

Nan Zhang

(470) 439-7958 | njz5124@psu.edu | [zn1010.github.io](https://github.com/njz5124)

EDUCATION

Pennsylvania State University, University Park, PA 08/2020-Present

- Ph.D. in Informatics (*Co-supervised by Prasenjit Mitra and Rui Zhang*) GPA: 3.91/4.0
- **Research Interests:** Language Model Compression, Text Summarization, Clinical NLP, Machine Learning

Georgia Institute of Technology, Atlanta, GA 08/2018-05/2020

- M.S. in Computational Science and Engineering GPA: 3.69/4.0

Worcester Polytechnic Institute (WPI), Worcester, MA 08/2013-05/2017

- BSc. in Computer Science & Industrial Engineering (double major) GPA: 3.5/4.0 (Honors with Distinction)

PUBLICATION

* denotes equal contribution

[5] **Nan Zhang***, Connor Heaton*, Sean Timothy Okonsky, Prasenjit Mitra, Hilal Ezgi Toraman. PEaCE: A Chemistry-Oriented Dataset for Optical Character Recognition on Scientific Documents. The 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING), 2024

[4] Guangji Bai, Zheng Chai, Chen Ling, Shiyu Wang, Jiaying Lu, **Nan Zhang**, Tingwei Shi, Ziyang Yu, Mengdan Zhu, Yifei Zhang, Carl Yang, Yue Cheng, Liang Zhao. Beyond Efficiency: A Systematic Survey of Resource-Efficient Large Language Models. Preprint.

[3] Yusen Zhang, **Nan Zhang**, Yixin Liu, Alexander Fabbri, Junru Liu, Ryo Kamoi, Xiaoxin Lu, Caiming Xiong, Jieyu Zhao, Dragomir Radev, Kathleen McKeown, Rui Zhang. Fair Abstractive Summarization of Diverse Perspectives. Preprint.

[2] **Nan Zhang**, Yusen Zhang, Wu Guo, Prasenjit Mitra, Rui Zhang. FaMeSumm: Investigating and Improving Faithfulness of Medical Summarization. Conference on Empirical Methods in Natural Language Processing (EMNLP), 2023

[1] **Nan Zhang**, Shomir Wilson, Prasenjit Mitra. STAPI: An Automatic Scraper for Extracting Iterative Title-Text Structure from Web Documents. Language Resources and Evaluation Conference (LREC), 2022.

PROFESSIONAL EXPERIENCE

Research Intern, NEC Labs America, Princeton, NJ 05/2023-08/2023

- Compressed Large Language Models (LLMs) for efficient deployment in the enterprise environment

Student Intern, BNP Paribas, New York, NY 10/2016-12/2016

- Devised the front-end via AngularJS and constructed the SQL database in Microsoft SQL Server

Software Development Intern, Fidelity Investments, Smithfield, RI 05/2016-08/2016

- Used Spring Framework to redesign a web application and created JUnit Test Cases for its Java utility classes

TALKS AND PRESENTATIONS

- “Medical Summarization on Factuality and Readability with Large Language Models”. Poster, The 10th Mid-Atlantic Student Colloquium on Speech, Language and Learning (MASC-SLL 2023). April 22, 2023
- “Transformer-based Optical Character Recognition Model for Chemistry Scholarly Papers”. Poster, Penn State University ICDS Fall 2022 Symposium. October 12, 2022.

SERVICES

- Program Committee of SUKI @ NAACL 2022
- Reviewer, ACL Rolling Review Dec. 2022
- Reviewer, EMNLP 2023
- Review Committee of CoNLL 2023