

RESEARCH INTERESTS

Natural Language Processing

Summarization, Generation, Low-Resource Languages

Generative AI

Large Language Models, Multimodal Foundation Models, Agentic AI

AI4Science

Healthcare, Linguistics, Sensing, Education

ACADEMIC EMPLOYMENT

University of North Carolina at Charlotte, School of Data Science

– Tenure-track Assistant Professor Aug. 2026–Present

University of Hawaii at Manoa, Department of Information and Computer Sciences

– Affiliate Faculty Aug. 2026–Present

– Tenure-track Assistant Professor Aug. 2024–Aug. 2026

EDUCATION

University of California, Davis

Ph.D. in Computer Science

Thesis: Building Intelligent and Reliable Summarization Systems

Jun. 2024
Advisor: Dr. Jiawei Zhang

Georgia Institute of Technology

M.S. in Computational Science and Engineering

M.S. in Electrical and Computer Engineering

Dec. 2018

University of Illinois Urbana-Champaign

B.E. in Electrical Engineering

May 2016

INDUSTRY EXPERIENCE

Amazon Web Services (AWS) AI Lab

Applied Scientist Intern

Santa Clara, CA

Jun. 2023–Sep. 2023

Megagon Labs

Research Intern

Mountain View, CA

Mar. 2023–Jun. 2023

Tencent AI Lab

Research Intern

Seattle, WA

Jun. 2022–Sep. 2022

Salesforce Research

Research Intern

Remote

May 2021–Aug. 2021

GRANT AND PROJECT FUNDING

Key Personnel, NOAA CIFARM	2026–2029
– Project: <i>Cooperative Institute Fostering Aquaculture Research and Markets</i>	
– Award Amount: \$13.5M (University of Hawai'i Share: \$1.73M; Personal Share: \$460K)	
Site PI, NSF Division of Information and Intelligent Systems (IIS)	2024–2028
– Project: <i>SCH: Personalized Machine Learning for Repeat Adverse Health Events using Novel Multimodal Self-Supervised Pretraining Methods</i>	
– Award Amount: \$1.2M (Personal Share: \$450K)	
co-I, American Medical Association ChangeMedEd	2025–2030
– Project: <i>Developing Culturally Responsive AI for Rural Health: Utilizing Precision Education and Precision Coaching to Strengthen Hawaii's Rural Physician Workforce</i>	
– Award Amount: \$1.1M	
Sole PI, OpenAI Researcher Access Program	2024–2025
– Computing Credits: \$7,500	
Sole PI, Google Cloud Research Credits Program	2024–2025
– Computing Credits: \$5,000	
Sole PI, NSF ACCESS Allocations Program	2025
– Project CIS250940: Exploring Multimodal Language Models for Wearable Signal Understanding and Prediction	
– Computing Credits: 1.5M	
Sole PI, NSF ACCESS Allocations Program	2024
– Project CIS240616: Exploring fine-grained domain adaptation for text summarization	
– Computing Credits: 400K	

AWARDS AND HONORS

Best Contribution Award, EMNLP MRL workshop	2025
UH Manoa Faculty Research Travel Award	2025
UC Davis Graduate Group in Computer Science (GGCS) Research Fellowship	2023
FSU Travel Award	2020
Adelaide D. Wilson Graduate Fellowship	2019
UIUC Russell E. Berthold Scholarship	2015

PUBLICATIONS

Citations: 1,583 (as of July 2026)

* *indicates equal contributions — underscore indicates my Ph.D. students.*

Conference Proceeding

- [1] Haohan Yuan and **Haopeng Zhang**, “Understanding LLM reasoning for abstractive summarization”, in *Findings of the 2026 Conference of the Association for Computational Linguistics (ACL 2026)*, 2026.
- [2] Haohan Yuan, Sukhwa Hong, and **Haopeng Zhang**, “StrucSum: Graph-structured reasoning for long document extractive summarization with LLMs”, in *Findings of the Association for Computational Linguistics: EACL 2026*, 2026, pp. 3708–3721.
- [3] Zizhong Li, **Haopeng Zhang**, and Jiawei Zhang, “Token-level precise attack on RAG: Searching for the best alternatives to mislead generation”, in *Findings of the Association for Computational Linguistics: EACL 2026*, 2026, pp. 3193–3206.
- [4] Yizhu Wen, Nan Zhang, Haohan Yuan, Xun Chen, **Haopeng Zhang**, and Hanqing Guo, “Position: On the risks of generative engine optimization in the era of LLMs”, in *The Forty-Third International Conference on Machine Learning (ICML 2026)*, 2026.
- [5] Haohan Yuan and **Haopeng Zhang**, “DomainSum: A hierarchical benchmark for fine-grained domain shift in abstractive text summarization”, in *Findings of the Association for Computational Linguistics: NAACL*, Albuquerque, New Mexico, Apr. 2025, pp. 2219–2231.
- [6] Haohan Yuan, Siu Cheung Hui, and **Haopeng Zhang**, “A structure-aware generative model for biomedical event extraction”, in *International Conference on Database Systems for Advanced Applications*, Springer, 2025, pp. 257–272.
- [7] Zizhong Li, **Haopeng Zhang**, and Jiawei Zhang, “Learning by ranking: Data-efficient knowledge distillation from black-box LLMs for information retrieval”, in *2025 International Joint Conference on Neural Networks (IJCNN)*, IEEE, 2025, pp. 1–10.
- [8] Kaiying Lin, Hsiyu Chen, and **Haopeng Zhang**, “FormosanBench: Benchmarking low-resource austronesian languages in the era of large language models”, in *Findings of the Association for Computational Linguistics: EMNLP 2025*, Association for Computational Linguistics, Nov. 2025, pp. 16 527–16 539.
- [9] Yili Ren*, **Haopeng Zhang***, Haohan Yuan, Jingzhe Zhang, and Yitong Shen, “Wi-Chat: Large language model-powered Wi-Fi-based human activity recognition”, in *Proceedings of the International Workshop on Environmental Sensing Systems for Smart Cities*, 2025, pp. 8–13.
- [10] **Haopeng Zhang**, Hayate Iso, Sairam Gurajada, and Nikita Bhutani, “XATU: A fine-grained instruction-based benchmark for explainable text updates”, in *Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING)*, Torino, Italia, May 2024, pp. 17 739–17 752.
- [11] Zizhong Li, **Haopeng Zhang**, and Jiawei Zhang, “Unveiling the magic: Investigating attention distillation in retrieval-augmented generation”, in *Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL)*, Mexico City, Mexico, Jun. 2024, pp. 745–754.
- [12] **Haopeng Zhang**, Sangwoo Cho, Kaiqiang Song, Xiaoyang Wang, Hongwei Wang, Jiawei Zhang, and Dong Yu, “Unsupervised multi-document summarization with holistic inference”, in *Findings of the Association for Computational Linguistics: IJCNLP-AACL*, Nusa Dua, Bali, Nov. 2023, pp. 123–133.
- [13] **Haopeng Zhang**, Xiao Liu, and Jiawei Zhang, “Extractive summarization via ChatGPT for faithful summary generation”, in *Findings of the Association for Computational Linguistics: EMNLP*, Singapore, Dec. 2023, pp. 3270–3278.
- [14] **Haopeng Zhang**, Xiao Liu, and Jiawei Zhang, “SummIt: Iterative text summarization via ChatGPT”, in *Findings of the Association for Computational Linguistics: EMNLP*, Singapore, Dec. 2023, pp. 10 644–10 657.

- [15] **Haopeng Zhang**, Xiao Liu*, and Jiawei Zhang, “DiffuSum: Generation enhanced extractive summarization with diffusion”, in *Findings of the Association for Computational Linguistics: ACL*, Toronto, Canada, Jul. 2023, pp. 13 089–13 100.
- [16] **Haopeng Zhang**, Xiao Liu, and Jiawei Zhang, “HEGEL: Hypergraph transformer for long document summarization”, in *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Abu Dhabi, United Arab Emirates, Dec. 2022, pp. 10 167–10 176.
- [17] **Haopeng Zhang**, Semih Yavuz, Wojciech Kryscinski, Kazuma Hashimoto, and Yingbo Zhou, “Improving the faithfulness of abstractive summarization via entity coverage control”, in *Findings of the Association for Computational Linguistics: NAACL*, Seattle, United States, Jul. 2022, pp. 528–535.
- [18] **Haopeng Zhang** and Jiawei Zhang, “Text graph transformer for document classification”, in *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, Online, Nov. 2020, pp. 8322–8327.

Workshop Paper

- [19] Zizhong Li, **Haopeng Zhang**, and Jiawei Zhang, “MMViR: A multi-modal and multi-granularity representation for long-range video understanding”, in *1st International Workshop on Sustainable and Efficient Language, Vision, and Action Models (ACL SELVA)*, 2026.
- [20] **Haopeng Zhang**, “Bridging multimodal and video summarization: A unified survey”, in *Proceedings of The 5th New Frontiers in Summarization Workshop*, 2025, pp. 157–171.
- [21] Yiqiao Huang, Zhaorun Chen, He Qi, **Haopeng Zhang**, Hanchao Yu, and Zhuokai Zhao, “Autonomous multimodal reasoning via implicit chain-of-vision”, in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, 2025, pp. 2988–2997.
- [22] Chengqi Zheng, Jianda Chen, Yueming Lyu, Wen Zheng Terence Ng, **Haopeng Zhang**, Yew-Soon Ong, Ivor Tsang, and Haiyan Yin, “MermaidFlow: Redefining agentic workflow generation via safety-constrained evolutionary programming”, in *Multi-Agent Systems in the Era of Foundation Models workshop at ICML (MAS)*, 2025.
- [23] **Haopeng Zhang**, Xiao Liu, and Jiawei Zhang, “Contrastive hierarchical discourse graph for scientific document summarization”, in *Proceedings of the 4th Workshop on Computational Approaches to Discourse at ACL*, Toronto, Canada, Jul. 2023, pp. 37–47.

Journal

- [24] Ali Kargarandehkordi, Aditi Jaiswal, Agnik Banerjee, Yang Qian, Christopher Slade, Yinan Sun, Tanvir Islam, Hiwot Tadesse, Alexander Kostrinsky-Thomas, Chanhyun Park, Urmimala Sarkar, Elaine Khoong, Nhung Nguyen, Xuhai Xu, Kristina Phillips, Roberto Benzo, **Haopeng Zhang**, Finale Doshi-Velez, and Peter Washington, “Personalised modeling of daily stress and blood pressure reactivity using wearable-based biosignals”, *npj Mental Health Research*, 2026.
- [25] **Haopeng Zhang**, Philip S. Yu, and Jiawei Zhang, “A systematic survey of text summarization: From statistical methods to large language models”, *ACM Computing Surveys*, vol. 57, no. 11, Jun. 2025.
- [26] Kaiying Lin, Abdur Rasool, Saimourya Surabhi, Cezmi Mutlu, **Haopeng Zhang**, Dennis Paul Wall, and Peter Washington, “Aiding LLMs with clinical scoresheets: No improvement in neurobehavioral diagnostic classification from text using basic prompting”, *JMIR AI*, 2025.
- [27] William O’Grady, **Haopeng Zhang**, and Miseon Lee, “The negative concord mystery: Insights from a language model”, *Information*, vol. 16, no. 8, p. 710, 2025.
- [28] **Haopeng Zhang**, “Building intelligent and reliable summarization systems”, Ph.D. dissertation, University of California, Davis, 2024.

Poster

- [29] Bardia Konh, Rex Imanaka, **Haopeng Zhang**, and Mitchell Kamrava, “Haptic simulation workstation for prostate brachytherapy training”, *Brachytherapy*, vol. 24, no. 4, S48, 2025.
- [30] Zizhong Li, **Haopeng Zhang**, and Jiawei Zhang, “Token-level precise attack on RAG: Searching for the best alternatives to mislead generation”, in *Proceedings of the 41st Annual Computer Security Applications Conference (ACSAC) Poster*, 2025.
- [31] Yili Ren*, **Haopeng Zhang***, Haohan Yuan, Jingzhe Zhang, and Yitong Shen, “Wi-Chat: Large Language Model Powered Wi-Fi Sensing”, in *Proceedings of the 23rd ACM International Conference on Mobile Systems, Applications, and Services (MobiSys) Posters*, ACM, 2025.

Preprint

- [32] Ethan Chung, Chuanjun Zheng, Jasper Tan, Jingxi Li, **Haopeng Zhang**, and Huaijin Chen, “Does AI understand imaging? a systematic benchmark of agentic AI for computational imaging tasks”, *arXiv preprint:2607.07189*, 2026.
- [33] Mengqi Shi and **Haopeng Zhang**, “HAVEN: Hierarchically aligned multimodal benchmark for unified video understanding”, *arXiv preprint:2605.19223*, 2026.
- [34] Shouju Wang and **Haopeng Zhang**, “MPCI-Bench: A benchmark for multimodal pairwise contextual integrity evaluation of language model agents”, *arXiv preprint*, 2026.
- [35] MRL 25 authors, “Global PIQA: Evaluating physical commonsense reasoning across 100+ languages and cultures”, *arXiv preprint*, 2025.
- [36] Zizhong Li, **Haopeng Zhang**, and Jiawei Zhang, “A revisit of fake news dataset with augmented fact-checking by ChatGPT”, *arXiv preprint*, 2023.
- [37] Jiawei Zhang, **Haopeng Zhang**, Li Sun, and Congying Xia, “Graph-BERT: Only attention is needed for learning graph representations”, *arXiv preprint*, 2020.

Patent

- [38] **Haopeng Zhang**, Semih Yavuz, Wojciech Kryscinski, Kazuma Hashimoto, and Yingbo Zhou, *Systems and methods for abstractive document summarization with entity coverage control*, US Patent 11741142, Aug. 2023.

PROFESSIONAL SERVICE

University of Hawaii Affiliation and Service

– Cooperating Graduate Faculty, Department of Linguistics	2025–2026
– Affiliate Member, University of Hawaii Cancer Center	2025–2026
– Member, Hawaii Data Science Institute	2024–2026
– Member, NSF EPSCoR Change Hawaii (HI) Initiative	2024–2026
– Graduate Studies Committee Member, ICS department	2024–2026
– Search Committee Member, Department of Information and Computer Sciences	2025–2026
– Search Committee Member, Department of Oceanography	2025–2026
– Data Science Committee, Communication and Information Sciences (CIS) Program	2024–2026

Panel and Proposal Review Service

– Panelist, NSF CISE Program	2025, 2024
------------------------------	------------

Conference and Workshop Service

- Organizer, EMNLP 2025 Workshop on New Frontiers in Summarization (NewSumm) 2025
- Area Chair, ACL Rolling Review (ARR) 2026, 2025
- Area Chair, Conference on Neural Information Processing Systems (NeurIPS) 2026
- Area Chair, Conference on Empirical Methods in Natural Language Processing (EMNLP) 2025
- Area Chair, Annual Meeting of the Association for Computational Linguistics (ACL) 2025
- Minitrack Chair, Americas Conference on Information Systems (AMCIS) 2025

Journal Reviewer

- IEEE/ACM Transactions on Audio, Speech, and Language Processing
- IEEE Transactions on Big Data
- IEEE Transactions on Computational Social Systems
- IEEE Transactions on Circuits and Systems for Video Technology
- IEEE Computer Magazine
- ACM Transactions on Embedded Computing Systems
- ACM Transactions on Social Computing
- Computer Speech and Language (Elsevier)
- Applied Soft Computing Journal
- Artificial Intelligence Review (Springer)
- Expert Systems with Applications (Elsevier)
- Engineering Applications of Artificial Intelligence (Elsevier)
- PeerJ Computer Science
- Frontiers in Artificial Intelligence

Conference Reviewer & Program Committee Member

- Annual Conference of the North American Chapter of the Association for Computational Linguistics (NAACL): 2025, 2024
- Annual Meeting of the Association for Computational Linguistics (ACL): 2024, 2023
- Conference on Empirical Methods in Natural Language Processing (EMNLP): 2024, 2023
- ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD): 2024, 2023
- International Joint Conference on Neural Networks (IJCNN): 2024, 2023
- SIAM International Conference on Data Mining (SDM): 2024
- International Conference on Pattern Recognition (ICPR): 2024
- ACL Rolling Review (ARR): 2024, 2023
- Conference of the European Chapter of the Association for Computational Linguistics (EACL): 2023
- EMNLP 2023 Workshop on New Frontiers in Summarization (NewSumm): 2023
- ACL 2023 Workshop on Computational Approaches to Discourse (CODI): 2024
- AAAI Conference on Artificial Intelligence (AAAI): 2022, 2021

TEACHING EXPERIENCE

[†] denotes graduate-level courses.

Instructor, University of Hawaii at Manoa

- ICS 691G[†] : Natural Language Processing and Large Language Models Fall 2025

- ICS 635[†] : Machine Learning
- ICS 435: Machine Learning Fundamentals
- ICS 661[†] : Advanced Artificial Intelligence

Spring 2025
Spring 2026, Spring 2025
Fall 2024

Teaching Assistant, University of California, Davis

- ECS 189G: Deep Learning

Winter 2022

Teaching Assistant, Florida State University

- COP 5725[†] : Advanced Database Systems
- COP 4531: Complexity and Analysis of Data Structures and Algorithms
- COP 4710: Theory and Structure of Databases

Fall 2021, Spring 2021
Spring 2020
Fall 2020, Fall 2019

Teaching Assistant, Georgia Institute of Technology

- ECE 3450: Semiconductor Devices

Fall 2017, Spring 2017

MENTORSHIP

Underscore indicates my Ph.D. students.

Haohan Yuan

Fall 2024–present

Ph.D. Student, University of Hawaii at Manoa
Research Area: Information Extraction, Summarization
Award: AI at Hawaii Data Science Institute Fellowship
Collaborated Publications: [6], [5], [2], [1]

Cade Kane

Fall 2025–present

Ph.D. Student, University of Hawaii at Manoa
Research Area: Hawaiian language NLP

Shouju Wang

Fall 2025–present

Ph.D. Student, University of Hawaii at Manoa
Internship at Zoom Inc.
Research Area: Agentic AI privacy
Collaborated Publications: [34]

Mengqi Shi

Fall 2025–present

Ph.D. Student, University of Hawaii at Manoa
Research Area: Video Summarization

Kaiying Lin

2024–present

Assistant Research Fellow, Institute of Linguistics, Academia Sinica
Research Area: Low-resource Languages
Collaborated Publication: [8]

Zizhong Li

2023–present

Ph.D. Student, University of California, Davis
Research Area: Retrieval-augmented Generation
Collaborated Publications: [3], [7], [11], [36]

Jolie Ching

Spring 2026–present

Master Student, University of Hawaii at Manoa
Master capstone project

Xiao Liu*2022–2024*

Ph.D. Candidate, University of California, Davis

Research Area: Text Summarization

Collaborated Publications: [14], [13], [15], [23], [16]

Samantha Mallari*2025–2026*

Master Student, University of Hawaii at Manoa

Master project: Benchmarking Vision-Language Models on Cross-View Camera Orientation Estimation