

刘亚慧

☎ +86 178 6380 7812 ✉ yahuiliu.nlp@foxmail.com

年龄: 30 岁 性别: 女 政治面貌: 中共党员

研究方向: 自然语言处理 英语: CET-6: 487 分

个人主页: <https://yahui19960717.github.io/>



个人简介

刘亚慧, 苏州大学计算机科学与技术学院博士生, 师从张民教授 (**ACL Fellow**、**国家杰青**) 和李正华教授, 所在 HLT 团队是国内 NLP 领域具有重要影响力的研究团队之一, 长期深耕语义分析、机器翻译等 NLP 核心方向。本人研究方向聚焦于**句法语义分析**、**语料库构建与标注数据错误检测**。学术上, 致力于通过精细化的语料库建设推动模型性能的提升, 从负责制定 **53 页** 的中文语义角色标注规范, 到构建覆盖 **6 个领域**、**3.1 万个句子**、**9.2 万个谓词** 的 MuCPAD 数据集, 再到探索基于 KL 散度的标注错误检测方法, 形成了从语料规范制定、数据集构建到标注质量控制、模型评估与性能提升的研究闭环。相关工作以第一作者发表于 **ESWA (SCI 一区)**、**NAACL (CCF-B)**、**EMNLP (CCF-B)** 等期刊会议, 并在 CCL 汉语框架语义分析评测中连续三年获得第一名、第二名和第三名, 发布相关技术报告。此外, 作为**青椒模拟器**团队核心成员, 深度参与该 **AI 互动娱乐落地项目** 的平台维护、内容迭代与社区运营。该项目以高校青年教师学术生涯为主题, 累计访问用户超过 **95 万**, 游戏内累计生成大模型文本数据超过 **330B tokens**。相关经历使本人积累了大模型应用落地、AI 互动娱乐产品运营与用户社区建设经验。同时, 参与 NICE 学术社区内容传播与活动组织, 该社区由 50+ 海内外学者共同发起, 相关渠道累计覆盖学者 13 万 +。

教育背景

苏州大学 计算机科学与技术学院	2018.09 - 至今 硕博连读
计算机科学与技术	
山东农业大学 信息科学与工程学院	2014.09 - 2018.06 本科
计算机科学与技术	

教学经历

《计算语言学》课程助教	苏州大学, 博士期间
- 协助课程教师完成课程教学支持工作, 包括课程资料整理、学生答疑、课程通知发布与作业反馈等。 - 在答疑与作业反馈中, 结合语义分析和语料库构建等研究经历, 帮助学生理解相关基础概念与方法。	

论文成果

已发表论文

- Yahui Liu, Zhenghua Li, Chen Gong, Shilin Zhou, Min Zhang. Annotation Error Detection in Painstakingly Annotated Data: Part-of-Speech Tagging as a Case Study. (一作 2025.05 **ESWA SCI 一区 IF:7.5**)
论文内容: 这篇文章重新标注了 CTB7E 的测试集, 为标注错误检测任务提供一个真实的错误检测数据集, 并提出在 MC Dropout 设定下基于 KL 散度的标注错误检测方法。
- Yahui Liu, Haoping Yang, Chen Gong, Qingrong Xia, Zhenghua Li, Min Zhang. MuCPAD: A Multi-Domain Chinese Predicate-Argument Dataset. (一作 2022.07 **NAACL CCF-B**)
论文内容: 论文构建了跨领域语义角色标注数据集 MuCPAD, 该数据集覆盖 **6 个领域**、**3.1 万句子**、**9.2 万个谓词** 实例。基于该数据集, 我们构建了一个语义角色标注基线模型, 并提出两种策略进一步

提升基线模型性能。

• Lijie Wang[†], **Yahui Liu[†]**, Zhenghua Li, Yue Zhang. LLM-assisted Data Quality Improvement: SRL as a Case Study. (共一, 本人为学生第一作者 2026.08 Findings of EMNLP CCF-B)

论文内容：这篇论文系统分析了 LLM 在 SRL 上表现弱于传统监督模型的原因，发现现有 SRL 测试集中存在边界错误、漏标、多合法片段和可选标注等标注质量问题，导致模型能力被低估。为解决这一问题，论文提出一个大模型辅助的数据质量优化框架：利用 LLM 生成候选论元、判断谓词-论元对的有效性，并结合人工审核修正测试集和标注规范。实验表明，在修正后的基准数据集上，传统监督模型和基于大模型的方法均取得明显性能提升，说明原始测试集的标注质量问题影响了 SRL 模型的可靠评估。

• **刘亚慧**, 杨浩苹, 李正华, 张民. 一种轻量级的汉语语义角色标注规范. (一作 2020.12 中文信息学报 CCF-T1 中文期刊)

论文内容：这篇文章论文提出了一个面向多源多领域的[中文谓词论元标注规范](#) (53 页)。

在投论文

• **Yahui Liu**, Zhenghua Li, Min Zhang. Corpus-level Metric-oriented Post-processing for Cross-domain Span-based Semantic Role Labelling. (一作 **IPM SCI 一区**在投)

论文内容：针对跨域基于片段的语义角色标注中局部解码与语料级别的 F1 值评价不一致的问题，这篇文章提出一种基于 Semi-CRF 边缘概率的语料级后处理方法，通过删除低置信度预测和添加高置信度候选来优化预测集合。实验表明，该方法在 CoNLL-2012 跨域设置下稳定提升 SRL 性能，并发现性能增益主要来自低置信度假阳性片段 (false-positive spans) 的过滤。

评测技术报告

• **Yahui Liu**, Ziheng Qiao, Chen Gong, and Min Zhang. 2025. System Report for CCL25-Eval Task 2: Enhanced Chinese Frame Semantic Parsing with Pre-trained Model and Linguistic Features. In Proceedings of the 24th CCL (Volume 3: Evaluations), pages 47-54.

• **Yahui Liu**, Chen Gong, and Min Zhang. 2024. Leveraging LLMs for Chinese Frame Semantic Parsing. In Proceedings of the 23rd CCL (Volume 3: Evaluations), pages 21-31.

• **刘亚慧**, 李正华, 张民. 2023. CCL23-Eval 任务 3 系统报告: 苏州大学 CFSP 系统. In Proceedings of the 22nd CCL (Volume 3: Evaluations), pages 84-93.

竞赛经历

中国计算语言学大会 (CCL) 是国内自然语言处理领域权威性最高、规模和影响最大的学术会议。依托该学术平台，聚焦汉语框架语义分析任务，连续三年将相关语义理论方法应用到汉语框架语义分析评测比赛中，与国内高校 (北航、中科大、国防科大等) 及研究机构同台竞技，分别斩获第二名 (2025 年)、第一名 (2024) 及第三名 (2023 年)：

• CCL 2025 CFSP(汉语框架语义分析) (2/156)	苏州大学	第二名	2025.04 - 2025.08
• CCL 2024 CFSP(汉语框架语义分析) (1/79)	苏州大学	第一名	2024.03 - 2024.07
• CCL 2023 CFSP(汉语框架语义分析) (3/21)	苏州大学	第三名	2023.03 - 2023.08

项目经历

1. 阿里达摩院语义分析项目2019.03 - 2020.09
- 参与大规模汉语语义分析语料库构建，负责标注规范撰写与迭代优化。
 - 组织并协调标注团队，管理多轮数据标注与质量控制流程，保障语料一致性与标注质量。
 - 与阿里团队定期沟通任务需求与数据反馈，跟踪项目进展并维护相关技术文档。
2. 国家自然科学基金面上项目

知识驱动的汉语网络文本依存句法分析 (61876116), 参与。2019.01 - 2022.12

- 相关成果: 发表 1 篇一作论文 (CCF-T1);
- 参与苏大句法树库 CODT 标注工作。

3. 国家自然科学基金面上项目

融合多源知识的跨领域汉语句子语义分析 (62176173), 参与。2022.01 - 2025.12

- 相关成果: 发表 2 篇一作论文 (SCI 一区、CCF-B), 发布 1 篇一作技术报告 (EI 检索);
- 相关数据资源: 构建跨领域汉语谓词论元数据集 MuCPAD, 覆盖 6 个领域、9.2 万个谓词实例;
- 相关评测: 参加 CCL 2023 汉语框架语义分析评测并获第三名。

4. 国家自然科学基金青年科学基金项目

汉语词汇级和句子级依存结构一体化表示与分析方法研究 (62306202), 参与。2024.12 - 至今

- 相关成果: 发表 2 篇一作技术报告 (EI 检索);
- 相关评测: 参加 2 次 CCL 汉语框架语义分析评测比赛 (第 23 届和第 24 届), 并分别获第一名和第二名。

5. 国家自然科学基金面上项目

基于大语言模型定制的汉语复杂文本错误纠正与解释 (62576237), 参与。2025.12 - 至今

- 相关成果: 1 篇一作论文 (CCF-B), 一篇在投 (SCI 一区)。

6. AI 互动娱乐落地项目

青椒模拟器 (项目网址: <https://tenure-plus.feedscription.com>), 核心成员。2025.12 - 至今

青椒模拟器是一款以高校青年教师学术生涯为主题、基于大模型的 AI 互动娱乐产品, 访问用户数已超过 95 万, 游戏内累计生成大模型文本数据超过 330B tokens。所在团队围绕 AI 互动娱乐方向持续开发多款产品, 相关项目见团队主页: <https://wortou.ai/>。

- 参与平台功能测试、内容迭代与运营维护, 利用大模型工具辅助问题定位、内容生成与体验优化。
- 负责社区生态建设与用户运营, 从 0 到 1 运营官方小红书账号“青椒模拟器”, 累计粉丝 4400+; 建立 10 个用户社群, 覆盖 4700+ 用户; 相关话题累计获得 483.8 万次浏览、2.1 万条讨论。
- 参与承办苏州首场“AI+ 小游戏”主题 Hackathon (抖音 AI 创变者计划主办), 负责活动策划、宣传与组织协调, 推动 AI 互动娱乐与开发者社区交流。

学术服务与社区活动

1. NICE 国际 AI 学术社区内容传播与活动组织2025.09 - 至今

NICE 是由 50+ 海内外顶尖学者发起, 采取“线上常态化 + 线下全球化”的模式, 触达 13 万 + 行业关注者, 辐射中、美、欧的国际化 AI 前沿社区。NICE 主页: <https://nice-intl.github.io/zh.html>。

- 负责 AI/NLP 前沿讲座与学术活动的小红书“NICE 学术”内容撰写与传播, 累计发布 80 余篇推文。
- 负责 AI/NLP 前沿讲座与学术活动的公众号“NICE 学术”内容撰写: 193 和 195 期。
- 深度参与苏州 EMNLP Startup Night 等线下学术交流活动的组织与协调工作。