

## DOCTORAL THESIS

# Human-centered Design and Evaluation of Conversational Recommender Systems

CAI, Wanling

*Date of Award:*  
2022

[Link to publication](#)

### General rights

Copyright and intellectual property rights for the publications made accessible in HKBU Scholars are retained by the authors and/or other copyright owners. In addition to the restrictions prescribed by the Copyright Ordinance of Hong Kong, all users and readers must also observe the following terms of use:

- Users may download and print one copy of any publication from HKBU Scholars for the purpose of private study or research
- Users cannot further distribute the material or use it for any profit-making activity or commercial gain
- To share publications in HKBU Scholars with others, users are welcome to freely distribute the permanent URL assigned to the publication

# Abstract

With the advancement of chatbot technology in recent years, there has been a surge in interest in conversational recommender systems (CRSs), which mimic human conversations to help users find recommendations (e.g., movies, music, and hotels) through multi-turn dialogues. Compared with traditional recommender systems that present a ranked list of items in one direction to users, CRSs can support more flexible interactions – mixed-initiative interaction – between users and the system, allowing users to ask questions or give feedback on recommendations. Yet, the majority of existing research has focused on optimizing system effectiveness and efficiency (i.e., improving recommendation quality and reducing the number of interaction turns), while many human-related issues such as user intent (e.g., feedback intents), user experience, and user trust have received insufficient attention, which are crucial for designing CRSs that can better serve users' needs and help them make confident decisions. In this dissertation, we intend to study CRSs from a human-centered perspective, with the aim of answering the main question: **How to design conversational recommender systems that can enhance user experience and inspire user trust?**

To address this question, we place our emphasis on the three concrete problems: 1) understanding and predicting user intents and satisfaction with conversation recommendations; 2) designing and evaluating conversational recommender systems with critiquing techniques; and 3) analyzing the factors influencing user trust in conversational recommender systems.

To develop a multi-turn conversational recommender system, it is important to understand and predict users' intents behind their utterances and their satisfaction with the recommendation, so as to allow the system to incrementally refine the user preference model and adjust its dialogue strategy. Therefore, we first aim to gain a deeper understanding of how users interact with recommenders through conversations. To this end, we analyzed a set of over 300 human-human recommendation dialogues, based on which two hierarchical taxonomies for classifying user intents and recommender actions, respectively, were established using a grounded theory approach. Then, we defined various categories of features considering content, discourse, sentiment, and context to predict users' intents and satisfaction using machine learning techniques.

Our above dialogue analysis indicates that users may have different needs (e.g., finding items of interest, exploring something different) during the conversation. Given that CRSs can support user feedback intents such as providing critiquing feedback (e.g., "*I want songs with a faster tempo!*"), we considered incorporating critiquing techniques (i.e., user-initiated critiquing (UC) and system-suggested critiquing (SC)) into the system to facilitate users' feedback provision and their exploration of recommendations. Specifically, we designed two kinds of system-suggested critiquing technique, progressive system-suggested critiquing (Progressive SC) and cascading system-suggested critiquing (Cascading SC), to enhance user exploration in two different ways, and implemented three prototype conversational music recommenders featuring three different critiquing techniques respectively: UC, Progressive SC and Cascading SC. We then conducted a user study to compare the three critiquing-based recommenders with regards to music exploration in terms of user experience, showing that incorporating Progressive SC and Cascading SC in the recommender can increase users' perceived diversity and serendipity, respectively.

Considering different task contexts may influence user interaction with the system, we carried out another user study and presented our evaluation of the three proposed critiquing-based conversational recommenders in a typical user task in recommender systems: basic recommendation task (BRT, i.e., looking for items according to the user's preferences), and compared it with our evaluation results in the exploration-oriented task (EOT, i.e., exploring different types of items). Our experimental results verified the impact of task type on user experience with the CRS; EOT stimulated more user interaction, while BRT resulted in higher user satisfaction.

While the preceding studies have demonstrated several advantages of CRSs, little is known about user trust toward CRSs, which strongly affects users' intention to adopt such systems in real-world situations. Grounded on the framework of Hoff and Bashir's three-layered trust model, we investigated the effects of three types of factors (user-related, system-related, and context-related) on user trust in two types of CRS, i.e., user-initiative and mixed-initiative. Our empirical study highlighted the impact of the user-related factors (personal characteristics), showing that users' trust propensity and domain knowledge positively influenced their trust in conversational recommender systems, and that users with high conscientiousness tended to trust the mixed-initiative system.

In the end, we provide practical implications for designing conversational recommender systems based on our experimental findings. They should be useful for researchers and practitioners to develop a human-centered conversational recommender system, making it capable of better fulfilling users' needs, adapting to different task contexts, and fostering individual trust toward the system, which may ultimately maximize the benefits of conversational recommender systems.

**Keywords:** Conversational recommender systems, human-centered design, user intent, critiquing technique, user exploration, user evaluation, task context, user trust, personal characteristic