Conversation Starters

How Can We Misunderstand AI Better?

Nicenboim, Iohanna; Venkat, Shruthi; Rustad, Neva Linn; Vardanyan, Diana; Giaccardi, Elisa; Redström, Johan

DOI
10.1145/3544549.3583914

Publication date
2023

Document Version
Final published version

Published in
CHI 2023 - Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems

Citation (APA)
Nicenboim, I., Venkat, S., Rustad, N. L., Vardanyan, D., Giaccardi, E., & Redström, J. (2023). Conversation Starters: How Can We Misunderstand AI Better? In *CHI 2023 - Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems* Article 431 (Conference on Human Factors in Computing Systems - Proceedings). ACM. https://doi.org/10.1145/3544549.3583914

Conversation Starters: How Can We Misunderstand AI Better?

Iohanna Nicenboim
Delft University of Technology
I.nicenboim@tudelft.nl

Shruthi Venkat
Delft University of Technology
S.Venkat@student.tudelft.nl

Neva Rustad
Delft University of Technology
N.L.Rustad@student.tudelft.nl

Diana Vardanyan
Delft University of Technology
D.E.Vardanyan@student.tudelft.nl

Elisa Giaccardi
Delft University of Technology
E.Giaccardi@tudelft.nl

Johan Redström
Umeå Institute of Design
johan.redstrom@umu.se

ABSTRACT

Conversation Starters is a series of interactive prototypes that probe how to design explainable interactions with AI in everyday life. Taking a more-than-human approach, we explore how 'failures' could be transformed into opportunities for situated understandings of AI. We describe the process of designing fictional artifacts and scenarios about conversational agents that can grow at home. While overall the project suggests that misunderstandings could help people develop sensitivities for knowing when to trust AI systems, the metaphor of 'growing an AI' (which positions training as a matter of care), highlights that practices of sharing and experimenting could be valuable starting points for designing explainable and trustworthy interactions with of AI.

CCS CONCEPTS

• **Human-centered computing** → Human computer interaction (HCI); HCI theory, concepts and models.

KEYWORDS

More-than-human Design, Conversational Agents, Conversational User Interfaces, Artificial Intelligence, Explainability of AI

ACM Reference Format:
Iohanna Nicenboim, Shruthi Venkat, Neva Rustad, Diana Vardanyan, Elisa Giaccardi, and Johan Redström. 2023. Conversation Starters: How Can We Misunderstand AI Better?. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems (CHI EA '23), April 23–28, 2023, Hamburg, Germany*. ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3544549.3583914

1 INTRODUCTION

"Hello, I am Starter, a growing conversational AI"

That is what visitors of the exhibit hear when they open the interactive jars of Conversation Starters, a series of prototypes that explore how to design explainable interactions with Artificial Intelligence (AI) (Figure 1).

The aim of this project is to address a key challenge for designers working with AI, which is to make interactions explainable [2]. While there are significant efforts within the HCI community to develop explainable AI systems, how to support people in understanding AI agents through interactions remains a challenge [7]. Designers need to find new ways to support that process, especially because technical explanations are not effective in everyday life where people need situated understandings that can be adapted over time [4, 12, 14]. In this project, we take the case of conversational agents to explore how designers could support situated and contextual understandings of AI during interactions. We focus on misunderstandings as opportunities to support Explainability by helping people understand the system's limitations in a contextual way [1, 7].

Digital assistants such as Alexa or Siri can disclose the origin of their responses, but they frequently fall short in providing an explanation for not responding. For example, both instances of not understanding a query and not being able to find an answer result in the same reply: "Sorry, I don't know that." With that in mind, we ask: How can misunderstandings be designed to support situated understandings of AI? To explore that, we engaged in a design process in which we got inspired by the practices of growing organisms at home (fermentation, brewing, and bread-making) and used the concept of the 'starter' as a metaphor for designing fictional scenarios and interactive prototypes. Through the prototypes, we could reflect on how living with a simple organism, that cannot be fully understood, is in some ways like living with an intelligent agent that cannot be explained. By recognizing the differences between those two agents, we point to new opportunities for designing misunderstandings to support AI Explainability.

2 BACKGROUND

The domain of conversational AI has evolved and taken many forms over the years, such as chatbots and digital assistants. But more broadly, the field has made significant progress in recent years with the development of sophisticated natural language processing techniques and large language models like ChatGPT3. Despite that, conversational agents still fall short because they are designed with narrow purposes and particular contexts in mind, which leads to their inability to maintain complex conversations. Moreover, these technologies have been critiqued because they endanger human autonomy [17] and amplify racial and gender biases through their interactions [19].

To address those challenges, it is important that AI systems are explainable. However, Explainability of AI is a complex challenge. While technical explanations might be effective for experts to interpret and account for the outcomes of AI systems, they are not useful for people interacting with AI agents in everyday life, because they

CHI EA '23, April 23–28, 2023, Hamburg, Germany

ACM ISBN 978-1-4503-9422-2/23/04.
https://doi.org/10.1145/3544549.3583914

Figure 1: Conversation Starters is a series of prototypes that explore how to design explainable interactions with AI. By opening the lids, visitors can listen to the fictional agent and smell the organisms growing in the jars.

need to be better situated [4, 12, 14]. While interactive approaches have emerged in Explainable AI [20] not much attention has been paid to conversational approaches for explanations [10] nor to the active role that users and agents have in those processes [14].

To address that gap and discover new opportunities for Explainability in everyday situations, we engaged in a design process. Methodologically we build on previous experiences that have pointed at the value of design fiction [6, 18] and more-than-human design [9, 21] to imagine alternative interactions with technology. More-than-human design supported this project not only methodologically, as we used the methods of Thing ethnography [8], Conversations with agents [15], and Noticing [3, 16], but also conceptually, as we took a posthumanist approach to Explainability of AI, acknowledging the positionality and agency of both people and AI agents in constructing situated understandings [14]. Although our prototypes draw from Biodesign aesthetically, that is not the main area that this work contributes to. Instead, this work aims to contribute to the communities of scholars and designers working with AI, especially to the ones interested in Human-Centered Explainability of AI (HCXAI) and more-than-human design approaches.

3 DESIGNING CONVERSATION STARTERS

The process of designing Conversation Starters included several activities, i.e., doing ethnographic research, designing scenarios, and making interactive prototypes. The design team was composed of a PhD candidate and three master's students from the "Design for Interactions" program at Delft University of Technology. The prototyping phase was done in collaboration with the design studio Cream on Chrome and the university maker's lab Studio Make.

In the first phase, the students lived with a digital assistant for a week (Alexa and Google Home) and conducted different ethnographic research.

One student chose to focus on bread-making practices using Thing Ethnography, another noticed sounds in a local park with the method of Noticing.

In the second phase, we did two exercises to conceptualize the design fiction, using the methods of world-building [5] and new metaphors [11, 13]. The world-building activity helped us to align on a common goal and visualize possible directions for the scenarios. The exercise on metaphors helped us to find concrete directions for the design of conversational agents that grow at home. The last activity also helped us refine the idea of 'growing things', as we tried different metaphors (AI is a spider web, a kid, and a starter). We chose the metaphor of the starter because it highlighted a relationship of care between the human and the AI agent. It also showed that understandings could be related to the environment and history of the agent.

Taking the starter as a metaphor, we designed three fictional scenarios and developed their prototypes and props (Figure 2). For every scenario, we defined a human and a non-human character, a goal, something that went wrong, and a way for the characters to deal with it. The first scenario was about a parent and a son growing a conversational AI to tell bedtime stories with the sounds of the home including the non-human co-habitants; the second was about a busy woman who bought a starter to grow a cloud in a jar to predict the weather more accurately, the third was about a roommate in a shared flat who had found a starter to grow an AI in her fridge and used it to make shopping lists for her plant-caring diet. The scenarios were explored in short videos, which were spoken in different languages, Spanish, English, and Dutch.

In the first scenario, there was a misunderstanding represented as the parent discussing with Starter the appropriateness of a story the agent told the child. The story took the perspective of an animal in the forest and described how humans destroyed its home. The agent, programmed not to tell lies, questioned why it had to change the perspective of the story. In the second scenario, the agent responded

to a simple question such as "What is the weather?" by referring to different temporalities and describing the Anthropocene. In the third scenario, Starter ordered food for the plants in the home because it had misunderstood the Dutch word for plant-based diet (the literal translation of plantvriendelijk to English is plant-friendly).

The electronics were designed to simulate the interaction with a conversational agent with sound and light (Figure 3). When people open the lid, the light and sound play. That encourages people to get closer to the jars to listen to the conversation, which in turn invites them to have a closer look at the living-like creature and sense the smell, which makes the experience stronger. People are encouraged to open the lids by mimicking other visitors and by simple messages written on the table/tiles that state "Open the lid." At the same time, visitors have full control over stopping the interaction by closing the lid. In the big jar the fictional agent Starter describes the project and the scenario in a conversational tone. It also invites people to open the lid of the other three smaller jars, which play bedtime stories, related to the three contexts where the jars were positioned in the fiction. The stories are also presented at the exhibition in a digital book.

Figure 2: The first iteration of the prototypes was designed for short movies to depict and explore the fictional scenarios. In that first iteration, we controlled the light manually to simulate the interaction with a conversational agent.

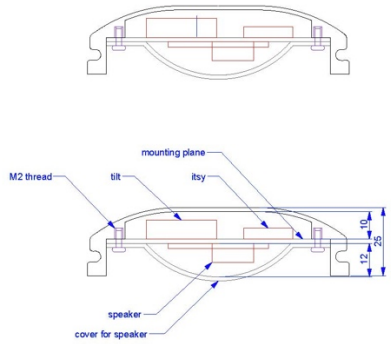

Figure 3: The design of the living-like organisms was done in wax and silicone and inserted in gelatin mixed with self-made kombucha. The jars contain a ring of LEDs at the bottom and a speaker and tilt sensor inside the lid, all connected to an itsy bitsy board. When the tilt sensor communicates to the board that the lid is open, the board activates the light and plays the recording. The jar lids were 3D printed to house all the electronics.

4 DISCUSSION

To address the question "How can we misunderstand AI better?" we unpack two opportunities for situated (mis)understandings that were opened by the project.

The first opportunity is related to the metaphor of 'growing an AI.' Using that metaphor and looking particularly at practices of fermentation and bread-making, allowed us to imagine alternative interactions with current conversational agents, especially in relation to how Explainability is approached. In Conversational Starters, users were positioned as active participants in understanding the capabilities and limitations of the system, and those interactions were based on relationships of cultivation and care. The practices of fermentation and bread-making, in which people often get a starter from someone they trust, highlighted that also sharing and experimenting could be interesting directions to design explainable human-AI interactions. On the other hand, the agents were also positioned as active participants. Designing the agents as changing things, allowed us to speculate for example, that the fine-tuning of a language model could happen partly at home as a daily practice, or to imagine that a conversational agent could be taken out for a walk or left hidden in the forest to be exposed to new sounds. The second opportunity created by our project was in how the misunderstandings were designed in the scenarios, as moments for people to reflect on the role of these agents in their lives. In this case, we wanted people to reflect on the agency of Starter as a non-human and its environment as a more-than-human ecosystem.

Overall, the project is an example of how Explainability of AI could be addressed through design, and how misunderstandings could be designed as provocations to expose the limitations of AI and make people reflect on the implications that those limitations have for their lives.

5 CONCLUSION

While technical explanations of AI systems are essential for accountability, in everyday life, people develop their own understanding of when and how to trust the products that are part of those systems. This interactive demo explores how designers could support situated understanding(s) of AI in everyday life. Specifically, the project illustrates how everyday misunderstandings of AI could allow people to develop sensitivities for grasping agents' capabilities and limitations. We hope that the interactive prototypes can promote reflection on how misunderstandings might be designed contextually, to ultimately inspire designers in HCI to find ways to support situated understandings of AI.

ACKNOWLEDGMENTS

We want to thank Aadjan van der Helm, Martin Havranek, and Adriaan Bernstein from Studio Make for their assistance with the

electronics; Cream on Chrome studio for their contribution to the design of the final prototypes, and for providing images of the process; Marica de Michele for creating the final video; to Marco Rozendaal, Dave Murray-Rust and Marijke Idema for their support in showcasing our prototypes at Dutch Design Week; and Twycer for allowing us to use photographs of the exhibition. We also thank Aparna Pallod for her input in the ideation phase. This project was developed as part of a PhD funded by a Microsoft Research Scholarship.

REFERENCES

[1] Saleema Amershi. 2020. Toward Responsible AI by Planning to Fail. Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, Association for Computing Machinery, 3607.

[2] Saleema Amershi, Kori Inkpen, Jaime Teevan, *et al.* 2019. Guidelines for Human-AI Interaction. Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems - CHI '19, ACM Press, 1–13.

[3] Heidi R. Biggs, Jeffrey Bardzell, and Shaowen Bardzell. 2021. Watching Myself Watching Birds: Abjection, Ecological Thinking, and Posthuman Design. In Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems. Association for Computing Machinery, New York, NY, USA, 1–16.

[4] Hans de Bruijn, Martijn Warnier, and Marijn Janssen. 2022. The perils and pitfalls of explainable AI: Strategies for explaining algorithmic decision-making. Government information quarterly 39, 2: 101666.

[5] Paul Coulton, Joseph Lindley, Miriam Sturdee, and Mike Stead. 2017. Design Fiction as World Building. Research through Design, unknown.

[6] Audrey Desjardins, Afroditi Psarra, and Bonnie A. Whiting. 2021. Voices and Voids: Subverting Voice Assistant Systems through Performative Experiments. In Creativity and Cognition. Association for Computing Machinery, New York, NY, USA, 1–10.

[7] Upol Ehsan, Philipp Wintersberger, Q. Vera Liao, *et al.* 2022. Human-Centered Explainable AI (HCXAI): Beyond Opening the Black-Box of AI. Extended Abstracts of the 2022 CHI Conference on Human Factors in Computing Systems, Association for Computing Machinery, 1–7.

[8] Elisa Giaccardi, Nazli Cila, Chris Speed, and Melissa Caldwell. 2016. Thing Ethnography: Doing Design Research with Non-Humans. Proceedings of the 2016 ACM Conference on Designing Interactive Systems, ACM, 377–387.

[9] Elisa Giaccardi and Johan Redström. 2020. Technology and more-than-human design. Design Issues 36, 4: 33–44.

[10] Diana C. Hernandez-Bocanegra and Jürgen Ziegler. 2021. Conversational review-based explanations for recommender systems: Exploring users' query behavior. Proceedings of the 3rd Conference on Conversational User Interfaces, Association for Computing Machinery, 1–11.

[11] Dan Lockton, Devika Singh, Saloni Sabnis, Michelle Chou, Sarah Foley, and Alejandro Pantoja. 2019. New Metaphors: A Workshop Method for Generating Ideas and Reframing Problems in Design and Beyond. Proceedings of the 2019 on Creativity and Cognition, Association for Computing Machinery, 319–332.

[12] Tim Miller. 2019. Explanation in artificial intelligence: Insights from the social sciences. Artificial intelligence 267: 1–38.

[13] Dave Murray-Rust, Iohanna Nicenboim, and Dan Lockton. 2022. Metaphors for designers working with AI. DRS Biennial Conference Series.

[14] Iohanna Nicenboim, Elisa Giaccardi, and Johan Redström. 2022. From explanations to shared understandings of AI. DRS Biennial Conference Series.

[15] Iohanna Nicenboim, Elisa Giaccardi, Marie Louise Juul Søndergaard, *et al.* 2020. More-Than-Human Design and AI: In Conversation with Agents. Companion Publication of the 2020 ACM Designing Interactive Systems Conference, Association for Computing Machinery, 397–400.

[16] Doenja Oogjes and Ron Wakkary. 2022. Weaving Stories: Toward Repertoires for Designing Things. Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems, Association for Computing Machinery, 1–21.

[17] Supraja Sankaran, Chao Zhang, Mathias Funk, Henk Aarts, and Panos Markopoulos. 2020. Do I have a say? Using conversational agents to re-imagine human-machine autonomy. Proceedings of the 2nd Conference on Conversational User Interfaces, Association for Computing Machinery, 1–3.

[18] Marie Louise Juul Søndergaard and Lone Koefoed Hansen. 2018. Intimate Futures: Staying with the Trouble of Digital Personal Assistants through Design Fiction. DIS '18 Proceedings of the 2018 Designing Interactive Systems Conference, ACM, 869–880.

[19] Yolande Strengers and Jenny Kennedy. 2020. The Smart Wife: Why Siri, Alexa, and Other Smart Home Devices Need a Feminist Reboot. The MIT Press.

[20] Q. Vera Liao and Kush R. Varshney. 2021. Human-Centered Explainable AI (XAI): From Algorithms to User Experiences. arXiv [cs.AI]. Retrieved from http://arxiv.org/abs/2110.10790.

[21] Ron Wakkary. 2021. Things We Could Design: For More Than Human-Centered Worlds (Design Thinking, Design Theory). The MIT Press.