

Dake Zhang

PHD STUDENT IN COMPUTER SCIENCE

DC 3301, 200 University Ave. W., Waterloo, Ontario, Canada N2L 3G1

✉ zhangdake1998@gmail.com | 🌐 zhangdake.tech | 📄 DakeZhang1998 | 📺 zhangdake | 🐦 @ZhangDake1998 | 📄 Google Scholar

Education

David R. Cheriton School of Computer Science, University of Waterloo

PhD Student in Computer Science, supervised by Prof. Mark D. Smucker
Master of Mathematics in Computer Science, supervised by Prof. Mark D. Smucker

GPA: 91.5 / 100

Sep. 2022 - May 2026 (Exp.)

Sep. 2020 - Aug. 2022

School of Computer Science, Wuhan University

Bachelor of Engineering in Software Engineering, supervised by Prof. Bo Han

GPA: 91.4 / 100

Sep. 2016 - May 2020

Academic Experience

Information Retrieval Group & Data Systems Group, University of Waterloo

Graduate Research Assistant supervised by Prof. Mark D. Smucker

Waterloo, Canada

Sep. 2020 - Present

- Developing benchmark datasets, evaluation metrics, and systems to support people evaluating online information [1, 4].
- Co-organized the TREC 2024 Lateral Reading Track with Prof. Mark D. Smucker and Prof. Charles L. A. Clarke [1].
- Modeled the impact of modern search engines on people's health decision-making with a decision prediction AUC of 0.84 [2].
- Created a system to derive correct answers to health questions from the Web and reduce misinformation in search results by 52% [5, 6].
- Helped organize and participated in the TREC 2020-2022 Health Misinformation Track to mitigate misinformation in health searches [7, 8].

Big Data Group, Wuhan University

Undergraduate Research Assistant supervised by Prof. Bo Han

Wuhan, China

Mar. 2018 - May 2020

- Combined Convolutional Neural Networks and Topological Data Analysis techniques to identify unique writing styles of different authors for the authorship attribution task, particularly excelling in the analysis of lengthy texts.
- Utilized Topological Data Analysis to create an Early Warning System for identifying structural disturbances within the Chinese stock market and predicting potential stock market crashes.

EdTeKLA Research Group, University of Alberta

Mitacs Globalink Undergraduate Research Intern supervised by Prof. Carrie Demmans Epp

Edmonton, Canada

Jun. 2019 - Sep. 2019

- Enhanced the visualization of student activities on an online learning platform.
- Designed and implemented an extractive summarization model capable of selecting key summary sentences from posts and replies in the discussion forum, promoting efficient interactions within the learning community.

Professional Experience

Bing Whole Page Optimization (WPO) Team, Microsoft

Applied Scientist Intern supervised by Dr. Peter Xiong

Remote

Apr. 2024 - Aug. 2024

- Streamlined data pipelines with robust validation processes, for training and evaluating recommender systems.
- Analyzed Bing widget user interaction data to inform iterative product design improvements.

Accelerated Neural Technology (ANT) Team, HUAWEI Noah's Ark Lab

Research Scientist Intern mentored by Dr. Michael R. Metel

Montreal, Canada

Jun. 2023 - Aug. 2023

- Implemented quantization methods on HUAWEI PanGu models, achieving inference speed-up with reduced GPU memory usage.
- Designed 4-bit KV-Cache quantization algorithms, reducing GPU memory usage by 40% with performance drops under 5%.

TikTok Recommender Systems Group, ByteDance

Machine Learning Intern mentored by Keyang Zhang and Xuecheng Yang

Beijing, China

May 2020 - Sep. 2020

- Collaborated on adapting the recommender system for the launch of the friend tab.
- Implemented personalization strategies to boost user interactions (e.g., like, comment, and share) among friends.
- Developed machine learning algorithms to predict user intimacy, leading to a 10% increase in friend interactions.

Publications

[1] Overview of the TREC 2024 Lateral Reading Track

Dake Zhang, Mark D. Smucker, Charles L. A. Clarke

To Appear in Text REtrieval Conference (TREC), 2024

- [2] **A Study of Web Search Assisted Health Decision-Making**
Kamyar Ghajar, **Dake Zhang**, Mark D. Smucker, Robin Cohen
Under Review, 2024
- [3] **UWaterlooMDS at TREC 2023: Deep Learning Track and Tip-of-the-Tongue Track**
Dake Zhang
Text REtrieval Conference (TREC), 2023
- [4] **ReadProbe: A Demo of Retrieval-Enhanced Large Language Models to Support Lateral Reading**
Dake Zhang, Ronak Pradeep
arXiv preprint arXiv:2306.07875, 2023
- [5] **Reducing Health Misinformation in Search Results**
Dake Zhang
UWSpace Theses, 2022
- [6] **Learning Trustworthy Web Sources to Derive Correct Answers and Reduce Health Misinformation in Search**
Dake Zhang, Amir Vakili Tahami, Mustafa Abualsaud, Mark D. Smucker
Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval, 2022
- [7] **UWaterlooMDS at the TREC 2022 Health Misinformation Track**
Amir Vakili Tahami, **Dake Zhang**, Mark D. Smucker
Text REtrieval Conference (TREC), 2022
- [8] **UWaterlooMDS at the TREC 2021 Health Misinformation Track**
Mustafa Abualsaud, Irene Xiangyi Chen, Kamyar Ghajar, Linh Nhi Phan Minh, Mark D. Smucker, Amir Vakili Tahami, **Dake Zhang**
Text REtrieval Conference (TREC), 2021

Presentation

TREC 2024 Lateral Reading Track	Rockville, U.S.
Presentation and Workshop Session at TREC 2024	Nov. 2024
TREC 2024 Lateral Reading Track - New Track Planning	Rockville, U.S.
Presentation at TREC 2023	Nov. 2023
Learning Trustworthy Web Sources to Derive Correct Answers and Reduce Health Misinformation in Search	Madrid, Spain
Poster at SIGIR 2022	July 2022

Service

2025	Graduate Student Representative , School Advisory Committee on Appointments	Waterloo, Canada
2025	Co-organizer of Eval4RAG: Workshop on Evaluation of RAG Systems , ECIR 2025	Lucca, Italy
2023, 2024	Graduate Mentor , Computing Research Association UR2PhD Program	Waterloo, Canada
2024	Co-organizer of the TREC 2024 Lateral Reading Track , TREC 2024	Rockville, U.S.
2024	Reviewer , ACM CIKM 2024	Online
2023	Reviewer , ACM ICTIR 2023	Online
2022	Conference Volunteer , ACM SIGIR 2022	Madrid, Spain

Honors & Awards

HONORS

2020	Outstanding Graduate , Wuhan University	Wuhan, China
2017, 2018, 2019	Outstanding Student , Wuhan University	Wuhan, China

AWARDS

2023	First Prize (2k CAD) , 2023 Canadian #AI Misinformation Hackathon	Online
2022	Student Travel Grant (350 EUR) , SIGIR 2022	Madrid, Spain
2021	Globalink Graduate Fellowship (15k CAD) , Mitacs	Vancouver, Canada
2019	Globalink Research Internship (5.7k CAD) , China Scholarship Council & Mitacs	China & Canada
2017, 2019	Scholarship for Exemplary Student (1k CNY) , Wuhan University	Wuhan, China
2018	Scholarship for Outstanding Student (2k CNY) , Wuhan University	Wuhan, China

Skills

Programming Python (Advanced), Java, C++, R, C, JavaScript, PHP, C#
Frameworks PyTorch, Transformers, DeepSpeed
Tools Git, Linux, Jupyter, Slurm, Docker, \LaTeX
Languages Chinese (Native), English (Fluent)