

Dake Zhang

DC 3301, 200 University Avenue West, Waterloo, ON, Canada N2L 3G1

zhangdake1998@gmail.com | zhangdake.tech | [DakeZhang1998](https://github.com/DakeZhang1998) | [zhangdake](https://www.linkedin.com/in/zhangdake) | [Google Scholar Profile](https://scholar.google.com/citations?user=...)

Summary

- PhD candidate in Computer Science at the University of Waterloo focusing on trustworthy AI, especially Retrieval-Augmented Generation (RAG). Extensive experience leading large-scale research initiatives and shipping production-level AI at four leading technology companies.
- Architect and lead two NIST-sponsored TREC evaluation campaigns to benchmark RAG trustworthiness, coordinating 10+ global research teams in the RAG community.
- Translate research to production through four internships at Amazon, Microsoft, Huawei, and ByteDance (recommender systems, model development and optimization, and large-scale data pipelines).

Education

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

GPA: 91.5/100

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

Sep. 2022 - Aug. 2026 (Exp.)

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

Sep. 2020 - Aug. 2022

School of Computer Science, Wuhan University

GPA: 91.4/100

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

Sep. 2016 - May 2020

Professional Experience

Everyday Essential Product Science, Amazon

Vancouver, BC, Canada

Applied Scientist Intern (Mentor: Dr. Amin Banitalebi)

May 2025 - Aug. 2025

- Designed and deployed a SageMaker-based human annotation platform, processing 20K+ product evaluations for product-understanding models serving millions of customers.
- Developed a lightweight ML classifier combining product embeddings with internal signals to identify everyday essential products across Amazon's marketplaces for downstream teams, achieving Claude-Opus-level accuracy with orders-of-magnitude lower inference cost.

Bing Whole Page Optimization (WPO), Microsoft

Remote

Applied Scientist Intern

Apr. 2024 - Aug. 2024

- Engineered a streamlined data validation pipeline for Bing recommender systems (10M+ daily interactions), eliminating a class of data integrity errors that impacted downstream model serving.
- Analyzed 20M+ user sessions across Bing widget interactions to identify strategies for UX and algorithmic improvements.

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

Montreal, QC, Canada

Research Scientist Intern (Mentor: Dr. Michael R. Metel)

Jun. 2023 - Aug. 2023

- Implemented quantization methods on Huawei PanGu models, achieving inference speedups and reduced GPU memory usage.
- Designed and implemented 4-bit KV-cache quantization algorithms for PanGu models, reducing GPU memory footprint by 40% during inference with a negligible (<5%) performance degradation on key benchmarks.

TikTok Recommender Systems, ByteDance

Beijing, China

Machine Learning Intern (Mentors: Keyang Zhang and Xuecheng Yang)

May 2020 - Sep. 2020

- Collaborated on adapting recommender systems for the launch of the Friends tab.
- Implemented personalization strategies to boost friend-to-friend interactions (likes, comments, shares).
- Developed ML models for user intimacy, leading to a 10% increase in friend interactions in A/B tests.

Research Leadership: Building Benchmarks for AI Evaluation

TREC 2024 Lateral Reading Track & TREC 2025 DRAGUN Track

Rockville, MD, USA

Co-organizer with Prof. Mark D. Smucker and Prof. Charles L. A. Clarke

Nov. 2023 - Present

- Pioneer and direct two NIST-sponsored international research campaigns (TREC Tracks) to establish benchmarks for RAG systems to help people assess news trustworthiness. Secured competitive NIST sponsorship through a novel research proposal, defining the tasks and methodologies that guide research for 10+ participating teams worldwide.
- Architected and deployed the end-to-end evaluation infrastructure, including the creation of benchmark datasets (80 rubrics, 5K+ human annotations), development of evaluation metrics, and the release of an open-source baseline system adopted by four research groups.
- Orchestrate the complete research lifecycle from ideation to publication, authoring calls for participation, managing coordination with NIST assessors, presenting results at conferences, and writing comprehensive overview papers that synthesize community-wide findings and set future research directions.

Academic Research

Information Retrieval Group & Data Systems Group, University of Waterloo

Waterloo, ON, Canada

Graduate Research Assistant supervised by Prof. Mark D. Smucker

Sep. 2020 - Present

- Conducted a user study to characterize the questions people have when assessing news trustworthiness, informing the track design [1].
- Built an interpretable predictive model for health decision-making using search interaction features, achieving an AUC of 0.86 for predicting user choices, demonstrating the measurable impact of search engines on critical health decisions [3].
- Created a health information verification system that automatically derives correct answers from the web, reducing exposure to health misinformation in search results by 52% [7].
- Helped organize the TREC Health Misinformation Track (2020-2022), building benchmark datasets for IR systems to reduce misinformation.
- Reviewed papers for AAAI, CIKM, ICTIR, SIGIR, SIGIR-AP, WSDM, and The Web Conference.
- Mentored four undergraduate students through the Computing Research Association UR2PhD Program.


Selected Publications & Manuscripts


- [1]


Evaluation of Question Generation in Support of Trustworthiness Assessment
Dake Zhang, Mark D. Smucker, Charles L. A. Clarke
Under review, 2025
- [2]


Overview of the TREC 2025 DRAGUN Track
Dake Zhang, Mark D. Smucker, Charles L. A. Clarke
To appear at the Text REtrieval Conference (TREC), 2025
- [3]

A Study of Web Search Assisted Health Decision-Making
Kamyar Ghajar, **Dake Zhang**, Mark D. Smucker, Robin Cohen
Under review, 2025
- [4]

Workshop on Evaluation of Retrieval-Augmented Generation Systems 
Eugene Yang, Ronak Pradeep, **Dake Zhang**, Sean MacAvaney, Maria Maistro, Mohammad Aliannejadi
47th European Conference on Information Retrieval, 2025
- [5]

Overview of the TREC 2024 Lateral Reading Track 
Dake Zhang, Mark D. Smucker, Charles L. A. Clarke
Text REtrieval Conference (TREC), 2024
- [6]

ReadProbe: A Demo of Retrieval-Enhanced Large Language Models to Support Lateral Reading 
Dake Zhang, Ronak Pradeep
arXiv:2306.07875, 2023
- [7]

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

Full publication list: zhangdake.tech/publications

Honors & Awards

2023	First Prize (2K CAD): ReadProbe [6], 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
2020	Outstanding Graduate , Wuhan University	Wuhan, China
2019	Globalink Research Internship (5.7K CAD) , China Scholarship Council & Mitacs	China & Canada
2017, 2018, 2019	Outstanding Student , Wuhan University	Wuhan, China
2017, 2019	Scholarship for Exemplary Student (1K CNY) , Wuhan University	Wuhan, China
2018	Scholarship for Outstanding Student (2K CNY) , Wuhan University	Wuhan, China

Technical Skills

Programming & Data	Python, Java, C++, R, SQL, Pandas, NumPy
ML Libraries	PyTorch, Transformers, DeepSpeed, Scikit-learn
LLM & RAG Expertise	Fine-tuning, Information Retrieval, Prompt Engineering, Quantization, Multi-agent Systems
Research & Evaluation	Experimental Design, Statistical Analysis, User Studies, Human Evaluation & Alignment, Interpretability
Platforms & Tools	AWS, Azure, Slurm, Git, Docker, Jupyter, LaTeX