

刘柱

· juniperliuzhu@163.com · <https://juniperliuzhu.netlify.app>

教育背景

清华大学, 中国语言文学 (计算语言学方向), 在读博士研究生 2022.9 - 2026.6

导师: 刘颖

研究方向: 语言学启发的模型评估与可解释性, 计算语义理解, 计算语言学, 数字人文, AI+ 教育

罗马第一大学, 自然语言处理方向, 访问学生 2025.3 - 2025.9

导师: Roberto Navigli (ACL 2025 大会主席);

研究方向: 词义理解, 多语数据增强

南方科技大学, 计算机科学与技术, 硕士研究生 2019.9 - 2022.6

导师: 郑锋

研究方向: 视频字幕标注, 概率生成模型, 多模态深度学习

大连理工大学, 软件工程, 学士 2015.9 - 2019.6

工作经历

上海大学, 中国语言文学 (计算语言学方向), 讲师 2026.8 - 至今

实习经历

乐学无限初创公司, 研发实习生 2026.4 - 2026.6

研发内容: 人工智能智能体辅助外语学习

机器之心, 编译评论员 2022.1 - 2022.6

研发内容: 人工智能相关的内容编译和评论

论文发表

注: CCF 分类指中国计算机学会 (中国计算机科学领域权威学术组织) 的评级。其中, ACL、EMNLP 和 NAACL 被国际公认为自然语言和人工智能领域的顶级会议。

- **Zhu Liu**, Teng Wang, Jinrui Zhang, Feng Zheng, Wenhao Jiang, Ke Lu (2022). Show, Tell and Rephrase: Diverse Video Captioning via Two-Stage Progressive Training. *IEEE Transactions on Multimedia*. (CCF A, JCR&SCI Q1, IF: 9.7, 他引: 5)
- **Zhu Liu**, Cunliang Kong, Ying Liu, and Maosong Sun. Fantastic Semantics and Where to Find Them: Investigating Which Layers of Generative LLMs Reflect Lexical Semantics. In *Findings of ACL 2024*, pages 14551–14558, Bangkok, Thailand. (CCF A, EI, 他引: 22)
- **Zhu Liu** and Ying Liu. Ambiguity Meets Uncertainty: Investigating Uncertainty Estimation for Word Sense Disambiguation. In *Findings of ACL 2023*, pages 3963–3977, Toronto, Canada. (CCF A, EI, 他引: 8)
- **Zhu Liu**, Cunliang Kong, Ying Liu, and Maosong Sun. A Top-down Graph-based Tool for Modeling Classical Semantic Maps: A Crosslinguistic Case Study of Supplementary Adverbs. *NAACL 2025*, pages 4567–4576, Albuquerque, New Mexico. Association for Computational Linguistics. (CCF B, EI)
- **Zhu Liu**, Zhen Hu, Ying Liu. JuniperLiu at CoMeDi Shared Task: Models as Annotators in Lexical Semantics Disagreements. CoMeDi workshop in *COLING 2025*, Abu Dhabi, UAE. (CCF B, EI, 他引: 1)
- 刘柱, 孔存良, 刘颖, 等. 基于自上而下图论算法的语义地图自动构建与评估——以“重复义副词”语义地图为例 [J]. *清华语言学*, 2026, (1): 194–213.
- Kangyang Luo, Yuzhuo Bai, Shuzheng Si, Cheng Gao, Zhitong Wang, Yingli Shen, Wenhao Li, **Zhu Liu**, Yufeng Han, Jiayi Wu, Cunliang Kong, Maosong Sun. ImCoref-CeS: An Improved Lightweight Pipeline for Coreference Resolution with LLM-based Checker-Splitter Refinement. *ACL 2026* (CCF A)
- Kangyang Luo, Yuzhuo Bai, Cheng Gao, Shuzheng Si, Yingli Shen, **Zhu Liu**, Zhitong Wang, Cunliang Kong, Wenhao Li, Yufei Huang, Ye Tian, Xuantang Xiong, Lei Han, Maosong Sun. GLTW: Joint Improved Graph Transformer and LLM via Three-Word Language for Knowledge Graph Completion. *ACL 2025 Findings*. (CCF A, EI, 他引: 5)

- Teng Wang, **Zhu Liu**, Feng Zheng, Zhichao Lu, Ran Cheng, and Ping Luo. Semantic-Aware Pretraining for Dense Video Captioning. arXiv preprint arXiv:2204.07449 (2022). (CVPR 2021 ActivityNet 工作坊技术报告, 他引: 5)
- Jian Ma, **Zhu Liu**, Yurong Shi, et al. 基于太阳黑子群数据的多模态太阳耀斑预报模型 [J]. 航天器环境工程, 2021,38(3):256-262.

科研经历

1. 模型评估与可解释性

词义消歧的不确定估计和分析

2022.9-2023.1

词义消歧旨在确定某个词汇在上下文中的含义。传统的词义消歧模型将这个过程视作一个分类任务：所有可能的词义选项视作相互语义正交的候选类别，然后从中选择一个最佳选项。但真实世界中的词汇语义是一个连续概念，无法确定唯一的最佳释义。本次研究试图考虑词汇语义的不确定性和模糊性，并试图寻找背后的词汇语言学因素。本次研究成果获得“清华大学第四届语言文字学青年学术论坛”优秀论文三等奖，并被顶会 ACL 2023 接收。此外，不确定性也会带来标注者之间的分歧，如何预测多样标注中的平均值和方差，是 CoMeDi@Coling 2025 的一项核心任务，本研究对该问题的探索亦被该工作坊收录。

大语言模型跨层表征的语义理解性能评估

2023.10-

生成式大语言模型具有强大的通用性能，但模型内部对于语义的理解仍需评估。本研究着眼于该类模型内部表征的可解释，重点从表征方面探究模型学习到哪些词汇语义知识，并与传统的较小规模模型、双向模型进行对比分析。该研究在清华大学自然语言处理与社会人文计算实验室科研访问期间展开，合作导师为孙茂松老师，目前已被 Findings of ACL 2024 录用。

2. 数字人文 | AI+ 教育

基于自上而下图算法的语义图模型构建与应用

2023.10-

语义地图模型是语言类型学中研究多功能语言形式的重要工具，现有的模型构建都是在小规模的数据量下由专家手工标注的，因而难以拓展。本文开发了一种自上而下的算法，将语义地图经典的连接性假设进行拓展，从而可以在大规模数据下高效地自动构建语义地图。同时，本文还设计针对图的评价指标对生成的语义地图进行评估。该工作获 NAACL 2025 录用及清华大学语言文字学青年论坛优秀论文一等奖，工具持续开发并已投稿顶会展示赛道。

壁画自动描述和交互聊天应用

2023.6-2023.11

壁画作为传统民间艺术存在于很多古建筑中，其艺术价值受到研究者很大的关注，如何将壁画数字化、文字化以及智能化，并吸引更多的观众观看对于保护和传播优秀文化至关重要。这个项目以广西玉林高山村的优秀壁画为素材，开发了可以自动描述壁画内容以及根据壁画的局部信息来与观众交流的应用，提升游览者兴趣的同时也让针对壁画内容知识问答更加精细。该项目可以辅助智慧展览、通识教育等场景。这一项目获得了第八届上海图书馆开放数据竞赛的优秀创意论文奖。

3. 多模态

基于变分自编码模型的视频字幕多样性生成

2021.4-2022.9

视频字幕生成 (Video Captioning) 用语言描述一段视频。多样性字幕生成旨在发掘从视频到多个描述语句的这种一对多的映射关系，从而更加逼近自然状态下的分布。本次研究把这一问题建模为条件隐变量模型，利用条件变分自编码器来刻画这种多样性。发表于国际顶刊：IEEE Transactions on Multimedia 上。

空间太阳风暴事件预测

2019.9-2020.4

该项目通过太阳表面的磁图预测太阳风暴事件。我们利用了一些先进的图片分类模型和缓解样本不平衡的方法。其间团队参加了太阳耀斑的预测比赛，并获得第一名。后续该项目继续得到军工委资助，作为二作在中文核心期刊《航天器环境工程》中发表一篇文章。

专利

郑锋, 刘柱. 视频内容描述模型的训练方法、应用方法、设备及介质: 中国, CN114386480B [P]. 2025-07-18.

学术报告

一种用于建模经典语义地图的自上而下图工具：以补充性副词的跨语言研究为例

展示形式：第二届 CCF 人文智能大会，中国上海，2026 年 7 月；NAACL 2025 线上海报展示，美国新墨西哥州；清华大学第六届语言文字学青年学术论坛，中国北京，2024 年 10 月，口头报告。

CoMeDi 共享任务：一个利用语义分歧标注者的模型	2025.4
展示形式：COLING 2025 英文海报展示，阿联酋阿布扎比。	
探究生成式大语言模型中哪些层反映词汇语义	2024.8
展示形式：ACL 2024 英文海报展示，泰国曼谷。	
词义消歧中的不确定性量化	2023.7
展示形式：ACL 2023 线上英文海报展示，加拿大多伦多；COLING 2025，阿联酋阿布扎比，英文海报展示；清华大学第四届语言文字学青年学术论坛，2023 年 10 月，中文口头报告。	
基于大语言模型的主宾互逆句的及物性分析	2023.10
展示形式：清华大学第五届语言文字学青年学术论坛，中国北京，中文口头报告。	

教学经历

《留学生现代汉语》课程助教	2026.2-2026.7
负责内容：辅导约 40 名外国留学生，辅助任课老师讲解现代汉语语法，负责课后答疑。	
《Python 与数字人文》课程助教	2023.3-2023.7
负责内容：辅导约 30 名学生课程作业，设计并讲解数字人文的经典案例，课后答疑。	
《C 和 C++ 编程设计》课程助教	2020.9-2021.1
负责内容：辅导约 120 名学生，协助主讲老师完成课程，课后答疑，批改作业。	

荣誉获奖

清华大学第四/五/六届语言文字学青年学术论坛	2022/2023/2024.11
一年一度的校内博士生论坛，分别获得优秀论文三等奖、一等奖、一等奖。	
清华大学校级综合奖学金一等奖	2023.10/2025.10
第八届上海图书馆开放数据竞赛	2023.10
“案例分析：文化遗产图像自动描述和交互聊天——以广西高山村壁画为例”获得优秀创意论文奖。	
ActivityNet：大规模视频行为密集视频标注	2021.5-6
由计算机视觉顶级会议 CVPR 工作坊组织，最终获得 第二名 。	
太阳风暴识别和预警人工智能挑战赛	2020.10
该比赛依托阿里天池，是中科院空间所发起的空间预报比赛，我们队伍最终获得 第一名（1/70） ，成功获得相关方的 30 万 科研资助，并在之后顺利结题。该项目同时申请了学校的竞赛奖学金资助。	

参与项目

2018 年 11 月-2025 年 3 月，国家哲学社会科学基金重大项目，“基于大数据技术的古代文学经典文本分析与研究” (18ZDA238), 80 万，优秀 2020 年 1 月-2023 年 11 月，清华大学人文社科振兴项目，“基于大数据技术的明清小说分析和研究” (2019THZWJC38), 10 万 2017 年 7 月-2022 年 4 月，教育部人文社科，“语体特征的自动提取和研究” (17YJAZH056), 10 万 2020 年 10 月-2021 年 9 月，科技创新特区，空间环境事件关联与预报的深度模型研究，50 万 2020 年 11 月-2021 年 3 月，科技创新特区，多模态深度学习太阳耀斑预测模型，30 万	
---	--

学术服务

- 担任 ACL 2025 学生研究论坛（SRW）的联合主席。 - 担任 ARR May 2026 的领域主席。 - 担任 EACL SRW 2026、ACL 2025、NAACL SRW 2025、CoMeDi 2024、EMNLP 2023 和 Rep4NLP workshop 的审稿人。 - 担任 Coling 2025、NAACL 2025 的会议组织志愿者。	
--	--