

# LIN DONGDING

📍 Hong Kong SAR, China • 📞 (+86) 137 5006 5371 • 📩 22037064r@connect.polyu.hk

🔗 [github.com/DongdingLin](https://github.com/DongdingLin) • 🎓 Google Scholar • **Target Role:** LLM Algorithm Engineer / Applied Research Scientist

## SUMMARY

Ph.D. student at the Department of Computing, The Hong Kong Polytechnic University. Research focuses on **LLM reasoning, dialogue systems, multimodal LLMs, and situated conversational recommendation**. Experienced in the full R&D lifecycle from task formulation and data construction to model design and experimental evaluation, with strong capability in engineering deployment and iterative optimization. Publications appear in ACM MM, ACL, AAAI, TOIS, and TNNLS.

## EDUCATION

<b>The Hong Kong Polytechnic University</b>	Sept. 2022 – Present
PhD Candidate, Department of Computing (NLP Group)	Hong Kong, China
<b>Sun Yat-sen University</b>	Sept. 2017 – Jul. 2020
M.Eng. in Computer Technology, GPA: 3.9/4.0 (Recommended Admission)	Guangzhou, China
<b>Sun Yat-sen University</b>	Sept. 2013 – Jul. 2017
B.Eng. in Software Engineering, GPA: 3.8/4.0 (Rank: 37/433)	Guangzhou, China

## EXPERIENCE

<b>Huawei Hong Kong Research Center (HKRC), Fermat Lab</b>	Aug 2025 – Present
Research Intern (Outstanding Intern)	Hong Kong SAR, China

- **Wireless online-training data pruning (base-station resource optimization):** For online training settings, designed local data selection and out-of-distribution (OOD) filtering strategies. Combined a distillation-based discriminator with GPU-KMeans clustering to reduce training compute by **4×** with  $\leq 3\%$  performance degradation. Delivered a technical report and production-ready implementation, including reusable configurations and a logging loop for rapid rollback and reproducible experiments.
- **Automatic generalization + formal verification for math problems (Math500 / Competitive-math benchmarks (e.g., AIME/HMMT) / K12):** Built an automated pipeline: variable extraction → parameter generalization → Mathematica solving and consistency checking, producing structured annotations for extensible problem types and batch generation. Improved Math500 generalization rate from **42.74% to 58.6%**, and new-item accuracy from **87.4% to 99%**. Improved competition-problem generalization from **7.8% to 63.95%**. For multi-type K12 settings, introduced multi-branch generalization strategies to raise overall generalization to **42%**. Packaged core capabilities as an **SDK** for deployment.
- **In-distribution (IID) dialogue data selection:** Constructed high-quality, in-distribution dialogue data to better emulate real user queries for model evaluation. Built a hybrid cleaning pipeline that integrates rule-based heuristics with sentence-embedding retrieval and clustering to filter ASR noise, refusal patterns, and outliers, producing a high-quality IID dialogue set for reliable evaluation.

<b>NLP Group, The Hong Kong Polytechnic University</b>	Dec 2022 – Present
Ph.D. Research	Hong Kong SAR, China

- **Situated conversational recommendation (SCR) task modeling and evaluation:** Systematically formulated the interaction pipeline involving scenarios, context, and user preferences, and defined task setups and evaluation metrics. Implemented an end-to-end pipeline for data processing, training, and evaluation to support a series of papers with reproducible experiments.
- **SiPeR: scenario transfer + Bayesian preference inference:** Addressed “Where/What” decision-making in SCR by jointly modeling scenario transfer and target-scenario prediction, with a calibrated transition decision module. Estimated whether the current scenario meets user needs and guided transfer when necessary. Leveraged MLLM likelihood for Bayesian inverse inference to characterize implicit preferences and filter visual distractors, achieving average gains of **10.9%/10.6%** on SIMMC 2.1 and SCREEN.
- **Re2A: rubric-driven preference reasoning and dual-aligned generation:** Cast SCR as a reason-then-align process: automatically generated rubrics to provide multi-dimensional reward signals, enabling explicit preference-state inference **without human annotation**. Then applied Preference-conditioned DPO to jointly align preference satisfaction and scenario consistency, reducing hallucination rate from **25.4% to 5.2%** and weak-match rate from **22.6% to 7.4%** on SIMMC 2.1.

- **Target-oriented dialogue planning and generation (framework engineering):** Designed planning and generation methods around target constraints, dialogue states, and user preferences; incorporated look-ahead and feedback-based correction to improve multi-turn controllability and goal achievement. Modularized preprocessing, intent modeling, planner, and generator with unified training/inference interfaces and evaluation scripts, supporting rapid strategy comparison and reproducible experiments.

### **Collective Intelligence Systems Lab, Sun Yat-sen University**

*Feb. 2018 – Jul. 2020*

Guangzhou, China

- **Machine reading comprehension and multi-hop reasoning (hierarchical propagation + memory flow):** Proposed hierarchical information propagation and memory-update mechanisms for evidence aggregation in long documents and cross-sentence relation modeling. Built training/evaluation pipelines and conducted systematic ablations and error analyses, with reproducible code and configurations.

## **SELECTED PUBLICATIONS**

---

- **Dongding Lin**, Jian Wang, Chak Tou Leong, Wenjie Li. SCREEN: A Benchmark for Situated Conversational Recommendation. *ACM MM 2024* (CCF-A).
- Jian Wang, **Dongding Lin**, Wenjie Li. Target-constrained Bidirectional Planning for Generation of Target-oriented Proactive Dialogue. *TOIS 2024* (CCF-A).
- Jian Wang, Chak Tou Leong, Jiashuo Wang, **Dongding Lin**, Wenjie Li, Xiao-Yong Wei. Instruct once, chat consistently in multiple rounds: An efficient tuning framework for dialogue. *ACL 2024* (CCF-A).
- Jian Wang, Yi Cheng, **Dongding Lin**, et al. Target-oriented proactive dialogue systems with personalization. *EMNLP 2023* (CCF-A).
- **Dongding Lin\***, Jian Wang\*, Wenjie Li. COLA: Improving Conversational Recommender Systems by Collaborative Augmentation. *AAAI 2023* (CCF-A).
- Jian Wang\*, **Dongding Lin\***, Wenjie Li. Dialogue Planning via Brownian Bridge Stochastic Process for Goal-directed Proactive Dialogue. *ACL Findings 2023* (CCF-A).

## **ACADEMIC SERVICES & TEACHING**

---

- Reviewer: ACL Rolling Review (ARR), ACL, EMNLP, ACM MM.
- Teaching Assistant: *Natural Language Processing* (2024/25 S2, 2023/24 S2), *Data Structures and Database Systems* (2024/25 S1), *Mobile Computing* (2023/24 S1).

## **SKILLS**

---

<b>Programming</b>	Python, C/C++, Java, JavaScript, SQL, MATLAB
<b>ML Stack</b>	PyTorch, TensorFlow, Hugging Face, Scikit-learn
<b>Research Areas</b>	LLM reasoning, dialogue systems, conversational recommendation, multimodal understanding

## **AWARDS**

---

- Ranked **4/750** in 2021 Language and Intelligence Challenge (Baidu).
- Ranked **4/119** in Materials Quality Prediction Kaggle Competition.
- Outstanding Graduate Award (**Top 3%** of Sun Yat-sen University), Excellent Graduation Thesis (**Top 3%** of Sun Yat-sen University).
- First/Second/Third Prize Scholarships during undergraduate and postgraduate study.