

Quality Assessment Framework for Conversational Agent Applications Related to Alzheimer's Disease and Related Dementias

Diva Smriti, MS¹, Nadien Zerban¹, Katelyn Moore¹, Laura Gitlin, PhD, FGSA, FAAN¹, Aleksandra Sarcevic, PhD, MLIS¹, Erjia Yan, PhD, MSIS¹, Rose Ann DiMaria-Ghalili, PhD, RN, FASPEN, FAAN, FGSA¹, Jina Huh-Yoo, PhD, MHCI¹
¹Drexel University, Philadelphia, PA, USA

Abstract

Alzheimer's disease and related dementias (ADRD) is a chronic progressive condition with no cure; hence most solutions have been focused on developing interventions to enhance the quality of life for persons with ADRD and their caregivers. Although there is an abundance of information online related to ADRD for self-management, the information is not synthesized and delivered in a manner that aligns with the needs and goals of persons with ADRD and their caregivers. Conversational agents, such as Amazon Alexa and Google Home, are receiving increasing attention for supporting the health of older adults. However, no evaluation framework exists to understand the quality of these mechanisms for older adults. In this work, we aim to develop a quality assessment framework on commercially available voice-based applications related to ADRD based on the synthesis of existing related frameworks and iterative testing. The resulting framework will generate novel contributions to improving the quality of information delivered to persons with ADRD and their caregivers.

Introduction

The number of older adults living with Alzheimer's disease and related dementias (ADRD) is projected to increase from 5.8 million in 2020 to 13.8 million in 2050¹. ADRD is a chronic progressive condition, which progresses from mild cognitive impairment to late stage dementia with complete dependency in all activities of daily living¹. Because ADRD cannot be cured, most interventions have been focusing on enhancing quality of life for patients and their caregivers². Although a myriad of health information about dementia self-management and interventions are available online, these data sources are not synthesized and delivered in a manner that addresses the needs, comorbidities, and disease stages of individuals living with ADRD^{3,4}. Conversational Agents (CA) such as Amazon Alexa and Google Assistant are increasingly gaining attention for supporting older adults' health⁵. Early evidence suggests that CA may benefit those who experience challenges in operating existing technologies such as older adults. Home-grown CA designed for various research scenarios evaluated with individuals living with ADRD showed positive effects as a medication reminder and as a tool for sharing stories and promoting exercise. Little is known how CA can help users gather diverse, scattered information from the web. Most prior research on CA for ADRD focused on supporting only a few tasks critical to ADRD such as a reminder or behavior change promoting systems. Little research exists on addressing the needs of caregivers and across the disease trajectory; even less is known about the quality of commercially available applications related to dementia or Alzheimer's disease for use with CA. In this work, we aim to develop a quality assessment framework for voice-based applications on dementia. We will engage and collaborate with stakeholders including persons with ADRD and their informal and formal caregivers for the evaluation of the framework. This project will contribute to the knowledge on assessing the quality of information available for persons with dementia and their caregivers for better design of the information technologies and deliveries for dementia care.

Methods

We first synthesized a framework to assess content and user interaction of commercially available free applications related to dementia or Alzheimer's disease to be used with Amazon Alexa and Google Assistant. The synthesis came from existing assessment frameworks that have been adopted in the contexts of mobile applications on dementia, Amazon Alexa Health Skills, and general conversational agents. A team of researchers in gerontology, human-computer interaction, and information and library science performed the synthesis based on what is the most cited framework and whether the items contained are relevant to supporting ADRD (see Table 1). Two researchers then used the resulting framework to assess the voice assistant applications related to dementia or Alzheimer's disease.

Preliminary Findings

The synthesized framework includes a combination of the MARS⁶, One Mind Psyberguide assessment criteria⁷, and Design Heuristics for CA⁸. This synthesized framework was then used to evaluate current ADRD related applications for use with CA.

Table 1. Preliminary framework synthesized by the researchers.

Assessment Criteria	Reference
Credibility (Consumer Ratings, last updated) and Transparency (Privacy Policy, Data storage and collection policies)	One Mind Psyberguide ⁷
Information Quality (Accuracy of app description, Evidence base)	MARS ⁶
Design Heuristics for CA (Visibility of system status, Match between real world and system, User Control and freedom, Aesthetic and minimalist design, Flexibility and efficiency of use, Consistency, Recognition rather than recall, Error prevention, Personified agents)	Design Heuristics for CA ⁸

As of July 2020, a search using ADRD related keywords (“dementia”, “Alzheimer’s”, or “Alzheimer’s disease”) on Google Assistant and Amazon Alexa returned 425 applications for use with CA. No existing analysis is available on the efficacy of these products. When we excluded applications that did not mention the ADRD related keywords in their descriptions, we were left with only 50 applications for our analysis (see Table 2). These results show the critical challenge of people with ADRD and their caregivers to find relevant information. None of the 50 applications provided tailored information (Flexibility and Efficiency of Use), which is key to supporting caregivers as ADRD patients have different caregiving needs. We used the synthesized framework to evaluate usability and found that the majority applications lacked Error Prevention (e.g., how the application attempts to repair failed conversations), User Control and Freedom (e.g., cannot skip to another menu item), and a Recognition rather than Recall (e.g., making commands explicit to the user so they don’t have to remember which command to use for an action). Despite the opportunities that CA bring to ADRD care, many available CA applications have critical shortcomings in being effectively used by people with ADRD and their caregivers as assessed by our framework. The majority applications also did not provide a privacy policy or an evidence base for their suggestions as informed by the framework.

Table 2. Categorization of 50 ADRD-related shortlisted applications for use with Amazon Alexa and Google Home.

Application Purpose (Amazon Alexa and Google Assistant combined)	Count
Providing ADRD related information for people with ADRD and their caregivers: Broadcasting podcasts (21); Q&A (7); news related to ADRD (6)	34
Providing Games and other activities for people with ADRD: Memory games to improve color and arithmetic skills (8); play music, tell jokes, or set up daily tasks (2); self-reflection (1); locating objects in a home (1); locating family members (1); improving hearing (1)	8
Archiving personal memory of people with ADRD	2
Total	50

Future Work

To make the framework reflective of the people with ADRD and their caregivers, we plan to use a two-pronged approach. First, the framework will be introduced to two focus groups of people with ADRD, informal caregivers, and formal caregivers of ADRD (3-5 people each). In this focus group, we will ask whether the questions of the framework make sense and ways to improve the assessment instrument. We will then improve the framework and its assessment instrument using the results of the focus groups. Second, we will recruit people with ADRD, informal caregivers, and formal caregivers who will be asked to assess the same set of CA applications on ADRD using the revised framework. We will compare the results amongst the participants using inter-rater reliability and follow up interviews. This will result in a framework that will provide a baseline for designing CA-based applications for use by people with ADRD and their caregivers, thereby making access to online resources easier.

References

1. Alzheimer's Association. 2020 Alzheimer's disease facts & figures. 2020.
2. Gitlin LN, Hodgson NA. Better living with dementia: Implications for individuals, families, communities, and societies. Academic Press; 2018.
3. Fischer SH, David D, Crotty BH, Dierks M, Safran C. Acceptance and use of health information technology by community-dwelling elders. *Int J Med Inform* [Internet]. 2014;83(9):624–35. Available from: <http://dx.doi.org/10.1016/j.ijmedinf.2014.06.005>
4. Choi SK, Yelton B, Ezeanya VK, Kannaley K, Friedman DB. Review of the content and quality of mobile applications about alzheimer's disease and related dementias. *J Appl Gerontol*. 2018;
5. Carroll C, Chiodo C, Lin AX, Nidever M, Prathipati J. Robin: Enabling independence for individuals with cognitive disabilities using voice assistive technology. *Conf Hum Factors Comput Syst - Proc*. 2017;Part F1276:46–53.
6. Stoyanov SR, Hides L, Kavanagh DJ, Zelenko O, Tjondronegoro D, Mani M. Mobile app rating scale: a new tool for assessing the quality of health mobile apps. *JMIR mHealth uHealth*. 2015;3(1):e27.
7. One Mind Psyberguide [Internet]. 2020 [cited 2020 Oct 1]. Available from: <https://onemindpsyberguide.org/about-psyberguide/>
8. Shin JY, Huh-Yoo J. Designing everyday conversational agents for managing health and wellness: a study of alexa skills reviews. In *Proceedings of the 14th EAI International Conference on Pervasive Computing Technologies for Healthcare*. 2020.