

# Jiaan Wang

Second-year NLP master student at the Soochow University

✉ jawang1@stu.suda.edu.cn | 📧 krystalan | https://wangjiaan.cn

## Research Interests

(the **bold** denotes the tasks which I am most familiar with)

My research interests lie in Natural Language Processing (NLP) and Machine Learning. I have a broad interest in many NLP tasks. My recent work has focused on **Cross-Lingual Summarization** [1], **Dialogue Summarization** [1], **Sports Game Summarization** [2,4] and Story Generation [3] (see publications). Though most of my publications focus on Natural Language Generation (NLG), I also take interest in many Natural Language Understanding (NLU) tasks, such as machine reading comprehension and question answering (especially **multi-hop QA**, KBQA and commonsense QA). In my spare time, I also participated in my collaborators' work that spans Knowledge Graph Embedding [6] and Visual Dialogue [7].

## Education

### Soochow University

MASTER STUDENT, COMPUTER SCIENCE

Suzhou, China

Sept 2020 - Jun 2023 (Expected)

- Advisor: [Prof. Zhixu Li](#) (now at Fudan University)
- Ranking: 1/190

### Soochow University

B.ENG, COMPUTER SCIENCE & ENGINEERING

Suzhou, China

Sept 2016 - Jun 2020

- Thesis Advisor: [Prof. Zhixu Li](#)
- GPA: 3.7/4.0 (total), Ranking: 8/200+

## Publications

- [1] [CLIDSUM: A Benchmark Dataset for Cross-Lingual Dialogue Summarization](#)  
**Jiaan Wang**, Fandong Meng, Ziyao Lu, Duo Zheng, Zhixu Li, Jianfeng Qu and Jie Zhou  
*arXiv preprint 2022*. [\[Data&Code\]](#)
  - We introduce cross-lingual dialogue summarization (XLDS) task and present CLIDSUM, the first large-scale XLDS benchmark dataset.
  - We propose the mDIALBART model that extends mBART-50 via the second pre-training stage. Our mDIALBART achieves the state-of-the-art performance on CLIDSUM.
  - We elaborately build and evaluate various baselines of different paradigms and give multiple promising directions for future research.
- [2] [Knowledge Enhanced Sports Game Summarization](#)  
**Jiaan Wang**, Zhixu Li, Tingyi Zhang, Duo Zheng, Jianfeng Qu, An Liu, Lei Zhao and Zhigang Chen  
*In Proc. of WSDM 2022 (long paper)*. [\[Data&Code\]](#)
  - We introduce K-SportsSum which is currently the highest quality and largest sports game summarization dataset.
  - We propose knowledge-enhanced summarizer (KES) which take the information of knowledge corpus into account when generating sports news.
  - Our KES achieves a new state-of-the-art performance on both K-SportsSum and SportsSum datasets.
- [3] [Incorporating Commonsense Knowledge into Story Ending Generation via Heterogeneous Graph Networks](#)  
**Jiaan Wang**, Beiqi Zou, Zhixu Li, Jianfeng Qu, Pengpeng Zhao, An Liu and Lei Zhao  
*In Proc. of DASFAA 2022 (long paper)*. [\[Code\]](#)
  - We propose a Story Heterogeneous Graph Network for story ending generation. Our model explicitly considers the story context at different granularity levels and the multi-grained interactive relations among them.
  - We design two auxiliary tasks to facilitate story comprehension.
  - Extensive experiments on widely used ROCStories Corpus show that our model achieves new state-of-the-art performances.
- [4] [SportsSum2.0: Generating High-Quality Sports News from Live Text Commentary](#)  
**Jiaan Wang**, Zhixu Li, Qiang Yang, Jianfeng Qu, Zhigang Chen, Qingsheng Liu and Guoping Hu  
*In Proc. of CIKM 2021 (short paper)*. [\[Data\]](#)
  - SportsSum2.0 is the preliminary study of *Knowledge Enhanced Sports Game Summarization*.
  - In this work, we present SportsSum2.0 dataset together with a reranker-enhanced summarizer.
- [5] [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). [Code]

- We propose a novel task, i.e., Chorus Recognition, aiming at identifying the chorus of a given song. We model Chorus Recognition as a multi-modal task where both lyrics and tune of songs would be taken into account.
- We construct CHORD, the first chorus recognition dataset.
- We propose a novel multi-modal Chorus Recognition model, i.e., MCCR, which achieves the state-of-the-art performance on CHORD.

[6] *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).

[7] *Enhancing Visual Dialog Questioner with Entity-based Strategy Learning and Augmented Guesser*  
Duo Zheng, Zipeng Xu, Fandong Meng, Xiaojie Wang, **Jiaan Wang** and Jie Zhou  
In Proc. of *EMNLP Findings 2021* (long paper). [Code]

## Internships

### WeChat AI, Tencent Inc.

Beijing, China

NLP RESEARCH INTERN (FULL TIME)

2021.09 - Present

- Advised by [Dr. Fandong Meng](#)
- Research on cross-lingual summarization [1].

### Knowledge Works Lab, Fudan University

Remote

NLP RESEARCH INTERN (PART TIME)

2021.09 - Present

- Advised by [Prof. Zhixu Li](#)
- Research on knowledge graph and sports game summarization [2].

### Fuxi AI Lab, NetEase Game

Hangzhou, China

NLP ENGINEERING INTERN (FULL TIME)

2021.06 - 2021.09

- Work with Cong Zhang, Linjian Zhang, Guanying Wang and Beiqi Zou
- Research on story ending generation [3] and other NLP/IR subtasks that could be used in games to enhance the players' game experience.
- Participate in the intelligent dialogue system of *Justice Online* (a MMORPG PC game) and *Wang Chuan Feng Hua Lu* (a MMO mobile game). Specifically, players can raise "children" in *Justice Online* and chat with their "children" through intelligent dialogue systems. The reply content of "children" can be influenced by age, character, gender, history dialogues, etc. For *Wang Chuan Feng Hua Lu*, players can adopt cats and have conversations with them.

### iFLYTEK AI Research

Suzhou, China

NLP RESEARCH INTERN (FULL TIME)

2019.10 - 2020.08 & 2021.03 - 2021.06

- Advised by [Prof. Zhixu Li](#)
- Research on text summarization, sports game summarization [2,4] and music summarization [5].

## Awards

2021.10	<b>National Scholarship</b> , Ranking: 1/190	MOE, China
2020.10	<b>The Grand Prize Scholarship</b> , Ranking: 1/190	Soochow University
2020.06	<b>Excellent Graduates</b> , Top 5%	Soochow University
2019.09	<b>Postgraduate Recommendation Qualification</b> , Ranking: 8/200+	Soochow University
2019	<b>The First-grade Innovation Scholarship</b> , Top 5%	Soochow University
2019	<b>The First-grade Scholarship</b> , Top 10%	Soochow University
2018	<b>The Second-grade Scholarship</b> , Top 25%	Soochow University
2017	<b>The First-grade Scholarship</b> , Top 10%	Soochow University
2015	<b>Provincial First Prize at Chinese Mathematics Olympiad</b> , Ranking: 42nd in Shaanxi province	CMS
2014	<b>Provincial First Prize at Chinese Mathematics Olympiad</b> , Ranking: 16nd in Shaanxi province	CMS

## Skills

- Languages: Chinese (native), English (familiar), Korean (beginner)
- Programming: Python