Conversational Agents for Elderly Users

Risako Owan, Libby Ferland, Naome Etori, Wilma Koutstaal, and Maria Gini
Department of Computer Science and Engineering, University of Minnesota, Minneapolis, USA
{owan0002,ferla006,etori001,kouts003,gini}@umn.edu

***Abstract*—A growing global population of elderly dictates new technological approaches to societal needs. Conversational agents (CAs) are a promising technology to support older adults in maintaining independence. A growing body of research supports the potential of CAs in diverse applications. One of the primary factors motivating the use of CAs as social support is that combating issues common to older adults, such as feelings of loneliness and isolation, has strong lasting effects on their health and ability to stay at home longer. A primary challenge is fostering and sustaining long-term user engagement and adoption. We outline the design of a CA meant to converse with older adults about topics of interest, as well as the promising results of testing with older adult participants in a series of five interactions. Preliminary data indicate prolonged and consistent engagement with the system, as well as a willingness to continue use.**

***Index Terms*—conversational agents, older adults, elderly users, positive affect, system usability.**

I. Introduction

The unprecedented aging of the global population has created a critical turning point in the development of infrastructure to support the care needs of older adults. The WHO estimates that there will be 1.5 billion people in the 65 and older age group by the year 2050; however, this skew means a parallel decrease in the working-age population that will make it much more difficult to provide adequate care to older adults who may be in need of varying levels of support [1], [2], [3].

Intelligent systems have evolved rapidly in recent years to be well-poised to become part of the infrastructure to support an older population. These technologies can cover a variety of needs in gerontechnology, from sensored homes to physical assistance to social companionship. Designing interfaces to these technologies for older users is a difficult and core question that requires consideration of their specific needs [4], [5]. Conversational agents and voice interfaces are uniquely poised to address these needs due to their accessibility and ease of use [6], [7]; however, the persistent gaps between user experience and expectation often lead to poor reception and low adoption rates for long-term use [8], [9]. One of the best ways to design conversational technology is to involve older users in the design process from the ground up [4]. There has been a recent shift towards this paradigm by the community; however, it is still resource intensive to undergo this type of design process, and many gaps remain unaddressed.

The work described here is derived from the last phase of a multi-year NSF project to address these gaps. 'Calena' is a CA (conversational agent) designed to interact with older adults about relevant areas of their lives such as their families' hobbies, and daily activities. Calena's design involved older adults from the beginning – the design was developed and fine-tuned by previous experiments with older adults and guided by previous work with younger adults [10]. Calena itself is powered by deep learning, and is a fully functional prototype tested directly through conversations with older users.

II. Background

Designing With Older Users. Designing technologies specifically for elderly users requires a user-centered design approach that directly involves the target users throughout the process [11], [12], [13], [4], [14], [15]. Studies emphasize the importance of participatory design methods like interviews, and surveys with older adults to understand their preferences, needs, and expectations [16], [17], [18]. For example, elderly users may desire systems that are somewhat constrained to feel more predictable but can still support social conversation, so designers should carefully test different interaction flows and constraints [19]. When designing for older adults, any limitations in mobility, vision, hearing, or other areas should be considered and design activities adjusted accordingly [20].

Concerns of Older Users. Elderly users have unique concerns about using technology, which can impact their willingness to adopt and use technology, and their overall satisfaction with it. Privacy and security are major concerns, as they wish to maintain independence [21], dignity [22], and trust [23]. Technologies that are perceived as monitoring or surveillance can make users uneasy, even if no data is externally stored [24]. Older adults are apprehensive about external data storage and transmission [25], hence very skeptical [26]. Acceptance of new technologies is influenced by perceptions of usefulness and usability [27]. Special consideration for potential physical and cognitive limitations is needed without patronizing [20], instructions and feedback should be clear and simple[28]. Age-inclusive rather than age-exclusive design is ideal [15].

III. Approach

The core of this work was the development of a prototype spoken dialogue system deployed in experimental interactions with older adults. We used an iterative design process with a series of experiments, including a feasibility testing phase with older adults interacting with a Wizard of Oz (WoZ) system [29] (Phase I). Phase II was a small interview study with the same WoZ participants to gain better insight into their thoughts on such a system. The information gathered in Phase II pinpointed the selection of family and hobbies as conversational domains of interest to older users. Actual development (Phase III) started after completing these information gathering stages.

The prototype we developed, named "Calena" for its origins in a task-oriented system designed for scheduling functions, included various components meant to support key experimental goals, including robust support for domain-specific conversations, maintenance of personalized user knowledge, powerful speech processing and generation to increase usability, and protection of user privacy and any sensitive data. To do this, we combined a number of elements into one prototype.

Conversational Agent and Mindmeld. Cisco's MindMeld API [30] is a robust conversational agent API released in the late 2010s designed to create CAs with deep domain knowledge supported by standard language models such as BERT. Using MindMeld we created in-depth domain models with custom entities and training data. We included data from participants in previous experiments and data generated by members of the research team. Using MindMeld we had created prototypes for other projects, such as a calendar and scheduling [10], which enabled us to make a more capable system by integrating some features from those prototypes.

Personalization. Personalization has proven to be a powerful tool for encouraging older individuals to continue engaging with their CAs [14]. We chose to store any user information in a PKB (personalized knowledge base) that was accessible and structured to make it easy to use for future extensions to the project. Apache Gremlin [31] was chosen for several reasons. First, it is not rigidly structured like JSON or XML, and allows any number of defined entity types that aren't required to have every attribute filled in. Gremlin data is also stored as a connected graph, so it is easier to store both different pieces of information and the relationships between them, which makes traversing and reasoning with the PKB easier.

Speech Processing and Generation. We chose Microsoft Azure cognitive services for the text-to-speech and speech-to-text required for the dialogue system. In addition to best addressing concerns about data protection, Azure allows to set custom timeouts when listening, to account for a wider number of disfluencies in older users. Using an out-of-the-box solution for these two parts of the system gave us greater processing power and live input sanitization, which enabled us to exploit the domain structure and entities in MindMeld.

User Privacy and Security. Our previous work shows that data protection and privacy are of great importance to older users [8]; we have also found that regardless of system design, users will disclose a diverse amount of personal information even in task-oriented systems [32]. With this in mind, we considered privacy and data security to be a primary focus in our prototype design. This dictated our choice of speech services and of knowledge representation in Gremlin. We found that, unlike competitors, Microsoft Azure allows ownership of the data because our cloud storage is protected and private [33], [34]. Last, Gremlin is a temporary server – unless the graph data is explicitly written to disk at some point, any data will be lost when the Gremlin server is shut down or restarted [31]. This was done as an extra means of user security.

IV. Experimental design and preliminary results

The culmination of Phase III was the deployment of Calena in live testing with older adults. The number of participants was 26: 19 females and 7 males, aged 60-81, all native English speakers. We wanted to measure the impact of positive prompts used by the CA in the conversation because talking about positive past or anticipated future events is known to increase feelings of well-being and was therefore expected to promote user engagement.

Participants were divided into two groups A and B, both received balanced positive vs. neutral prompts in different days, with the total number of positive prompts equated across the groups. Over the five days, the positive prompts were 1-4-1-1-4 for group A and 1-1-4-4-1 for group B. Participants provided feedback after each interaction and at the conclusion of the study. The positive prompt determines participant's willingness to interact with CA. Prior to the study, all participants provided a 'technology profile' outlining their past and present technology experience and their self-assessed technology skills and interests [35], [36], [37].

We collected a wealth of information from participants throughout the study, including session-specific usability assessments via the System Usability Scale (SUS) [38] and impressions of Calena's behavior. Participants answered 10 questions on a 5-point scale about their impressions at the end of the fifth session. The scores were combined and scaled to produce a score between 0 and 100.

Preliminary analysis of the data is promising. SUS scores throughout the course of the experiment remained reasonably high and consistent, as seen in Figure 1. The average SUS score for studies is 68, so we did quite well. At the conclusion of the experiment, many participants indicated continued interest and willingness to interact with the system and were fairly positive about using similar systems for longer-term use.

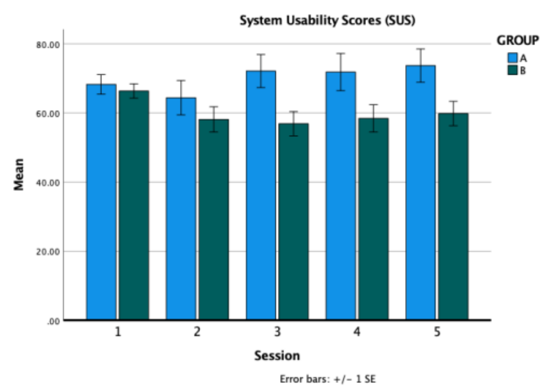

Fig. 1. Participant System Usability Scale scores across sessions

V. Conclusions

Our preliminary results with Calena are promising. CAs can become a useful way of providing social support to elderly people at a limited cost and with no need for complex instructions and training. To maintain engagement from the users, more work needs to be done to enrich the conversational ability of the CAs while ensuring the privacy is respected.

Acknowledgments

Partial funding for this work has been provided by the University of Minnesota Grand Challenges Research, NSF I/UCRC (IIP-1439728), NSF EAGER (IIS 1927190), and the Minnesota Robotics Institute.

REFERENCES

[1] "Ageing and health," https://www.who.int/news-room/fact-sheets/detail/ageing-and-health, 2022, [Accessed 21-08-2023].

[2] U. Nations, "Ageing," https://www.un.org/en/global-issues/ageing, [Accessed 21-08-2023].

[3] K. G. Kinsella and D. R. Phillips, "Global aging: The challenge of success," *Population Reference Bureau Washington, DC*, vol. 60, no. 1, 2005.

[4] M. Heerink, B. Kröse, V. Evers, and B. Wielinga, "Assessing acceptance of assistive social agent technology by older adults: the Almere model," *International Journal of Social Robotics*, vol. 2, pp. 361–275, 2010.

[5] F. Meiland *et al.*, "Technologies to support community-dwelling persons with dementia: a position paper on issues regarding development, usability, effectiveness and cost-effectiveness, deployment, and ethics," *JMIR Rehabilitation and Assistive Technologies*, vol. 4, no. 1, 2017.

[6] Z. Ashktorab, M. Jain, Q. V. Liao, and J. D. Weisz, "Resilient chatbots: Repair strategy preferences for conversational breakdowns," in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019, pp. 1–12.

[7] S. Pednekar, P. Dhirawani, R. Shah, N. Shekokar, and K. Ghag, "Voice-based interaction for an aging population: A systematic review," in *2023 3rd International Conference on Intelligent Communication and Computational Techniques (ICCT)*. IEEE, 2023, pp. 1–8.

[8] L. Ferland, T. Huffstutler, J. Rice, J. Zheng, S. Ni, and M. Gini, "Evaluating older users' experiences with commercial dialogue systems: Implications for future design and development," in *Workshop DEEP DIAL at AAAI*, 2019.

[9] M. Trajkova and A. Martin-Hammond, ""Alexa is a Toy": exploring older adults' reasons for using, limiting, and abandoning echo," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 2020, pp. 1–13.

[10] L. Ferland, J. Sauve, M. Lucke, R. Nie, M. Khadar, S. Pakhomov, and M. Gini, "Tell me about your day: designing a conversational agent for time and stress management," in *International Workshop on Health Intelligence (W3PHIAI) at AAAI*, 2020.

[11] ——, "Tell me about your day: designing a conversational agent for time and stress management," *Explainable AI in Healthcare and Medicine: Building a Culture of Transparency and Accountability*, pp. 297–303, 2021.

[12] L. Ferland and W. Koutstaal, "How's your day look? the (un) expected sociolinguistic effects of user modeling in a conversational agent," in *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, 2020, pp. 1–8.

[13] N. da Paixão Pinto, J. B. dos Santos França, H. P. de Sá Sousa, A. S. Vivacqua, and A. C. B. Garcia, "Conversational agents for elderly interaction," in *2021 IEEE 24th International Conference on Computer Supported Cooperative Work in Design (CSCWD)*. IEEE, 2021, pp. 1–6.

[14] N. d. P. Pinto, J. B. dos Santos Franca, H. P. de Sa Sousa, A. S. Vivacqua, and A. C. Bicharra Garcia, "Conversational agents for elderly interaction," in *Proceedings of the 2021 IEEE 24th International Conference on Computer Supported Cooperative Work in Design (CSCWD)*. IEEE, 2021, pp. 1–6.

[15] A. F. Newell and P. Gregor, "User sensitive inclusive design —in search of a new paradigm," in *Proceedings on the 2000 Conference on Universal Usability*, 2000, pp. 39–44.

[16] V. K. Jones, M. Hanus, C. Yan, M. Y. Shade, J. B. Boron, and R. M. Bicudo, "Reducing loneliness among aging adults: The roles of personal voice assistants and anthropomorphic interactions," *Frontiers in Public Health*, vol. 9, 2021.

[17] S. Sayago, B. B. Neves, and B. R. Cowan, "Voice assistants and older people: some open issues," in *Proceedings of the 1st International Conference on Conversational User Interfaces*, 2019, pp. 1–3.

[18] 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 *Proceedings of the 2019 on Designing Interactive Systems Conference*, 2019, pp. 1557–1569.

[19] L. Ferland, J. Sauve, M. Lucke, R. Nie, M. Khadar, S. Pakhomov, and M. Gini, "Tell me about your day: Designing a conversational agent for time and stress management," in *International Workshop on Health Intelligence (W3PHIAI) at AAAI*, 2020.

[20] A. D. Fisk, S. J. Czaja, W. A. Rogers, N. Charness, and J. Sharit, *Designing for older adults: Principles and creative human factors approaches*. CRC press, 2020.

[21] E.-M. Schomakers and M. Ziefle, "Privacy concerns and the acceptance of technologies for aging in place," in *Human Aspects of IT for the Aged Population. Design for the Elderly and Technology Acceptance: 5th International Conference, ITAP 2019, part of 21st HCI International Conference*. Springer, 2019, pp. 313–331.

[22] H. Umemuro, "Lowering elderly Japanese users' resistance towards computers by using touchscreen technology," *Universal Access in the Information Society*, vol. 3, pp. 276–288, 2004.

[23] J. Edwards and E. Sanoubari, "A need for trust in conversational interface research," in *Proceedings of the 1st International Conference on Conversational User Interfaces*, 2019, pp. 1–3.

[24] F. Kirchbuchner, T. Grosse-Puppendahl, M. R. Hastall, M. Distler, and A. Kuijper, "Ambient intelligence from senior citizens' perspectives: Understanding privacy concerns, technology acceptance, and expectations," in *Ambient Intelligence: 12th European Conference (AmI 2015)*. Springer, 2015, pp. 48–59.

[25] L. Ferland, Z. Li, S. Sukhani, J. Zheng, L. Zhao, and M. L. Gini, "Assistive ai for coping with memory loss." in *AAAI Workshops*, 2018, pp. 431–434.

[26] S. Lindsay, K. Brittain, D. Jackson, C. Ladha, K. Ladha, and P. Olivier, "Empathy, participatory design and people with dementia," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2012, pp. 521–530.

[27] K. Chen and A. H. Chan, "A review of technology acceptance by older adults." *Gerontechnology*, 2011.

[28] S. Kemper and J. C. Lacal, "Addressing the communication needs of an aging society," in *Technology for Adaptive Aging*. National Academies Press (US), 2004.

[29] N. Dahlbäck, A. Jönsson, and L. Ahrenberg, "Wizard of oz studies: why and how," in *Proceedings of the 1st International Conference on Intelligent User Interfaces*, 1993, pp. 193–200.

[30] A. Raghuvanshi, L. Carroll, and K. Raghunathan, "Developing production-level conversational interfaces with shallow semantic parsing," in *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing: System Demonstrations*, 2018, pp. 157–162.

[31] Apache. Apache TinkerPop Gremlin query language. [Online]. Available: https://tinkerpop.apache.org/gremlin.html

[32] L. Ferland and W. Koutstaal, "How's your day look? the (un)expected sociolinguistic effects of user modeling in a conversational agent," in *CHI EA '20: Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, 2020.

[33] Microsoft. Azure customer data protection. [Online]. Available: https://learn.microsoft.com/en-us/azure/security/fundamentals/protection-customer-data

[34] ——. Data management at Microsoft. [Online]. Available: https://www.microsoft.com/en-us/trust-center/privacy/data-management

[35] W. R. Boot, N. Charness, S. J. Czaja, J. Sharit, W. A. Rogers, A. D. Fisk, T. Mitzner, C. C. Lee, and S. Nair, "Computer proficiency questionnaire: assessing low and high computer proficient seniors," *The Gerontologist*, vol. 55, no. 3, pp. 404–411, 2015.

[36] A. Karnoe, D. Furstrand, K. B. Christensen, O. Norgaard, and L. Kayser, "Assessing competencies needed to engage with digital health services: development of the ehealth literacy assessment toolkit," *Journal of Medical Internet Research*, vol. 20, no. 5, p. e178, 2018.

[37] A. Poli, S. Kelfve, and A. Motel-Klingebiel, "A research tool for measuring non-participation of older people in research on digital health," *BMC Public Health*, vol. 19, no. 1, pp. 1–12, 2019.

[38] J. Brooke, "SUS: a "quick and dirty" usability scale," *Usability Evaluation in Industry*, vol. 189, no. 3, pp. 189–194, 1996.