

ZHOU, EJ

+(86) 188-0671-8188 ✧ +(44) 07919 955758

zhouej4810@gmail.com ✧ GitHub: [Cyber-E-J](#) ✧ Website: [cyber-e-j.github.io](#)

EDUCATION BACKGROUND

University of Oxford, Visiting for Year
St Hilda's College, majoring in Linguistics and Computer Science
Oxfordshire, UK
Oct 2023 - Jun 2024(Expected)

Cornell University, Exchange for semester
College of Engineering, GPA: 4.082/4.3
Ithaca, NY, USA
Jan 2023 - May 2023

Relevant Courses: Computational Linguistics (A)
Language and Information (A) Independent Reading & Research (A+)

Zhejiang University, Bachelor of Science (Honours) in Computer Science
Chu Kochen Honors College
Hangzhou, China
Sept 2020 - Jun 2024(Expected)

- Overall GPA: 3.99/4. Major GPA: 4.0/4.0
- Ranking: Top 2%
- Courses include: Probability and Mathematical Science (99) Fundamentals of Data Structure (99) ...

PUBLICATION

[1] **Yijie Zhou**, Kejian Shi, Wencai Zhang, Yixin Liu, Yilun Zhao, Arman Cohan: *ODSum: New Benchmarks for Open Domain Multi-Document Summarization*, in submission to EACL 2024.

[2] Yulong Chen, Huajian Zhang, **Yijie Zhou**, Xuefeng Bai, Yueguan Wang, Ming Zhong, Jianhao Yan, Yafu Li, Judy Li, Michael Zhu and Yue Zhang: *Revisiting Cross-Lingual Summarization: A Corpus-based Study and A New Benchmark with Improved Annotation*, **accepted by ACL 2023**.

EXPERIENCE

University of Oxford
Project: Hallucination in Multilingual NLP
Research Assistant
Nov 2023 - present
Oxford Internet Institute

- Ongoing research...

Yale University
Project: ODSum, new benchmark for Open-domain MDS [[paper](#)]
Research Assistant, Advisor: [Arman Cohan](#)
Apr 2023 - Oct 2023
YaleNLP Lab

- Worked on bench-marking the open-domain multi-document summarization task, proposed novel streamlined methodology, introduced two new benchmarks that concentrate on distinct use-cases and feature realistic queries coupled with a varied document index, closely mirroring real-world applications. Employed various models from pretrained models to LLMs and examined the performance, and proposed methods to better robustness.
- Submitted work as sole first author.

Cornell University
Project: Persuasive Chatbox
Research Assistant
Feb 2023 - Jun 2023
CommCollabTech Lab

- Designed an LLM-driven persuasive agent capable of detecting factual inaccuracies in conversations as in social media scenarios and providing accurate information supported by evidence, deployed on Telegram and Reddit.
- Collaborated with the social computing group. The research was taken for credit and achieved an A+ grade, recognized by professor for innovative project work.

Westlake University

Project: Building ConvSumX, an improved XLS corpus [[paper](#)]

Research Intern, Advisor: [Yue Zhang](#)

Sept 2022 - Feb 2023

WestLakeNLP Lab

- Led the Corpus-Annotation and Corpus-Evaluation team in the creation of the corpus. Conducted experiments by fine-tuning the mBART model using various corpora on NVIDIA A100 GPUs, while implementing shell scripts to significantly reduce workload.
- Delivered comprehensive analysis of existing XLS corpora, annotation requirements and experiment results, in terms of transformer-based model framework as well as multilingualism, morphology and lexicon rules.

Hangzhou City Brain

Intern

Jul 2022 - Sept 2022

Hangzhou, China

- Worked on the maintenance of IRS database system integral to the Hangzhou CityBrain, enabling real-time analysis, visualization and reports of city-wide traffic conditions.

CONTEST, EVENT

OxGen 2023 Summit & Oxford Generative AI Hackathon *Participant & Hacker*

Oct 2023

Developed a package utilizing LLMs to automatically generate summaries of news articles from websites. Implemented data visualization of extracted event details and casualties on an interactive map. Special Mention Award.

Polyglot Conference 2023 Budapest *Fellow Polyglot*

Oct 2023

Connected with amazing language-loving peers. Attended talks and discussed on how to utilize and further train LLMs in the sense of revitalise endangered languages and low-resource languages.

ACL 2023 *Participant*

Jul 2023

Showcased our work of *ConvSumX*.

Mini-SQL [GitHub](#) Coursework for *Database System*; score **98**/100

May 2022 - Jun 2022

Developed from scratch a compact single-user SQL engine (DBMS) supporting fundamental database operations.

The 9th INTEL CUP Parallel Application Challenge *Team Leader & Finalist*

Aug 2021

Optimized the Weighted Back-Projection algorithm by using techniques including multi-thread paralleling, compiler optimization, and cache alignment. Achieved 3,431% acceleration.

AWARDS AND HONORS

- First-class Scholarship, Zhejiang University 2022/2021
- Excellence Honors Scholarship, Zhejiang University (*Top 1% in Honors College*) 2021
- Hengyi Scholarship (*5/~24,000 in ZJU*) 2021

MISC

Polyglot

- Learning languages is my passion and what drew my interest in NLP. I am native in *Wu* and *Mandarin*, TOEFL **115**/120 for *English*, TFU 84/100 for *French*, conversation level in *Japanese* and this year I'm devoting myself to *Russian*.
- I'm part of the global polyglot community. I was a member of *Cornell Language Expansion Program* and I attend annually the *Polyglot Conference* and the *Language Event*.

Cinemaphile, Viewed over 1k films. Attended the *76e Festival de Cannes* and other film festivals. Write film critiques on self-media and *Tide*, a student-run magazine. One of my [film reviews](#). I was in the film *In Some Nights*.

Music: Trained in *WestLake Qin Society* on Guqin since 10. Served as the Vice President of the *ZJU Qin Society*, meanwhile worked as a Guqin teacher for one year. Also like indie-rock and techno.

Leadership: Vice president of the *CKC Student Union* in 2021. Director of the *2021 CKC New Year's Gala*.

Sports: Medal-winner in the university tennis tournament (2021). Member of Oxford Kendo Society.