Haopeng Zhang

haopengz@hawaii.edu
309C, 1680 East-West Rd
Honolulu, HI 96822
404-376-1886
website
Google Scholar

RESEARCH INTERESTS

- **Natural Language Processing**
  Summarization, Generation, Information Retrieval

- **Generative AI**
  Large Language Models, Multi-modal Language Models

- **NLP for Social Good**
  Indigenous Culture, Climate Change, Healthcare

ACADEMIC EMPLOYMENT

University of Hawaii at Manoa   Aug. 2024 - Present
Tenure-track Assistant Professor in Information and Computer Sciences
Director of Advanced Language Processing and Human Analytics (ALOHA) Lab

EDUCATION

University of California, Davis   Jun. 2024
Ph.D. in Computer Science   Advisor: Dr. Jiawei Zhang
Thesis: Building Intelligent and Reliable Summarization Systems

Georgia Institute of Technology   Dec. 2018
M.S. in Computational Science and Engineering
M.S. in Electrical and Computer Engineering

University of Illinois Urbana-Champaign   May 2016
B.E. in Electrical Engineering

INDUSTRY EXPERIENCE

Amazon Web Services (AWS) AI Lab   Santa Clara, CA
Applied Scientist Intern (Mentors: Dr. Lan Liu, Dr. Zhiheng Huang)   Summer 2023

Megagon Labs   Mountain View, CA
Research Intern (Mentors: Dr. Hayate Iso)   Spring 2023

Tencent AI Lab   Seattle, WA
Research Intern (Mentors: Dr. Kaiqiang Song, Dr. Yu Dong)   Summer 2022

Salesforce Research   Remote
Research Intern (Mentors: Dr. Semih Yavuz, Dr. Yingbo Zhou)   Summer 2021

Teaching Experience (* Graduate Courses)

- **Instructor@UH Manoa**

ICS 661* - Advanced Artificial Intelligence	Fall 2024
ICS 491 - Generative Artificial Intelligence	Spring 2025

- **Teaching Assistant@UC Davis**

ECS189G - Deep Learning	Winter 2022

- **Teaching Assistant@FSU**

COP 5725* - Advanced Database Systems	Fall 2021, Spring 2021
COP 4531 - Complexity and Analysis of Data Structures and Algorithms	Spring 2020
COP 4710 - Theory and Structure of Databases	Fall 2020, Fall 2019

- **Teaching Assistant@Georgia Tech**

ECE 3450 - Semiconductor Devices	Fall 2017, Spring 2017

Mentorship (* My Students)

- **Haohan Yuan***

  Current: Ph.D. student, UH Manoa
  Research Topic: Information Extraction
  Collaborated Papers: [12]

- **Ziyi Zhuang***

  Current: Ph.D. student, UH Manoa
  Research Topic: Video Summarization

- **Zizhong Li**

  Current: Ph.D. student, UC Davis
  Research Topic: Information Retrieval
  Collaborated Papers: [2], [14], [15]

- **Xiao Liu**

  Current: Ph.D. student, UC Davis
  Research Topic: Summarization
  Collaborated Papers: [4], [5], [7], [8], [9]

Publications

[1] **H. Zhang**, "Building intelligent and reliable summarization systems", Ph.D. dissertation, University of California, Davis, 2024.

[2] Z. Li, **H. Zhang**, and J. Zhang, "Unveiling the magic: Investigating attention distillation in retrieval-augmented generation", in *2024 Annual Conference of the North American Chapter of the Association for Computational Linguistics (**NAACL 2024**)*, 2024.

[3] **H. Zhang**, H. Iso, S. Gurajada, and N. Bhutani, "Xatu: A fine-grained instruction-based benchmark for explainable text updates", in *The 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (**LREC-COLING 2024**)*, 2024.

[4] **H. Zhang**, X. Liu, and J. Zhang, "Summit: Iterative text summarization via chatgpt", in *Findings of the 2023 Conference on Empirical Methods in Natural Language Processing (**EMNLP 2023**)*, Singapore: Association for Computational Linguistics, 2023.

[5] **H. Zhang**, X. Liu, and J. Zhang, "Extractive summarization via chatgpt for faithful summary generation", in *Findings of the 2023 Conference on Empirical Methods in Natural Language Processing (**EMNLP 2023**)*, Singapore: Association for Computational Linguistics, 2023.

[6] **H. Zhang**, S. Cho, K. Song, X. Wang, J. Zhang, and Y. Dong, "Unsupervised multi-document summarization with holistic inference", *The 13th International Joint Conference on Natural Language Processing and the 3rd Conference of the Asia-Pacific Chapter of the Association for Computational Linguistics (**IJCNLP-AACL 2023**)*, 2023.

[7] **H. Zhang**, X. Liu, and J. Zhang, "DiffuSum: Generation enhanced extractive summarization with diffusion", in *Findings of the Association for Computational Linguistics: (**ACL 2023**)*, Toronto, Canada: Association for Computational Linguistics, Jul. 2023, pp. 13 089–13 100.

[8] **H. Zhang**, X. Liu, and J. Zhang, "Contrastive hierarchical discourse graph for scientific document summarization", in *Proceedings of the 4th Workshop on Computational Approaches to Discourse (**ACL CODI 2023**)*, Toronto, Canada: Association for Computational Linguistics, Jul. 2023, pp. 37–47.

[9] **H. Zhang**, X. Liu, and J. Zhang, "HEGEL: Hypergraph transformer for long document summarization", in *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing (**EMNLP 2022**)*, Abu Dhabi, United Arab Emirates: Association for Computational Linguistics, Dec. 2022, pp. 10 167–10 176.

[10] **H. Zhang**, S. Yavuz, W. Kryscinski, K. Hashimoto, and Y. Zhou, "Improving the faithfulness of abstractive summarization via entity coverage control", in *Findings of the Association for Computational Linguistics: (**NAACL 2022**)*, Seattle, United States: Association for Computational Linguistics, Jul. 2022, pp. 528–535.

[11] **H. Zhang** and J. Zhang, "Text graph transformer for document classification", in *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (**EMNLP 2020**)*, Online: Association for Computational Linguistics, Nov. 2020, pp. 8322–8327.

Preprints

[12] H. Yuan*, S. C. Hui, and **H. Zhang**, "A structure-aware generative model for biomedical event extraction", *arXiv preprint arXiv:2408.06583*, 2024.

[13] **H. Zhang**, P. S. Yu, and J. Zhang, "A systematic survey of text summarization: From statistical methods to large language models", *arXiv preprint arXiv:2406.11289*, 2024.

[14] Z. Li, **H. Zhang**, and J. Zhang, "Intermediate distillation: Data-efficient distillation from black-box llms for information retrieval", *arXiv preprint arXiv:2406.12169*, 2024.

[15] Z. Li, **H. Zhang**, and J. Zhang, "A revisit of fake news dataset with augmented fact-checking by chatgpt", *arXiv preprint:2312.11870*, 2023.

[16] **H. Zhang** and J. Zhang, "Centrality meets centroid: A graph-based approach for unsupervised document summarization", *arXiv preprint:2103.15327*, 2021.

[17] J. Zhang, **H. Zhang**, L. Sun, and C. Xia, "Graph-bert: Only attention is needed for learning graph representations", *arXiv preprint:2001.05140*, 2020.

Scholarships and Awards

- Computing credits, NSF ACCESS Allocations program   2024
- UC Davis Graduate Group in Computer Science (GGCS) Research Fellowship   2023
- FSU Travel Award   2020
- FSU Adelaide D. Wilson Graduate Fellowship Endowment Fund   2019
- UIUC Russell E. Berthold Scholarship   2015

Professional Service

- Organizer, the 5th New Frontiers in Summarization Workshop   2024
- Graduate Studies Committee Member, UH Manoa   2024
- Member, Hawaii EPSCoR Change Hawaii (HI)   2024
- Member, Hawaii Data Science Institute   2024
- Panelist, NSF Accelerating Computing-Enabled Scientific Discovery BioMed panel   2024
- **Journal Reviewer**

  IEEE/ACM Transactions on Audio, Speech, and Language Processing
  IEEE Transactions on Big Data
  Artificial Intelligence Review
  Computer Speech & Language
  PeerJ Computer Science
  Frontiers in Artificial Intelligence

- **Conference Reviewer & Program Committee Member**

  SDM 24, EMNLP 23/24, IJCNN 23/24, ACL 23/24, EACL 23, KDD 23/24, AAAI 21/22, NewSumm 23, NAACL 24, CODI 24, ACL Rolling Review 23/24, ICPR 24