents of nursing homes. Journal of physical therapy science. 2016;28(9):2486-90. doi: 10.1589/jpts.28.2486.
63. Sakthivadivel V, Geetha J, Gaur A, Kaliappan A. Performance-Oriented Mobility Assessment test and Timed Up and Go test as predictors of falls in the elderly–A cross-sectional study. Journal of Family Medicine and Primary Care. 2022 Nov;11(11):7294. doi: 10.4103/jfmpc.jfmpc_1294_22.
64. Nazareth DS. Emotion recognition in dementia: advancing technology for multimodal analysis of emotion expression in everyday life. In2019 8th International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW) 2019 Sep 3 (pp. 45-49). IEEE. doi: 10.1109/ACIIW.2019.8925059.
65. Siddiqui A, Khanna P, Kumar S, Pragya. Progression Analysis and Facial Emotion Recognition in Dementia Patients Using Machine Learning. InInternational Conference on Network Security and Blockchain Technology 2023 Mar 24 (pp. 489-500). Singapore: Springer Nature Singapore.https://doi.org/10.1007/978-981-99-4433-0_41
66. Adeli V, Korhani N, Sabo A, Mehdizadeh S, Mansfield A, Flint A, Iaboni A, Taati B. Ambient monitoring of gait and machine learning models for dynamic and short-term falls risk assessment in people with dementia. IEEE journal of biomedical and health informatics. 2023 Apr 14. doi: 10.1109/JBHI.2023.3267039.

67. Epstein-Lubow G, Gaudiano BA, Hinckley M, Salloway S, Miller IW. Evidence for the validity of the American Medical Association's caregiver self-assessment questionnaire as a screening measure for depression. Journal of the American Geriatrics Society. 2010 Feb. doi: 10.1111/j.1532-5415.2009.02701.x.
68. Kokkonen TM, Cheston RI, Dallos R, Smart CA. Attachment and coping of dementia care staff: The role of staff attachment style, geriatric nursing self-efficacy, and approaches to dementia in burnout. Dementia. 2014 Jul;13(4):544-68. doi: 10.1177/1471301213479469.
69. J. Saldaña, *The coding manual for qualitative researchers*, 2nd ed. Los Angeles: SAGE, 2013.
70. Saldaña, J. (2021). The coding manual for qualitative researchers.
71. Shenton, A. K. (2004). Strategies for ensuring trustworthiness in qualitative research projects. *Education for information*, *22*(2), 63-75.
72. Duan H, Zhao Y, Chen K, Lin D, Dai B. Revisiting skeleton-based action recognition. InProceedings of the IEEE/CVF conference on computer vision and pattern recognition 2022 (pp. 2969-2978).
73. Poria, S., Hazarika, D., Majumder, N., Naik, G., Cambria, E., & Mihalcea, R. (2018). Meld: A multimodal multi-party dataset for emotion recognition in conversations. *arXiv preprint arXiv:1810.02508*.
74. Jalali L, Huo D, Oh H, Tang M, Pongpaichet S, Jain R. Personicle: personal chronicle of life events. InWorkshop on Personal Data Analytics in the Internet of Things (PDA@ IOT) at the 40th International Conference on Very Large Databases (VLDB), Hangzhou, China 2014 Sep 5.

Supplementary Files

Figures

Design Components.