

---

# Conversational Agents: Acting on the Wave of Research and Development

**Richard Jacques**

Microsoft  
Redmond, WA, 98052, USA  
rjacques@microsoft.com

**Elizabeth Gerber**

Northwestern University  
Evanston, Illinois, 60208, USA  
egerber@northwestern.edu

**Ewa Luger**

University of Edinburgh  
Edinburgh, UK  
ewa.luger@ed.ac.uk

**Dakuo Wang**

IBM Research  
Yorktown Heights, NY, 10598, USA  
dakuo.wang@ibm.com

**Asbjørn Følstad**

SINTEF  
Oslo, Norway  
asbjorn.folstad@sintef.no

**Jonathan Grudin**

Microsoft and Univ. of Washington  
Redmond, WA, 98052, USA  
jgrudin@microsoft.com

**Andrés Monroy-Hernández**

Snap Inc. and Univ. of Washington  
Seattle, WA, 98121, USA  
amh@snap.com

**ABSTRACT**

In the last five years, work on software that interacts with people via typed or spoken natural language, called chatbots, intelligent assistants, social bots, virtual companions, non-human players, and so on, increased dramatically. Chatbots burst into prominence in 2016. Then came a wave of research, more development, and some use.

---

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

*CHI '19 Extended Abstracts, May 4–9, 2019, Glasgow, Scotland, UK.*

© 2019 Copyright is held by the author/owner(s).

ACM ISBN 978-1-4503-5971-9/19/05. <https://doi.org/10.1145/3290607.3299034>

**CCS CONCEPTS**

Human-centered computing: Ubiquitous and mobile computing: Ubiquitous and mobile devices: Personal digital assistants  
 Human-centered computing: Human computer interaction (HCI): Interaction paradigms: Natural language interfaces

**KEYWORDS**

Conversational agents; conversational user interfaces; intelligent assistants; chatbots; virtual companions; social bots.

**ACM Reference format:**

R. Jacques, A. Følstad, E. Gerber, J. Grudin, E. Luger, A. Monroy-Hernández, D. Wang, 2019. Conversational Agents: Acting on the Wave of Research and Development. In *CHI Conference on Human Factors in Computing Systems Extended Abstracts (CHI'19 Extended Abstracts)*, May 4–9, 2019, Glasgow, Scotland, UK. ACM, New York, NY, USA. 8 pages.  
<https://doi.org/10.1145/3290607.3299034>

The time is right to assess what we have learned from endeavouring to build conversational user interfaces that simulate quasi-human partners engaged in real conversations with real people. This workshop brings together people who developed or studied various conversational agents, to explore themes that include what works (and hasn't) in home, education, healthcare, and work settings, what we have learned from this about people and their activities, and social or ethical possibilities for good or risk.

**1 BACKGROUND**

Our goal is to organize and synthesize work on conversational user interfaces. The concept of an artificial entity capable of conversing in language shifted from fiction to scientific pursuit when Alan Turing [21] forecast machines that could converse like humans. It took a major step in 1964 with Joseph Weizenbaum's [23] ELIZA. Progress was slow for 30 years. ELIZA's 1995 successor ALICE inspired the 2013 film *Her*. New approaches to language understanding were enabled by low-cost memory, greater processing speed, and machine learning. Work on non-conversational software agents, such as recommender systems, picked up in the 1990s [16].

Intelligent assistants splashed on the scene in 2011 with Siri, followed by Cortana and Alexa in 2014 and Google Assistant in 2016. Extended dialogue gave way to quick exchanges. Rather than mimicking humans in long conversations, intelligent assistants are distinctly non-human and focus on short retrieval or notification tasks across a broad range of content. They also act on command, as when they play music or make phone calls.

Another type of conversational agent became a focus of frenzied commercial activity in 2016: task-focused chatbots. Many were for customer service [23] or conversational interfaces to frequently asked question lists [13]. Their conversations have more depth than those of intelligent assistants, but a narrower range.

Other ELIZA's successors, such as Cleverbot, progressed in the background until a new generation of 'virtual companions' appeared. Xiaoice [19] is a quirky hip chatbot with hundreds of millions of users and a family of siblings; Hugging Face specializes in recognizing the mood of a human conversational partner.

In 2016, Facebook and Microsoft launched chatbot platforms and report producing hundreds of thousands of task-focused chatbots. Slack launched an investment fund for chatbots. In 2017 Amazon released Lex, a chatbot tool that interfaces to Alexa. Gartner and others predicted that these friendlier-than-voice-answering interfaces would soon take over. For example, "By 2020, 80% of new enterprise applications will use Chatbots." [7]

By 2018, intelligent agents had appreciative users, but progressing beyond basic tasks was challenging [15]. Even the most successful conversational agents often failed to meet users' expectations [12]. Chatbots were not meeting the predictions of 2016; thoughtful articles asked questions such as "Chatbots: What happened?" [10] The difficulties of designing bots that fully met user and investor expectations slowed the momentum and generated interest in identifying and overcoming challenges.

Studies of user motivation indicated that although efficiency and easy access to services and data are the main attraction of chatbots, experiential and social aspects are also important [6, 15]. Chatbot developers must better understand user needs and experiences to meet expectations and obtain wide use [5].

Tasks that seem simple and natural fits for chatbots, such as buying airline tickets, making a dinner reservation, or scheduling a meeting proved surprisingly complex. This led to innovative approaches such as escalating difficult cases to human assistants [8], sometimes called ‘human-in-the-loop AI.’ A challenge to learning from others’ experiences is that efforts that do not meet expectations are not attractive candidates for publication, for authors or for reviewers who reject arguing “they should have known better.”

In summary, chatbots have moved along the Gartner hype cycle curve, having passed the Peak of Inflated Expectations and in some cases fallen into a Trough of Disillusionment. This workshop is focused on climbing the Slope of Enlightenment that can lead to the Plateau of Productivity. Some conversational agents are in use; one of our goals is to understand why some users’ expectations are met and some are not.

Work on conversational agents is of broad significance for user-centered design. It pushes the need to understand humans and their activities to the limit when it sets out to simulate a human that has meaningful conversations with one or more real human beings on topics that are relevant to them. Significant challenges in bot use have arisen. Social bots that pretend to be human are launched to interfere with political movements and elections [1]. Identifying fake Twitterbots became a significant undertaking [11]. Microsoft’s virtual companion Tay was lured into making offensive remarks. Google Duplex, which convinced people it was human by inserting “uh”s and “mhm”s, led to a healthy discussion of whether chatbots should reveal themselves to be non-human in all or some circumstances. An effort to explore the use of chatbots for social good was a topic of a CHI 2018 Special Interest Group organized by some of us. This workshop brings together what is emerging from these efforts.

Work published in peer-reviewed sources and thoughtful essays and media reports, much of it since 2015, needs to be digested, systematically organized, and integrated. A workshop cannot do it all. We propose to organize a few key facets of this space and make progress on setting out what is known and the implications, perhaps including design guidance for developers in some areas and in other areas identifying impacts of design decisions that can be conveyed to prospective developers.

## 2 THEMES

Our interests and those of the submitters will be used to identify a few significant areas in which to address challenges and consider opportunities for future research. Below are examples of possible themes, only a few of which can be explored in a one-day workshop.

- **Assistant/agent goals:** open-ended dialog (e.g. ELIZA), narrow task focus (customer service), broad quick-response (Siri): How do we design for efficiency and engagement in different contexts?
- **Assistant/agent capabilities:** How can users learn about limitations, capabilities, upgrades?
- **Usage scenarios:** How do design assumptions shift when an agent operates behind organizational firewalls or supports patients and doctors in hospitals?
- **Usage and engagement:** How do we measure success when use can be regular or occasional, short-duration or long?
- **User perceptions:** How are assistants understood and experienced by users?
- **User-chatbot behavior:** How is conversational behavior affected when talking to machines?
- **Skills:** When should designers add a skill to an existing agent or create a new bot?
- **Context:** How much context—conversation history or software application—can and should bots make use of? What issues arise when crossing boundaries, for example between home and work?
- **Platforms:** Standalone apps, embodied in boxes or robots, or software platforms (SMS, email, Messenger...) What is effective for what purposes?
- **Modalities:** When should agents support voice, text, images, video, menus as input or output?
- **Personality:** When is personality beneficial? How should gender be selected?
- **Ethics and trust:** How do we design responsible experiences? Manage bias? Should agents reflect societal norms or embody aspirational goals?
- **Bots in social networks:** What changes when chatbots join networks of humans and bots?
- **Humans as partners:** When should agents escalate requests to humans?
- **Training and assessment:** How should training data be collected and curated? How can agents best be tested prior to general release?

### 3 ORGANIZERS

*Richard Jacques* is a Principal UX Researcher in Microsoft's AI & Research organization. He works on Search experiences and how AI is influencing the way people retrieve and use information from search engines, agents and assistants. He is particularly interested in the opportunities and challenges this brings for interaction designers.

*Asbjørn Følstad* is a Senior Researcher at SINTEF. His background is in psychology and human-computer interaction. Particular research interests include chatbot user experience and chatbot interaction design, and he leads two Norwegian chatbot research projects. He has co-organized the Chatbots for Social Good SIG at CHI 2018 and the CONVERSATIONS workshops on chatbot research at INSCI 2017 and 2018.

**Elizabeth Gerber** is Charles Deering McCormick Professor of Teaching Excellence in the Mechanical Engineering and Communication Studies Departments at Northwestern University where she directs the Design Research Cluster. Formally trained in Product Design and Management Science, she researches the future of work. Specifically, she designs and studies open, ubiquitous, sociotechnical systems to support collaborative problem solving. She teaches user centered service design which increasingly involves understanding the role of chatbots. She also participated in the 2018 Dagstuhl workshop.

**Jonathan Grudin** is a Principal Researcher at Microsoft. He has worked on chatbot development projects, co-authored an *Interactions* cover article on human-computer integration [9] that included conversational agent examples, co-organized a panel and a SIG on AI agents at CHI 2017, and participated in a Dagstuhl workshop on human-computer integration in 2018.

**Ewa Luger** is a Chancellor's Fellow in Digital Arts and Humanities (Intelligible data-driven systems) and a Fellow of the Alan Turing Institute (Ethical AI by Design: Formalising an HCI agenda). Her work explores applied ethical issues within the sphere of machine intelligence and data-driven systems, including conversational agents. This encompasses practical considerations such as data governance, consent, privacy, explainable AI, and how intelligent networked systems might be made intelligible to the user, with a particular interest in distribution of power and spheres of digital exclusion. Recent cognate research includes investigation of user experiences of task-based conversational agents [2], exploration of human-agent collaboration in Minecraft [3, 4] and the privacy implications arising from conversation as platform [14].

**Andrés Monroy-Hernández** is Lead Research Scientist at Snap Inc., and an affiliate professor at the University of Washington. He led the incubation of Microsoft's Calendar.help: a bot that helps people schedule meetings over email [8]. He has also co-authored several articles about bots, including one on a Twitter chatbot for supporting activists [17] and a study of chatbot-mediated task management for teams [20].

**Dakuo Wang** is a Research Scientist at IBM T.J. Watson Research Center. He is particularly interested in human-agent collaboration. He has worked on the design and development of conversational agents including an embodied intelligent assistant [18]. He has helped organize or participated actively in a CSCW2017 workshop on robots in groups and teams, a CHI 2018 panel on human-agent collaboration, and the 2018 Dagstuhl human-computer integration workshop.

### 3 WEBSITE

The website <https://convagents.org/> is being used to promote the workshop and could become a public repository for position papers and reports.

## 4.2 Workshop Activities

<i>Time</i>	<i>Activity</i>
8:30-9:00	Arrival, poster set up and viewing
9:00-9:15	Welcome and introduction to the goals, schedule
9:15-10.00	Opening discussion of themes. Participant introductions
10.00-10.45	Pre-determined groups form, adjust, start discussion
10.45-11.00	Coffee break and poster viewing
11.00-11.45	Groups resume discussion
11.45-12.30	Plenary: report out by groups, discuss afternoon
12.30-13.45	Lunch
13.45-15.15	Group or new group discussions
15.15-15.30	Coffee break and poster viewing
15.30-16.45	Plenary: report out and look for syntheses
16.45-17.00	Future plans and wrap up

## 4 WORKSHOP PLANS

### 4.1 Pre-Workshop

Participants will be recruited from those who have worked or published on this topic and those who participated in previous relevant events, some of which are listed below. We anticipate a balanced mix of participants from academia and industry and from work in different sub-genres. Our call identifies thematic interests of the organizers, but the workshop content will be shaped by the position papers and study results that are submitted. Submissions will be clustered into tentative themes to form the basis of breakout groups. Position papers will be shared in advance and not presented at the workshop.

The final 90 minutes will focus on conclusions, possible syntheses or divergences as we consider what a comprehensive view of conversational user interfaces might look like, and assess what would be helpful moving forward, both for the workshop output and for the field.

### 4.3 Post-Workshop

This workshop follows a handful of CHI SIGs, panels, courses, and paper sessions on related topics, as well as similar discussions in related fields. The workshop will pull some of this work together and get the word out. The timing is appropriate as other events are being lined up to take on some of these issues. Interactions, Communication of the ACM, a special issue of a journal, a guideline for practitioners, or a CHI 2020 paper are possible products we could pursue.

## 5 CALL FOR PARTICIPATION

Position paper submission date: Feb 12, 2019

Submission length: 1500 – 2500 words

Email submissions to [chi19convai@outlook.com](mailto:chi19convai@outlook.com)

Chatbots, conversational agents, intelligent assistants, AI systems—many names are applied to a range of conversational user interfaces. This intensive one-day workshop will bring together researchers and developers to organize and start synthesizing the flood of work that has appeared in the last few years. Some agents like Siri have a broad focus, some have a narrow task focus, and others want to engage at length in entertaining discussions. All represent a supremely user-centred effort: to create simulated humans that can converse with real humans about real activities. We can accelerate progress by sharing results and experiences across these endeavours.

The workshop will select a manageable set of themes based on submissions. Examples could include design guidelines, user experiences, clear success cases, lessons learned from unsuccessful projects, use in enterprise settings, and ethical issues.

The submission should include a description of its contribution to the workshop topic and identify one or more thematic interests of the author(s). Quality and fit will determine acceptance and contribute to selecting themes for the workshop to focus on.

CHI requires that one author register for the workshop and at least one day of CHI 2019. Depending on the submission numbers, possibly only one participant per accepted submission can attend.

More information about the workshop is at <https://convagents.org/>

## 6 WORKSHOP AND SPECIAL INTEREST GROUPS

The organizers have been involved in the past workshops and special interest groups:

CSCW 2017 workshop on conversational agents in collaborative action.

CHI 2017 panel on human-computer integration.

CHI 2017 panel on human-agent collaboration.

CHI 2017 SIG on human-computer integration.

CHI 2018 SIG on chatbots for social good.

Conversations 2018 workshop on chatbot research and design.

2018 Dagstuhl workshop on human-computer integration.

## REFERENCES

- [1] Abokhodair, N., Yoo, D. & McDonald, D.W. (2015). Dissecting a social botnet: Growth, content, and influence in Twitter. Proc. CSCW 2015, 839–851.
- [2] Allison, F., Luger, E., Hofmann, K (2018) How players speak to an intelligent game character using natural language messages. Proc. DiGRA Special Edition.
- [3] Allison, F., Luger, E., & Hoffman, K. (2017) Player expectations of a learning AI companion in Minecraft. In Proc. DiGRA.
- [4] Allison, F., Luger, E., Hoffman, K. (2017) Spontaneous interactions with a virtually embodied intelligent assistant in Minecraft. CHI'17 Extended Abstracts.
- [5] Brandtzaeg, P. B., & Følstad, A. (2017). Why people use chatbots. Proc. International Conference on Internet Science, 377-392.
- [6] Brandtzaeg, P. B., & Følstad, A. (2018). Chatbots: changing user needs and motivations. Interactions, 25, 5, 38-43.
- [7] Chatbotinsider (2016). Gartner predicts enterprise Chatbots by 2020. Nov 4. <http://www.chatbotinsider.ai/industry-news/gartner-predicts-enterprise-chatbots-2020/>
- [8] Cranshaw, J., Elwany, E., Newman, T., Kocielnik, R., Yu, R., Soni, S., Teevan, J. & Monroy-Hernandez, A. 2017. Calendar.help: Designing a workflow-based scheduling agent with humans in the loop. Proc. CHI 2017, 2382-2393. <https://dl.acm.org/citation.cfm?id=3025780>
- [9] Farooq, U. & Grudin, J. (2016) Human-computer integration. ACM Interactions, 23, 6, 26-32.
- [10] Feldman, D. (2018). Chatbots: What happened? Chatbots life, April 10. <https://chatbotslife.com/chatbots-what-happened-dcc3f91a512c>
- [11] Ferrara, E., Varol, O., Davis, C., Menczer, F. & Flammini, A. (2016). The Rise of Social Bots. Communications of the ACM, 96-104. <https://arxiv.org/abs/1407.5225>

- [12] Jain, M., Kumar, P., Kota, R. & Patel, S.N. (2018). Evaluating and Informing the Design of Chatbots. Proc. DIS 2018, 895-906. <https://dl.acm.org/citation.cfm?id=3196735>
- [13] Liao, Q.V., Hussain, M.M., Chandar, P., Davis, M., Khazaen, Y., Crasso, M.P., Wang, D., Muller, M., Shami, N.S. & Geyer, W. (2018). All work and no play? Conversations with a question-and-answer chatbot in the wild. Proc. CHI 2018, paper 3.
- [14] Luger, E. & Rosner, G. (2017) Considering the privacy design issues arising from conversation as platform. In Leenes, R., van Brakel, R., Gutwirth, S., De Hert, P. (Eds.), Data protection and privacy: The age of intelligent machines. Computers, Privacy and Data Protection (10), Hart Publishing.
- [15] Luger, E. & Sellen, A. "Like having a really bad PA": The gulf between user expectation and experience of conversational agents. Proc. CHI 2016, 5286-5297
- [16] Maes, P. 1994. Agents that reduce work and information overload. Comm. ACM, 37, 7, 30-40. <https://dl.acm.org/citation.cfm?doid=176789.176792>
- [17] Savage, S., Monroy-Hernandez, A. & Höllerer, T. 2016. Botivist: Calling volunteers to action using online bots. Proc.CSCW 16, 813-822. DOI: <https://doi.org/10.1145/2818048.2819985>
- [18] Shamekhi, A., Liao, Q.V., Wang, D., Bellamy, R.K. & Erickson, T. 2018. Face value? Exploring the effects of embodiment for a group facilitation agent. Proc. CHI 2018, paper 391, 13 pages. DOI: <https://doi.org/10.1145/3173574.3173965>
- [19] Shum, H.-Y., He, X., & Li, D. (2018). From Eliza to Xiaolce: challenges and opportunities with social chatbots. Journal of Zhejiang University Science C, 19, 1, 10-26. <https://arxiv.org/abs/1801.01957>
- [20] Toxtli, C., Monroy-Hernández, A. & Cranshaw, J. 2018. Understanding chatbot-mediated task management. Proc. CHI 2018, Paper 58. DOI: <https://doi.org/10.1145/3173574.3173632>
- [21] Turing, A. 1949. London Times letter to the editor, June 11.
- [22] Xu, A., Liu, Z., Guo, Y., Sinha, V., & Akkiraju, R. (2017). A new chatbot for customer service on social media. Proc. CHI 2017, 3506-3510.
- [23] Weizenbaum, J. (1966.) ELIZA — A computer program for the study of natural language communication between man and machine," Comm. ACM, 9, 1, 36-45.