

王佳安

15370021939 | wangjiaan98@hotmail.com
https://wangjiaan.cn/



教育经历

苏州大学	2020年09月 - 2023年06月
计算机科学与技术 硕士 计算机科学与技术学院	
研究方向：跨语言摘要（详见论文[1-3]），体育赛事摘要（详见论文[4,5]），故事生成（详见论文[6]）等	
排名：1/190	
苏州大学	2016年09月 - 2020年06月
计算机科学与技术 本科 计算机科学与技术学院	
GPA：3.7/4.0，排名：8/175，2019年9月获得推免资格，推免至苏州大学	

实习经历

米哈游·鹿鸣Lumi AI	2022年06月 - 至今
NLP文本生成研究员 上海	
• 参与鹿鸣项目的预研（落地向）	
腾讯·微信模式识别中心	2021年09月 - 2022年06月
NLP研究员 北京	
• 跨语言摘要：详见论文[1-3]	
网易游戏·伏羲AI Lab	2021年06月 - 2021年09月
NLP算法工程师 杭州	
• 信息检索：将DPR模型在中文检索数据集上应用，并利用检索结果在对话中引导话题。	
• 故事结尾生成：详见论文[6]	
• 开放式指令预研：游戏NPC对于玩家输入文本产生相应的动作反馈（意图识别、对话系统、情感分析等），详见专利[4-6]	
• Parameter Efficiency：参与讨论以及论文修改，详见论文[8]	
科大讯飞·苏州研究院	2021年03月 - 2021年06月
NLP研究员 苏州	
• 体育赛事摘要：详见论文[4,5]与专利[1]	
• 多跳阅读理解：感兴趣的研究点但未有产出，有过调研： [Github] [Slides]	
科大讯飞·苏州研究院	2019年10月 - 2020年08月
NLP研究员 苏州	
• 文本摘要：本科毕业设计，为体育新闻生成摘要，设计了一个抽取式与生成式相结合的文本摘要系统。详见专利[2-3]	
• 音乐摘要：详见论文[7]	
• 数据爬取：使用python爬虫爬取互联网资源，并构建知识库	

论文

[1] A Survey on Cross-Lingual Summarization

Jiaan Wang, Fandong Meng, Duo Zheng, Yunlong Liang, Zhixu Li, Jianfeng Qu and Jie Zhou

[代表论文] arXiv preprint 2022.03 (已被TACL有条件接受, i.e., acceptance subject to specific changes) | [\[Paper\]](#)

- 我们撰写了第一篇跨语言摘要综述，其中包括对现有工作的梳理以及对未来工作的展望

[2] ClidSum: A Benchmark Dataset for Cross-Lingual Dialogue Summarization

Jiaan Wang, Fandong Meng, Ziyao Lu, Duo Zheng, Zhixu Li, Jianfeng Qu, Jie Zhou

[代表论文] arXiv preprint 2022.02 (Under Review EMNLP 2022) | [\[Paper\]](#) | [\[Data&Code\]](#) | [\[Huggingface Models\]](#)

- 我们提出了跨语言对话摘要任务并构建了ClidSum基准数据集
- 我们提出了mDialBART预训练语言模型并在ClidSum取得了最好的效果

[3] Pisces: A Pre-Trained Model for Many-to-Many Summarization

Jiaan Wang, Fandong Meng, Duo Zheng, Yunlong Liang, Zhixu Li, Jianfeng Qu, Jie Zhou

[未公开] **Under Review EMNLP 2022**

- 我们提出了多语言对多语言摘要任务：模型输入多种语言文档，输出多种语言摘要。
- 我们提出了Pisces预训练语言模型

[4] Knowledge Enhanced Sports Game Summarization

Jiaan Wang, Zhixu Li, Tingyi Zhang, Duo Zheng, Jianfeng Qu, An Liu, Lei Zhao, Zhigang Chen

[代表论文] In Proc. of **WSDM 2022** (long paper, CCF-B) | [\[Paper\]](#) | [\[Data&Code\]](#)

- 我们提出了K-SportsSum数据集，现有最大且质量最高的体育赛事摘要数据集
- 我们首次考虑到了体育赛事摘要中的知识差距（knowledge gap）问题，并提出了KES模型，在生成体育赛事摘要的过程中引入外部知识库中的信息
- 我们在K-SportsSum与此前广泛使用的SportsSum数据集上实现了最先进的（state-of-the-art）性能
- 该工作中的数据集获得了国内会议**CCKS 2021**的**最佳论文提名奖**

[5] SportsSum2.0: Generating High-Quality Sports News from Live Text Commentary

Jiaan Wang, Zhixu Li, Qiang Yang, Jianfeng Qu, Zhigang Chen, Qingsheng Liu, Guoping Hu

In Proc. of **CIKM 2021** (short paper, CCF-B) | [\[Paper\]](#)

- 发表文章[4]的前身工作
- 提出了体育赛事摘要的三步pipeline解决方案：RES

[6] Incorporating Commonsense Knowledge into Story Ending Generation via Heterogeneous Graph Networks

Jiaan Wang, Beiqi Zou, Zhixu Li, Jianfeng Qu, Pengpeng Zhao, An Liu, Lei Zhao

In Proc. of **DASFAA 2022** (long paper, CCF-B) | [\[Paper\]](#) | [\[Code\]](#)

- 我们提出了故事异质图神经网络模型，将故事上下文建模成不同粒度（文档级、句子级、单词级与概念级）并显式考虑其中的多粒度关系
- 我们首次将多任务学习应用于故事结尾生成任务
- 在被广泛应用的ROCStories数据集上，我们实现了最先进的（state-of-the-art）性能。

[7] Multi-Modal Chorus Recognition for Improving Song Search

Jiaan Wang, Zhixu Li, Binbin Gu, Tingyi Zhang, Qingsheng Liu, Zhigang Chen

In Proc. of **ICANN 2021** (long paper, CCF-C) | [\[Paper\]](#) | [\[Code\]](#)

- 提出了多模态副歌抽取任务，为一首歌抽取其副歌部分的歌词
- 提出了MMCR模型来解决该任务，同时考虑了歌曲的音频与歌词信息

[8] LayerConnect: Hypernetwork-Assisted Inter-Layer Connector to Enhance Parameter Efficiency

Haoliang Shi, Rongsheng Zhang, **Jiaan Wang**, CEN WANG, Guandan Chen, Yinhe Zheng and Tetsuya Sakai

Under Review COLING 2022 (short paper)

[9] RT-KGD: Relation Transition Aware Knowledge-Grounded Dialogue Generation

Kexin Wang, Zhixu Li, **Jiaan Wang**, Jianfeng Qu, Ying He, An Liu and Lei Zhao

In Proc. of **ISWC 2022** (long paper, CCF-B)

[10] Aligning Internal Regularity and External Influence of Multi-Granularity for Temporal Knowledge Graph Embedding

Tingyi Zhang, Zhixu Li, **Jiaan Wang**, Jianfeng Qu, Lin Yuan, An Liu, Lei Zhao and Zhigang Chen

In Proc. of **DASFAA 2022** (long paper, CCF-B)

[11] Enhancing Visual Dialog Questioner with Entity-based Strategy Learning and Augmented Guesser

Duo Zheng, Zipeng Xu, Fandong Meng, Xiaojie Wang, **Jiaan Wang**, Jie Zhou

In Proc. of **EMNLP Findings 2021** (long paper, CCF-B) | [\[Paper\]](#) | [\[Code\]](#)

专利

[1] 文本提取方法、文本提取系统、电子设备和存储装置

李直旭，郑新，支洪平，**王佳安**（学生一作）

CN113505218B | [\[Details\]](#)

[2] 一种体育文本摘要提取方法、装置、存储介质及设备 [发明]

王佳安，李直旭，陈志刚，何莹，郑新（学生一作）

CN111708878B | [\[Details\]](#)

[3] 文本摘要方法、装置、电子设备和存储介质

何莹，李直旭，陈志刚，张兆银，**王佳安**（学生一作）

CN111428025B | [\[Details\]](#)

[4] 剧情文本生成的方法、装置、终端和存储介质 [发明]

张林箭，**王佳安**，邹北琪，张聪，汪硕芃，宋有伟，范长杰，胡志鹏（学生一作）

CN114416921A | [\[Details\]](#)

[5] 意图识别的方法、装置及电子设备 [发明]

张林箭，**王佳安**，邹北琪，张聪，汪硕芃，宋有伟，范长杰，胡志鹏（学生一作）

CN114429142A | [\[Details\]](#)

[6] 人机交互方法、装置、电子设备及存储介质 [发明]

张林箭，邹北琪，**王佳安**，蔡泽锐，张聪，汪硕芃，宋有伟，范长杰，胡志鹏（学生二作）

CN114461775A | [\[Details\]](#)

荣誉奖项

国家奖学金（1/190+）	2021.10
苏州大学研究生学业优秀奖学金特等奖（1/190+）	2020.10
苏州大学本科优秀毕业生（15%）	2020.6
全国高中数学奥林匹克联赛（陕西赛区）一等奖（陕西省第16名）	2015