Intelligent Agents in Everyday Settings: Leveraging a Multi-Methods' Approach

Krithika Jagannath
University of California, Irvine
Irvine, CA 92617, USA
kjaganna@uci.edu

Eugenia Ha Rim Rho
University of California, Irvine
Irvine, CA 92617, USA
hrrho@uci.edu

Abstract
Conversational Agents (CAs) or Intelligent Personal Assistants (IPAs) (e.g., Apple's *Siri*, Microsoft's *Cortana*; Amazon's *Alexa* and Google's *Google Assistant*) are voice-based interfaces designed for tasks in everyday life including: retrieval of information (e.g., weather, traffic, news), streaming of music, online shopping, controlling of home appliances, and voice-calls within the home and automobiles. Continuous enhancements of their natural language processing abilities, seamless set up of miniaturized hardware, and large-scale cloud-based infrastructures render CAs as unobtrusive, artificially intelligent voice sensors. With CAs rapidly making their way into the home market, the social implications remain unclear. Some product companies have released open-source software platforms that allow third-party developers and the general public to contribute software towards the growth of CAs. However, research around user-interaction with CAs in social settings is still at a nascent stage. In this workshop paper, we unpack the methods used in our ongoing work on people's social interactions with CAs in order to generate discussion around how the research community can leverage various methodologies using both qualitative and quantitative techniques.

Author Keywords
Conversational Agents; Intelligent Personal Assistants; Everyday Technologies; Voice Sensors; Intelligent Agents; Social Intelligence; Ethnomethodology; Conversation Analysis; Machine Learning.

ACM Classification Keywords
H.4.0. Information Systems Applications; H.4.m Miscellaneous; H.5.m Information Interfaces and presentation (e.g., HCI): Miscellaneous.

Introduction

Recent studies in HCI (e.g., [7], [23], [39]) underscore the need to further examine mainstream CAs in everyday settings. CAs such as Alexa and Google Assistant are being incorporated into several "smart" products (e.g., cars, watches, portable speakers) which many people use on a daily or regular basis in private as well as public settings. These devices are capable of listening in to conversations even when not being addressed. Consequently, there are significant ramifications for social norms, privacy, and human-machine trust [cite]. Current CA studies deploy interviews [25] and ethnomethodology [41] and offer considerations for designers of CAs to improve user experiences. However, user experiences are also being shaped by an emerging trend where users and third-party developers contribute towards CA development (e.g., Amazon's Alexa [1]and Google's DialogFlow[2]) - a phenomenon that is being largely overlooked at the moment. As [12] caution, "heteromated labor" occurs as product companies may benefit from users' contribution towards "Alexa skills" and improved dialogues while users themselves incur the labor of time and (intellectual) effort. We would also argue that there is value in investigating the approaches that designers and developers of CAs currently use. The growth and evolution of CA technologies can be better understood by investigating the work of system builders (Hughes in [22])). A better understanding about CA designers' ongoing practices may be key in enabling researchers to assertively offer advice and bridge any gaps in understanding.

[1] https://developer.amazon.com/alexa-skills-kit

[2] https://dialogflow.com/

Additionally, there is a growing interest towards agent applications within a number of other fields such as Internet of Things (IOT) and Ambient Intelligence (AmI). IoT refers to the superset of everyday objects that are embedded with sensors and/or connectivity-based technologies (e.g., wearables, credit cards, home automation) which can communicate with each other and with users (e.g., [5]). Similarly, AmI is an emerging interdisciplinary area which seeks to create intelligent and responsive user interfaces and embed sensors in various environments (e.g., [27] & [41]). For example, Lugano [21] informs that CAs like *Cortana* (Microsoft), *Google Assistant* (Google), and *Siri* (Apple) are already available in cars. He explains that CAs or voice-based virtual assistants (VAs) have begun to pique the interest of consumer product manufacturers and the automotive industry. Lugano suggests that researchers and practitioners along with stakeholders in the vehicle industry would need to prepare for scenarios in which CAs may soon become "*virtual companions*" (p. 4) for drivers and/or passengers in automobiles.

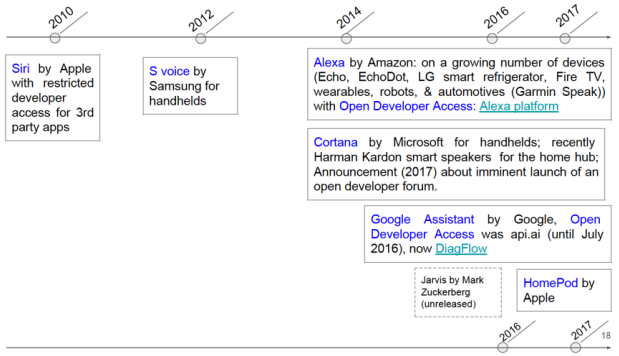

Figure 1: Timeline of Emerging Conversational Agents.

With CAs rapidly making their way into the home market (Fig. 1), the social implications due to CA use are unclear.

Research Questions of Interest

With CAs rapidly making their way into the home market (Fig. 1), the social implications due to CA use are unclear [36]. How should CAs function based on the setting (e.g., family, visitors, unrelated co-occupants) and social contexts (e.g., entertaining, private gatherings, home-office)? How do people perceive CAs in terms of agent intelligence and trust as CAs become more embedded into common objects in daily life?

In order to answer these questions, we need to understand what kinds of methods are most appropriate to tackle these questions holistically and effectively, given the nature of CAs as well as other factors such as the ways in which people interact with CAs (home versus public settings) as well as the availability of the CAs (e.g., infrastructure at home). In this workshop paper, we describe a current study in progress and break down the various methods we use in a multi-faceted approach to unpack the methodological implications of studying how people interact with CAs in everyday life. In the following sections, we first describe our current work in progress and unpack the various methods we employ. By doing so, we discuss the unique challenges and opportunities these methods pose and offer to better understand and improve how people interact with CAs through our ongoing work.

Current Work

In this section, we report on a three-phase research project in which we deploy multiple approaches to gather and analyze CA data - in-home deployment, ethnomethodology, and machine learning.

In-home Deployment

We first began our research through exploratory studies in which we analyzed a large corpora of tweets using a broad set of hashtags related to CAs (e.g., #Alexa, #Siri, #Cortana, #GoogleAssistant) between December 2016 and March 2017. We then interviewed several users (current and past) of CAs in their homes and also conducted several observational studies in our lab to learn how people interact with Alexa devices and the Alexa app. While initial findings from these exploratory studies provided useful preliminary insights, we decided to expand the scope of our user-observation through in-home deployment settings. HCI studies using in-home deployments can yield rich data over time by allowing researchers to observe how participants interact with technologies in the context of everyday life [15]. We recognize such importance of in-home deployment of CAs in our study given our goal of examining every-day interactions as opposed to conversations probed through lab settings with specific directions or tasks. The benefits of an in-situ approach at the participant's home allow for both use and non-use of the technology around certain contexts, which can elicit greater understanding behind the motivations of using CAs in everyday settings.

Conversation Analysis

Posited as an analytic method in HCI, conversation analysis involves analyzing audio transcripts or conversational logs without compromising on the "integrity" of the actual conversations ([37], p. 239). This method is useful for analyzing social talk in human-agent and/or human-human interactions (e.g.,

[3], [4], and [41]). According to [37], conversation analysis can aid the design of more sophisticated interactive agents as it enables researchers to discover the intricacies of talk, recognize user's patterns and detect shifts and problems during conversations. For instance, using conversational analysis, Aoki et al. [4] develop a vocabulary (e.g., primary and secondary participation, conversation floors, and participation sequence) to explain various mechanisms in simultaneous social talk. In our research, we use conversation analysis and qualitative techniques to analyze chat logs from users of CAs such as Alexa and Google Assistant as well, text corpora from communities on Reddit, as well as blog posts on Medium and CA developer websites[3].

Machine Learning

Researchers have used computational modeling and machine learning techniques to build agents that can personalize solutions better-tailored to individuals' preferences (e.g. [30]). Maes and Kozierok [26] argue that machine learning models built from such data can improve the capabilities of agents in a cost-effective and user-approved manner. In fact, conversation analysis and natural language processing techniques have been applied on large corpora of audio data to train models that can intelligently process technical aspects of speech. Researchers have leveraged such methods used by [1], [3], and [4] where conversation analysis can be applied to inform machine learning models to detect of overlaps [46], turn-taking, and floor assignment in conversations [3].

It is no surprise that CAs will produce a vast amount of text-based conversational data in the future. As researchers, we can certainly benefit from applying machine learning approaches to better understand and improve user-interaction with CAs. However, these approaches are not without challenges. For machine learning methods to work on agents, a number of expectations are being thrust on users. First, users must not only use the agent continuously, but also do so in newer ways for the models to learn new rules. Second, these methods also assume that users will be tolerant of errors as the agent learns new skills and willing to provide feedback to train the agent. Furthermore, another drawback is the effort and time required to train and calibrate models to achieve high levels of accuracy. As Nwana [38] cautions, the process of evaluating the most suitable learning technique(s) for a given set of desired outcomes involves arduous and time-consuming work. For example, Mitchell et al. [30] empirically determined a threshold (180 samples) for their model, and then gathered data from six participants' use over a time-frame of 16 months.

In our future work, we intend to use the corpora of data based on conversation analysis to extract temporal features using machine learning techniques. Our question of interest here remains: can CAs predict users' affective state or any signs of stress based on their tone, pitch, and words used? Going forth, we posit that findings from multiple approaches would enable us to answer such questions, revealing deeper insights into user experiences and design strategies for future work with CAs.

[3] While chat logs reveal insights about challenges that users experience and their typical current CA interactions, online content on Reddit forums and blog posts enable us to follow various communities of users and developers over time and raise pertinent questions around shifts that impact the adoption of CA technologies.

References

1. Mark S.Ackerman, Brian Starr, Debby Hindus, and Scott D. Mainwaring.1997. Hanging on the 'wire: a field study of an audio-only media space. ACM Transactions on Computer-Human Interaction (TOCHI) 4, no. 1: 39-66.
2. Mark Andrejevic. 2007. Surveillance in the digital enclosure. The Communication Review 10, no. 4: 295-317.
3. Paul M. Aoki, Matthew Romaine, Margaret H. Szymanski, James D. Thornton, Daniel Wilson, and Allison Woodruff. 2003. The mad hatter's cocktail party: a social mobile audio space supporting multiple simultaneous conversations. In Proceedings of the SIGCHI conference on human factors in computing systems, pp. 425-432. ACM.
4. Paul M.Aoki, Margaret H. Szymanski, Luke Plurkowski, James D. Thornton, Allison Woodruff, and Weilie Yi. 2006. Where's the party in multi-party?: Analyzing the structure of small-group sociable talk. In Proceedings of the 2006 20th anniversary conference on Computer supported cooperative work, pp. 393-402. ACM.
5. Luigi Atzori, Antonio Iera, and Giacomo Morabito. 2010. The internet of things: A survey. Computer networks 54, no. 15: 2787-2805
6. John M Carroll. 2000. Five reasons for scenario-based design. Interacting with computers 13, no. 1 (2000): 43-60.
7. Justine Cassell, Joseph Sullivan, Scott Prevost, and Elizabeth. F. Churchill. 2000. Embodied conversational agents. MIT Press.
8. Phil Cohen, Adam Cheyer, Eric Horvitz, Rana El Kaliouby, and Steve Whittaker. 2016. On the future of personal assistants. In Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems, pp. 1032-1037. ACM.
9. Ronald A Cole. 2000. Developing and evaluating conversational agents. Embodied conversational agents (2000): 287.
10. Harry Collins and Martin Kusch. 1999. The shape of actions: What humans and machines can do. MIT press.
11. Benjamin R. Cowan, Nadia Pantidi, David Coyle, Kellie Morrissey, Peter Clarke, Sara Al-Shehri, David Earley, and Natasha Bandeira. 2017. What can i help you with?: infrequent users' experiences of intelligent personal assistants. In Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services, p. 43. ACM.
12. Hamid R. Ekbia and Bonnie A. Nardi. 2017. Heteromation, and Other Stories of Computing and Capitalism. MIT Press.
13. R. Stuart Geiger and David Ribes. 2011. Trace ethnography: Following coordination through documentary practices. In System Sciences (HICSS), 2011 44th Hawaii International Conference on, pp. 1-10. IEEE.
14. Nicholas R. Jennings and Michael J. Wooldridge. 1998. Applications of intelligent agents: 3-28.
15. Tejinder K. Judge and Carman Neustaedter. 2014. Studying and Designing Technology for Domestic Life: Lessons from Home. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.)
16. James Everett Katz, ed. 2002. Machines that become us: The social context of personal communication technology. Transaction Publishers.
17. Jenny Kitzinger. 1995. Qualitative research. Introducing focus groups. BMJ: British medical journal 311, no. 7000: 299.
18. Kari Kuutti. 1996. Activity theory as a potential framework for human-computer interaction research. Context and consciousness: Activity theory and human-computer interaction 1744.

19. Bart P. Knijnenburg and Martijn C. Willemsen. 2016. Inferring capabilities of intelligent agents from their external traits. ACM Transactions on Interactive Intelligent Systems (TiiS) 6, no. 4: 28.
20. Tina Klüwer. 2011. I Like Your Shirt-Dialogue Acts for Enabling Social Talk in Conversational Agents. In Intelligent Virtual Agents, pp. 14-27. Springer Berlin/Heidelberg.
21. Kaushik Lakshmish, Abhijeet Sangwan, and John HL Hansen. 2015. Automatic audio sentiment extraction using keyword spotting. In Sixteenth Annual Conference of the International Speech Communication Association.
22. John Law. 1990. Introduction: monsters, machines and sociotechnical relations. The Sociological Review 38, no. S1: 1-23.
23. Giuseppe Lugano. 2017. Virtual assistants and self-driving cars. In ITS Telecommunications (ITST), 2017 15th International Conference on, pp. 1-5. IEEE.
24. Ewa Luger and Tom Rodden. 2013. An informed view on consent for UbiComp. In Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing, pp. 529-538. ACM.
25. Ewa Luger and Abigail Sellen. 2016. Like having a really bad PA: the gulf between user expectation and experience of conversational agents. In Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, pp.5286-5297. ACM.
26. Pattie Maes and Robyn Kozierok. 1993. Learning interface agents. In AAAI, vol. 93, pp. 459-465.
27. Spyros Makridakis. 2017. The Forthcoming Artificial Intelligence (AI) Revolution: Its Impact on Society and Firms. Futures.
28. Stephen Marsh and Pamela Briggs. 2009. Examining trust, forgiveness and regret as computational concepts. In Computing with social trust, pp. 9-43. Springer London.
29. Stefano Marzano. 2003. The new everyday: Views on ambient intelligence. 010 Publishers.
30. Tom M. Mitchell, Rich Caruana, Dayne Freitag, John McDermott, and David Zabowski. 1994. Experience with a learning personal assistant. Communications of the ACM 37, no. 7: 80-91.
31. Bonnie A. Nardi. 1996. ed. Context and consciousness: activity theory and human-computer interaction. MIT Press.
32. Clifford Nass, Jonathan Steuer, and Ellen R. Tauber. 1994. Computers are social actors. In Proceedings of the SIGCHI conference on Human factors in computing systems, pp. 72-78. ACM.
33. Clifford Nass and Li Gong. 2000. Speech interfaces from an evolutionary perspective. Communications of the ACM 43, no. 9: 36-43.
34. Carman Neustaedter, Kathryn Elliot, and Saul Greenberg. 2006. Interpersonal awareness in the domestic realm. In Proceedings of the 18th Australia conference on Computer-Human Interaction: Design: Activities, Artefacts and Environments, pp. 15-22. ACM.
35. Christena E. Nippert-Eng. 2008. Home and work: Negotiating boundaries through everyday life. University of Chicago Press.
36. Donald A Norman. 1994. How might people interact with agents. Communications of the ACM 37, no. 7 (1994): 68-71.
37. Michael A. Norman and Peter J. Thomas. 1991. Informing HCI design through conversation analysis. International journal of man-machine studies 35, no. 2: 235-250.
38. Hyacinth S Nwana. 1996. Software agents: An overview. The knowledge engineering review 11, no. 3: 205-244.

39. Gary M. Olson and Judith S. Olson. 2003. Human-computer interaction: Psychological aspects of the human use of computing. Annual review of psychology 54, no. 1: 491-516.
40. Judith S. Olson and Wendy A. Kellogg, eds. 2014. Ways of Knowing in HCI. Springer Science & Business.
41. Martin Porcheron, Joel E. Fischer, and Sarah Sharples. 2016. Do animals have accents?: talking with agents in multi-party Conversation.
42. Martin Porcheron, Joel E. Fischer, Moira McGregor, Barry Brown, Ewa Luger, Heloisa Candello, and Kenton O'Hara. Talking with conversational agents in collaborative action. In Companion of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing, pp. 431-436. ACM, 2017.
43. Yves Punie. 2005. The future of ambient intelligence in Europe: the need for more everyday life. Communications and Strategies 57, no. 1: 141-165.
44. Tom A. Rodden, Joel E. Fischer, Nadia Pantidi, Khaled Bachour, and Stuart Moran. 2013. At home with agents: exploring attitudes towards future smart energy infrastructures. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 1173-1182. ACM.
45. Yvonne Rogers. 2004. New theoretical approaches for HCI. Annual review of information science and technology 38, no. 1: 87-143.
46. Elizabeth Shriberg, Andreas Stolcke, and Don Baron. 2001. Observations on overlap: findings and implications for automatic processing of multi-party conversation. In INTERSPEECH, pp. 1359-1362.
47. Lucy Suchman. 1998. Human/machine reconsidered. Cognitive Studies 5, no. 1: 1_5-1_13.
48. Lucy A Suchman. 1987. Plans and situated actions: The problem of human-machine communication. Cambridge university press.
49. Deborah Tannen. 1984. Conversational style: Analyzing talk among friends. Oxford University Press.
50. Peter-Paul Verbeek. 2009. Ambient intelligence and persuasive technology: The blurring boundaries between human and technology. Nanoethics 3, no. 3: 231.
51. Judy Wajcman. 2014. Pressed for time. University of Chicago Press.