Voice-Based Conversational Agents for Older Adults

Janet Johnson
Khalil Mrini
Computer Science and Engineering, UC San Diego
La Jolla, CA 92093, USA
janetjohnson@ucsd.edu
khalil@ucsd.edu

Michael Hogarth
Biomedical Informatics
UC San Diego
La Jolla, CA 92093, USA
mihogarth@ucsd.edu

Alison Moore
Geriatrics and Gerontology
UC San Diego
La Jolla, CA 92093, USA
alm123@ucsd.edu

Ndapa Nakashole
Computer Science and Engineering, UC San Diego
La Jolla, CA 92093, USA
nnakashole@eng.ucsd.edu

Nadir Weibel
Computer Science and Engineering, UC San Diego
La Jolla, CA 92093, USA
weibel@ucsd.edu

Emilia Farcas
Qualcomm Institute
UC San Diego
La Jolla, CA 92093, USA
efarcas@ucsd.edu

CHI'20,, April 25–30, 2020, Honolulu, HI, USA
ACM 978-1-4503-6819-3/20/04.
https://doi.org/10.1145/3334480.XXXXXXX

Abstract
Most older adults are reluctant to move out of their homes as they age, but multiple chronic conditions and reduced cognitive and physical abilities make living independently extremely challenging. With our population aging rapidly, enabling older adults to take an active role in their health and remain independent while staying at home has become a crucial societal goal. This position paper discusses designing and building VOLI, a voice-based conversational agent for older adults that makes monitoring their well-being and connecting with their doctors more accessible. VOLI builds a pipeline from older adults' smart speakers at home to the healthcare system and their Electronic Health Records, and allows them to interact with it through dialogue-based interactions. We highlight the challenges and opportunities that our community faces when creating conversational agents for older adults at the intersection of HCI, computational linguistics, and medical informatics.

Author Keywords
Conversational Agents; Older Adults; Electronic Health Records; Quality of Life; VOLI

VOLI: A Voice Assistant for Quality of Life and Healthcare Improvement in Aging Populations

Technical Aim: Investigate how voice-based assistants, NLP, and machine learning can produce meaningful health-related conversations by leveraging population- and patient-level data from EHRs.

Social, Behavioral, and Cognitive Aim: Design services for older adults based on their needs to support independence and investigate the acceptability of digital assistants for the aging population.

Clinical Aim: Investigate how voice-based assistants can help detect new symptoms and correlate them with medication side effects, medication interactions, worsening of existing conditions, or onset of a new illness or allergy.

Introduction

According to the 2017 National Population Projections, there are at least 49.2 million people aged 65 and older in the US alone. This number is projected to increase to 77 million by 2034, making it the first time older adults would outnumber children in the US history [2].

Despite multiple chronic conditions, reduced cognitive capabilities, and difficulty performing daily activities [5, 13], most older adults typically wish to age in place and don't want to move out of homes. Enabling our aging population to stay in their homes and remain independent is an important societal goal in our future. However, reducing the onset and impact of disability requires high quality care and can prove to be challenging for the healthcare system, stressful for the caregivers, and fiscally unsustainable. Technology offers an opportunity to support and enhance the daily lives of older adults: it can protect and augment their ability to perform essential functions, communicate with others, access information, and take an active role in their healthcare.

Voice is a natural interaction method for humans and could be a powerful tool for older adults, especially in their homes. Previous research [3] has shown that spoken language was the preferred interaction modality for elderly users, especially those unfamiliar with technology. Conversational agents have also been shown to be effective in providing health education, monitoring, and also promoting health behavior changes [7]. Smart speakers with integrated voice-based conversational agents (like Google Home or Amazon Echo) are already used by older adults, with commonly cited uses including entertainment, companionship, home control, reminders, and emergency communication [8]. However, while these devices are relatively ubiquitous and affordable today, they aren't designed with the unique needs of older adults in mind.

VOLI: Voice-Based Assistant for Older Adults

We are building a personalized and context-aware voice-based digital assistant, called VOLI, to connect older adults at home with the health system and their Electronic Health Records (EHR), and improve their quality of life (Fig. 1). The development of VOLI will follow a human-centered process with need-finding revolving around three major populations: the older adults, the healthcare professionals they interact with, and their caregivers. This position paper discusses the challenges and opportunities that our community faces when creating conversational agents at the intersection of older adults and EHRs.

Integrating context from Electronic Health Records

EHRs holds longitudinal patient health information and are commonly accompanied by portals that patients can use to message their doctors. The growing amount of information in EHRs provides an unprecedented opportunity for it to directly support self-care and behavior change. VOLI aims to integrate EHR data and combine it with information from reliable public health knowledge-bases to help maintain personalized dialogues and answer specific questions based on the history of clinical visits.

VOLI exploits two types of information contained in patient EHRs: (1) structured information (demographics, diagnosis, lab tests, medications, and procedures), and (2) unstructured information (free text documents such as clinical notes and messages between patients and doctors). The structured EHR data provides us with the context required to personalize conversations to the individual, their medications, and their current health status. Unstructured data aids the detection of more colloquial terms that are commonly used to refer to medications, diagnoses, and symptoms, helping us present medical information in familiar terms that older adults can better understand.

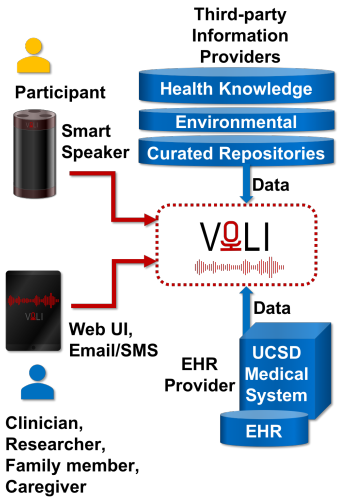

Figure 1: Overview of the VOLI system, where older adults interact with a smart speaker to ask questions, access health-related information and services, and report on their health condition. VOLI uses deep learning for natural language understanding backed by information from EHRs, as well as reliable public health knowledge-bases (like those authored by federal government agencies or professional organizations). This information will help personalize the experience and bring in the specific health context of each individual.

Supporting daily living activities and independence
The reduced physical, cognitive, and social capabilities of our aging population makes it hard for them to maintain a daily routine of timely meals, medication, and exercise. External events like meeting friends and family, or appointments like a doctors visit are typically quite challenging. However, helping older adults follow a structured daily schedule has been shown to allow many of them to live autonomously [14] by encouraging the formation of healthy habits and a lifestyle that preserves independence.

A conversational agent designed for older adults could not only provide utility for common tasks (reminders, emergency communication, entertainment), but could also enhance its utility by taking into account their unique context. Including information gleaned from their EHRs will allow us to both personalize and dynamically compose the dialogue, and encourage behaviors beneficial for them. Such an agent would provide consistent health information, and allow the older adults as much time as they need to fully understand in a low-pressure environment where questions could be asked multiple times. This could also potentially increase older adults' ability to understand their health better, allowing them to help detect new symptoms and correlate them with medication side effects, worsening of existing conditions, or onset of new ones.

Enabling conversations with healthcare providers
Since visits with doctors typically happen in intimate face-to-face settings, conversations with older adults are tailored to their individual needs and include opportunities to ask clarifying questions. However, a major problem with such in-person encounters is that the healthcare environment only allows for limited time per patient. The time pressure makes people feel scared to ask questions, or to ask for information to be repeated [1].

While it is important to encourage older adults to work with their doctors to manage chronic diseases and/or maintain their health and well-being, constant in-person visits can both be infeasible and expensive. Concern over new symptoms or a change in behavior could also result in a costly, often unnecessary trip to the emergency room. In providing health information, conversational agents can help reduce unnecessarily frequent doctors visits and make communication with doctors through patient portals more seamless and accessible. A companion interface dedicated for healthcare providers can also be designed to provide any additional details and context that is needed to monitor and help reduce the onset and impact of disability in older adults.

Finally, many people do not have easy access to healthcare professionals and, even with a potentially lower efficacy than with in-person visits, conversational agents (especially on existing devices in the market), have the potential to reach a much greater portion of the population [1].

Empowering the older adults' caregivers
Routine tasks such as monitoring and reminding older adults of daily activities often constitute a sizable chunk of caregiver tasks, taking time away from ambulatory care or time that could be better spent in uninterrupted in-depth interactions [14].

Using a conversational agent opens up new opportunities to reduce caregiver burden, whether they are professionally employed, a family member, or a friend. A conversational agent is naturally designed for reminders, and can be used as a companion interface that can help caregivers, including those geographically distant, to monitor and be more easily involved in the health and care of an older adult.

NLP and Speech Technologies for Voice Assistants

The process to create voice assistants involves two main parts: a module for voice-to-text and text-to-speech translation, and a module for language understanding and generation.

Voice <–> Text: Amazon's Lex is a chatbot-building service that includes Automatic Speech Recognition (ASR) as well as language understanding capacities, using similar technologies as Alexa. Other examples of ASR systems include Baidu Research's Deep Speech [6], and Facebook's wav2letter [9].

Language Understanding and Generation: Facebook's DrQA is a question-answering (QA) chatbot trained to extract answers from Wikipedia [6]. Unlike QA chatbots, knowledge-grounded chatbots can take context into account: the Wizard of Wikipedia is able to hold a conversation [4].

Design and Technical Challenges

While a personalized, context-aware conversational agent might indeed offer many advantages to older adults, designing such a system can be extremely challenging. In VOLI we rely on a variety of innovations in computational linguistics and medical informatics, as well as on a thorough human-centered design process.

A key challenge that faces Machine Learning (ML) today is generalization; modern ML methods have achieved success on various Natural Language Processing (NLP) problems (e.g., question answering and speech recognition), in part due to advances in training deep neural networks. To robustly estimate model parameters, training deep learning systems requires large quantities of labeled data. In the absence of enough labeled data, such models generalize poorly to new test datasets. For this reason, ML models are trained on datasets from domains with large quantities of data. Conversational agents are typically trained on Twitter conversations, or movie dialogues, while question-answering systems are trained on Wikipedia. In addition, speech recognition systems are usually trained on data obtained from middle-aged people, and not older adults.

We see the design of reliable conversational agents for older adults as a significant challenge, since older adults have additional characteristics that are not typically considered in current work. For instance older adults preserve their vocabulary and could have a richer, more varied vocabulary compared to young adults despite their declining cognitive abilities [11]. They are also more likely to have trouble finding a particular word and/or stutter, leading to a conversation with more unexpected pauses, 'tip of the tongue' experiences, and different patterns of speech [10]. Finally, older adults often attempt social interaction with systems and use expressions more appropriate in human-human interaction (e.g., 'thanks' or 'goodbye'), making systems trained on younger adults an ill-fit for older adults [12].

Conversations around health and well-being also present a major challenge due to two reasons: (1) Deriving relevant information from EHR data is non-trivial since the data is heterogeneous, sparse, and noisy; (2) Health-related dialogues have unique features and require great attention to detail, as the dialogues need to reflect the inferred health goals of the user at that particular moment. More research is needed on NLP methods that generalize to domains of healthcare without requiring new massive datasets.

While the lack of complex on-screen interactions is beneficial for the aging population, voice-based interaction also inherently lacks signifiers and design cues. This means that it is typically hard for users to understand what a conversational agent can or cannot do. The aging population is highly diverse, and age itself is a very poor predictor of an individuals physical and cognitive health, making it difficult to create an interaction model or dialogue architecture that accounts for this. This calls for a comprehensive, more in-depth human-centered design process to outline needs, preferences and abilities of older people.

Conclusion

Our VOLI system envisions a tighter connection of the older adult home with the healthcare system and its EHR through conversational agents. We believe that the challenges and opportunities presented in this position paper are important, and need to be further discussed to advance the work of our community. We are confident that by having the HCI, computational linguistics, and medical informatics communities come together to design and build systems like VOLI, we will be able to truly enhance the quality of life of the growing population of older adults.

Acknowledgements

We gratefully acknowledge the award from NIH/NIA R56AG067393. Michael Hogarth has an equity interest in LifeLink Inc. and also serves on the company's Scientific Advisory Board. The terms of this arrangement have been reviewed and approved by the University of California, San Diego in accordance with its conflict of interest policies.

REFERENCES

[1] Timothy Bickmore and Toni Giorgino. 2006. Health dialog systems for patients and consumers. *Journal of biomedical informatics* 39, 5 (2006), 556–571.

[2] US Census Bureau. 2018. Older people projected to outnumber children for first time in US history. *RELEASE NUMBER CB18-41* (2018).

[3] Amy Cheng, Vaishnavi Raghavaraju, Jayanth Kanugo, Yohanes P Handrianto, and Yi Shang. 2018. Development and evaluation of a healthy coping voice interface application using the Google home for elderly patients with type 2 diabetes. In *2018 15th IEEE Annual Consumer Communications & Networking Conference (CCNC)*. IEEE, 1–5.

[4] Emily Dinan, Stephen Roller, Kurt Shuster, Angela Fan, Michael Auli, and Jason Weston. 2018. Wizard of wikipedia: Knowledge-powered conversational agents. *arXiv preprint arXiv:1811.01241* (2018).

[5] Cother Hajat and Emma Stein. 2018. The global burden of multiple chronic conditions: A narrative review. *Preventive medicine reports* 12 (2018), 284–293.

[6] Awni Hannun, Carl Case, Jared Casper, Bryan Catanzaro, Greg Diamos, Erich Elsen, Ryan Prenger, Sanjeev Satheesh, Shubho Sengupta, Adam Coates, and others. 2014. Deep speech: Scaling up end-to-end speech recognition. *arXiv preprint arXiv:1412.5567* (2014).

[7] Catriona M Kennedy, John Powell, Thomas H Payne, John Ainsworth, Alan Boyd, and Iain Buchan. 2012. Active assistance technology for health-related behavior change: an interdisciplinary review. *Journal of medical Internet research* 14, 3 (2012), e80.

[8] Katherine O'Brien, Anna Liggett, Vanessa Ramirez-Zohfeld, Priya Sunkara, and Lee A Lindquist. 2020. Voice-Controlled Intelligent Personal Assistants to Support Aging in Place. *Journal of the American Geriatrics Society* 68, 1 (2020), 176–179.

[9] Vineel Pratap, Awni Hannun, Qiantong Xu, Jeff Cai, Jacob Kahn, Gabriel Synnaeve, Vitaliy Liptchinsky, and Ronan Collobert. 2018. wav2letter++: The fastest open-source speech recognition system. *arXiv preprint arXiv:1812.07625* (2018).

[10] Meredith A Shafto, Deborah M Burke, Emmanuel A Stamatakis, Phyllis P Tam, and Lorraine K Tyler. 2007. On the tip-of-the-tongue: neural correlates of increased word-finding failures in normal aging. *Journal of cognitive neuroscience* 19, 12 (2007), 2060–2070.

[11] Paul Verhaeghen. 2003. Aging and vocabulary score: A meta-analysis. *Psychology and aging* 18, 2 (2003), 332.

[12] Ravichander Vipperla, Maria Wolters, Kallirroi Georgila, and Steve Renals. 2009. Speech input from older users in smart environments: Challenges and perspectives. In *International Conference on Universal Access in Human-Computer Interaction*. Springer, 117–126.

[13] Brian W Ward, Jeannine S Schiller, and Richard A Goodman. 2014. Peer reviewed: multiple chronic conditions among us adults: a 2012 update. *Preventing chronic disease* 11 (2014).

[14] Ramin Yaghoubzadeh, Marcel Kramer, Karola Pitsch, and Stefan Kopp. 2013. Virtual agents as daily assistants for elderly or cognitively impaired people. In *International workshop on intelligent virtual agents*. Springer, 79–91.