# Discovering Behavioral Patterns Using Conversational Technology for In-Home Health and Well-Being Monitoring

Maria R. Lima, *Graduate Student Member, IEEE*, Ting Su, Melanie Jouaiti, *Member, IEEE*, Maitreyee Wairagkar, Paresh Malhotra, Eyal Soreq, Payam Barnaghi, *Senior Member, IEEE*, and Ravi Vaidyanathan, *Member, IEEE*

***Abstract*—Advancements in conversational artificial intelligence (AI) have created unparalleled opportunities to promote the independence and well-being of older adults, including people living with dementia (PLWD). However, conversational agents have yet to demonstrate a direct impact in supporting target populations at home, particularly with long-term user benefits and clinical utility. We introduce an infrastructure fusing in-home activity data captured by Internet of Things (IoT) technologies with voice interactions using conversational technology (Amazon Alexa). We collect 3103 person-days of voice and environmental data across 14 households with PLWD to identify behavioral patterns. Interactions include an automated well-being questionnaire and ten topics of interest, identified using topic modeling. Although a significant decrease in conversational technology usage was observed after the novelty phase across the cohort, steady state data acquisition for modeling was sustained. We analyze household activity sequences preceding or following Alexa interactions through pairwise similarity and clustering methods. Our analysis demonstrates the capability to identify behavioral patterns, changes in those patterns and the corresponding time periods. We further report that households with PLWD continued using Alexa following clinical events (e.g., hospitalizations), which offers a compelling opportunity for proactive health and well-being data gathering related to medical changes. Results demonstrate the promise of conversational AI in digital health monitoring for aging and dementia support and offer a basis for tracking health and deterioration as indicated by household activity, which can inform healthcare professionals and relevant stakeholders for timely interventions. Future work will use the bespoke behavioral patterns extracted to create more personalized AI conversations.**

***Index Terms*—Behavioral patterns, conversational artificial intelligence (AI), dementia care, digital health monitoring, smart home technology.**

Manuscript received 30 October 2022; revised 12 June 2023; accepted 20 June 2023. Date of publication 29 June 2023; date of current version 24 October 2023. This work was supported in part by The U.K. Dementia Research Institute which receives its funding from U.K. DRI Ltd., funded by the U.K. Medical Research Council, Alzheimer's Society and Alzheimer's Research U.K., under Grant UKDRI-7003, and in part by the Research England Grand Challenge Research Fund (GCRF) through Imperial College London, the Imperial Biomedical Research Centre, and the Imperial College London's President's Ph.D. Scholarships. *(Payam Barnaghi, Eyal Soreq, and Ravi Vaidyanathan contributed equally to this work.) (Corresponding author: Ravi Vaidyanathan.)*

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the Surrey Borders Research Ethics Committee.

Maria R. Lima, Ting Su, Melanie Jouaiti, and Ravi Vaidyanathan are with the Department of Mechanical Engineering, Imperial College London, SW7 1AL London, U.K., and also with the U.K. Dementia Research Institute Care Research and Technology Centre, W12 0BZ London, U.K. (e-mail: r.vaidyanathan@ic.ac.uk).

Maitreyee Wairagkar is with the Department of Neurological Surgery, University of California at Davis, Davis, CA 95616 USA.

Paresh Malhotra, Eyal Soreq, and Payam Barnaghi are with the Department of Brain Sciences, Imperial College London, SW7 1AL London, U.K., and also with the U.K. Dementia Research Institute Care Research and Technology Centre, W12 0BZ London, U.K.

This article has supplementary downloadable material available at https://doi.org/10.1109/JIOT.2023.3290833, provided by the authors.

Digital Object Identifier 10.1109/JIOT.2023.3290833

## I. Introduction

TODAY, more than 55 million people live with dementia worldwide [1]. The aging population is set to double by 2050 [2], and the number of people affected by dementia is predicted to reach 139 million by then [1]. Global care costs of dementia are projected to surpass U.S.$ 2.8 trillion by 2030 [1]. In the U.K. alone, 25% of hospital beds are occupied due to a dementia-related condition [3]. This global health crisis has been exacerbated by the COVID-19 pandemic, with vulnerable populations facing unprecedented isolation, experiencing worsened mental health conditions, and receiving limited care [4], [5]. With limited resources for home care services and no immediate cure in sight, the global socioeconomic burden on healthcare systems is only expected to become more critical with time. This, in turn, places an increased psychological burden and strain on family members and caregivers [6]. Dementia is one of the world's major public health challenges [7]. Advancements in Internet of Things (IoT) technologies enable frequent and contextually rich interactions between people and the environment [8]. Several studies have been conducted on creating smart environments, such as smart homes, and eventually smart cities for urban living [9]. Furthermore, the integration of artificial intelligence (AI) and IoT for smart healthcare systems is growing dramatically, particularly for behavioral, physical and mental health monitoring, welfare interventions, or incident detection [10], [11], [12], [13].

The development of home-based assistive technology, particularly IoT technologies and social robotics, has been at the forefront of much research effort to date to promote independence, well-being, and quality of life of older adults, including people affected by dementia [14], [15], [16]. Monitoring an individual's home environment and daily routines—such as motion activity in the house, meal preparation, and physiological readings—combined with machine learning (ML) models for detecting anomalous behavior (i.e., deviations from the baseline routine) can provide an effective means to alert healthcare professionals or relevant stakeholders to potential risks, such as a fall, illness, or social isolation [14], [17], [18], [19]. In this context, detecting changes in the daily routines of target populations can offer important insights into their physical and mental health status [20], [21], [22]. This way, caregivers can be better informed on the expected changes in the patient's behavior, health status, and disease progression, which can help mitigate further deterioration through early intervention.

Smart homes equipped with IoT technologies for activity monitoring and habit assessment gather information in a passive way, in that they do not directly interact or engage with end users. Conversational technology, however, may give insights into subjective experiences and feelings by directly querying users and engaging in conversations, which could encourage behavioral changes. Research in conversational AI technology, including smart speakers integrated into the living environment, has experienced prolific growth in recent years [23]. Commercially available instances include devices, such as Amazon Echo and Google Home, with constantly evolving AI capabilities to understand human intent and provide relevant responses. For example, these devices can be used to set up medication reminders, self-management of daily activities, provide entertainment (e.g., playing music or games), or answer general questions as frequently as needed (e.g., the current time, date, and weather). This way, conversational AI technology holds potential to promote the independence of older adults, including people living with dementia (PLWD) and help reduce the burden on carers. Despite increasing interest in IoT monitoring systems deployed in smart environments and conversational AI in their respective fields, to the best of our knowledge, no research to date has combined voice with in-home activity data to inspect behavioral patterns. Furthermore, conversational agents have yet to produce research results addressing utility from user benefit and health monitoring perspectives, particularly in dementia care at home.

This study aims to investigate the integration of conversational agents in smart environments. We argue the potential of conversational agents for utility in health and well-being monitoring to support households with people affected by dementia. We discuss the role of conversational AI in health and well-being monitoring, particularly for aging and dementia support, and highlight future directions to address current challenges inhibiting long-term engagement and user benefits. We believe the ability to map individual behavior in smart environments and detect deviations or changes from previously observed patterns forms a strong baseline to personalize interactions. Furthermore, user-initiated interactions with conversational agents often indicate wants, needs or overall interests which could be mapped over time. For instance, the conversational agent could proactively engage with end users at appropriate times to remind them of an activity of interest or encourage behavior. We argue the potential of conversational technology to trace household behavior, directly query users for subjective perceptions of health and well-being in the event of household activity changes, and inform relevant stakeholders so that appropriate intervention can be activated if necessary. We examine the daily contexts in which 14 households with PLWD interact with a smart speaker (Amazon Alexa) by fusing home activity data captured by IoT technologies and remote health monitoring devices with regular interactions with Alexa. Broadly, our analysis inspects: 1) the use of Alexa in households with PLWD over time, particularly to assess compliance with a daily well-being questionnaire and prevalence of topics of interest beyond the novelty phase; 2) activity sequences in the 10-m period preceding or following user-initiated interactions with Alexa to identify behavioral patterns, changes in those patterns, and the corresponding time periods; and 3) Alexa usage in the week after health events occurred (information logged by a monitoring team, as elaborated in Section III-A). The contributions of this article are as follows.

1) We introduce an infrastructure fusing environmental and voice data using conversational technology to trace behavior. While our target in this study is health and well-being monitoring in the living environment, we argue our approach could be implemented in other smart environments to give insights into users' behavior.
2) We demonstrate technical feasibility to identify behavioral patterns and their corresponding time periods by analyzing sequences of household activities which precede or follow user-initiated interactions with conversational AI.
3) We offer the approach as a basis to adapt automated interactions aimed at providing personalized and proactive support for PLWD. This includes automated dialogues on health and well-being (e.g., sleep quality, mood, agitation, and anxiety) to obtain medically relevant data and sustain user engagement.

The remainder of this article is organized as follows. Section II reviews related works and identifies the main gaps. Section III describes the research questions that motivate this work, the experimental design and analysis methods. Section IV presents results in the form of user case studies. Section V discusses the utility, limitations and future directions of the proposed data-driven approach. Finally, conclusions are drawn in Section VI.

## II. Related Work

### A. Conversational AI for Aging and Dementia Care

There has been emerging interest in applying conversational AI technology in healthcare applications [24], including for home support of older populations [23]. By understanding and responding to natural spoken language, conversational agents present a versatile, intuitive, and natural user interface with potential to promote and monitor health [25]. In light of commercial viability, increased worldwide adoption, and expanding AI capabilities, conversational technology—including conversational agents and ubiquitous smart speakers—holds significant promise to assist older people and those affected by dementia in home settings. Commercially available smart speakers, such as Amazon Alexa and Google Home, have recently been explored as assistive tools to promote the independence of older populations for routine management [26], remote caring [27], or self-management of diabetes [28]. Voice skills for Alexa have been proposed to support older adults complete daily tasks, including medication reminders [27], depression screening and dressing assistance [29], as well as to send fall alerts to the caregiver [30].

There has been a growing interest in exploring how conversational AI is used in home settings by target populations. Qualitative studies have revealed initial insights into user experience with voice technology, patterns of daily use, and why interest is oftentimes lost after the novelty effect [31], [32], [33]. Current barriers inhibiting long-term adoption by older adults have been reported in the literature. These encompass the need for intelligent adaptation to user needs and cognitive abilities over time, limitations in speech recognition for effective verbal interaction, and privacy concerns related to voice data gathering [23], [34]. Furthermore, voice technology and analysis techniques from interactions, specifically looking at linguistic and speech patterns, have been investigated as a baseline for health monitoring and assessment of cognitive decline, including dementia progression [35], [36], [37], [38].

In light of the increased access to individual information from IoT technologies in smart homes, the use of conversational AI systems able to engage with end-users in natural interactions holds very strong promise to support aging and dementia care. However, the feasibility and utility of these tools for health and well-being monitoring at home remain largely untapped. Research to date lacks longitudinal data collection from real-world contexts, e.g., people's homes. Few investigations have demonstrated a direct impact on supporting the care needs of target populations. Additionally, the use of home-based conversational technology in combination with ML analysis for tracking behavior and cognitive changes over time remains underexplored. Further research with longitudinal depth of analysis is needed, particularly addressing: 1) adaptive interactions based on individual needs and changing health conditions; 2) end-user long-term engagement with conversational technology beyond the novelty phase; and 3) clinical utility.

### *B. Activity Monitoring in Smart Homes*

Opportunities in the use of smart home technology for older populations and PLWD are well noted in the literature: from diagnostic assessment to tailored care, health monitoring, cognitive support, and completion of activities of daily living (ADL) [14], [17], [39], [40]. Advances in IoT technologies have spurred significant progress in activity recognition, habit assessment, and anomaly detection within smart home environments [41], [42], [43], [44], [45], including monitoring systems, that aim to support independent living of older adults and people affected by dementia [19], [21], [46], [47], [48]. Despite the growing interest and potential for enhancing dementia care using remote monitoring systems, if not designed carefully and with end-user involvement, these can be perceived as complex and intrusive and may raise ethical concerns regarding privacy [23], [40], [49], [50], [51]. These factors have been investigated through user-centered design approaches focusing on fulfilment from the stakeholder perspective [52]. Models capable of recognizing behavioral patterns are of particular relevance to our study. Specifically, recognition of individual behavioral patterns can be achieved by using ADL data to capture regular activity sequences with temporal and spatial information (e.g., what a user does every morning between 10:00 and 12:00) [53], [54]. Such analysis can be used to detect behavioral changes over time that could indicate changes in lifestyle, functional abilities, and potentially cognitive decline [42], [55].

Recent work has proposed habit representation methods using activity data collected from smart environments, focusing on the sequence and duration of activities [53]. Along similar lines, a real-time monitoring framework has been proposed to recognize habits and detect anomalies with the aim of supporting seniors living alone [54], yet no results were obtained from data collection with target users. In [42], sequence comparison and clustering methods have been applied to activity vectors to obtain regular daily routines. The authors argued the future potential of a support system for individuals who may require assistance with ADL, including older populations. Further studies have analyzed abnormal behavior of PLWD, identifying differences in routine patterns within daily living contexts [47]. However, this analysis was conducted using a limited data set of three households. A common gap identified across these studies is the lack of longitudinal data collected from target populations in real-world contexts. Furthermore, researchers have investigated correlations between changes in daily routine and alterations in cognitive and physical health [56]. The authors evaluated the approach using continuous smart home sensor data collected from 18 senior residents. While there has been work in each of these areas individually, to the best of our knowledge, no research to date has correlated in-home activity captured by IoT technologies to voice interactions with conversational technology to track behavior.

The investigation of ADL patterns to understand human behavior comprises sequence comparison and clustering approaches to identify typical or unusual patterns from the recognized activity sequences captured by IoT technologies [42], [57]. Sequence mining algorithms have been successfully applied in bioinformatics to investigate related gene sequences [58], [59]. Different similarity measures have been studied in the context of sequence analysis, which can be categorized as follows: distances between probability distributions, counts of common attributes, and optimal matching between sequences by considering the necessary operations to transform one sequence into the other [60], [61]. The choice of a suitable similarity measure for comparing activity sequences to uncover patterns largely depends on the sequential features being considered, e.g., temporal information, duration, and order [18], [42]. Popular distance metrics are used to calculate the similarity between pairs of categorical sequences, such as the Hamming distance, which calculates the positionwise comparison of pairs of sequences of equal length [62], and the Levenshtein distance, given by the smallest number of edit operations needed to turn one sequence into another [63].

TABLE I

13 Behavioral Events Considered in This Analysis, Captured by IoT Technologies, Remote Health Monitoring Devices, and the Alexa Smart Speaker. Each Behavioral Event Indicates an Activity. Note That We Grouped All Physiological Readings Into the Same Vitals Event and Considered Three Alexa Behavioral Events

| Behavioural event | Devices | Activity |
|---|---|---|
| Lounge | Passive infrared sensors | Motion |
| Kitchen | Passive infrared sensors | Motion |
| Bedroom | Passive infrared sensors | Motion |
| Bathroom | Passive infrared sensors | Motion |
| Hallway | Passive infrared sensors | Motion |
| Bed in | Sleep mat | Getting into bed |
| Bed out | Sleep mat | Getting out of bed |
| Front door | Door sensor | Door opening/closing |
| Back door | Door sensor | Door opening/closing |
| Vitals | Pulse oximeter, scale, thermometer, blood pressure cuff | Taking vitals |
| Start questionnaire | Amazon Alexa | Interacting with voice agent |
| End questionnaire | Amazon Alexa | Interacting with voice agent |
| Random interactions | Amazon Alexa | Interacting with voice agent |

These, however, are not suitable when dealing with temporal event sequences where common and consecutive elements ought to be considered. An alternative approach has been proposed in [64] aimed at capturing the sequentiality of events.

### C. Topic Modeling Techniques

Topic modeling, an unsupervised learning technique used to identify hidden patterns from a text corpus [65], can be applied to analyze text-based interactions with conversational agents and further inspect user interests and preferences from conversation topics over time. Conventional models, typically based on latent Dirichlet allocation (LDA), employ a bag-of-words model, wherein each unique word is modeled independently from the others [66]. These models, however, involve simplistic assumptions and preprocessing steps that often dismiss semantic relationships between words, especially when analyzing short texts, resulting in the learned topics being less coherent and interpretable [67], [68]. With recent developments in natural language processing, pretrained language models have been proposed to capture semantic and contextual information from text more effectively (e.g., BERT [69] and GPT-3 [70]). Similarly, Top2Vec [71] and BERTopic [72] have been proposed to infer topics while keeping the original structure of text with high efficacy [73], [74].

## III. Methods

### A. Preliminary

The U.K. Dementia Research Institute Care Research and Technology Centre (U.K. DRI-CR&T) has created a unique infrastructure for gathering environmental data from households with PLWD to enhance independence and safety at home. The U.K. DRI-CR&T brings together a multidisciplinary team of doctors, engineers, and scientists that develop and study new technologies for effective use in smart homes, deploy them in real-world evaluation studies following iterative user-centered design approaches, and deliver them to PLWD and their carers. A range of systems are studied to track a person's behavior and health at home, predict when problems might arise, and provide intervention solutions while allowing continuous interaction between PLWD, caregivers, and medical professionals.

In the context of this study, we define a *behavioral event* as a sensor trigger captured in a smart home with an associated timestamp. Each behavioral event indicates an activity (e.g., motion in the house, taking vitals, interacting with a voice agent). Following a previous approach for activity data collection in households with PLWD [75], we analyzed 13 behavioral events (outlined in Table I) captured by IoT technologies, remote health monitoring devices, and the Amazon Alexa smart speaker from 14 households with PLWD. As part of our recruitment and deployment protocol, a monitoring team and a design team maintained communication with participants to clarify the purpose of data collection and the capabilities of the devices deployed in order to mitigate potential ethical concerns regarding data privacy. The design team encouraged PLWD to complete a daily well-being questionnaire.[1] The questionnaire comprised six questions assessing the subjective perception of mood, agitation, anxiety, sleep quality, tiredness, and activity plans. PLWD were further encouraged to interact with Alexa freely (e.g., ask for the weather, news, or entertainment). Additionally, as part of our experimental design, a monitoring team in regular contact with participants noted individual health events (e.g., falls, infections, and hospitalizations).

We conducted a household analysis and did not identify individuals to protect privacy. We were interested in investigating behavior in households with PLWD using conversational technology. We consider an *activity sequence* as a sequence of ordered behavioral events that occur in the 10-m period (defined based on domain knowledge) preceding or following Alexa use, with an associated start and end timestamp. We refer to a *behavioral pattern* as a set of activity sequences with a high degree of similarity (see details on the similarity approach used to quantify the degree of similarity between pairs of activity sequences in Section III-E) that occur in the household for a period of at least three weeks, determined by domain knowledge (e.g., a user takes vitals in the morning before interacting with Alexa for a month). When a new set of activity sequences emerges by changing the previously observed pattern (evidenced by a lower degree of similarity

[1]An Alexa Skill was developed for the purpose of the ongoing research conducted by the U.K. DRI-CR&T.

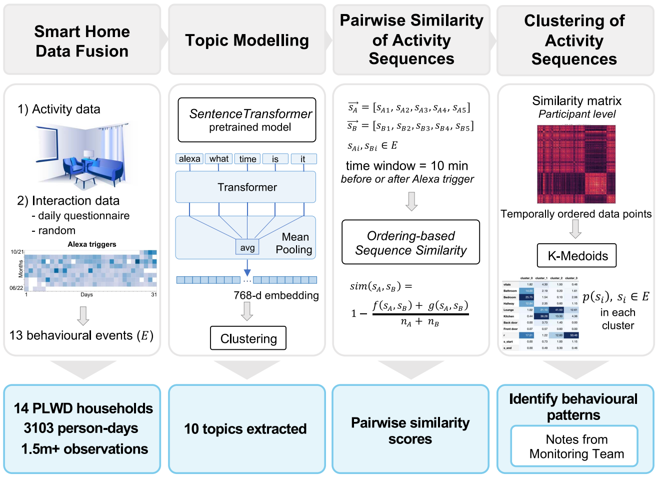



Fig. 1. Pipeline illustration with the main steps used in the study and the mathematical formulations. This includes: 1) the longitudinal fusion of environmental and voice data, comprising a set of 13 distinct behavioral events, as outlined in Table I; 2) the topic modeling approach employing sentence embeddings and clustering techniques; 3) the analysis of activity sequences on a participant level using a pairwise similarity measure; and 4) clustering to discern behavioral patterns. Additionally, notes from the monitoring team are used to explain and validate the quantitative findings.

compared to adjacent activity sequences), we consider it a *behavioral change* (e.g., a user stops taking vitals before talking to Alexa, which had previously been identified as a behavioral pattern). We are particularly interested in discovering behavioral patterns beyond the novelty phase, i.e., after the first three months, in line with other studies using home-based robotic technology [76]. The proposed pipeline of analysis is illustrated in Fig. 1.

## B. Research Questions

This study is driven by the following research questions.

1) *RQ1:* Will end-users sustain engagement with conversational technology beyond the novelty phase (i.e., the first three months)? We examine Alexa usage in households with PLWD over time. Particularly, we assess compliance regarding the sustained use of a daily well-being questionnaire and apply topic modeling to inspect the prevalence of topics over time.
2) *RQ2:* Can behavioral patterns in households with PLWD be traced by mapping in-home activity and voice interactions with conversational technology? We examine pairwise similarity of activity sequences preceding or following Alexa use. We identify behavioral patterns, changes in these patterns, and their corresponding time periods.
3) *RQ3:* Do participants continue using Alexa following clinical outcomes (e.g., falls, infections, and hospitalizations)? We inspect whether households with PLWD continue interacting with Alexa in the week following a health event.

## C. Study Sample

In this research, we collected 3103 person-days of interactions with Alexa and in-home activity data captured by IoT technologies across 14 households with PLWD as part of ongoing research in dementia care conducted by the U.K. DRI-CR&T. Participants (75-94 years, 4 females, 10 males) lived in the U.K., had a diagnosis of dementia or mild cognitive impairment (MCI), and were living in their own homes with a caregiver during the time of data collection. Table II lists full participant and data collection information. The total timeframe of data collection varied across participants due to two different recruitment stages followed by Alexa device deployment. The study was ethically approved by the Surrey Borders Research Ethics Committee.

## D. Technology and Data Overview

This study fuses in-home activity data and voice interactions with conversational technology to analyze behavior in households with PLWD. Each household included a range of IoT technologies and remote health monitoring devices, namely, passive infrared sensors installed in the bedroom, the lounge/living room, the kitchen, the bathroom, and the hallway; door sensors placed on the front door and the back door to detect when a door was opened or closed; a sleeping mat to collect information about a person getting in or out of bed; physiological devices to take vital signs, including a pulse oximeter, scale, thermometer, and blood pressure cuff; and the smart speaker Amazon Echo Show. The activity data related to motion, taking vitals, opening/closing doors, getting in/out of bed (as presented in Table I) was extracted offline from DCARTE [77], a framework that allows continuous and anonymized data access by DRI-CR&T researchers. The interaction data comprises text utterances of what users said to Alexa and the corresponding timestamp. We consider two types of data from Alexa interactions.

1) *Questionnaire:* Participants were encouraged to trigger a well-being questionnaire on a daily basis. This data type was used specifically to inspect the frequency and time of questionnaire completeness over time. Therefore, we extracted the timestamps of the start and end of the questionnaire and excluded answers to each question.
2) *Random Interactions:* All the Alexa interactions excluding the start, end, and answers to the well-being questionnaire. This includes free use of the smart speaker across different topics, e.g., asking for the news, weather, time, setting reminders, and playing music, among others. This type of Alexa interactions was used to examine participants' topics of interest over time.

The activity and interaction data sets were aggregated, grouped for each participant, and sorted by timestamp, resulting in a total of over 1.5 million unique observations collected over 3103 person-days (see details in Table II). Each observation represented one event with a timestamp. The analysis encompassed a total of 13 different behavioral events (listed in Table I) related to user-initiated interactions with Alexa, location in the house, bed in/out information, opening or closing of front/back door, and vitals.[2] Furthermore, we used the dates of individual health events logged by a monitoring team (e.g., falls, infections, and hospitalizations) to examine

[2]Note all physiological measurements were grouped into one vitals event.

TABLE II

PARTICIPANT COHORT AND DATA ACQUISITION DETAILS. TOTAL DAYS REFERS TO THE DURATION OF DATA COLLECTION, FROM THE FIRST TO THE LAST DATE OF ALEXA INTERACTIONS, HOWEVER, PARTICIPANTS DID NOT USE THE DEVICE DAILY. THE TOTAL NUMBER OF ALEXA TRIGGERS ENCOMPASSES BOTH THE QUESTIONNAIRE TRIGGER AND RANDOM INTERACTIONS. UNIQUE EVENTS COMPRISE THE AGGREGATE RAW DATA OF BEHAVIORAL EVENTS FOR EACH PARTICIPANT INCLUDING ALEXA INTERACTIONS

| Participant | Gender | Diagnosis | Total days | Start date | End date | Alexa triggers | Unique events |
|---|---|---|---|---|---|---|---|
| P1 | M | Vascular dementia | 169 | 2021-05-07 | 2021-10-23 | 202 | 100436 |
| P2 | M | Dementia in Parkinson's | 388 | 2021-05-13 | 2022-06-05 | 1824 | 239611 |
| P3 | M | Mild Cognitive Impairment | 381 | 2021-05-16 | 2022-06-01 | 96 | 214879 |
| P4 | M | Alzheimer's Disease | 387 | 2021-05-14 | 2022-06-05 | 448 | 210271 |
| P5 | M | Alzheimer's Disease | 244 | 2021-09-28 | 2022-05-30 | 368 | 105326 |
| P6 | F | Alzheimer's Disease | 260 | 2021-09-08 | 2022-05-26 | 600 | 95762 |
| P7 | F | Alzheimer's Disease | 263 | 2021-09-08 | 2022-05-29 | 555 | 89475 |
| P8 | M | Alzheimer's Disease | 23 | 2021-09-08 | 2021-10-01 | 19 | 9716 |
| P9 | F | Alzheimer's Disease | 102 | 2021-09-08 | 2021-12-19 | 102 | 33569 |
| P10 | F | Alzheimer's Disease | 107 | 2021-09-20 | 2022-01-05 | 69 | 75 |
| P11 | M | Alzheimer's Disease | 73 | 2021-09-25 | 2021-12-07 | 72 | 15687 |
| P12 | M | Lewy Body Dementia | 220 | 2021-10-28 | 2022-06-05 | 2125 | 131032 |
| P13 | M | Alzheimer's Disease | 244 | 2021-10-05 | 2022-06-06 | 610 | 152162 |
| P14 | M | Alzheimer's Disease | 242 | 2021-10-05 | 2022-06-04 | 656 | 125112 |

whether users continued using Alexa in the week following a health event.

### E. Pairwise Similarity of Activity Sequences

This study investigates behavioral patterns by analyzing in-home activity data and voice interactions using conversational AI technology. The analysis focused on calculating the pairwise similarity of activity sequences near (i.e., in the 10-m period preceding or following) Alexa triggers. We further identify behavioral patterns by grouping activity sequences with higher or lower similarity scores. We considered an activity sequence, $\vec{s}$, as a vector of temporally ordered sensor triggers (i.e., behavioral events)

$$\vec{s} = [s_1, \ldots, s_n], s_i \in E$$

where $E$ is the finite set of 13 behavioral events considered in this study (see Table I), and $s_i$ is the event in position $i$ of the sequence.

Each activity sequence includes temporal window and duration parameters, in addition to a defined target event, i.e., the Alexa type of interaction: *Start* stands for the trigger of the well-being questionnaire; *End* denotes the end of the questionnaire; *Random* stands for other utterances from free use of the smart speaker (note different topics were considered, as elaborated in Section III-F). Through exploratory data analysis, we chose an optimal window of five consecutive events ($n = 5$) and filtered sequences by a maximum duration of 10 min near the target trigger. Note the duration of each event varies. Therefore, activity sequences may comprise repetitive events as long as the total vector size is five (i.e., $n = 5$) and the maximum duration is 10 min. Below are examples of equal-length sequences that preceded the trigger of the daily questionnaire (i.e., $s_n =$ "Start") for a given participant:

$$\textit{Lounge} > \textit{Kitchen} > \textit{Lounge} > \textit{Lounge} > \textit{Start}$$
$$\textit{Vitals} > \textit{Vitals} > \textit{Kitchen} > \textit{Lounge} > \textit{Start}.$$

To quantify the level of similarity between activity sequences composed of chronologically ordered behavioral events, $\vec{s}$, we calculated pairwise sequence similarity. We were particularly interested in computing similarity by capturing the temporal sequence of events. Therefore, we applied Ordering-based Sequence Similarity [64], a categorical sequence mining technique which considers the number of common elements and their order in the sequence. Let $\vec{s_A} = [s_{A1}, \ldots, s_{An}]$ and $\vec{s_B} = [s_{B1}, \ldots, s_{Bn}]$ be two equal-length activity sequences. The similarity score between $s_A$ and $s_B$ (the vector symbol was omitted for simplicity) was calculated as follows:

$$\text{sim}(s_A, s_B) = 1 - \frac{f(s_A, s_B) + g(s_A, s_B)}{n_A + n_B} \quad (1)$$

where $f(s_A, s_B)$ quantifies the similarity in the position of elements in the sequence (i.e., the order), $g(s_A, s_B)$ counts the number of noncommon elements, and $n_A$ and $n_B$ denote the vector size of $s_A$ and $s_B$, respectively.

For a behavioral event $e \in E$ and activity sequences $s_A$ and $s_B$, let $L_A^e$ be the number of times $e$ appears in $s_A$, and $s_A^e(k)$ the $k^{th}$ position of $e$ in $s_A$. $C_{AB}$ denotes the set of common events in $s_A$ and $s_B$. $U_{AB}$ denotes the set of events that appear in $s_A$ but not in $s_B$. Then, $f(s_A, s_B)$ and $g(s_A, s_B)$ are calculated as follows:

$$f(s_A, s_B) = \frac{\sum_{e \in C_{AB}} \left( \sum_{k=1}^{K_{AB}^e} |s_A^e(k) - s_B^e(k)| \right)}{\max(n_A, n_B)} \quad (2)$$

and

$$g(s_A, s_B) = \sum_{e \in U_{AB}} L_A^e + \sum_{e \in U_{BA}} L_B^e \quad (3)$$

where $K_{AB}^e = \min(L_A^e, L_B^e)$.

Using the two activity sequences shown above as an example with simplified notation $s_A = \{L, K, L, L, S\}$, $s_B = \{V, V, K, L, S\}$: since $L$, $K$, and $S$ appear in both sequences, $C_{AB} = \{L, K, S\}$. Looking at the position of common events in each sequence, $s_A^L = \{0, 2, 3\}$, $s_B^L = \{3\}$, $s_A^K = \{1\}$, $s_B^K = \{2\}$, $s_A^S = s_B^S = \{4\}$, therefore $f(s_A, s_B) = (|0 - 3| + |1 - 2| + |4 - 4|)/5 = 0.8$. Calculating the noncommon events, $V$ appears twice in $s_B$, hence $U_{BA} = \{V\}$ and $g(s_A, s_B) = 2$. Following (1), $sim(s_A, s_B) = 1-\ (0.8 + 2)/10 = 0.72$.

We predefined the target event, $s_5$, to analyze activity sequences in the 10-m temporal window preceding Alexa use (e.g., for the activity sequences preceding the trigger of the questionnaire, $s_5 = \textit{Start}$). We further analyzed similarity matrices, on a participant level, based on the pairwise similarity scores computed. Higher values within these matrices indicate a higher degree of similarity between pairs of activity sequences. Subsequently, similarity matrices were used to cluster activity sequences by grouping those with higher or lower similarity scores. Given that the data points to cluster (i.e., the activity sequences) are not in a vector space, we applied $K$-Medoids clustering, a method based on the partition around medoids algorithm [78]. The silhouette method was used to determine the number of clusters. Altogether, our approach involved computing pairwise similarity and performing clustering of activity sequences to identify behavioral patterns and examine changes in behavior using conversational technology at home.

### *F. Topic Modeling*

We applied topic modeling methods to analyze the Alexa interactions of type random (i.e., Alexa usage that is not related to the well-being questionnaire, as described in Section III-D). Specifically, we used the pretrained language model *SentenceTransformer* [79] to embed each user utterance into a 768-D vector. We applied $K$-Means clustering[3] in two iterations on the obtained utterance vectors and used the silhouette method to choose the number of clusters.[4] In the first round, we applied the $K$-Means clustering model to cluster the vectors into 16 clusters. Manually inspecting these clusters, we identified a set of generic utterances (the *undefined* topic). Additionally, we combined clusters with similar topics, resulting in a total of eight clusters, including the *undefined* topic cluster. In the second round, we specifically focused on the *undefined* cluster from the first round and further clustered it into 15 topics using $K$-Means. We merged similar topics from the second round of $K$-Means clustering with those identified in the first round, integrated two newly emerged clusters and identified the new *undefined* topic cluster. Fig. 2 shows the identified topics and the most prevalent words in each topic. In total, we identified ten topics from Alexa interactions across the cohort of participants as follows. 1) *Answers:* Participants may be prompted to confirm Alexa actions or speech recognition. 2) *Control commands:* Voice commands used to control Alexa (e.g., change the volume and start or stop actions). 3) *Entertainment:* Participants ask Alexa to play music, radio, or games. 4) *Timers:* Participants ask Alexa to set timers. 5) *Weather:* Participants ask for weather information. 6) *Questionnaire attempt:* Participants attempt to start the daily questionnaire, but Alexa does not recognize participants' speech correctly. 7) *Reminders, time, and date:* Participants ask about the current time, date, or day of the week, to set reminders or alarms. 8) *News:* Participants ask for general news or headlines of the day (e.g., "tell me the latest news"), news from specific channels (e.g., "what is on BBC One tonight"), or news on specific themes (e.g., "news on prince harry"). 9) *Greetings:* Participants greet Alexa. 10) *Undefined:* All remaining Alexa interactions of type random.

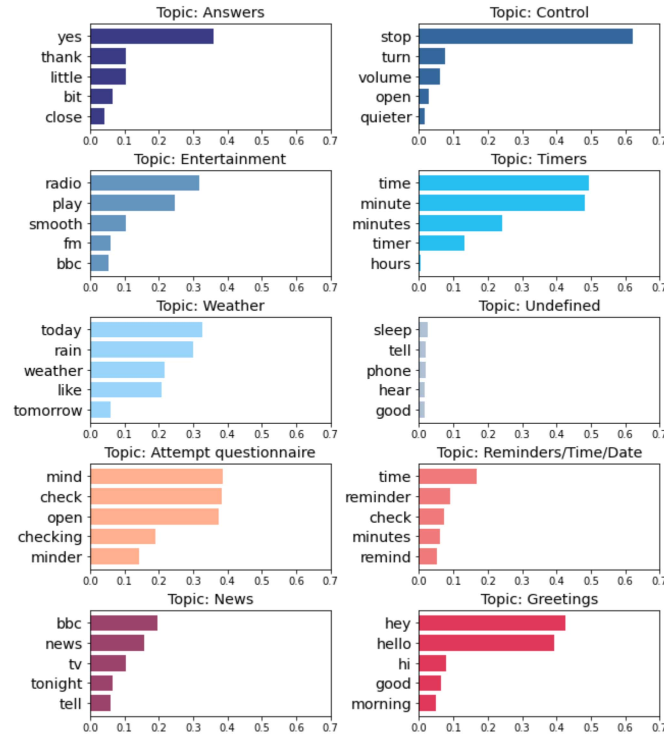



Fig. 2. Top five words in each topic. These words were selected based on their proportion relative to all words within their respective topic. The $x$-axis measures this proportion. Note the *undefined* topic contains many different utterances, thus, all words among this topic have low-term frequency–inverse document frequency (TF-IDF) scores.

## IV. Results

We analyzed over 1.5 million events captured from IoT technologies, remote health monitoring devices, and the Amazon Alexa smart speaker over 3103 person-days across a unique cohort of 14 households with PLWD. We first investigated trends in conversational technology usage over time across the 14 households, particularly beyond the novelty phase. We selected four participants (i.e., P2, P6, P12, and P14) that interacted the most with Alexa (see the total number of Alexa triggers in Table II) to report a series of case studies. We examined the similarity of activity sequences preceding or following user-initiated interactions with Alexa to identify behavioral patterns. We also inspected Alexa usage in the week following the occurrence of health events across the cohort.

### *A. Prevalence of Interactions With Conversational AI Beyond the Novelty Phase*

We aimed to examine the use of Alexa in households with PLWD over time. We inspected the novelty effect across the cohort and which topics of interest prevailed after the first three months of usage (see RQ1 in Section III-B). Specifically, we analyzed the weekly average number of Alexa interactions both during and beyond the novelty phase and examined the prevalence of the ten identified topics (see Section III-F) in participants' interactions over time.

[3] We conducted preliminary research on topic modeling and found nonnegative matrix factorization (NMF) and LDA produced unsatisfactory results

[4] We compared the performance of different clustering methods by computing the silhouette score as a measure of coherence. $K$-Means marginally outperformed Hierarchical Clustering, Gaussian Mixture Models, and Spectral Clustering (see details in Supplementary Table I in the Appendix).

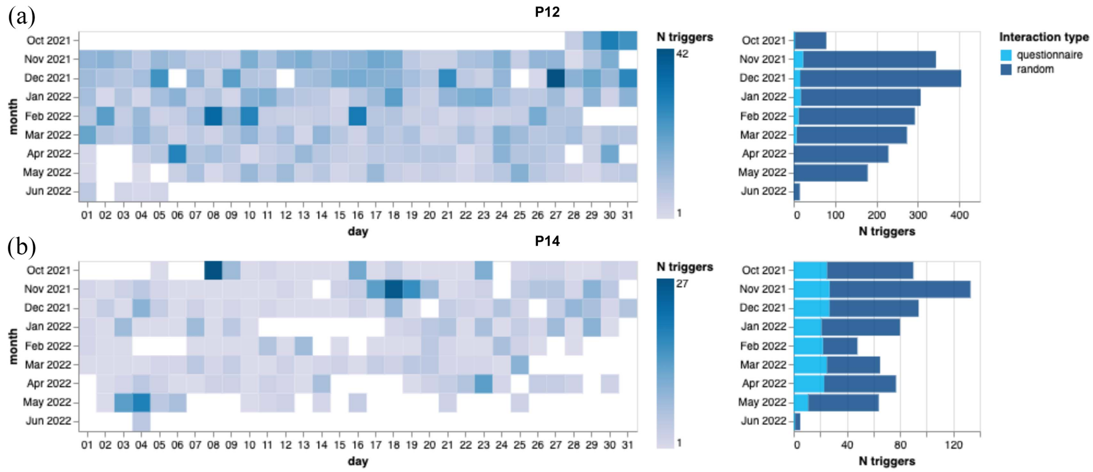



Fig. 3. Alexa interaction data across the total usage timeframe for (a) P12 and (b) P14. The left plots show the daily counts of Alexa events per day. The right plots show monthly usage of both types of interactions considered—questionnaire triggers and random interactions—over the duration of Alexa use. *N triggers* quantifies the total number of interaction triggers.

Fig. 3 shows an overview of Alexa interaction data across the total usage timeframe for P12 and P14. P12 interacted with Alexa consistently over time, showing an increased interest in its capabilities during the novelty phase, as evidenced by a rise in random interactions. However, there was a gradual decline in overall Alexa usage during the post-novelty phase (i.e., after the first three months). P12's daily interactions with Alexa peaked in December 2021 (N triggers = 42) during the novelty period. While P14 used Alexa consistently over time, there were noticeable intervals of consecutive noninteraction days. P14's daily triggers peaked in November 2021 (N triggers = 27), recorded on the second day of using Alexa. Furthermore, participants' engagement with the well-being questionnaire varied across the cohort. For instance, P14 consistently completed the questionnaire on a monthly basis, with a decrease in engagement only noticeable after seven months of use, in May 2022. Conversely, P12 gradually reduced the frequency of questionnaire triggers and stopped completing it after March 2022.

We compared the weekly average number of Alexa interactions (of both questionnaire and random types) in the three months of the novelty phase to the weekly average number of interactions in the post-novelty phase. Thus, we only considered participants with a total Alexa usage time of at least four months and evaluated the novelty effect across a total of 11 participants. We observed a significant decrease in overall Alexa usage after the novelty period across participants (Wilcoxon signed-rank, $W = 5$, $p_{\text{corr}} = 0.02$, CLES = 0.66).[5] We further observed a significant decrease in compliance with the daily well-being questionnaire in the post-novelty period across participants (Wilcoxon signed-rank: $W = 0$, $p_{\text{corr}} =$ 1.95e-03, CLES = 0.72). Notably, of the participants who continued using Alexa beyond the novelty phase, two stopped completing the well-being questionnaire after the first month of usage. Focusing on the four participants who sustained the use of Alexa for at least five months beyond the novelty phase (i.e., total technology usage exceeding eight months), Fig. 4 shows an overall decline in Alexa usage in the post-novelty period for these households. Notably, the weekly number of Alexa triggers peaked during the novelty phase. For example, P2's weekly interactions with Alexa reached a peak in the first week of usage (13 May, 2021, N triggers = 302, as highlighted in a gray box in Fig. 4). However, there was a noticeable decrease in usage after the first month. Interestingly, interactions ceased for a 4-week period in February 2022. Health notes from the monitoring team indicate that P2 was hospitalized between 3 February and 23 February 2022, which explains the absence of Alexa usage at home during this period. We further observed an increased number of weekly interactions in April 2022. This surge coincided with P2's return home on 31 March 2022 after a period of hospitalization (information logged by the monitoring team). Overall, these findings indicate a decline in the use of conversational technology after the novelty effect across the cohort.

We next investigated the prevalence of topics within Alexa interactions of type random (see the interaction data types used in this study in Section III-D), particularly beyond the novelty effect. Fig. 5 illustrates the proportion of topics triggered for P2, P6, P12, and P14. After the novelty phase, P2 used Alexa more frequently to request weather information in the morning (85.26% of weather triggers were observed in the morning during this period; see details in Supplementary Table II in the Appendix) and set timers during the morning (53.7%), particularly after January 2022 (refer to Fig. 5). Topics related to news, entertainment, and reminders were also frequently triggered over time. P6 showed increased interest in using Alexa for entertainment in the morning, both during (55.56%) and after (51.85%) the novelty period, and asking about current date, time, and reminders in the morning during the first three months of usage (66.67%), a trend that prevailed in the post-novelty phase (52.63%). P12 and P14 showed consistent interest in utilizing Alexa for entertainment. During the post-novelty phase, P12 reduced the usage of reminders while

[5]False discovery rate was applied, hence the corrected $p$-values are compared against the significance level $\alpha = 0.025$. We also report the common-language effect size (CLES).

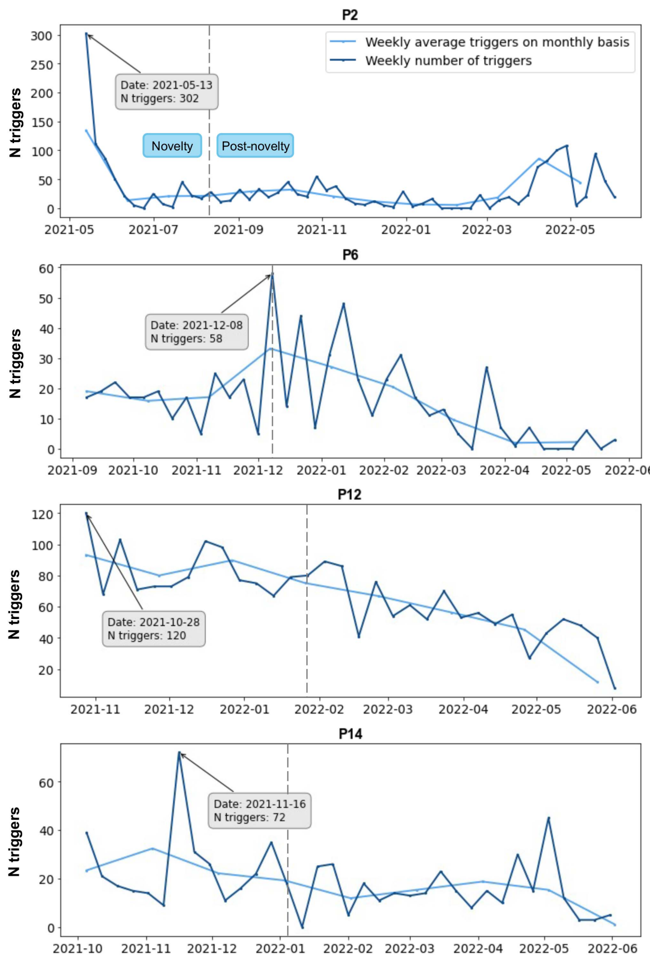



Fig. 4. Weekly average number of Alexa interactions on a monthly basis and total weekly interactions for P2, P6, P12, and P14. *N triggers* quantifies the total number of triggers. The vertical dashed line separates the novelty phase (first three months of usage) from the post-novelty period for each participant. Grey boxes in each plot represent the peak with the highest number of weekly interactions and total Alexa triggers. Note the start date and Alexa usage periods varied across participants (as outlined in Table I), therefore, we standardized usage based on the first Alexa interaction for each participant.

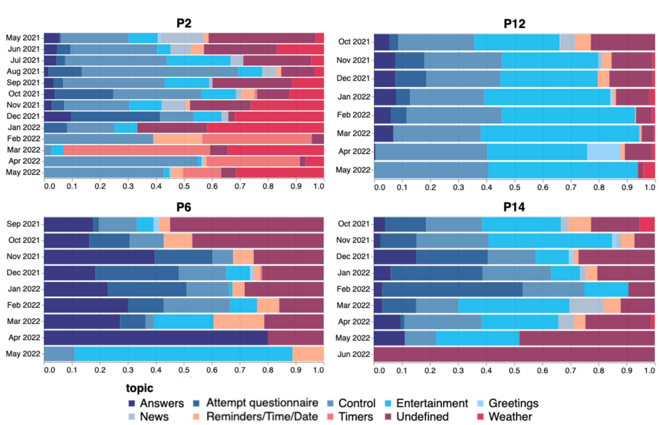



Fig. 5. Proportion of topics triggered among P2, P6, P12, and P14 for the different months of Alexa usage.

increasing the frequency of weather-related queries, particularly in the morning (62.5%). Furthermore, P14 frequently prompted news-related topics in the morning, which prevailed after the novelty phase (66.67%).

| | P2 | | | P6 | | | P12 | | | P14 | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | A(%) | B(%) | C(%) | A(%) | B(%) | C(%) | A(%) | B(%) | C(%) | A(%) | B(%) | C(%) |
| **Bed in** | 0.00 | 0.00 | 0.07 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.22 | 0.00 | 0.00 | 0.00 |
| **Bed out** | 0.00 | 0.00 | 0.07 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.12 | 0.00 | 0.00 | 0.00 |
| **Vitals** | 16.91 | 10.10 | 2.69 | 7.19 | 3.04 | 3.05 | 0.00 | 0.00 | 0.29 | 7.67 | 16.62 | 5.02 |
| **Bathroom** | 1.01 | 1.92 | 6.82 | 0.31 | 0.34 | 1.13 | 2.73 | 2.50 | 3.70 | 0.71 | 0.15 | 3.03 |
| **Bedroom** | 0.72 | 1.28 | 14.54 | 0.62 | 2.03 | 1.02 | 7.73 | 5.42 | 12.62 | 1.14 | 1.50 | 2.77 |
| **Hallway** | 0.87 | 3.37 | 6.00 | 19.06 | 21.62 | 17.65 | 11.82 | 27.92 | 23.73 | 0.00 | 0.00 | 0.00 |
| **Lounge** | 39.88 | 46.47 | 25.52 | 0.00 | 0.00 | 0.00 | 4.55 | 8.75 | 8.38 | 63.49 | 62.13 | 53.46 |
| **Kitchen** | 28.90 | 29.65 | 18.89 | 50.94 | 58.78 | 47.74 | 47.73 | 40.00 | 38.31 | 17.61 | 11.08 | 11.51 |
| **Back door** | 1.01 | 1.28 | 1.61 | 5.31 | 5.41 | 6.11 | 0.45 | 0.00 | 0.71 | 2.27 | 1.05 | 1.99 |
| **Front door** | 0.43 | 0.96 | 0.45 | 0.31 | 3.38 | 0.68 | 1.36 | 1.25 | 2.82 | 0.00 | 0.00 | 0.26 |
| **Random** | 8.24 | 3.04 | 22.01 | 15.62 | 4.73 | 12.10 | 22.27 | 12.50 | 6.74 | 5.97 | 6.44 | 12.63 |
| **Start** | 0.43 | 1.76 | 0.86 | 0.00 | 0.68 | 8.82 | 0.00 | 1.25 | 1.62 | 0.14 | 0.90 | 5.36 |
| **End** | 1.59 | 0.16 | 0.48 | 0.62 | 0.00 | 1.70 | 1.36 | 0.42 | 0.74 | 0.99 | 0.15 | 3.98 |

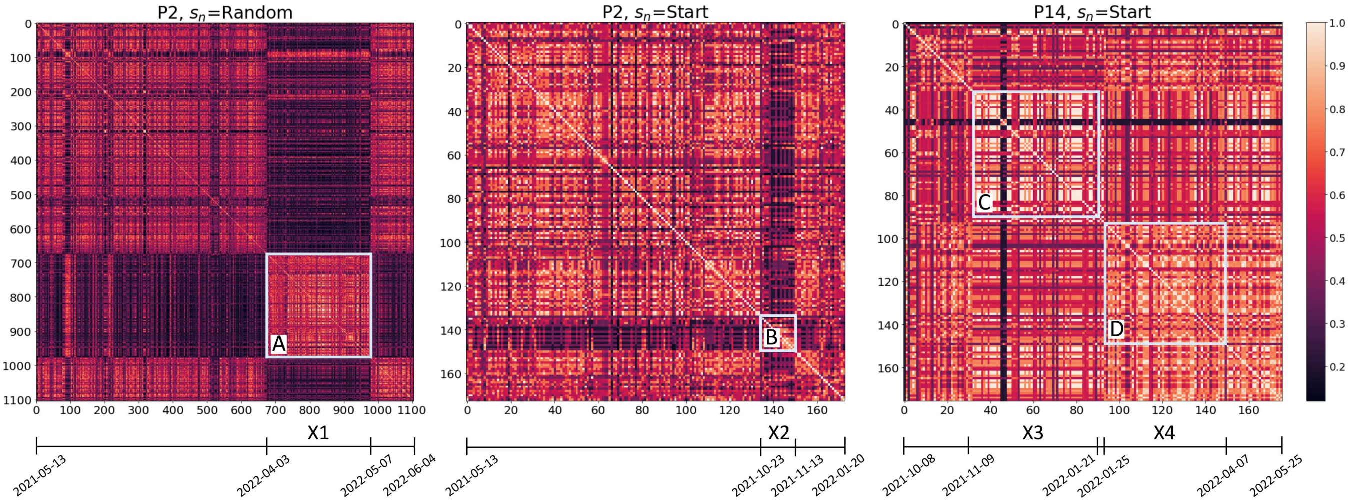
Fig. 6. Probability (%) of each behavioral event (13 in total) in the 10-m period preceding or following Alexa use for P2, P6, P12, and P14. For each participant, we looked at three target Alexa events represented as columns in the table, A: comprises activity sequences that preceded the start of the well-being questionnaire $s_5=$ Start; B: comprises activity sequences that followed the completeness of the well-being questionnaire $s_1=$ End; C: comprises activity sequences that preceded any Alexa interaction of type random $s_5 =$ Random. For each participant and activity sequence type (A, B, and C) the probability of an event $s_i \in E$ (the finite set of 13 events) is calculated by summing occurrences of event $s_i$ in all sequences and dividing by the total number of possible occurrences. Note the total number of occurrences excludes the first or last event of each sequence, $s_1$ or $s_5$, since it was predefined and equal for all sequences in consideration.

Overall, despite the variations in usage across the cohort, participants continued to find certain capabilities of Alexa engaging, using them regularly. While the verified decrease in Alexa usage—oftentimes due to unmet expectations or perceived lack of utility [80]—aligns with previous research involving older demographics [32], [33], results suggest future improvements are needed in the design of home-based conversational agents. Specifically, addressing long-term user engagement and facilitating more personalized interactions could foster better integration in households with PLWD. Based on these findings, we discuss challenges and design opportunities for future conversational agents in Section V.

## *B. Behavior Discovery From Activity Sequences*

We aimed to investigate the daily contexts in which Alexa was triggered in households with PLWD to identify behavioral patterns, changes in those patterns, and the corresponding time periods (see RQ2 in Section III-B). We analyzed activity sequences in the 10-m period preceding or following Alexa triggers in households with PLWD. We first computed the probability of behavioral events in the activity sequences considering the two types of Alexa triggers, i.e., the questionnaire and random interactions. Fig. 6 shows, for four representative users (i.e., P2, P6, P12, and P14), the probability of each behavioral event, captured by IoT technologies, in sets of activity sequences in the 10-m temporal window considered (see details in Section III-E).

Our design team reached out to participants to ask for the device location(s) in the house. By inspecting the probability of behavioral events in the activity sequences, we were able to corroborate the primary locations of the Alexa smart speaker




(a) (b) (c)

Fig. 7. Similarity matrices of activity sequences near the trigger of Alexa. (a) Similarity of activity sequences that preceded interactions of type random ($s_5$ = *Random*) from May 2021 until June 2022 for P2. (b) Similarity of activity sequences that preceded the activation of the well-being questionnaire ($s_5$ = *Start*) from May 2021 until January 2022 for P2. (c) Similarity of activity sequences that preceded the activation of the well-being questionnaire ($s_5$ = *Start*) from October 2021 until May 2022 for P14. The axis numbers denote activity sequences preceding Alexa use which are chronologically ordered. The dates of selected periods are represented below each matrix as well as the dates of the first and last activity sequences analyzed. Lighter tones represent a high degree of similarity and darker tones represent a low degree of similarity. The color bar applies to all matrices.

in participants' households. P2 and P14 had Alexa placed in the lounge, as indicated in Fig. 6 by higher probabilities of lounge events before and after Alexa triggers (see darker areas for P2 and P14). P6 and P12 had Alexa located in the kitchen, which is supported by higher probabilities of kitchen events (see darker areas for P6 and P12). The higher probability values found for the aforementioned behavioral events held true for the three types of activity sequences considered, specifically: A) activities preceding the trigger of the questionnaire; B) activities following the completion of the questionnaire; and C) activities prior to random Alexa interactions, as depicted in Fig. 6. In addition to primary Alexa locations, the frequent occurrence of certain events, such as kitchen for P2 or Hallway for P12, suggests that these activities often occurred near the time of Alexa interactions (i.e., within the 10-m period before or after triggering Alexa). For example, P2's most common sequence of activities preceding the activation of the well-being questionnaire ($s_5$ = Start, column A in Fig. 6) was as follows: *Kitchen* > *Lounge* > *Kitchen* > *Lounge* > *Start*. This indicates the questionnaire was typically completed in the lounge after P2 had moved from the kitchen to the lounge.

We were further interested in identifying behavioral patterns and the associated time periods (on a participant level) through in-home use of conversational technology. We computed the pairwise similarity for each pair of activity sequences, as detailed in Section III-E. This allowed us to form a distinct similarity matrix for each participant. Each data point in the matrix corresponds to an activity sequence of five behavioral events that occurred in the 10-m period preceding or following Alexa use. Pairwise similarity scores were subsequently used to cluster the activity sequences, using the *K*-Medoids method, to identify groups of similar activity sequences and inspect differences between clusters. Because data points in each similarity matrix are chronologically ordered, we were able to identify common household behavior during specific time periods, i.e., groups of activity sequences with a high degree of similarity within the same time period.

Fig. 7(a) shows the similarity matrix with activity sequences that preceded P2's random Alexa interactions from May 2021 until June 2022. The selected border A was further inspected. We used the *K*-Medoids clustering method to identify clusters in the similarity matrix. We identified four distinct clusters in P2's activity sequences. One of these clusters (denoted as $cl_1$ for clarity) became more prominent during a 6-week period, labeled as X1. During X1, 84.69% of activity sequences belonged to the dominant cluster, $cl_1$. Moreover, outside of period X1, only 11.88% of activity sequences belonged to $cl_1$. Further details are presented in Supplementary Table III in the Appendix. This indicates that during X1, there was a distinct pattern in P2's activity sequences. Notably, the probability of bedroom events in activity sequences within $cl_1$ increased to 25.75% compared to a maximum of 2.06% in other clusters (Supplementary Table IV in the Appendix). Additionally, the probability of lounge events declined from a maximum of 41.50% in other clusters to 1.02% in $cl_1$. Furthermore, during X1, 61.06% of household activity preceding Alexa interactions of type random occurred in the morning.

These findings suggest a shift in device location from the lounge to the bedroom in P2's household and a trend toward triggering Alexa for random interactions in the morning. This observation aligns with the higher probability of bedroom events for activity sequences preceding Alexa interactions of type random, as observed in Fig. 6 (14.54%, column C). Importantly, our design team verified with P2 that the Alexa device had been moved to the bedroom during period X1 due to health and caregiving circumstances.

We extend the aforementioned approach and interpretation to other individual similarity matrices. Fig. 7(b) shows the similarity matrix containing P2's activity sequences that preceded the questionnaire trigger from May 2021 until January

2022. The selected border, B, was further analyzed. Note that even though P2 stopped completing the daily questionnaire in late January 2022, the usage of Alexa for random interactions remained consistent until the end of our data collection period, as observed in the similarity matrix previously analyzed [Fig. 7(a)]. One of the three clusters identified using $K$-Medoids (denoted as $cl_3$) was found predominant during a 4-week period, labeled as X2. During X2, 77.78% of data points belonged to cluster $cl_3$. Outside of X2, the proportion of activity sequences associated with $cl_3$ was 21.29% (Supplementary Table III in the Appendix). Furthermore, our analysis revealed 94.74% of household activity preceding the questionnaire trigger during X2 took place in the morning. The increased probability of Vitals events in activity sequences within $cl_3$ (47.87% compared to a maximum of 5.45% in other clusters, see details in Supplementary Table V in the Appendix) suggests a behavioral pattern of taking vitals before completing the daily questionnaire in the morning.

Similarly, Fig. 7(c) shows the similarity matrix representing P14's household behavior before starting the Alexa well-being questionnaire. The selected borders, C and D, were further inspected. We identified four clusters using $K$-Medoids, two of which (denoted as $cl_1$ and $cl_4$) were prominent across two different 11-week time periods, denoted as X3 and X4. During X3, 64.81% of data points fell within cluster $cl_4$, and the proportion of activity sequences associated with $cl_4$ outside of X3 was 29.51%. During X4, 75% of data points were associated with cluster $cl_1$, while only 24.17% of data points outside of X4 belonged to $cl_1$ (Supplementary Table III in the Appendix). Furthermore, all activity sequences within both X3 and X4 were concentrated in the morning hours. Period X3 suggests P14 was consistently in the lounge before completing the questionnaire (the probability of lounge events within $cl_4$ is 89.08%). On the other hand, period X4 indicates the predominance of activity in both the lounge (50% within $cl_1$) and kitchen (38.73% within $cl_1$) (Supplementary Table VI in the Appendix). Overall, these two periods suggest a behavioral pattern related to the use of conversational technology for the purpose of completing a daily questionnaire. Specifically, P14 routinely triggered the questionnaire in the morning, while at the lounge, or following activity in the kitchen. Furthermore, a noticeable darker area within period X3 on the matrix suggests a change in behavior, spanning nine days in early December 2021. During this period, we verified P14's activity sequences consistently incorporated the Vitals event, e.g., *Lounge* > *Lounge* > *Vitals* > *Vitals* > *Start*. This suggests that P14 routinely took vitals in the 10 min before completing the Alexa questionnaire, a pattern that is captured as highly dissimilar compared to the surrounding activity sequences.

By fusing in-home activity data with voice interactions using conversational technology, our analysis demonstrated technical capability in establishing behavioral patterns in households with PLWD, changes in those patterns and the corresponding time periods. We believe the ability to map behavioral trends forms a basis to personalize future interactions. By tracking the user's daily contexts at home, the conversational agent could proactively initiate conversations about relevant domains at appropriate times. Moreover, by detecting changes in behavior (compared to previously identified patterns), the conversational agent could inform healthcare professionals and relevant stakeholders, ultimately enhancing health and well-being monitoring of PLWD at home.

### C. *Use of Conversational AI in Smart Homes Following Clinical Outcomes*

We also investigated whether households with PLWD continued using Alexa during the week following the occurrence of health events (see RQ3 in Section III-B). We used the dates of individual health events (e.g., falls, infections, and hospitalizations) logged by a monitoring team in regular contact with participants. We considered user-initiated triggers of both the questionnaire and random Alexa interactions, specifically in the seven days after the occurrence of a health event. A total of 38 health events were evaluated across the cohort, which comprised events that took place on a day at home (e.g., falls), events of longer duration than a day during which PLWD stayed at home (e.g., Covid-19 infections), and events of longer duration than a day out of home (e.g., hospitalizations). Furthermore, the total events considered in this study corresponded to ten participants (no health events were logged by the monitoring team for the remaining participants).

We observed participants continued interacting with Alexa even after exhibiting a clinical event (Wilcoxon signed rank: $W = 561$, $p_{\text{corr}} = 8.06\text{e-}07$).[6] The same observation was verified when inspecting single dates and periods of health events during which PLWD stayed at home (Wilcoxon signed rank: $W = 435$, $p_{\text{corr}} = 1.92\text{e-}06$). Interestingly, participants continued using Alexa in the 7-day period after returning home from hospitalizations (Wilcoxon signed rank: $W = 10$, $p_{\text{corr}} = 3.39\text{e-}02$). Despite the limited number of observations analyzed across the cohort, results suggest the potential for conversational agents to provide personalized assessments of health and well-being after the occurrence of health events, which we discuss in the next section.

## V. Discussion

In this study, we fused in-home activity data captured by IoT technologies and remote health monitoring devices with interactions with conversational technology. We analyzed 3103 person-days of environmental and voice data across a unique cohort of 14 households with PLWD or MCI. In this section, we summarize the main findings of the study, their implication for future research and real-world translation of conversational agents for utility in digital health monitoring. We also outline the limitations of our investigation and highlight future directions.

### A. *Summary of Findings*

We investigated the integration of conversational AI technology in smart environments. Our target in this study was health and well-being monitoring within smart homes to

[6] To assess whether participants used Alexa after clinical outcomes, we used a one-sample Wilcoxon test. False discovery rate was applied, hence the corrected $p$-values are compared against the significance level $\alpha = 0.017$.

support households with PLWD. We present a method to identify behavioral patterns, changes in those patterns, and the corresponding time periods using conversational technology. Specifically, we analyzed 13 behavioral events (outlined in Table I) related to in-home activity (e.g., motion and taking vitals) and voice interactions with Alexa, both the trigger of a daily well-being questionnaire and other topics of interest (listed in Section III-F).

We first explored the use of Alexa in households with PLWD over time by inspecting the prevalence of interactions beyond the novelty phase (i.e., after the first three months of usage). While a significant decrease in Alexa usage was verified after the novelty phase across the cohort, some topics of interest prevailed in users' daily routines in the post-novelty period. Moreover, results showed a significant decrease in compliance with the daily well-being questionnaire after the novelty phase. We argue this decline in engagement is likely due to a perceived lack of utility and personalization of interactions. One potential explanation is the fact that the questionnaire participants completed throughout the data collection period included the same set of questions each day. The development of an adaptive questionnaire able to proactively check-in for health and well-being self-assessments and follow up on previous user responses represents a very promising area of future research, which our ongoing work is addressing.

Next, we investigated the daily contexts in which Alexa was triggered in the households by analyzing activity sequences (i.e., sequences of ordered behavioral events) in the 10-m period preceding or following Alexa use (see details in Section III-E). By integrating longitudinal in-home activity data captured by IoT technologies and Alexa voice interactions, our research demonstrated the technical capability of identifying behavioral patterns. The analysis of activity sequences led us to confirm the most common location of Alexa in participants' households. We presented a series of case studies for selected participants, which reported different behavioral patterns before user-initiated interactions with Alexa and the corresponding time periods. Reported examples include: a change of the Alexa device location to the patient's bedroom for a duration of six weeks, which was further confirmed by our design team as a response to evolving health and caregiving needs; a pattern of taking vitals before completing the well-being questionnaire in the morning over a period of four weeks; and a consistent sequence of morning activity in the kitchen, followed by triggering the Alexa questionnaire for a period of 11 weeks. We further reported an example of a detected change in household behavior preceding the questionnaire trigger, represented in the similarity matrix as highly dissimilar compared to the surrounding activity sequences (see details in Section IV-B).

Moreover, we found end-users continued using Alexa in the week following clinical outcomes, including after returning home from hospitalizations. While the number of health events analyzed across the cohort prevents broader conclusions with depth of clinical outcome, these preliminary findings indicate an opportunity for proactive and personalized health and well-being check-ins after the occurrence of health events. Overall, although further investigation is needed with a larger cohort over longer periods of time, these findings offer a firm basis for the integration of conversational AI technology in smart environments to monitor user behavior over time. We address design, deployment, engagement, data acquisition, and analysis over 6+ months on average. Results indicate promise to incorporate adaptive conversational agents in smart home contexts. Ultimately, these systems could act on user information captured by IoT technologies, such as nonverbal indicators of physical or mental state, by directly verifying symptoms with end-users.

### *B. Promise of Conversational AI in Digital Health Monitoring*

We believe conversational agents should be inherently integrated with the living environment for a direct impact in supporting aging and dementia care at home. If well incorporated in the home context, including integration with IoT technologies, conversational agents could map and learn common behavioral patterns and act on flagged changes in household activities by initiating automated dialogues. Conversational technology could, for instance, directly query users to verify symptoms or mental state, encourage behavioral changes or provide personalized assistance, such as suggesting an activity or a drink for hydration. This provides a strong foundation for the verification of changes in health and behavior captured by IoT technologies. Conversational agents could ultimately offer verbal support in the event of perceived agitation or confusion and promptly notify relevant stakeholders, which may help mitigate further deterioration through early intervention. Furthermore, more bespoke and engaging interactions could be targeted based on end-users' topics of interest, such as suggesting entertainment activities, pointing out the time of a favorite television program, or reporting news of interest. We envision conversational agents playing an instrumental role in proactive and personalized dementia care. Particularly, conversational agents hold promise to: 1) administer adaptive questionnaires to verify symptoms of deterioration captured by changes in household activity; 2) proactively query users for self-assessments of health and well-being after the occurrence of specific clinical events; and 3) trigger automated alerts to inform healthcare professionals and relevant stakeholders, facilitating timely interventions.

### *C. Limitations*

The current sample size limits broader conclusions with depth of clinical outcome or related to the effect of different demographics in the data collected. Nevertheless, the depth of longitudinal data combined with the unique smart home data collected from households with PLWD enables us to scope our analysis and demonstrate the technical capability to identify behavioral patterns. While a larger sample size will be required to apply these findings to a wider and more diverse population, our study demonstrates the veracity of information gathering and analysis by combining voice with in-home monitoring data. To the best of our knowledge, no study to date has combined interactions with conversational AI with continuous in-home monitoring data, specifically targeting support of older adults and PLWD.

The use of in-home monitoring technologies in real-world evaluation studies poses considerable challenges. For this study, participants needed to have a range of monitoring technologies and an additional interactive device in their homes. One major drawback identified relates to end-users' ethical concerns around personal voice data gathering, which limited the number of participants willing to incorporate Alexa in their homes. While monitoring systems and automated interventions hold very strong promise to improve health and well-being, if not designed with end-user involvement and engagement, they can be perceived as overly complex or intrusive, aside from the associated concerns around data privacy and protection. Furthermore, ensuring long-term acceptability and sustained engagement with conversational technology remains a significant challenge, particularly for older populations with cognitive impairment. These challenges, as discussed in this study, must be addressed in the earliest stages of the design, recruitment, and technology deployment cycle with close end-user feedback to ensure that the technology infrastructure remains noninvasive and privacy aware.

The use of commercial smart speakers brings inherent limitations, such as the 8-s window restriction for the user to respond to Alexa. This limits the applicability of larger scale models, such as large language models (LLMs), in identifying potential indicators of behavioral changes or cognitive decline through acoustic and linguistic features. Another limitation to consider is the need for ground truth. Obtaining ground truth is a complex challenge as it requires finding an appropriate balance to have sufficient training data to validate prediction models or identified behavioral patterns while avoiding intrusion into people's privacy. We are addressing this in our ongoing research with people affected by dementia through user-centered design methods, including user workshops, allowing us to understand their needs, lived experiences, perceived benefits, and concerns while iteratively refining our study design.

### *D. Future Work*

For real-world translation and scalability in smart living environments, we recommend future research on conversational agents aimed at supporting target populations to address intelligent adaptation of interactions, including automated questionnaires. This should account for historical user responses concerning subjective perceptions of health and well-being, individual cognitive abilities, and detected changes from routine behavior at home. The behavioral analysis findings presented in this study could be incorporated into automated and personalized conversations (e.g., prompting users about changes in behavior) to obtain medically relevant data and sustain user engagement.

Future work could address end-user long-term engagement with conversational AI technology. While in general participants used Alexa for long periods of time (note the total days of data collection varied across participants, as outlined in Table II), there was an overall decrease in usage over time, particularly in the post-novelty phase. We believe this is likely due to a perceived lack of utility and adaptation of interactions. We argue that to effectively engage with and support the well-being of target populations in smart environments, conversational technology should be: 1) easy to use; 2) adequately integrated with the environment to proactively respond to contextual cues and provide personalized support; 3) adapt to individual needs, preferences, and cognitive abilities over time; 4) promote the autonomy of the carer, so they can undertake their caring tasks while still benefiting from forms of relief; and 5) facilitate meaningful human connections (e.g., between PLWD, carers, and clinicians).

Furthermore, collecting larger datasets across wider and more diverse populations represents a very promising area of future work, particularly to uncover what type of real-world use cases can be effectively addressed in smart environments from both user benefit and clinical perspectives. Key use cases we identify for future investigation include: 1) tracking mental health and cognitive decline from the use of language when interacting with conversational AI and 2) assessing neuropsychiatric symptoms, such as the risk of agitation and predicting health outcomes from changes in the signature of home activity. In our study, we have analyzed in-home activity data at the household level. This could be extended to a more individualized and personalized approach. However, in instances where multiple individuals (e.g., visitors at home) interact with Alexa, the analysis of activity patterns may be biased. Speaker recognition techniques could help address this issue. Additionally, future work could investigate feature engineering with various parameters of regular interactions with conversational AI, such as their frequency, time of day, changes in vocabulary usage, or sentiment in user utterances, to train ML models toward predicting clinical outcomes.

## VI. Conclusion

With speech and language being natural interaction modalities, conversational agents are promising tools for integration into smart environments to provide insights into users' behavior over time. Yet, the integration of conversational technology in smart homes to trace household behavior, identify patterns or changes in routine, and trigger health-related alerts remains largely untapped. We introduced an infrastructure combining in-home activity data with voice interactions using conversational technology to trace household behavior in smart environments. Our longitudinal data collection spanned 3103 person-days across a unique cohort of 14 households with PLWD. We investigated sustained engagement with Alexa and found a significant decrease in usage after the novelty phase across the cohort. We argue this is likely due to a perceived lack of utility and personalization of interactions and further propose future directions to address current barriers inhibiting longer term user engagement and scalability in smart environments. Our results demonstrated technical capability in establishing behavioral patterns, changes in those patterns and the corresponding time periods using conversational technology. We offer the approach as a basis to personalize future interactions. Moreover, results revealed that participants continued using Alexa following clinical events, which suggests a future opportunity to proactively initiate conversations to monitor health and well-being.

In conclusion, we believe the adequate integration of conversational technology in smart environments—including smart homes with PLWD—holds very strong promise in digital health monitoring. Key to realizing this potential is the development of adaptive AI that can direct conversations to automatically query changes in household behavior and user health, as captured by IoT technologies and remote health monitoring devices. Such systems could encourage behavior through verbal prompts and suggestions tailored to the changing needs of end-users, as well as trigger alerts that inform healthcare professionals and relevant stakeholders for timely interventions. We are unaware of other works with a comparable longitudinal depth of analysis and number of households with PLWD. Plans for longer term evaluation studies across a larger and more diverse cohort, addressing adaptive questionnaires and proactive interactions based on environmental output, are underway.

## Acknowledgment

The authors gratefully acknowledge the ongoing guidance, support and vision of the U.K. Dementia Research Institute. The authors further express their sincerest appreciation to the participants who took part in the study and the U.K. DRI-CR&T monitoring team for the valuable work and insights provided.

## References

[1] "Dementia key facts (2021)." World Health Organization. Sep. 2022. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/dementia

[2] "Aging and health (2021)." World Health Organization. Sep. 2022. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health

[3] "U.K. dementia statistics." Alzheimers Research U.K. Sep. 2022. [Online]. Available: https://www.dementiastatistics.org/statistics/hospitals/

[4] H. Wang et al., "Dementia care during COVID-19," *Lancet*, vol. 395, no. 10231, pp. 1190–1191, 2020.

[5] A. Mehra and S. Grover, "COVID-19: A crisis for people with dementia," *J. Geriatric Mental Health*, vol. 7, no. 1, pp. 1–3, 2020.

[6] H. Brodaty and M. Donkin, "Family caregivers of people with dementia," *Dialogues Clin. Neurosci.*, vol. 11, no. 2, pp. 217–228, 2022.

[7] *Dementia: A Public Health Priority*, World Health Org., Geneva, Switzerland, 2012.

[8] J. Venkatesh, B. Aksanli, C. S. Chan, A. S. Akyurek, and T. S. Rosing, "Modular and personalized smart health application design in a smart city environment," *IEEE Internet Things J.*, vol. 5, no. 2, pp. 614–623, Apr. 2018.

[9] Ş. Kolozali et al., "Observing the pulse of a city: A smart city framework for real-time discovery, federation, and aggregation of data streams," *IEEE Internet Things J.*, vol. 6, no. 2, pp. 2651–2668, Apr. 2019.

[10] A. Solanas et al., "Smart health: A context-aware health paradigm within smart cities," *IEEE Commun. Mag.*, vol. 52, no. 8, pp. 74–81, Aug. 2014.

[11] F. Firouzi et al., "Harnessing the power of smart and connected health to tackle COVID-19: IoT, AI, robotics, and blockchain for a better world," *IEEE Internet Things J.*, vol. 8, no. 16, pp. 12826–12846, Aug. 2021.

[12] I. de la Torre Díez, S. G. Alonso, S. Hamrioui, E. M. Cruz, L. M. Nozaleda, and M. A. Franco, "IoT-based services and applications for mental health in the literature," *J. Med. Syst.*, vol. 43, no. 1, pp. 1–6, 2019.

[13] S. Mehrdad, Y. Wang, and S. F. Atashzar, "Perspective: Wearable Internet of Medical Things for remote tracking of symptoms, prediction of health anomalies, implementation of preventative measures, and control of virus spread during the era of COVID-19," *Front. Robot. AI*, vol. 8, Apr. 2021, Art. no. 610653.

[14] W. Moyle, "The promise of technology in the future of dementia care," *Nat. Rev. Neurol.*, vol. 15, no. 6, pp. 353–359, 2019.

[15] M. R. Lima, M. Wairagkar, N. Natarajan, S. Vaitheswaran, and R. Vaidyanathan, "Robotic telemedicine for mental health: A multimodal approach to improve human-robot engagement," *Front. Robot. AI*, vol. 8, Mar. 2021, Art. no. 618866.

[16] S. Majumder et al., "Smart homes for elderly healthcare—Recent advances and research challenges," *Sensors*, vol. 17, no. 11, p. 2496, 2017.

[17] K. Qian, Z. Zhang, Y. Yamamoto, and B. W. Schuller, "Artificial intelligence Internet of Things for the elderly: From assisted living to health-care monitoring," *IEEE Signal Process. Mag.*, vol. 38, no. 4, pp. 78–88, Jul. 2021.

[18] P. Rashidi, D. J. Cook, L. B. Holder, and M. Schmitter-Edgecombe, "Discovering activities to recognize and track in a smart environment," *IEEE Trans. Knowl. Data Eng.*, vol. 23, no. 4, pp. 527–539, Apr. 2011.

[19] A. Lotfi, C. Langensiepen, S. M. Mahmoud, and M. J. Akhlaghinia, "Smart homes for the elderly dementia sufferers: Identification and prediction of abnormal behavior," *J. Ambient Intell. Humanized Comput.*, vol. 3, no. 3, pp. 205–218, 2012.

[20] P. N. Dawadi, D. J. Cook, and M. Schmitter-Edgecombe, "Modeling patterns of activities using activity curves," *Pervasive Mobile Comput.*, vol. 28, pp. 51–68, Jun. 2016.

[21] Z. K. Shahid, S. Saguna, and C. Åhlund, "Detecting anomalies in daily activity routines of older persons in single resident smart homes: Proof-of-concept study," *JMIR Aging*, vol. 5, no. 2, 2022, Art. no. e28260.

[22] J. Dahmen and D. J. Cook, "Indirectly supervised anomaly detection of clinically meaningful health events from smart home data," *ACM Trans. Intell. Syst. Technol.*, vol. 12, no. 2, pp. 1–18, 2021.

[23] M. R. Lima et al., "Conversational affective social robots for aging and dementia support," *IEEE Trans. Cogn. Development, Syst.*, vol. 14, no. 4, pp. 1378–1397, Dec. 2022.

[24] L. Laranjo et al., "Conversational agents in healthcare: A systematic review," *J. Amer. Med. Inform. Assoc.*, vol. 25, no. 9, pp. 1248–1258, 2018.

[25] T. Dingler, D. Kwasnicka, J. Wei, E. Gong, and B. Oldenburg, "The use and promise of conversational agents in digital health," *Yearbook Med. Inform.*, vol. 30, no. 1, pp. 191–199, 2021.

[26] C. Carroll, C. Chiodo, A. X. Lin, M. Nidever, and J. Prathipati, "Robin: Enabling independence for individuals with cognitive disabilities using voice assistive technology," in *Proc. CHI Conf. Extended Abstracts Human Factors Comput. Syst.*, 2017, pp. 46–53.

[27] "My carer by Alzheimer's Society." Alzheimers Society. Sep. 2022. [Online]. Available: https://www.mrm.com/en/work-case-studies/frankfurt-alzheimers-society-my-carer.html

[28] A. Cheng, V. Raghavaraju, J. Kanugo, Y. P. Handrianto, and Y. Shang, "Development and evaluation of a healthy coping voice interface application using the Google home for elderly patients with type 2 diabetes," in *Proc. 15th IEEE Annu. Consum. Commun. Netw. Conf. (CCNC)*, 2018, pp. 1–5.

[29] K. Tan, K. Sekhar, J. Wong, J. Holgado, M. Ameer, and G. Vesonder, "Alexa Eldercare toolbox: A smarthome solution for the elderly," in *Proc. 11th IEEE Annu. Ubiquitous Comput. Electron. Mobile Commun. Conf. (UEMCON)*, 2020, pp. 806–812.

[30] E. Lee, G. Vesonder, and E. Wendel, "Eldercare robotics—Alexa," in *Proc. 11th IEEE Annu. Ubiquitous Comput. Electron. Mobile Commun. Conf. (UEMCON)*, 2020, pp. 820–825.

[31] A. Sciuto, A. Saini, J. Forlizzi, and J. I. Hong, "'Hey Alexa, what's up?': A mixed-methods studies of in-home conversational agent usage," in *Proc. Des. Interact. Syst. Conf.*, 2018, pp. 857–868.

[32] M. Trajkova and A. Martin-Hammond, "'Alexa is a toy': Exploring older adults reasons for using, limiting, and abandoning echo," in *Proc. CHI Conf. Human Factors Comput. Syst.*, 2020, pp. 1–13.

[33] S. Kim and A. Choudhury, "Exploring older adults' perception and use of smart speaker-based voice assistants: A longitudinal study," *Comput. Human Behav.*, vol. 124, Nov. 2021, Art. no. 106914.

[34] D. Jakob, "Voice controlled devices and older adults—A systematic literature review," in *Proc. Int. Conf. Human-Comput. Interact.*, 2022, pp. 175–200.

[35] D. Beltrami, G. Gagliardi, R. R. Favretti, E. Ghidoni, F. Tamburini, and L. Calzà, "Speech analysis by natural language processing techniques: A possible tool for very early detection of cognitive decline?" *Front. Aging Neurosci.*, vol. 10, p. 369, Nov. 2018.

[36] B. Roark, M. Mitchell, J.-P. Hosom, K. Hollingshead, and J. Kaye, "Spoken language derived measures for detecting mild cognitive impairment," *IEEE Trans. Audio, Speech, Language Process.*, vol. 19, no. 7, pp. 2081–2090, Sep. 2011.

[37] A. Satt, R. Hoory, A. König, P. Aalten, and P. H. Robert, "Speech-based automatic and robust detection of very early dementia," in *Proc. 15th Annu. Conf. Int. Speech Commun. Assoc.*, 2014, pp. 2538–2542.

[38] M. Asgari, J. Kaye, and H. Dodge, "Predicting mild cognitive impairment from spontaneous spoken utterances," *Alzheimer's Dementia Transl. Res. Clin. Interv.*, vol. 3, no. 2, pp. 219–228, 2017.

[39] A. J. Astell et al., "Technology and dementia: The future is now," *Dementia Geriatric Cogn. Disord.*, vol. 47, no. 3, pp. 131–139, 2019.

[40] W. Moyle, J. Murfield, and K. Lion, "The effectiveness of smart home technologies to support the health outcomes of community-dwelling older adults living with dementia: A scoping review," *Int. J. Med. Inform.*, vol. 153, Sep. 2021, Art. no. 104513.

[41] D. J. Cook, N. C. Krishnan, and P. Rashidi, "Activity discovery and activity recognition: A new partnership," *IEEE Trans. Cybern.*, vol. 43, no. 3, pp. 820–828, Jun. 2013.

[42] M. S. Maučec and G. Donaj, "Discovering daily activity patterns from sensor data sequences and activity sequences," *Sensors*, vol. 21, no. 20, p. 6920, 2021.

[43] K. A. Alaghbari, M. H. M. Saad, A. Hussain, and M. R. Alam, "Activities recognition, anomaly detection and next activity prediction based on neural networks in smart homes," *IEEE Access*, vol. 10, pp. 28219–28232, 2022.

[44] J. Yin, Q. Zhang, and M. Karunanithi, "Unsupervised daily routine and activity discovery in smart homes," in *Proc. 37th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, 2015, pp. 5497–5500.

[45] S. Deep, X. Zheng, C. Karmakar, D. Yu, L. G. C. Hamey, and J. Jin, "A survey on anomalous behavior detection for elderly care using dense-sensing networks," *IEEE Commun. Surveys Tuts.*, vol. 22, no. 1, pp. 352–370, 1st Quart., 2020.

[46] D. Arifoglu and A. Bouchachia, "Detection of abnormal behavior for dementia sufferers using convolutional neural networks," *Artif. Intell. Med.*, vol. 94, pp. 88–95, Mar. 2019.

[47] D. Arifoglu and A. Bouchachia, "Activity recognition and abnormal behavior detection with recurrent neural networks," *Procedia Comput. Sci.*, vol. 110, pp. 86–93, Jul. 2017.

[48] N. Camp et al., "Technology used to recognize activities of daily living in community-dwelling older adults," *Int. J. Environ. Res. Public Health*, vol. 18, no. 1, p. 163, 2021.

[49] P. Novitzky et al., "A review of contemporary work on the ethics of ambient assisted living technologies for people with dementia," *Sci. Eng. Ethics*, vol. 21, no. 3, pp. 707–765, 2015.

[50] G. Chimamiwa, A. Giaretta, M. Alirezaie, F. Pecora, and A. Loutfi, "Are smart homes adequate for older adults with dementia?" *Sensors*, vol. 22, no. 11, p. 4254, 2022.

[51] S. Alkhatib, J. Waycott, G. Buchanan, and R. Bosua, "Privacy and the Internet of Things (IoT) monitoring solutions for older adults: A review," in *Connecting System to Enhance Practitioner Consumer Experience Healthcare*. Amsterdam, The Netherlands: IOS Press, 2018, pp. 8–14.

[52] F. Tiersen et al., "Smart home sensing and monitoring in households with dementia: User-centered design approach," *JMIR Aging*, vol. 4, no. 3, 2021, Art. no. e27047.

[53] J. Lee and N. Melo, "Habit representation based on activity recognition," *Sensors*, vol. 20, no. 7, p. 1928, 2020.

[54] L. Meng, C. Miao, and C. Leung, "Towards online and personalized daily activity recognition, habit modeling, and anomaly detection for the solitary elderly through unobtrusive sensing," *Multimedia Tools Appl.*, vol. 76, no. 8, pp. 10779–10799, 2017.

[55] D. Zekri, T. Delot, M. Thilliez, S. Lecomte, and M. Desertot, "A framework for detecting and analyzing behavior changes of elderly people over time using learning techniques," *Sensors*, vol. 20, no. 24, p. 7112, 2020.

[56] D. J. Cook, M. Schmitter-Edgecombe, and P. Dawadi, "Analyzing activity behavior and movement in a naturalistic environment using smart home techniques," *IEEE J. Biomed. Health Inform.*, vol. 19, no. 6, pp. 1882–1892, Nov. 2015.

[57] P. Lago, C. Jiménez-Guarín, and C. Roncancio, "Contextualized behavior patterns for change reasoning in ambient assisted living: A formal model," *Expert Syst.*, vol. 34, no. 2, 2017, Art. no. e12163.

[58] K. Malde, E. Coward, and I. Jonassen, "Fast sequence clustering using a suffix array algorithm," *Bioinformatics*, vol. 19, no. 10, pp. 1221–1226, 2003.

[59] S. Mantaci, A. Restivo, and M. Sciortino, "Distance measures for biological sequences: Some recent approaches," *Int. J. Approx. Reason.*, vol. 47, no. 1, pp. 109–124, 2008.

[60] M. Studer and G. Ritschard, "What matters in differences between life trajectories: A comparative review of sequence dissimilarity measures," *J. Roy. Stat. Soc. A, Stat. Soc.*, vol. 179, no. 2, pp. 481–511, 2016.

[61] H. Saneifar, S. Bringay, A. Laurent, and M. Teisseire, "S$^2$MP: Similarity measure for sequential patterns," in *Proc. Aust. Data Min. Conf.*, vol. 87, 2008, pp. 95–104.

[62] R. W. Hamming, "Error detecting and error correcting codes," *Bell Syst. Tech. J.*, vol. 29, no. 2, pp. 147–160, 1950.

[63] V. I. Levenshtein et al., "Binary codes capable of correcting deletions, insertions, and reversals," *Sov. Phys. Doklady*, vol. 10, no. 8, pp. 707–710, 1996.

[64] C. Gómez-Alonso and A. Valls, "A similarity measure for sequences of categorical data based on the ordering of common elements," in *Proc. Int. Conf. Model. Decis. Artif. Intell.*, 2008, pp. 134–145.

[65] L. Hong and B. D. Davison, "Empirical study of topic modeling in Twitter," in *Proc. 1st Workshop Soc. Media Anal.*, 2010, pp. 80–88.

[66] U. Chauhan and A. Shah, "Topic modeling using latent Dirichlet allocation: A survey," *ACM Comput. Surveys*, vol. 54, no. 7, pp. 1–35, 2021.

[67] D. Newman, E. V. Bonilla, and W. Buntine, "Improving topic coherence with regularized topic models," in *Proc. Int. Conf. Adv. Neural Inf. Process. Syst.*, vol. 24, 2011, pp. 496–504.

[68] Y. Chen, H. Zhang, R. Liu, Z. Ye, and J. Lin, "Experimental explorations on short text topic mining between LDA and NMF based schemes," *Knowl. Based Syst.*, vol. 163, pp. 1–13, Jan. 2019.

[69] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of deep bidirectional transformers for language understanding," in *Proc. NAACL-HLT*, 2019, pp. 4171–4186.

[70] C. Raffel et al., "Exploring the limits of transfer learning with a unified text-to-text transformer," *J. Mach. Learn. Res.*, vol. 21, no. 1, pp. 5485–5551, 2020.

[71] D. Angelov, "Top2Vec: Distributed representations of topics," 2020, *arXiv:2008.09470*.

[72] M. Grootendorst, "BERTopic: Neural topic modeling with a class-based TF-IDF procedure," 2022, *arXiv:2203.05794*.

[73] C. Li, Y. Duan, H. Wang, Z. Zhang, A. Sun, and Z. Ma, "Enhancing topic modeling for short texts with auxiliary word embeddings," *ACM Trans. Inf. Syst.*, vol. 36, no. 2, pp. 1–30, 2017.

[74] R. Egger and J. Yu, "A topic modeling comparison between LDA, NMF, Top2Vec, and BERTopic to demystify twitter posts," *Front. Sociol.*, vol. 7, May 2022, Art. no. 886498.

[75] S. Enshaeifar et al., "Health management and pattern analysis of daily living activities of people with dementia using in-home sensors and machine learning techniques," *PLoS One*, vol. 13, no. 5, 2018, Art. no. e0195605.

[76] J. Sung, H. I. Christensen, and R. E. Grinter, "Robots in the wild: Understanding long-term use," in *Proc. 4th ACM/IEEE Int. Conf. Human Robot Interact.*, 2009, pp. 45–52.

[77] E. Soreq. "U.K.-DRI CRT DCARTE." 2021. [Online]. Available: https://zenodo.org/record/4592930

[78] L. Kaufman and P. J. Rousseeuw, *Finding Groups in Data: An Introduction to Cluster Analysis*. Hoboken, NJ, USA: Wiley, 2009.

[79] N. Reimers and I. Gurevych, "Sentence-BERT: Sentence embeddings using siamese BERT-networks," 2019, *arXiv:1908.10084*.

[80] M. Cho, S.-S. Lee, and K.-P. Lee, "Once a kind friend is now a thing: Understanding how conversational agents at home are forgotten," in *Proc. Design. Interact. Syst. Conf.*, 2019, pp. 1557–1569.


**Maria R. Lima** (Graduate Student Member, IEEE) received the M.Sc. degree in advanced mechanical engineering from Imperial College London, London, U.K., in 2019.

She is a Ph.D. Researcher with Imperial College London and The U.K. Dementia Research Institute Care Research and Technology Centre, London, and holds the President's Ph.D. Scholarship. Her research focuses on conversational AI applications, social robotics, and machine learning for aging and dementia support.

**Ting Su** received the Ph.D. degree in deep neural network and information retrieval from the University of Glasgow, Glasgow, U.K., in 2022.

She is a Postdoctoral Researcher with Imperial College London, London, U.K., and The U.K. Dementia Research Institute Care Research and Technology Centre, London. Her research interests include natural language processing, development and applications of large language models, and chatbots for healthcare.

**Melanie Jouaiti** (Member, IEEE) received the Ph.D. degree in computer science from the Université de Lorraine, Nancy, France, in 2020.

She is a Postdoctoral Researcher with Imperial College London, London, U.K., and The U.K. Dementia Research Institute Care Research and Technology Centre, London. Her research interests include social and assistive robotics for children, development of assistive technologies, machine learning, and speech processing for disordered speech.

**Maitreyee Wairagkar** received the Ph.D. degree in cybernetics from the University of Reading, Reading, U.K, in 2019.

She is a Postdoctoral Researcher with the University of California at Davis, Davis, CA, USA, and BrainGate Consortium, Providence, RI, USA. Her research focuses on breaking barriers between humans and technology by developing intuitive modes of interaction with technology using brain signals, movement, and natural language.

**Paresh Malhotra** received the Ph.D. degree in clinical neurosciences from the University of London, London, U.K, in 2007.

He is a Professor of Clinical Neurology and the Head of the Division of Neurology with Imperial College London, London, U.K. He is an Honorary Consultant Neurologist with Imperial College Healthcare NHS Trust and an Associate Member of The U.K. Dementia Research Institute, London. His research focuses on dementia, particularly AlzheimerŠs disease and vascular dementia, with the aim of developing better treatments for patients.

**Eyal Soreq** received the Ph.D. degree in cognitive neuroscience from Imperial College London, London, U.K., in 2019.

He is a Data Scientist with the U.K. Dementia Research Institute, London. Specialized in neurodegeneration and dementia, he employs machine learning and signal processing to decipher functional neural networks. His primary mission is to identify and validate behavioral biomarkers from diverse data sources, ranging from population-level data to individual IoT devices, to enhance patient well-being and quality of life.

**Payam Barnaghi** (Senior Member, IEEE) received the Ph.D degree in computer science from the University of Malaya, Kuala Lumpur, Malaysia, in 2007.

He is a Professor of Machine Intelligence Applied to Medicine and the Deputy Head of the Division of Neurology, Department of Brain Science, Imperial College London, London, U.K. His research interests include machine learning, adaptive algorithms, digital health, and dementia.

**Ravi Vaidyanathan** (Member, IEEE) received the Ph.D. degree in bioinspired systems from Case Western Reserve University, Cleveland, OH, USA, in 2001.

He is a Professor in Biomechatronics with Imperial College London, London, U.K., and the Robotics Group Lead of The U.K. Dementia Research Institute Care Research and Technology Centre, London. His research in brain–robot interface has resulted in more than 100 refereed publications, six pending or issued patents, and supported the foundation of four new start-up ventures in neurorobotics.