

# Mingzhe Du

(+65) 9658 2486 / [dumingzhex@gmail.com](mailto:dumingzhex@gmail.com)  
Innovation 4.0, 3 Research Link, Singapore, 117902

## RESEARCH INTERESTS

---

Code Generation / Bias Evaluation and Mitigation / Preference Alignment / Continual Learning

## EDUCATION

---

### Nanyang Technological University

Singapore

*PhD Candidate in Computer Science, College of Computing and Data Science*

*Aug. 2022 – Present*

- Advisor: Prof. Anh Tuan Luu

### University of Melbourne

Melbourne, Australia

*Master of Information Technology, Faculty of Engineering and Information Technology* *Oct. 2017 – Dec. 2019*

- Specialization: Distributed Computing (Honorable Dissertation)
- NTU Exchange Scholarship.

### Central South University

Changsha, China

*Bachelor of Information Security, School of Computer Science and Engineering*

*Sep. 2013 – Jun. 2017*

- Excellent Scholarships (*GPA: 3.7/4.0*)

## PROFESSIONAL EXPERIENCE

---

### National University of Singapore, *Institute of Data Science*

*Aug. 2021 – Present*

- Research Associate (Advisor: Prof. See-Kiong Ng).
- Affiliated at Cisco-NUS Accelerated Digital Economy Corporate Laboratory.

### ByteDance, *TikTok*

*Dec. 2019 – Aug. 2021*

- Software Architect
- Implemented a high-performance NLP operator library in C++.
- Designed a general search engine, now powering the search services of TikTok and TouTiao.

### Microsoft, *Asia-Pacific Research and Development Group*

*Nov. 2018 – Mar. 2019*

- Research Intern (Mentor: Wei Liu).
- Contributed features to XiaoIce chatbot.
- Developed a text-to-image transformer that converts textual sentences into visually representative images.

### Alibaba, *Ant Finance Core Clearing Team*

*Jul. 2016 – Oct. 2016*

- Software Engineer Intern.
- Contributed to a core system that supports Alipay payment account clearing flow, merging millions of orders in a minute-level interval for exchanging data with bank partners.

### Tencent, *Social Network Group*

*Mar. 2016 – Jun. 2016*

- Software Engineer Intern.
- Implemented a URL flow processing platform to analyze and detect suspicious URL content in public message traffic, enabling the filtering of malicious information.

## ACADEMIC SERVICES

---

- **Conference Reviewer:** *AISTATS 2024, ICLR 2024, NeurIPS 2024, ECAI 2024, CLEF 2023*
- **Journal Reviewer:** *ACM Computing Survey*
- **Volunteer:** *WWW 2024, EMNLP 2024, AAAI 2023.*
- **STEM Mentor:** *University of Melbourne (2019-Present).*

## PUBLICATIONS

---

1. **Du, M.**, Luu, A. T., Ji, B., Liu, Q., & Ng, S. K. (2024). Mercury: A code efficiency benchmark for code large language models. In *The Thirty-Eighth Conference on Neural Information Processing Systems Datasets and Benchmarks Track (NeurIPS'24)*.
2. **Du, M.**, Luu, A. T., Ji, B., & Ng, S. K. (2024, March). From Static to Dynamic: Knowledge Metabolism for Large Language Models. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 38, No. 21, pp. 23784-23786) (**AAAI'24**).
3. Ji, B., Liu, H., **Du, M.**, & Ng, S. K. (2024, March). Chain-of-Thought Improves Text Generation with Citations in Large Language Models. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 38, No. 16, pp. 18345-18353) (**AAAI'24**).
4. Huang, D., Zhang, J. M., **Du, M.**, Harman, M., & Cui, H. (2024). Rethinking the Influence of Source Code on Test Case Generation. *arXiv preprint arXiv:2409.09464* (**Preprint**).
5. Gollapalli, S. D., Ang, B. H., **Du, M.**, & Ng, S. K. (2024). Counseling Responses for Mental Health Forum Questions with Early Maladaptive Schema Prediction. In *ECAI 2024* (pp. 2556-2563). IOS Press. (**ECAI'24**).
6. **Du, M.**, Hakam, M., Ng, S. K., & Bressan, S. (2023). Constituency-Informed and Constituency-Constrained Extractive Question Answering with Heterogeneous Graph Transformer. In *Transactions on Large-Scale Data-and Knowledge-Centered Systems LIII* (pp. 90-106). Berlin, Heidelberg: Springer Berlin Heidelberg. (**TLDK'23**).
7. Gollapalli, S. D., **Du, M.**, & Ng, S. K. (2023, April). Identifying Checkworthy Cure Claims on Twitter. In *Proceedings of the ACM Web Conference 2023* (pp. 4015-4019). (**WWW'23**).
8. Gollapalli, S. D., **Du, M.**, & Ng, S. K. (2023, September). Generating Reflective Questions for Engaging Gallery Visitors in ArtMuse. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 37, No. 13, pp. 16434-16436) (**AAAI'23**).
9. **Du, M.**, Gollapalli, S. D., & Ng, S. K. (2022). NUS-IDS at CheckThat! 2022: Identifying Check-worthiness of Tweets using CheckthaT5. In *CLEF (Working Notes)* (pp. 468-477) (**CLEF'22**).

## AWARDS

---

- **First Runner-Up Award** at the *Singapore Healthcare AI Datathon 2022*.
- **First Prize** at the *Catalyst 2019: TREKK* (A Path Recommendation System based on data aggregation and analysis).
- **First Prize** at the *Codebrew 2019: After Dark* (Finding safe and fast routes for pedestrians at night).
- **First Prize** at *The Mathematical Contest in Modeling (MCM, 2014)*: Study on the wreckage position of Malaysia Airlines flight *MH370*.
- **Second Prize** at the *8th National Information Security Competition (2015)*: An efficient express system based on hierarchical encryption.
- **Second Prize** at the *Second National Internet of Things Application Innovation Competition (2015)*: Smart Dormitory Remote-Control System based on Raspberry Pi.
- **Technology Pioneer Award** at the *Cyberspace Security Talent Development and Entrepreneurship Forum (2015)*.
- **First Prize** at the *National Internet of Things Competition (2015)*.
- **First Prize** at the *Central South University ACM School Competition (2013)*.

## MISCELLANEOUS

---

- **Personal Website:** <https://mingzhe.space>
- **LinkedIn:** *Mingzhe's LinkedIn*
- **WeChat:** *spirit\_song*
- **GitHub:** <https://www.github.com/elfsong>
- **Work Email:** [mingzhe@nus.edu.sg](mailto:mingzhe@nus.edu.sg) / [mingzhe001@ntu.edu.sg](mailto:mingzhe001@ntu.edu.sg)