Research Statement

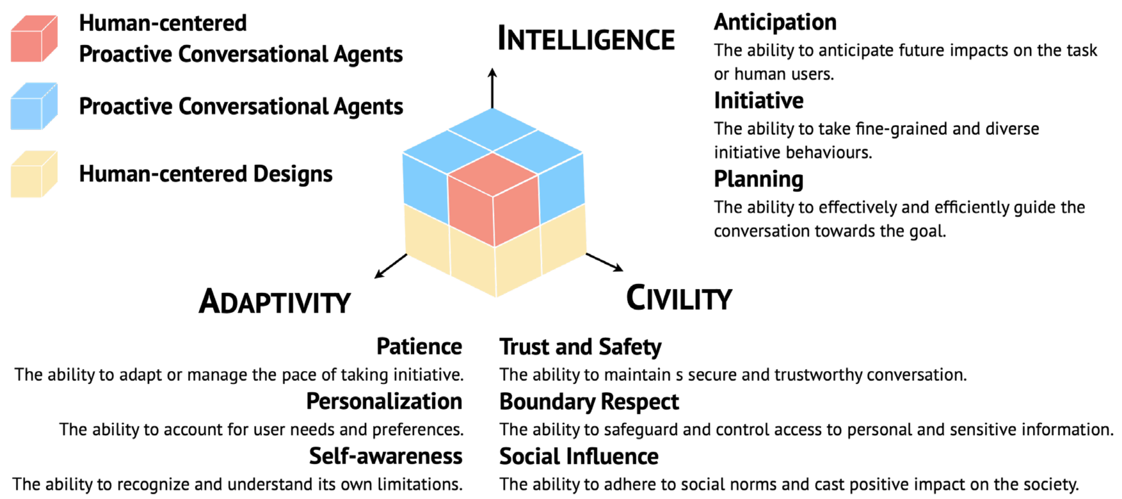
DENG Yang
School of Computing and Information Systems, Singapore Management University
Tel: (65) 6828-4800; Email: ydeng@smu.edu.sg
20 (Day) 12 (Month) 2024 (Year)

Background

Conversational AI agents are envisioned to provide social support or functional service to human users via natural language interactions. Conversational agent research typically centers around a system's response capabilities, such as understanding the context of dialogue and generating appropriate responses to user requests. The popularity of conversational agents has grown unprecedentedly with the advent of ChatGPT, which showcases exceptional proficiency in the capabilities of context understanding and response generation with large language models (LLMs). However, typical conversational systems are built to follow instructions, which means that the conversation is led by the user, and the system simply follows the user's instructions or intents.

My research endows the conversational agent with the capabilities of creating or controlling the conversation to achieve the conversational goals by taking initiative and anticipating impacts on themselves or human users, namely *Proactive Conversational Agents*. Proactive conversational agents can not only largely improve user engagement and service efficiency in the conversation, but also empower the system to handle more complicated interactive tasks that involve strategic and motivational interactions.

Key aims of my work include tackling the challenges of building human-centered proactive conversational agents [1] from the following perspectives:



Research Areas

1. Intelligence – Proactive Conversational and Interactive Systems

Proactive features in dialogue systems [2] have the potential to enhance user engagement and service efficiency across a wide range of conversational contexts. Additionally, they enable these systems to effectively navigate intricate conversational tasks, which encompass strategic and motivational interactions. Recognizing the manifold advantages of proactivity, my research is consistently dedicated to the advancement of proactive dialogue systems across the advent of large language models (LLMs):

1) Proactive conversational systems in the pre-LLM era

To improve the efficiency and effectiveness of conversational recommender systems (CRS), I proposed a novel reinforcement learning (RL) paradigm to reformulate the decision making of recommending items and asking preference-eliciting questions into a unified policy learning problem [3], which is adopted as the standard backbone for most of the following studies of RL-based CRS [4]. Apart from RL-based approaches, another main-stream line of approaches was corpus-based learning. I proposed one of the earliest approaches [5] to unify all the natural language understanding problems in CRS into the sequence-to-sequence problems to be solved by generative pre-trained language models. Moreover, this was also the early attempt for multi-task instruction-tuning. My follow-up study [6] also applies this framework into proactive conversational question answering in finance domain.

2) Proactive conversational systems in the era of LLMs

With the advent of LLMs, the paradigm of building dialogue systems has been revolutionized. I conducted the first comprehensive evaluation [7] of LLM-based dialogue systems in handling various proactive dialogue systems, including clarification in information-seeking dialogues, target-guided dialogues, and non-collaborative dialogues. I have studies different in-context learning approaches [7,8] for proactive dialogue problems. Furthermore, I introduced a new dialogue policy planning paradigm [9] to strategize LLMs for proactive dialogue problems with a tunable language model plug-in as a plug-and-play dialogue policy planner, which can be supervisedly fine-tuned over available human-annotated data as well as conduct reinforcement learning from goal-oriented AI feedback with dynamic interaction data collected by the LLM-based self-play simulation. This framework is further applied into various applications, such as target-guided conversational recommendation [31], asking clarification questions in conversational information seeking [25], and non-collaborative dialogues [34]. Another promising solution is to leverage LLMs for data augmentation [26,33].

2. Adaptivity – User-centric Information Seeking

User-centric information seeking involves designing and developing systems in a way that involves human needs and preferences into information-seeking systems, rather than solely focusing on functional capabilities.

The system should explore and identify the human user's needs, preferences, and values, and should be able to leverage the user information to enhance the future interactions. Interactive systems must efficiently understand about interaction context, including the history of the interaction, online and offline user information beyond the language. My previous works focused on incorporating various types of user information into E-Commerce question answering systems for better aligning with individual user preferences and needs in product-related questions, thereby improving the overall user experience in online shopping environments. This included integrating a range of user opinions [10], constructing models that capture detailed user preferences [11], and estimating user satisfaction [12]. Furthermore, I also investigated some practical issues in exploiting user persona for LLM-based personalized dialogues, including the robustness of prompting with different orders of persona [13] and the source planning capability of LLMs with multi-source knowledge [14]. In addition, I also developed dialogue systems that strategically provide emotional support [15], focusing on enhancing the mental well-being of human users.

3. Civility – Trust and Reliability of Large Language Models

As Large Language Models (LLMs) serve as foundation of the conversational agents, the trust and reliability of LLMs becomes utmost important. We need to identify and understand the potential causes and mechanisms of unintended behaviors in the LLM-powered conversational agents and develop techniques to reduce the likelihood of such behaviors occurring and the potential harm that may be caused by them. LLMs often struggle with different trust and reliability issues, including generating factually incorrect content [16,17,30] and producing toxic or disruptive content [18]. Specifically, I investigate the knowledge boundary of LLMs, such as mitigating unknown questions [36] and supplementing additional knowledge [35]. Furthermore, it is also crucial to ensure transparency in system decision-making and reasoning processes [19-21] for explainable AI.

4. Future Works

1) Embodied Language Agents. Embodied agent is an artificial intelligence system that is designed to interact with a specific environment, while embodied conversational agent can further interact with the human user via natural language. These agents may integrate various capabilities, such as spatial reasoning for navigating physical environment [22,23], multimodal understanding for more natural and intuitive interactions [24], tool using for accomplishing real-world planning [29,32]. Additionally, there are different types of environments, including physical and virtual environments. Embodied conversational agents are expected to be capable of interacting with various environments.

2) Proactive Interaction beyond Human-Agent Interaction. Proactive interactions not only are beneficial to human-agent conversations, but also contribute to various human-human and agent-agent interaction applications. As for human-human interactions, proactive AI mentor systems can proactively educate or train human to learn social skills, rather than just passively addressing user questions. As for agent-

agent interactions, proactive multi-agent systems can proactively interact with other agents to achieve communicative objectives, such as collaborative tasks or society simulation, rather than just passively following user instructions.

3) Applications of Proactive Conversational Agents in Vertical Domains. My dialogue research has also developed novel applications in various vertical domains such as finance [6], mental health [9,15,26], education [9,28]. For example, the proactive conversational question answering system [6] can initiate clarification questions for clarifying the ambiguity or uncertainty in financial information seeking. The proactivity of emotional support dialogue systems [9,15,26] lies in planning a sequence of mixed-initiative emotional support strategies.

Selected Publications and Outputs

[1] **Yang Deng**, Lizi Liao, Zhonghua Zheng, Grace Hui Yang, Tat-Seng Chua. Towards human-centered proactive conversational agents. In ***SIGIR 2024***, 2024.
[2] **Yang Deng**, Wenqiang Lei, Wai Lam, and Tat-Seng Chua. A survey on proactive dialogue systems: Problems, methods, and prospects. In ***IJCAI 2023***, 2023.
[3] **Yang Deng**, Yaliang Li, Fei Sun, Bolin Ding, and Wai Lam. Unified conversational recommendation policy learning via graph-based reinforcement learning. In ***SIGIR 2021***, 2021.
[4] **Yang Deng**, Yaliang Li, Bolin Ding, and Wai Lam. Leveraging long short-term user preference in conversational recommendation via multi-agent reinforcement learning. *IEEE Trans. Knowl. Data Eng. (**TKDE**)*, 35(11), 2023.
[5] **Yang Deng**, Yaliang Li, Wenxuan Zhang, Bolin Ding, and Wai Lam. Toward personalized answer generation in e-commerce via multi-perspective preference modeling. *ACM Trans. Inf. Syst. (**TOIS**)*, 40(4), 2022.
[6] **Yang Deng**, Wenqiang Lei, Wenxuan Zhang, Wai Lam, and Tat-Seng Chua. PACIFIC: towards proactive conversational question answering over tabular and textual data in finance. In ***EMNLP 2022***, 2022.
[7] **Yang Deng**, Lizi Liao, Liang Cheng, Hongru Wang, Wenqiang Lei, and Tat-Seng Chua. Prompting and evaluating large language models for proactive dialogues: Clarification, target-guided, and non-collaboration. In *Findings of the Association for Computational Linguistics: **EMNLP 2023***, 2023.
[8] Hongru Wang, Rui Wang, Fei Mi, **Yang Deng**, Zezhong Wang, Bin Liang, Ruifeng Xu, and Kam-Fai Wong. Cue-cot: Chain-of-thought prompting for responding to in-depth dialogue questions with llms. In Findings of the Association for Computational Linguistics: ***EMNLP 2023***, 2023.
[9] **Yang Deng**, Wenxuan Zhang, Wai Lam, See-Kiong Ng, and Tat-Seng Chua. Plug-and-play policy planner for large language model powered dialogue agents. In ***ICLR 2024***, 2024.
[10] **Yang Deng**, Wenxuan Zhang, and Wai Lam. Opinion-aware answer generation for review-driven question answering in e-commerce. In ***CIKM 2020***, 2020.
[11] **Yang Deng**, Yaliang Li, Wenxuan Zhang, Bolin Ding, and Wai Lam. Toward personalized answer generation in e-commerce via multi-perspective preference modeling. *ACM Trans. Inf. Syst. (**TOIS**)*, 40(4), 2022.
[12] **Yang Deng**, Wenxuan Zhang, Wai Lam, Hong Cheng, and Helen Meng. User satisfaction estimation with sequential dialogue act modeling in goal-oriented conversational systems. In ***WWW 2022***, 2022.
[13] Liang Chen, Hongru Wang, **Yang Deng**, Wai-Chung Kwan, Zezhong Wang, and Kam-Fai Wong. Towards robust personalized dialogue generation via order-insensitive representation regularization. In Findings of the Association for Computational Linguistics: ***ACL 2023***, 2023.
[14] Hongru Wang, Minda Hu, **Yang Deng**, Rui Wang, Fei Mi, Weichao Wang, Yasheng Wang, Wai Chung Kwan, Irwin King, and Kam-Fai Wong. Large language models as source planner for personalized knowledge-grounded dialogues. In Findings of the Association for Computational Linguistics: ***EMNLP 2023***, 2023.
[15] **Yang Deng**, Wenxuan Zhang, Yifei Yuan, and Wai Lam. Knowledge-enhanced mixed-initiative dialogue system for emotional support conversations. In ***ACL 2023***, 2023.
[16] Yuexiang Xie, Fei Sun, **Yang Deng**, Yaliang Li, and Bolin Ding. Factual consistency evaluation for text summarization via counterfactual estimation. In Findings of the Association for Computational

Linguistics: ***EMNLP 2021***, 2021.
[17] Liang Chen, **Yang Deng**, Yatao Bian, Zeyu Qin, Bingzhe Wu, Tat-Seng Chua, and Kam-Fai Wong. Beyond factuality: A comprehensive evaluation of large language models as knowledge generators. In ***EMNLP 2023***, 2023.
[18] Boyi Deng, Wenjie Wang, Fuli Feng, **Yang Deng**, Qifan Wang, and Xiangnan He. Attack prompt generation for red teaming and defending large language models. In Findings of the Association for Computational Linguistics: ***EMNLP 2023***, 2023.
[19] **Yang Deng**, Wenxuan Zhang, and Wai Lam. Multi-hop inference for question-driven summarization. In ***EMNLP 2021***, 2020.
[20] Weiwen Xu, **Yang Deng**, Huihui Zhang, Deng Cai, and Wai Lam. Exploiting reasoning chains for multi-hop science question answering. In Findings of the Association for Computational Linguistics: ***EMNLP 2021***, 2021.
[21] **Yang Deng**, Wenxuan Zhang, Weiwen Xu, Ying Shen, and Wai Lam. Nonfactoid question answering as query-focused summarization with graph-enhanced multihop inference. *IEEE Transactions on Neural Networks and Learning Systems (**TNNLS**)*, 2023.
[22] Shuaiyi Li, **Yang Deng**, and Wai Lam. Depwignn: A depth-wise graph neural network for multi-hop spatial reasoning in text. In Findings of the Association for Computational Linguistics: ***EMNLP 2023***, 2023.
[23] **Yang Deng**, Shuaiyi Li, and Wai Lam. Learning to ask clarification questions with spatial reasoning. In ***SIGIR 2023***, 2023.
[24] Haohao Luo, Ying Shen, and **Yang Deng**. Unifying text, tables, and images for multimodal question answering. In Findings of the Association for Computational Linguistics: ***EMNLP 2023***, 2023.
[25] Yue Chen, Chen Huang, **Yang Deng**, Wenqiang Lei, Dingnan Jin, Jia Liu, Tat-Seng Chua. STYLE: Improving Domain Transferability of Asking Clarification Questions in Large Language Model Powered Conversational Agents. In Findings of the Association for Computational Linguistics: ***ACL 2024***, 2024.
[26] Zhonghua Zheng, Lizi Liao, **Yang Deng**, Libo Qin, Liqiang Nie. Self-chats from Large Language Models Make Small ChatPal Better. In ***ACL 2024***, 2024.
[27] Tong Zhang, Peixin Qin, **Yang Deng**, Chen Huang, Wenqiang Lei, Junhong Liu, Dingnan Jin, Hongru Liang, Tat-Seng Chua. CLAMBER: A Benchmark of Identifying and Clarifying Ambiguous Information Needs in Large Language Models. In ***ACL 2024***, 2024.
[28] Haohao Luo, **Yang Deng**, Ying Shen, See-Kiong Ng, Tat-Seng Chua. Chain-of-Exemplar: Enhancing Distractor Generation for Multimodal Educational Question Generation. In ***ACL 2024***, 2024.
[29] Yang Deng, Xuan Zhang, Wenxuan Zhang, Yifei Yuan, See-Kiong Ng, Tat-Seng Chua. On the Multi-turn Instruction Following for Conversational Web Agents. In ***ACL 2024***, 2024.
[30] Peixin Qin, Chen Huang, **Yang Deng**, Wenqiang Lei, Tat-Seng Chua. Beyond Persuasion: Towards Conversational Recommender System with Credible Explanations. In Findings of the Association for Computational Linguistics: ***EMNLP 2024***, 2024.
[31] Huy Quang Dao, **Yang Deng**, Khanh-Huyen Bui, Dung D. Le, Lizi Liao. Experience as Source for Anticipation and Planning: Experiential Policy Learning for Target-driven Recommendation Dialogues. In Findings of the Association for Computational Linguistics: ***EMNLP 2024***, 2024.
[32] Xuan Zhang, **Yang Deng**, Zifeng Ren, See-Kiong Ng, Tat-Seng Chua. Ask-before-Plan: Proactive Language Agents for Real-World Planning. In Findings of the Association for Computational Linguistics: ***EMNLP 2024***, 2024.
[33] Zhonghua Zheng, Lizi Liao, **Yang Deng**, Ee-Peng Lim, Minlie Huang, Liqiang Nie. Thoughts to Target: Enhance Planning for Target-driven Conversation. In ***EMNLP 2024***, 2024.
[34] Tong Zhang, Chen Huang, **Yang Deng**, Hongru Liang, Jia Liu, Zujie Wen, Wenqiang Lei, Tat-Seng Chua. Strength Lies in Differences! Improving Strategy Planning for Non-collaborative Dialogues via Diversified User Simulation. In ***EMNLP 2024***, 2024.
[35] Shuaiyi Li, **Yang Deng**, Deng Cai, Hongyuan Lu, Liang Chen, Wai Lam. Consecutive Batch Model Editing with HooK Layers. In ***EMNLP 2024***, 2024.
[36] **Yang Deng**, Yong Zhao, Moxin Li, See-Kiong Ng, Tat-Seng Chua. Don't Just Say "I Don't Know"! Self-aligning Large Language Models for Responding to Unknown Questions with Explanations. In ***EMNLP 2024***, 2024.