

Weiwei Sun

✉ sunnweiwei@gmail.com 🐦 @sunnweiwei12 🌐 sunnweiwei
🌐 https://sunnweiwei.github.io/ 🎓 google scholar



Research Interests: Natural Language Processing, Information Retrieval

Education

- 2021 – 2023 📖 **M.E. Shandong University** Computer Science
Advisor: Prof. Zhaochun Ren
- 2017 – 2021 📖 **B.E. with honours Shandong University** Computer Science
GPA: 3.8 / 4.0 · Advisor: Prof. Zhaochun Ren

Work/Research Experience

- 2022 – 2023 📖 **Research Intern.** Search Science Team, Baidu Inc.
Research on Large Language Models and Generative Information Retrieval.
- 2021 – 2022 📖 **Research Intern.** Data Science Lab, JD.com.
Research on Question Answering Systems.
- 2019 – 2021 📖 **Research Assistant.** Information Retrieval Lab, Shandong University.
Research on Dialogue Generation and Evaluation.

Publications /Google Scholar /Semantic Scholar

- 1 **Weiwei Sun**, L. Yan, Z. Chen, S. Wang, H. Zhu, P. Ren, Z. Chen, D. Yin, M. de Rijke, and Z. Ren, “Learning to tokenize for generative retrieval,” in *NeurIPS*, 2023.
- 2 **Weiwei Sun**, L. Yan, X. Ma, S. Wang, P. Ren, Z. Chen, D. Yin, and Z. Ren, “Is chatgpt good at search? Investigating large language models as re-ranking agents,” in *EMNLP*, 2023, **outstanding paper award**.
- 3 Y. Zhao, L. Yan, **Weiwei Sun**, C. Meng, S. Wang, Z. Cheng, Z. Ren, and D. Yin, “Diqad: A benchmark dataset for end-to-end open-domain dialogue assessment,” in *Findings of EMNLP*, 2023.
- 4 **Weiwei Sun**, H. Cai, H. Chen, P. Ren, Z. Chen, M. de Rijke, and Z. Ren, “Answering ambiguous questions via iterative prompting,” in *ACL*, 2023, pp. 7669–7683.
- 5 Z. Shi, **Weiwei Sun**, S. Zhang, Z. Zhang, P. Ren, and Z. Ren, “Rade: Reference-assisted dialogue evaluation for open-domain dialogue,” in *ACL*, 2023, pp. 12 856–12 875.
- 6 S. Guo, S. Zhang, **Weiwei Sun**, P. Ren, Z. Chen, and Z. Ren, “Towards explainable conversational recommender systems,” in *SIGIR*, 2023, pp. 2786–2795.
- 7 **Weiwei Sun**, P. Ren, and Z. Ren, “Generative knowledge selection for knowledge-grounded dialogues,” in *Findings of EACL*, 2023, pp. 2077–2088.
- 8 **Weiwei Sun**, Z. Shi, S. Gao, P. Ren, M. de Rijke, and Z. Ren, “Contrastive learning reduces hallucination in conversations,” in *AAAI*, 2023, pp. 13 618–13 626.
- 9 **Weiwei Sun**, S. Guo, S. Zhang, P. Ren, Z. Chen, M. de Rijke, and Z. Ren, “Metaphorical user simulators for evaluating task-oriented dialogue systems,” *ACM Transactions on Information Systems*, vol. 42, 2023.
- 10 **Weiwei Sun**, S. Zhang, K. Balog, Z. Ren, P. Ren, Z. Chen, and M. de Rijke, “Simulating user satisfaction for the evaluation of task-oriented dialogue systems,” in *SIGIR*, 2021, pp. 2499–2506.
- 11 **Weiwei Sun**, C. Meng, Q. Meng, Z. Ren, P. Ren, Z. Chen, and M. de Rijke, “Conversations powered by cross-lingual knowledge,” in *SIGIR*, 2021, pp. 1442–1451.

- 12 C. Meng, P. Ren, Z. Chen, **Weiwei Sun**, Z. Ren, Z. Tu, and M. de Rijke, “Dukenet: A dual knowledge interaction network for knowledge-grounded conversation,” in *SIGIR*, 2020, pp. 1151–1160.
- 13 **Weiwei Sun**, Z. Chen, X. Ma, L. Yan, S. Wang, P. Ren, Z. Chen, D. Yin, and Z. Ren, “Instruction distillation makes large language models efficient rankers,” in *GenRec workshop at CIKM*, 2023.
- 14 T. Ju, **Weiwei Sun**, W. Du, X. Yuan, Z. Ren, and G. Liu, “How large language models encode context knowledge? A layer-wise probing study,” in *COLING submission*, 2023.
- 15 Y. Zhao, L. Yan, **Weiwei Sun**, G. Xing, S. Wang, C. Meng, Z. Cheng, Z. Ren, and D. Yin, “Improving the robustness of large language models via consistency alignment,” in *COLING submission*, 2023.
- 16 Y. Zhao, L. Yan, **Weiwei Sun**, G. Xing, C. Meng, S. Wang, Z. Cheng, Z. Ren, and D. Yin, “Knowing what llms do not know: A simple yet effective self-detection method,” *ArXiv*, vol. abs/2310.17918, 2023.

Awards and Achievements

- 2023
 - 📌 **Outstanding Paper Award**, EMNLP 2023.
 - 📌 **Baidu Scholarship (Global Top 10, \$28k)**, Baidu Inc.
 - 📌 **Presidential Scholarship (Top 0.2%)**, Shandong University.
 - 📌 **National Scholarship (Top 1%)**, Ministry of Education of the PRC.
- 2022
 - 📌 **Second Price**, College Academic Show, China Computer Federation.
- 2021
 - 📌 **Outstanding Graduate with an honorary degree (Top 2%)**, Shandong University.
- 2019
 - 📌 **National Second Price (Top 3%)**, Mathematical Contest in Modeling (CUMCM).
 - 📌 **Ranked No.2 out of 11395 participants**, CSP, China Computer Federation.
- 2018
 - 📌 **Outstanding Presentation**, Short-Term Visiting Program, Oxford University.
- 2017 – 2023
 - 📌 **College Academic Scholarship**, Shandong University.

Academic Service

- Reviewer 📌 EMNLP 2022, ACL 2023, SIGIR 2023, SIGIR-AP 2023, ECML/PKDD 2023, IPM 2023, WSDM 2023, EMNLP 2023, TALLIP 2023, SIGIR 2024.
- Volunteer 📌 YSSNLP 2022, NLPCC 2022, SIGIR-AP 2023.




Selected Projects

- 2023
 - 📌 **RankGPT** LLMs as re-ranking agents. (300 Stars)
<https://github.com/sunnweiwei/RankGPT> <https://arxiv.org/abs/2304.09542>
 - 📌 **Fuzi · Mingcha** LLMs for legal applications. (130 Stars, 5k Monthly Downloads, 130 Likes)
<https://huggingface.co/fireballoon> <https://huggingface.co/SDUIRLab>
- 2019
 - 📌 **Voice of AI** Scene-based music generation. (Outstanding presentation in DeeCamp 2019)
<https://voiceofai.github.io/> video demo

Resources Shared






- 2023
 - 📌 **NovelEval** IR test set with the novel queries and passages.
<https://github.com/sunnweiwei/RankGPT/tree/main/NovelEval>
 - 📌 **DiQAD** Long-form question answering dataset with detailed quality annotations.
https://github.com/yukunZhao/Dataset_Dialogue_quality_evaluation

Resources Shared (continued)





- 2022  **E-Redial** Conversational recommender dataset with high-quality explanations.
<https://github.com/Superbooming/E-Redial>
- 2021  **USS** Task-oriented dialogue dataset with multi-grain user **satisfaction** annotations.
<https://github.com/sunnweiwei/user-satisfaction-simulation>
-  **CKGC** Open-domain dialogue dataset with annotated cross-lingual knowledge.
<https://github.com/sunnweiwei/ckgc>

Miscellaneous Experience





Talks

- 2024  **Recent Advances in Generative Information Retrieval**, a tutorial at the Web Conference 2024, co-organized with Y. Tang, R. Zhang, J. Guo, and M. de Rijke · [Link](#).
- 2023  **Is ChatGPT Good at Search? Investigating LLMs as Re-Ranking Agents**, in EMNLP 2023, Singapore · [Link](#).
-  **Large Language Models for Information Retrieval**, Institute of Computing Technology, Chinese Academy of Sciences · [Link](#).
-  **Instruction Distillation of LLMs**, in the 1st GenRec Workshop at CIKM 2023 · [Link](#).
- 2021 – 2023  A collection of talks at the IRLab group meeting · [Link](#)

Teaching Experience

- 2022  **Introduction to Big Data Technology**, Shandong University · *Teaching Assistant*.
-  **Computer Organization and Design**, Shandong University · *Teaching Assistant*.
- 2021  **Summer School on Mathematical Modeling**, Shandong University · *Lecturer*.
- 2021 – 2023  Student Supervision in Shandong University: Shuyu Guo (PhD, Conversational Recommendation) · Zhengliang Shi (MSc, Dialogue Systems), Zheng Chen (BSc, Generative Retrieval) · Liu Yang (MSc, Dense Retrieval) · Minghui Tian (MSc, Dialogue Systems) · Yidan Wang (MSc, Generative Retrieval).

Skills

- Coding  Python (CSP score **400**), Java, C++, SQL, \LaTeX , ...
- Database  MySQL, Redis, Elasticsearch, MapReduce.
- Research  PyTorch, TensorFlow, HuggingFace, Deepspeed, OpenAI
- Web Dev  HTML, CSS, JavaScript.

References

Prof. Zhaochun Ren

renzhaochun.github.io
Associate Professor
Leiden University
z.ren@liacs.leidenuniv.nl

Prof. Maarten de Rijke

staff.fnwi.uva.nl/m.derijke
Professor
University of Amsterdam
M.deRijke@uva.nl

Dr. Dawei Yin

www.yindawei.com
Senior Director
Baidu Inc.
yindawei@acm.org

Dr. Shuo Zhang

imsure318.github.io
Senior Research Scientist
Bloomberg AI
szhang611@bloomberg.net

Prof. Pengjie Ren

pengjieren.github.io
Professor
Shandong University
renpengjie@sdu.edu.cn