关 健

j-guan19@mails.tsinghua.edu.cn | https://jianguanthu.github.io/ | +86 13051331318

清华大学，北京市海淀区双清路 30 号，100084

教育经历

博士研究生（五年级），计算机科学与技术系，清华大学　　　　　　　**2019年8月 - 至今**
导师：黄民烈 教授（NLP领域三位杰青之一）
GPA: 3.56/4.00

理学学士，物理系，清华大学　　　　　　　　　　　　　　　　　**2015年8月 - 2019年6月**
GPA: 3.50/4.00；排名：7/49

实习经历

研究科学家，[艾伦人工智能研究院 (AI2)](#)，西雅图，美国　　　　　　　**2023年1月 - 2023年4月**
科研实习，弗吉尼亚大学，夏洛茨维尔，美国　　　　　　　　　　　　**2018年6月 - 2018年9月**
科研实习，CoAI Group，北京，清华大学，中国　　　　　　　　　　**2017年6月 - 2019年6月**

研究亮点

1. 发表 **10+**篇一作论文（**ACL, AAAI, NeurIPS, EMNLP, NAACL, TACL**），引用 **1000+**，H-index **12**

2. 论文获 **ACL 2023 Area Chair Award**（共同一作）

3. 中文长文本预训练模型 **LongLM**（TACL2022），峰值每月 **10k+**次下载

4. 长期深耕**长文本生成和评价领域**，包括**数据资源构建**（**ACL 2021; NAACL 2022; TACL 2022**）、**提升长程连贯性**（**ACL 2021; AAAI 2023**）、**减少知识幻觉**（**AAAI 2019；TACL 2020**）、**提升生成可控性**（**ACL 202; NAACL 2022; ACL 2023**）、**文本生成评价算法**（**EMNLP 2020**）

 (1) 数据使用和申请者包括 **Amazon**、**Meta**、**AI2**、腾讯、网易、华为云等著名研究机构

 (2) 研究成果多次获得顶级名校教授的引用，包括 **Kevin Knight**（**ACL fellow**）、**Ruslan Salakhutdinov**（**CMU** 教授，引用 **10w+**）、**Yejin Choi**（**UW** 教授，引用 **4w+**）、**Noah Smith**（**UW** 教授，引用 **4w+**）等

5. 开源代码累计获得 **500+**星标

学术论文

研究兴趣： 自然语言生成。

论文列表（其中*表示共同一作）：

1. **Jian Guan***, Yansen Wang*, and Minlie Huang. Story ending generation with incremental encoding and commonsense knowledge. In AAAI 2019 (CCF-A 类会议).

2. Xueying Bai*, **Jian Guan***, Hongning Wang. A Model-Based Reinforcement Learning with Adversarial Training for Online Recommendation. In NeurIPS 2019 (CCF-A 类会议).

3. **Jian Guan**, Xiaoxi Mao, changjie fan, Zitao Liu, Wenbiao Ding, Minlie Huang. Long Text Generation by Modeling Sentence-Level and Discourse-Level Coherence. In ACL 2021 (CCF-A 类会议).

4. **Jian Guan**, Zhexin Zhang, Zhuoer Feng, Zitao Liu, Wenbiao Ding, Xiaoxi Mao, Changjie Fan, Minlie Huang. OpenMEVA: A Benchmark for Evaluating Open-ended Story Generation Metrics. In ACL 2021 (CCF-A 类会议).

5. **Jian Guan**, Zhenyu Yang, Rongsheng Zhang, Zhipeng Hu, Minlie Huang. Generating Coherent Narratives by Learning Dynamic and Discrete Entity States with a Contrastive Framework. In AAAI 2023 (CCF-A 类会议).

6. Xuekai Zhu*, **Jian Guan***, Minlie Huang, Juan Liu. StoryTrans: Non-Parallel Story Author-Style Transfer with Discourse Representations and Content Enhancing. In ACL 2023 (CCF-A 类会议). [Area Chair Award]

7. **Jian Guan**, Fei Huang, Zhihao Zhao, Xiaoyan Zhu, and Minlie Huang. A knowledge-enhanced pretraining model for commonsense story generation. In TACL 2020 (CCF-B 类期刊).

8. **Jian Guan**, Minlie Huang. UNION: An Unreferenced Metric for Evaluating Open-ended Story Generation. In EMNLP 2020 (CCF-B 类会议).

9. **Jian Guan**, Zhuoer Feng, Yamei Chen, Ruilin He, Xiaoxi Mao, Changjie Fan, Minlie Huang. LOT: A Benchmark for Evaluating Chinese Long Text Understanding and Generation. In TACL 2022 (CCF-B 类期刊).

10. **Jian Guan**, Ziqi Liu, Minlie Huang. A Corpus for Understanding and Generating Moral Stories. In NAACL 2022 (CCF-B 类会议).

11. **Jian Guan**, Jesse Dodge, David Wadden, Minlie Huang, Hao Peng. Language Models Hallucinate, but May Excel at Fact Verification. In NAACL 2024 (CCF-B 类会议).

12. **Jian Guan**, Minlie Huang. Mitigating the Learning Bias towards Repetition by Self-Contrastive Training for Open-Ended Generation. In Findings of ACL 2023 (Short Paper).

13. Zhexin Zhang*, **Jian Guan***, Guowei Xu, Yixiang Tian, Minlie Huang. Automatic Comment Generation for Chinese Student Narrative Essays. In EMNLP 2022 (Demo).

14. Pei Ke, **Jian Guan**, Minlie Huang, and Xiaoyan Zhu. Generating informative responses with controlled sentence function. In ACL 2018 (CCF-A 类会议).

15. Zhengyan Zhang, **Jian Guan,** et al. CPM: A large-scale generative Chinese pre-trained language model. In AI Open 2021.

16. Zhengyan Zhang, **Jian Guan**, et al. CPM-2: Large-scale Cost-efficient Pre-trained Language Models. In AI Open 2021.

17. Jiaxin Wen*, Zhexin Zhang*, **Jian Guan**, Minlie Huang. Persona-Guided Planning for Controlling the Protagonist's Persona in Story Generation. In NAACL 2022 (CCF-B 类会议).

18. Xiangzhe Kong*, Jialiang Huang*, Ziquan Tung*, **Jian Guan**, Minlie Huang. Stylized Story Generation with Style-Guided Planning. In Findings of ACL 2021 (Short Paper).

19. Zhexin Zhang, **Jian Guan**, Xin Cui, Yu Ran, Bo Liu and Minlie Huang. Self-Supervised Sentence Polishing by Adding Engaging Modifiers. In ACL 2023 (Demo).

Preprint

1. Chuanqi Cheng*, **Jian Guan***, et al. From the Least to the Most: Building a Plug-and-Play Visual Reasoner via Data Synthesis. arxiv 2024.

2. **Jian Guan**, Wei Wu, Zujie Wen, Peng Xu, Hongning Wang, Minlie Huang. AMOR: A Recipe for Building Adaptable Modular Knowledge Agents Through Process Feedback. arxiv 2024.

3. Fei Huang, **Jian Guan**, Pei Ke, Qihan Guo, Xiaoyan Zhu, Minlie Huang. A text GAN for language generation with non-autoregressive generator. arxiv 2020

4. Fei Huang, **Jian Guan**, et al. Cotk: An open-source toolkit for fast development and fair evaluation of text generation. arxiv 2020.

5. Yuan Yao, **Jian Guan,** et al. Cuge: A chinese language understanding and generation evaluation benchmark. arxiv 2021.

6. Jianzhu Yao, Ziqi Liu, **Jian Guan**, Minlie Huang. A Benchmark for Understanding and Generating Dialogue between Characters in Stories. arxiv 2022.

开源项目

1. [Must-Read Papers for Open-Ended Language Generation.](#) (Github 180+ Stars)
2. [Chinese Long Text Pretraining Models: LongLM.](#) (Download 1000+/month)

荣誉奖项

清华大学优秀博士论文	**2024年6月**
微软亚洲研究院微软学者提名奖	**2022年10月**
博士研究生国家奖学金	**2021年10月**

北京市优秀毕业生	**2019年6月**
清华大学优良毕业生	**2019年6月**
清华大学优秀学生干部	**2018年10月**
清华大学综合优秀奖	**2018年10月**
清华大学综合优秀奖	**2017年10月**
清华大学学业优秀奖	**2016年10月**
清华大学社会实践奖学金	**2016年10月**
清华大学新生奖学金	**2015年10月**

社工经历

清华大学物理系学生会主席	**2018年5月 - 2019年5月**
清华大学物理系基科52班团支书	**2017年9月 - 2018年9月**
清华大学赴重庆武隆支教支队支队长	**2016年6月 - 2016年7月**