

## References

1. Aliannejadi, M., Kiseleva, J., Chuklin, A., Dalton, J., Burtsev, M.: Building and evaluating open-domain dialogue corpora with clarifying questions. arXiv preprint arXiv:2109.05794 (2021)
2. Aliannejadi, M., Zamani, H., Crestani, F., Croft, W.B.: Asking clarifying questions in open-domain information-seeking conversations. In: Proceedings of the 42nd international acm sigir conference on research and development in information retrieval. pp. 475–484 (2019)
3. Banerjee, S., Lavie, A.: Meteor: An automatic metric for mt evaluation with improved correlation with human judgments. In: Proceedings of the acl workshop on intrinsic and extrinsic evaluation measures for machine translation and/or summarization. pp. 65–72 (2005)
4. Choi, J., Chae, K., Song, J., Jo, Y., Kim, T.: Model-based preference optimization in abstractive summarization without human feedback. In: Al-Onaizan, Y., Bansal, M., Chen, Y.N. (eds.) Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing. pp. 18837–18851. Association for Computational Linguistics, Miami, Florida, USA (Nov 2024). <https://doi.org/10.18653/v1/2024.emnlp-main.1048>, <https://aclanthology.org/2024.emnlp-main.1048>
5. Dhingra, B., Faruqui, M., Parikh, A., Chang, M.W., Das, D., Cohen, W.W.: Handling divergent reference texts when evaluating table-to-text generation. arXiv preprint arXiv:1906.01081 (2019)
6. Duong, S., Bronnec, F.L., Allauzen, A., Guigue, V., Lumbrieras, A., Soulier, L., Gallinari, P.: SCOPE: A self-supervised framework for improving faithfulness in conditional text generation. In: The Thirteenth International Conference on Learning Representations (2025), <https://openreview.net/forum?id=dTkqaCKLPp>
7. Erbacher, P., Nie, J.Y., Preux, P., Soulier, L.: Augmenting ad-hoc ir dataset for interactive conversational search. Transactions on Machine Learning Research (2024), <https://openreview.net/forum?id=z8d7nT1HWw>
8. Erbacher, P., Nie, J.Y., Preux, P., Soulier, L.: Paqa: toward proactive open-retrieval question answering. arXiv preprint arXiv:2402.16608 (2024)
9. Izacard, G., Grave, E.: Leveraging passage retrieval with generative models for open domain question answering. In: Merlo, P., Tiedemann, J., Tsarfaty, R. (eds.) Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume. pp. 874–880. Association for Computational Linguistics, Online (Apr 2021). <https://doi.org/10.18653/v1/2021.eacl-main.74>, <https://aclanthology.org/2021.eacl-main.74>
10. Krasakis, A.M., Yates, A., Kanoulas, E.: Corpus-informed retrieval augmented generation of clarifying questions. arXiv preprint arXiv:2409.18575 (2024)
11. Lee, D., Kim, S., Lee, M., Lee, H., Park, J., Lee, S.W., Jung, K.: Asking clarification questions to handle ambiguity in open-domain qa. arXiv preprint arXiv:2305.13808 (2023)
12. Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W.t., Rocktäschel, T., et al.: Retrieval-augmented generation for knowledge-intensive nlp tasks. Advances in neural information processing systems **33**, 9459–9474 (2020)
13. Li, Y.Z.Y.Y.R., Hu, Z.: Alignscore: Evaluating factual consistency with a unified alignment function
14. Lin, C.Y.: Rouge: A package for automatic evaluation of summaries. In: Text summarization branches out. pp. 74–81 (2004)

15. Lin, J., Ma, X., Lin, S.C., Yang, J.H., Pradeep, R., Nogueira, R.: Pyserini: A python toolkit for reproducible information retrieval research with sparse and dense representations. In: Proceedings of the 44th international ACM SIGIR conference on research and development in information retrieval. pp. 2356–2362 (2021)
16. Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., Stoyanov, V.: Roberta: A robustly optimized bert pretraining approach. arXiv preprint arXiv:1907.11692 (2019)
17. Lu, L., Meng, C., Ravenda, F., Aliannejadi, M., Crestani, F.: Zero-shot and efficient clarification need prediction in conversational search. In: European Conference on Information Retrieval. pp. 389–404. Springer (2025)
18. Mass, Y., Cohen, D., Yehudai, A., Konopnicki, D.: Conversational search with mixed-initiative - asking good clarification questions backed-up by passage retrieval. In: Feng, S., Wan, H., Yuan, C., Yu, H. (eds.) Proceedings of the Second DialDoc Workshop on Document-grounded Dialogue and Conversational Question Answering. pp. 65–71. Association for Computational Linguistics, Dublin, Ireland (May 2022). <https://doi.org/10.18653/v1/2022.dialdoc-1.7>, <https://aclanthology.org/2022.dialdoc-1.7/>
19. Ouyang, L., Wu, J., Jiang, X., Almeida, D., Wainwright, C., Mishkin, P., Zhang, C., Agarwal, S., Slama, K., Ray, A., et al.: Training language models to follow instructions with human feedback. Advances in neural information processing systems **35**, 27730–27744 (2022)
20. Owoicho, P., Dalton, J., Aliannejadi, M., Azzopardi, L., Trippas, J.R., Vakulenko, S.: Trec cast 2022: Going beyond user ask and system retrieve with initiative and response generation. In: TREC (2022)
21. Papineni, K., Roukos, S., Ward, T., Zhu, W.J.: Bleu: a method for automatic evaluation of machine translation. In: Proceedings of the 40th annual meeting of the Association for Computational Linguistics. pp. 311–318 (2002)
22. Radlinski, F., Craswell, N.: A theoretical framework for conversational search. In: Proceedings of the 2017 conference on conference human information interaction and retrieval. pp. 117–126 (2017)
23. Rafailov, R., Sharma, A., Mitchell, E., Manning, C.D., Ermon, S., Finn, C.: Direct preference optimization: Your language model is secretly a reward model. Advances in neural information processing systems **36**, 53728–53741 (2023)
24. Rao, S., Daumé, H.: Learning to ask good questions: Ranking clarification questions using neural expected value of perfect information. In: Annual Meeting of the Association for Computational Linguistics (2018)
25. Robertson, S., Zaragoza, H., et al.: The probabilistic relevance framework: Bm25 and beyond. Foundations and Trends® in Information Retrieval **3**(4), 333–389 (2009)
26. Sahay, R., Tekumalla, L.S., Aggarwal, P., Jain, A., Saladi, A.: Ask: Aspects and retrieval based hybrid clarification in task oriented dialogue systems. In: Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 6: Industry Track). pp. 881–895 (2025)
27. Sekulić, I., Aliannejadi, M., Crestani, F.: Towards facet-driven generation of clarifying questions for conversational search. In: Proceedings of the 2021 ACM SIGIR international conference on theory of information retrieval. pp. 167–175 (2021)
28. Siro, C., Yuan, Y., Aliannejadi, M., de Rijke, M.: AGENT-CQ: Automatic Generation and Evaluation of Clarifying Questions for Conversational Search with LLMs. arXiv preprint arXiv:2410.19692 (2024), <http://arxiv.org/abs/2410.19692>

29. Tang, A., Soulier, L., Guigue, V.: Clarifying ambiguities: on the role of ambiguity types in prompting methods for clarification generation. In: Proceedings of the 48th International ACM SIGIR Conference on Research and Development in Information Retrieval. pp. 20–30 (2025)
30. Wang, Z., Tu, Y., Rosset, C., Craswell, N., Wu, M., Ai, Q.: Zero-shot clarifying question generation for conversational search. In: Proceedings of the ACM web conference 2023. pp. 3288–3298 (2023)
31. Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., Cistac, P., Rault, T., Louf, R., Funtowicz, M., et al.: Transformers: State-of-the-art natural language processing. In: Proceedings of the 2020 conference on empirical methods in natural language processing: system demonstrations. pp. 38–45 (2020)
32. Yang, X., Tan, Z., Li, H.: Ipo: Iterative preference optimization for text-to-video generation. arXiv preprint arXiv:2502.02088 (2025)
33. Zhang, T., Kishore, V., Wu, F., Weinberger, K.Q., Artzi, Y.: Bertscore: Evaluating text generation with bert. arXiv preprint arXiv:1904.09675 (2019)