Using AI to Improve Product Teams' Customer Empathy

Valentina Grigoreanu, Monty Hammontree and Travis Lowdermilk

Microsoft Corporation, Redmond, WA 98053, USA
valeng@microsoft.com

Abstract

During customer conversations, it is important to know both *what* questions to ask at any point during the development cycle, and *how* to ask them. Asking the right questions to capture rich, accurate, and relevant customer feedback is not easy, and professionally-trained researchers cannot be a part of every customer conversation. To scale out researchers' knowledge, we built an artificial intelligence system, the VIVID whisper-bot, trained on three theories: the Hypothesis Progression Framework (contextual research questions for each product development phase), the VIVID grammar framework (asking who, what, why, how, where, how much, and when type questions to recreate rich stories), and the syntactical structure of biased and leading questions. The whisper-bot listens in on a customer conversation, highlights customers' key verbalization (e.g., pain points using the product), and suggests follow-up interview questions (e.g., removing bias or enriching a story). It thereby encourages good interview practices for everyone, which we believe will increase empathy on product development teams, and lead to improvements in the products' user experience.

Keywords

AI/ML · HCI · Cognitive Services · Design/Customer Research · User Experience · Empathy · Product Development · Software Development

1 Introduction and Background

One area at the intersection between AI and HCI is how artificial intelligence systems can help improve HCI research (or user research, customer experience research, usability engineering, design research, etc.). We will refer to this space as "AI for UX Research. It aims to explore: *How can the research skills of anyone doing customer, product, or business development be augmented through artificial intelligence systems?*

There are few publications on "AI for UX Research" to-date (e.g., [1, 3]). These examples applied AI to help analyze customer research data. However, at the time of this paper, we found no previous research on how AI can be used at the *data collection* stage, particularly for qualitative data. Our VIVID whisper-bot solution scales out research interview skills to *anyone* wanting to conduct customer interviews, no matter their level of UX research training. This is of interest to our Research team in the Developer Division at Microsoft, where customer conversations are happening on a larger scale than ever before. Our corporate vice-president credits our data-driven and customer-obsessed culture for a major increase from 2 million to 14 million active users in a handful of years [2]. With an estimate of more than 10,000 customer interviews conducted every year by our product teams, researchers split their time between conducting research and training product team members to have better customer conversations (e.g., through a distribution list, workshops, bootcamps, and a book).

We also took an AI approach to scale out UX research knowledge, by teaching a bot to "whisper" suggestions to product team members about what questions to ask during an interview and how to best ask them.

2 Results: Reasons for the Empathy Gap

Based on data collected during a survey and two focus groups, four themes emerged for improving our product teams' interviews:

1. Taking notes during customer conversations is challenging.
2. Patterns of leading, biased, and closed questioning during interviews are common.
3. It is hard to identify opportunities to probe for deeper insights.
4. It is difficult to share empathy post-interview, to get the organizations' attention.

We wondered whether we could design an artificial intelligence system that would achieve the goals above when a Researcher cannot be present during a customer conversation. The idea of the VIVID whisper-bot was born.

3 Working Prototype: The VIVID Whisper-bot

We hypothesized that we could teach an AI system the rules of asking rich, relevant, and unbiased questions – and that this would solve many of the gaps mentioned in the previous section.

1. Real-time speech-to-text transcription: Our first step was to build a component that transcribes speech-to-text, as accurately as possible. This provides the interviewer conversation notes and is also the venue for showing the agent's feedback real-time within the context of the conversation being held.

2. Trained LUIS (natural language processing) models to identify biased and closed questions: Training the models so that we could accurately identify closed and leading interview questions was mostly based on sentence structure. Most closed questions started with commonly used auxiliary verbs (such as, "do", "can", "would"). For leading questions, the adjectives and adverbs mattered, and emotions were more prevalent in the questions. Another tool for feedback about the quality of the interview was the *Conversation Mix*: an indicator that tracks how much the interviewer was talking in relation to the customer. Depending on the goal of the interview, this could be a useful reminder to the interviewer to leave adequate time for the customer to express their thoughts.

3. Trained models to identify opportunities to probe deeper: Two frameworks were core to our ability to train our VIVID whisper-bot *when* to probe deeper, and *how* to do so. (1) The first is the Hypothesis Progression Framework (HPF) [4], which we used to teach the model five product development stages: Customer, Product, Concept, Feature, and Business. Based on a combination of common sentence structures in each stage and key words that might be used within each, we taught the model to identify when a customer is talking about one of their responsibilities, or a problem they are encountering. (2) The second framework was the VIVID grammar [5]. We found that having vivid stories helps move an organization to inspired action, but that having such stories requires having vivid conversations to begin with. This framework ensured teams capture crucial elements of a vivid customer story: the who/what, how many, where, when, how, and why? We combined the HPF with VIVID grammar, so that the model was trained on sets of rich vivid questions at each phase of the HPF. The end-result was that the whisper-bot could now identify a job responsibility or a problem in the customer's verbalizations and suggest follow-up questions to probe deeper with VIVID questions to get a rich meaningful story.
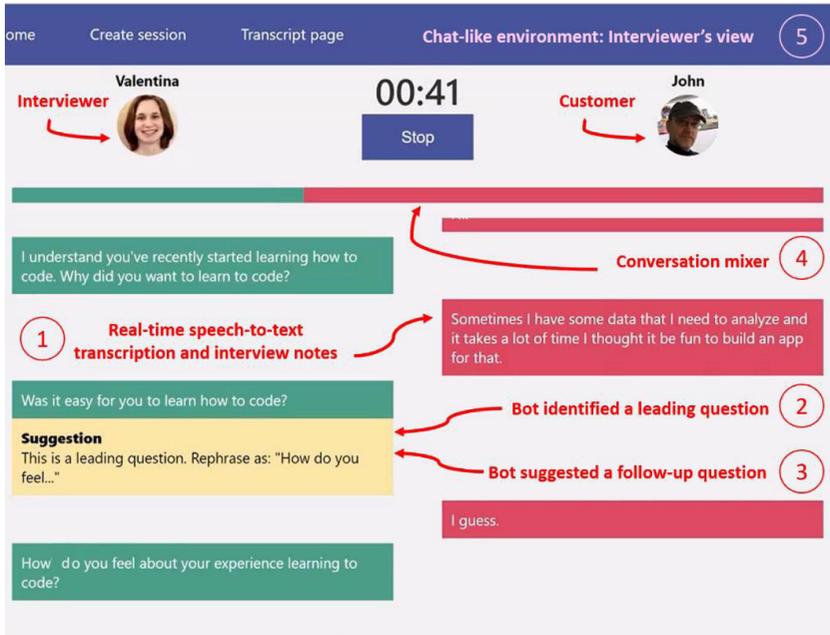
Fig. 1: Working proof-of-concept prototype of the 'VIVID whisper-bot.'

4. The whisper-bot interface: For our proof-of-concept, we created a webservice that mimicked an IM environment, showing a real-time transcript of the conversation to the interviewer (see Fig. 1) and its real-time feedback (based on the analysis of the interviewer's and the customer's verbalizations), to gently guide the interviewer to a richer conversation.

4 Conclusions and Future Work

We have introduced a framework for training an AI system to augment anyone's skills for conducting better interviews. Our whisper-bot prototype helps product team members ask the right questions, at the right time, and in the right way.

For the whisper-bot to move from proof-of-concept to a minimum viable solution, it would need to reach customers through their existing conversational tools (e.g., Microsoft Teams), and be updated with the latest advances in speech-to-text transcription (including some innovative methods, such as creating a training dictionary from the interview's discussion guide, to help identify domain-specific words and phrases).

How to surface the AI system's information to the user such a way that it is helpful and not disruptive is an important part of the future work. This will require improvements to the dialog management service, which manages what response to suggest next. This service has all the information to make the conversation as natural and productive as possible (e.g., how to respond for each category of feedback, how often, when to interject and when not to, as well as the specifics of the response). It is important that the whisper-bot is not leading to cognitive overload for the interviewer, but rather making his/her job easier.

Finally, this same concept can be extended for everything from planning conversations (e.g., using the whisper-bot framework to build the interview questions), all the way through to analyzing the data (e.g., easy filtering or high-lights where "problems" are stated), and sharing insights (e.g., video clips of the vivid stories the bot helped the interviewer unearth).

Acknowledgements

References

1. Fern, X., et al.: Mining problem-solving strategies from HCI data. ACM Transactions on Computer-Human Interactions 17(1), 22 pages (2010).
2. Greenwood, M.: Five years to the top: Microsoft's software boss Julia Liuson, https://techvibes.com/2019/04/09/five-years-to-the-top-microsofts-software-boss-julia-liuson, last accessed 2019/05/06.
3. Grigoreanu, V., et al.: "Gender Differences in End-User Debugging, Revisited: What the Miners Found." In: VISUAL LANGUAGES AND HUMAN-CENTRIC COMPUTING (VL/HCC'06), pp. 19–26, Brighton (2006).
4. Lowdermilk, T., Rich, J.: The customer-driven playbook: Converting customer feedback into successful products. O'Reilly Media (2017).
5. Roam, D.: Blah, blah, blah: What to do when words don't work. Portfolio/Penguin, New York (2011).