In the Hours of Transition from Web to Conversational Search

Sruthi Viswanathan*
Naver Labs Europe
Meylan, France
sruthi.viswanathan@naverlabs.com

Fabien Guillot*
Naver Labs Europe
Meylan, France
fabien.guillot@naverlabs.com

Antonietta Maria Grasso
Naver Labs Europe
Meylan, France
antonietta.grasso@naverlabs.com

Jean-Michel Renders
Naver Labs Europe
Meylan, France
jean-michel.renders@naverlabs.com

ABSTRACT

Are we ready yet? To start conversations with our personal computing systems, after three decades of browsing for information by typing, scrolling, and clicking, is unlikely to be an overnight adjustment. With the current error prone state of Conversational Agents (CAs), we observed the end-users to be unsupported, for switching from text to speech. In this position statement we discuss our theory that the end-users are currently situated in the hours of transition from web to conversational search. Our findings suggest that their expectations for a conversational interaction fluctuates between the features of web search and natural human-like conversation. Building on these expectations of the current users, we propose to design Conversational User Interfaces (CUIs) by joining the necessary features of GUIs and human conversations.

* Authors with equal contribution to this work

FROM WEB SEARCH ENGINES TO CONVERSATIONAL AGENTS

In recent years CAs have been transforming from 'voice-only' (smart speakers) to 'voice and visual' systems (such as Google Assistant, Nest Hub, and Amazon Echo Show), expanding towards multimodal capabilities. They have also come to be increasingly integrated into mobile phones and car infotainment systems, growing as multi-environment systems. However, these mobile and multimodal systems are far from end-users' expectations [2] and text queries reign as the user's favourites for interacting and exploring search results [1].

Our research focus is on conversational restaurant recommendation, with a CA integrated in smartphone, enabling a two-way multimodal dialogue system, where both the user and the system engage in voice and visual interaction. We chose to test the current state of CAs for restaurant recommendation activity in the three environments of home, work and street. We recruited 9 participants to perform the task of asking restaurant recommendations from a Google Assistant imbedded in a Google Pixel 3 smartphone. We reported our initial findings on the challenges and opportunities for a conversational recommender system in a poster at CUI 2020 [4]. Further analysing the data from this study, focusing on the user behaviour, we discovered their interactions with the CA to be influenced by years of interacting with computers. Our findings suggest that users interact with CAs enacting mental models which they have formerly developed from their use of web search engines. They expected a conversational interaction to naturally contain features of both, web search and human-like conversation.

While using graphical and interaction elements of the web in a conversation with the system was not successful in 1995 [5], after decades of co-evolving with technology, today a mixture of natural language and GUI buttons is found to be the best interaction mode with text-based recommender systems [3]. We propose interacting with the CAs in a simple manner, starting with what users know about interacting with web search engines. In this Workshop, "*CUI @ CHI'21 Let's Talk About CUIs: Putting Conversational User Interface Design Into Practice",* we look forward to discussing our research for building a multimodal conversational recommender system rooted in the understanding of current user needs. We hope to share our current theories, methods and future plans for building CUIs set in the hours of transition from web to conversation.

REFERENCES

[1] I. Guy. The characteristics of voice search: Comparing spoken with typed-in mobile web search queries. ACM Transactions on Information Systems (TOIS), 36(3):30:1–30:28, 2018.

[2] 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, 5286–5297. https://doi.org/10.1145/2858036.2858288

[3] Narducci, F., Basile, P., de Gemmis, M. et al. An investigation on the user interaction modes of conversational recommender systems for the music domain. User Model User-Adap Inter 30, 251–284 (2020). https://doi.org/10.1007/s11257-019-09250-7

[4] Sruthi Viswanathan, Fabien Guillot and Antonietta Maria Grasso. 2020.What is Natural? Challenges and Opportunities for Conversational Recommender Systems. In Proceedings of Conversational UserInterfaces (CUI'20). Bilbao, Spain, 4 pages.https://doi.org/10.1145/3405755.3406174

[5] Yankelovich, N., Levow, G.-A., & Marx, M. (1995). Designing SpeechActs: Issues in Speech User Interfaces. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 369–376). New York, NY, USA: ACM Press/Addison-Wesley Publishing Co. https://doi.org/10.1145/223904.223952.