Fan **Jiang**

(+61) 469-814-650 | fan.jiang1@student.unimelb.edu.au | fantabulous-j.github.io | fantabulous-j | Fan Jiang

Educations

The University of Melbourne — *Melbourne, Victoria, Australia*

PhD in Natural Language Processing — *Feb. 2022 - Sep. 2025 (expected)*

- Supervisors: Prof. Trevor Cohn & Prof. Tom Drummond

The University of Melbourne — *Melbourne, Victoria, Australia*

Master of Science (Computer Science) (with Distinction) — *Jul. 2019 - Jun. 2021*

- GPA: 91.9% (3.94/4.0) First Class Honours
- Supervisor: Prof. Trevor Cohn

Fuzhou University — *Fuzhou, Fujian, China*

Bachelor of Engineering in Software Engineering — *Sep. 2015 - Jun. 2019*

- GPA: 4.07/5.0 (Ranking: 1/155) Graduated with Honours

Research Interest

My primary research interests are centered around natural language processing. My research is motivated by the ultimate goal of enhancing information access for individuals across diverse backgrounds. Recently, I have been concentrating on achieving robust and general information access by overcoming linguistic and domain barriers:

- **General knowledge retrieval systems across domains** – Robust systems capable of retrieving reliable information across various formats and domains are crucial for meeting the information requirements of intricate real-world user queries across diverse fields.
- **Intelligent open-domain question-answering systems across languages** – It is crucial to enhance information accessibility for users from diverse linguistic backgrounds, particularly those who speak underrepresented languages. My emphasis is on developing sophisticated systems capable of aiding linguistic minorities in accessing essential information.

Publications

Preprints — **Franken-Adapter: Cross-Lingual Adaptation of LLMs by Embedding Surgery.**
Fan Jiang, Honglin Yu, Grace Chung and Trevor Cohn.
Under review by the International Conference on Machine Learning (ICML)

TACL′25 — **Few-Shot Multilingual Open-Domain QA from 5 Examples.**
Fan Jiang, Tom Drummond and Trevor Cohn.
Transactions of the Association for Computational Linguistics.

MRL′24 — **Language Bias in Multilingual Information Retrieval: The Nature of the Beast and Mitigation Methods.**
Jinrui Yang, **Fan Jiang**, and Timothy Baldwin.
In Proceedings of the Fourth Workshop on Multilingual Representation Learning.

EMNLP′24 — **Pre-training Cross-lingual Open Domain Question Answering with Large-scale Synthetic Supervision.**
Fan Jiang, Tom Drummond and Trevor Cohn.
In Proceedings of the 2024 Conference on Empirical Methods in Natural Language Processing.

EMNLP-Findings′23 — **Boot and Switch: Alternating Distillation for Zero-Shot Dense Retrieval.**
Fan Jiang, Qiongkai Xu, Tom Drummond and Trevor Cohn.
In Findings of the Association for Computational Linguistics: EMNLP 2023.

EMNLP-Findings′23 — **Noisy Self-Training with Synthetic Queries for Dense Retrieval.**
Fan Jiang, Tom Drummond and Trevor Cohn.
In Findings of the Association for Computational Linguistics: EMNLP 2023.

EACL′23 — **Don't Mess with Mister-in-Between: Improved Negative Search for Knowledge Graph Completion.**
Fan Jiang, Tom Drummond and Trevor Cohn.
Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics.

AAAI'22 **Incorporating Constituent Syntax for Coreference Resolution.**
Fan Jiang and Trevor Cohn.
Proceedings of the 36th AAAI Conference on Artificial Intelligence.

NAACL'21 **Incorporating Syntax and Semantics in Coreference Resolution with Heterogeneous Graph Attention Network.**
Fan Jiang and Trevor Cohn.
Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies.

Research Experience

Google Research Australia *Sydney, NSW, Australia*
Project: Franken-Adapter: Cross-Lingual Adaptation of LLMs by Embedding Surgery. *June. 2024 - Dec. 2024*
Student Researcher Hosts: Trevor Cohn & Honglin Yu

- A modular framework for language adaptation of LLMs by customizing tokenizers for target languages and relearning the embedding layer.
- Achieve zero-shot cross-lingual transfer through model composition: swapping embedding layers of post-trained LLMs.

The University of Melbourne *Melbourne, Victoria, Australia*
Ph.D. Thesis: Semi-Supervised Non-Parametric Information Access Across Languages and Domains *March. 2022 - Present*
Student Research Assistant Advisors: Prof. Trevor Cohn & Prof. Tom Drummond

- Structured knowledge: adapting neural retrievers to knowledge graphs. (EACL 2023)
- Domain adaptation: adapting neural retrievers to zero-resource domains with a novel self-training approach. (EMNLP-Findings 2023)
- General neural retriever: a bootstrapping algorithm for unsupervised retriever training across massive domains. (EMNLP-Findings 2023)
- A unified system for cross-lingual retrieval and multilingual ODQA: a two-stage self-supervised pre-training framework with large-scale synthetic data for strong zero-shot cross-lingual retrieval and multilingual QA performance. (EMNLP 2024)
- Few-shot multilingual ODQA: a data generation approach that employs LLMs for generating synthetic data from 5-shot in-language examples, and can be extended to to zero-shot setting to enable adaption to low-resource languages with only English supervision. (TACL 2025)

The University of Melbourne *Melbourne, Victoria, Australia*
Master Thesis: Towards Syntax and Semantics-Driven Neural Coreference Resolution *Aug. 2020 - Aug. 2021*
Student Research Assistant Advisor: Prof. Trevor Cohn

- Improved a strong neural coreference resolution model by incorporating dependency syntax and semantic role labels using heterogeneous graph attention networks. (NAACL 2021)
- Proposed a novel method to effectively incorporate constituent parse trees to enhance a strong neural coreference resolution model by utilising graph attention networks. (AAAI 2022)

Services

Conference Reviewers

- EACL, EMNLP 2023
- ACL Rolling Review 2024, 2025

Skills

Programming Languages Python, C/C++, Java
Deep Learning Framework PyTorch, JAX

Honors & Awards

Dean's Honours List 2021, Faculty of Science, The University of Melbourne Aug. 2022
AAAI-22 Student Scholarship, Association for the Advancement of Artificial Intelligence Feb. 2022
Melbourne Research Scholarship, Faculty of Engineering and Information Technology, The University of Melbourne Nov. 2021
Outstanding Graduate of 2019 Academic Year, Fuzhou University Jun. 2019