

Online Appendix to: A Survey on Recent Advances in Conversational Data Generation

HEYDAR SOUDANI, Radboud University, The Netherlands

ROXANA PETCU, University of Amsterdam, The Netherlands

EVANGELOS KANOULAS, University of Amsterdam, The Netherlands

FAEGHEH HASIBI, Radboud University, The Netherlands

Additional Key Words and Phrases: Conversational AI, Dialogue System, Data Augmentation, Conversation Generation, Task Oriented Dialogues, Open Domain Dialogues, Conversational Information Seeking

ACM Reference Format:

Heydar Soudani, Roxana Petcu, Evangelos Kanoulas, and Faegheh Hasibi. 2026. Online Appendix to: A Survey on Recent Advances in Conversational Data Generation. *ACM Comput. Surv.* 1, 1 (January 2026), 3 pages. <https://doi.org/10.1145/3795686>

A EXAMPLES OF DIALOGUE TYPES

Dialogue Type	Example Dialogue
Task-Oriented (TOD)	User: Book a restaurant in Orlando for 4 people. System: What type of food and price range should I look for? User: I'd like a moderately priced taiwanese restaurant. System: I found a restaurant in the center of Orlando. They have availability at 8PM and 8:30PM.
Open-Domain (ODD)	User 1: Hi there! Are you looking for someone to take care of your pet while you're out of town? User 2: Yes, I'm going on vacation for a week and I need someone to watch my cat. User 1: That's no problem at all! I love taking care of animals, and I have experience with cats. I'll be happy to take care of your cat while you're gone. User 2: That sounds great. How much do you charge?
Information-Seeking (IS)	User: I am interested to learn more about taiwanese cousine, it's my favourite. System: Of course! I can help with that. What would you like to know? User: I really like the five spice powder. Can you tell me what it is made of, in what proportions and what are common areas each gets cultivated? System: Five-spice powder typically includes star anise, cloves, Chinese cinnamon (cassia), Sichuan peppercorns, and fennel seeds. The exact proportions vary by region and recipe, but a common blend might use equal parts of each, or slightly more star anise and cinnamon for warmth and sweetness. In terms of cultivation, ..

Authors' addresses: Heydar Soudani, heydar.soudani@ru.nl, Radboud University, Nijmegen, The Netherlands; Roxana Petcu, r.m.petcu@uva.nl, University of Amsterdam, Amsterdam, The Netherlands; Evangelos Kanoulas, e.kanoulas@uva.nl, University of Amsterdam, Amsterdam, The Netherlands; Faegheh Hasibi, f.hasibi@cs.ru.nl, Radboud University, Nijmegen, The Netherlands.



This work is licensed under a Creative Commons Attribution 4.0 International License.

© 2026 Copyright held by the owner/author(s).

Manuscript submitted to ACM

Manuscript submitted to ACM

B EXAMPLES OF TASK-ORIENTED CONVERSATIONAL DATA GENERATION

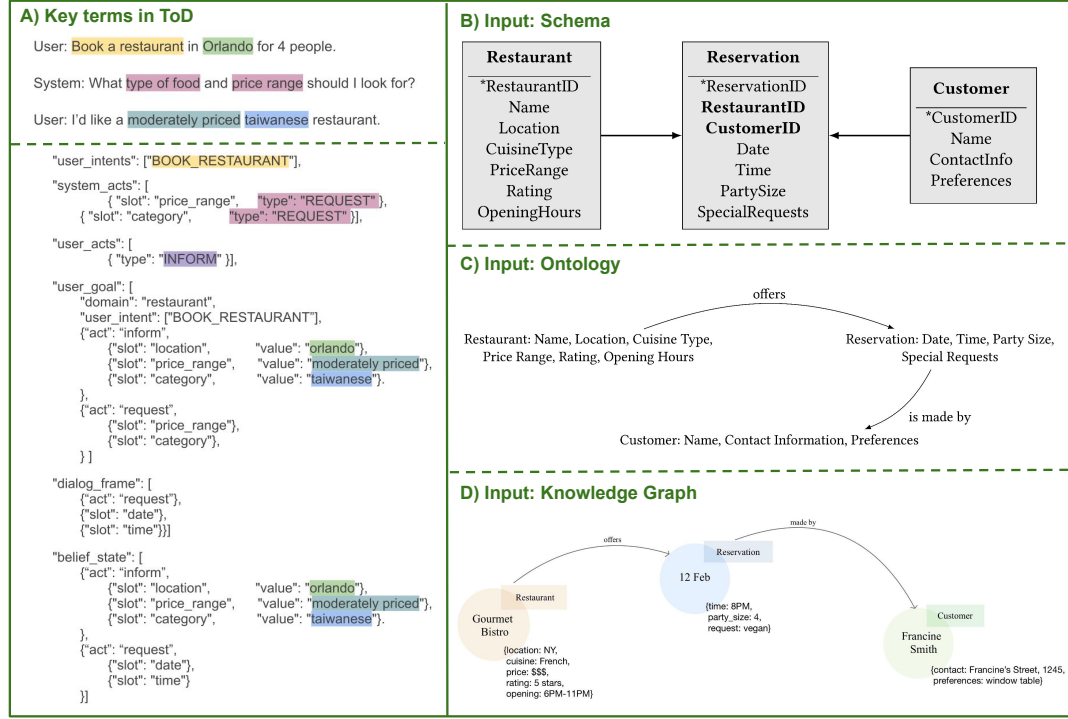


Fig. 1. This figure illustrates several aspects of a task-oriented dialogue (TOD): (A) key terms, slots, and associated values used in a TOD dialogue; (B) an example entity-attribute schema for the task of restaurant table reservation; (C) an example ontology with relational information for the same task; and (D) an example knowledge graph (KG) constructed for this task.

C EXAMPLES OF OPEN DOMAIN CONVERSATIONAL DATA GENERATION

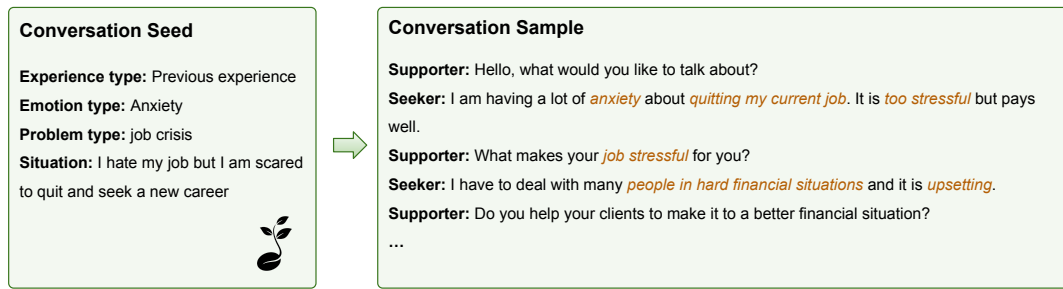


Fig. 2. Conversation seed and the corresponding dialogue from the ESConv Dataset [3]. This dataset focuses on emotional support conversations. The example shows a conversation seed that outlines the emotional problem and current situation of a support seeker. The corresponding dialogue develops around the information provided in the seed.

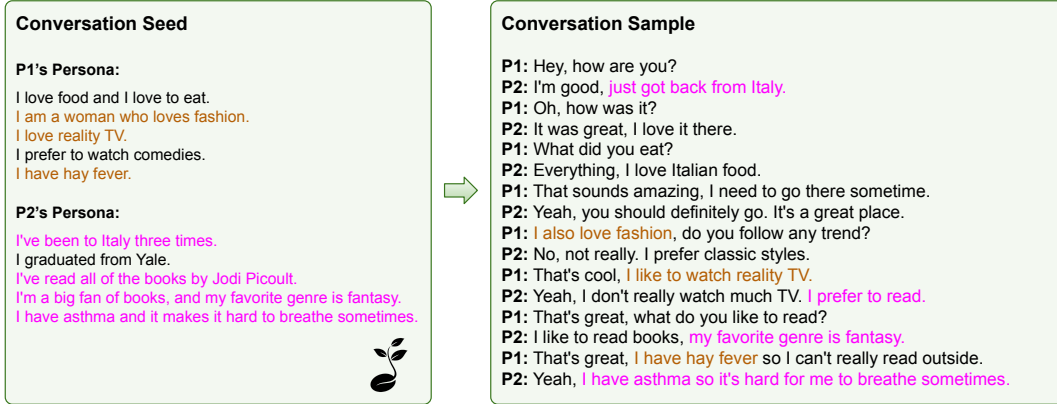


Fig. 3. Examples of generated dialogue from PersonaChatGen [2] based on a conversation seed containing two user profiles. Utterances in orange are directly related to Person 1, while those in magenta relate to Person 2.

D EXAMPLES OF INFORMATION SEEKING CONVERSATIONAL DATA GENERATION

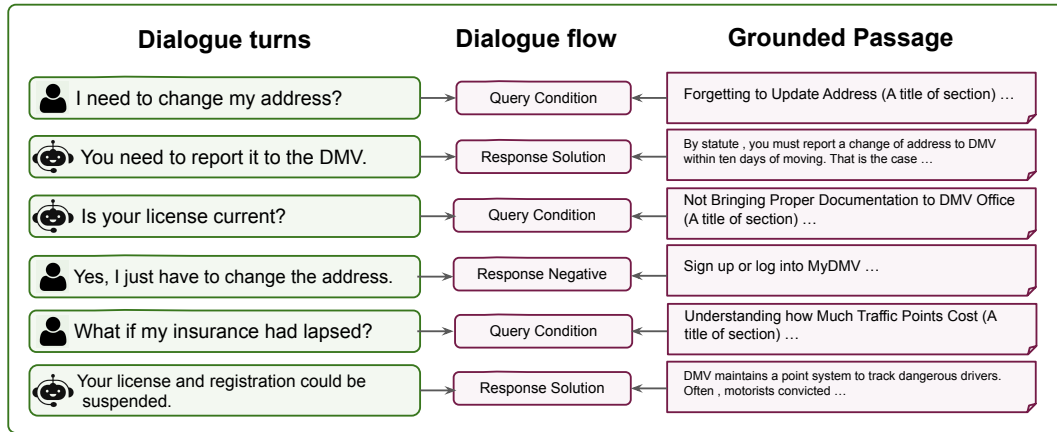


Fig. 4. Conversation Seed and Corresponding Dialogue from the Doc2Dial Dataset [1]. This dataset is dedicated to generating conversations based on governmental documents. The displayed example features a conversation seed comprising a title, a document, and a sequence of dialogue acts. In the corresponding conversation sample, each turn is labeled with a specific dialogue act.

REFERENCES

- [1] Song Feng, Kshitij P. Fadnis, Q. Vera Liao, and Luis A. Lastras. 2020. Doc2Dial: A Framework for Dialogue Composition Grounded in Documents. In *Proceedings of the Thirty-Fourth AAAI Conference on Artificial Intelligence (AAAI 2020)*. 13604–13605.
- [2] Young-Jun Lee, Chae-Gyun Lim, Yunsu Choi, Ji-Hui Lm, and Ho-Jin Choi. 2022. PERSONACHATGEN: Generating Personalized Dialogues using GPT-3. In *Proceedings of the 1st Workshop on Customized Chat Grounding Persona and Knowledge*. 29–48.
- [3] Siyang Liu, Chujie Zheng, Orianna Demasi, Sahand Sabour, Yu Li, Zhou Yu, Yong Jiang, and Minlie Huang. 2021. Towards Emotional Support Dialog Systems. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing (ACL/IJCNLP 2021)*. 3469–3483.