

# MINGZHE DU

(+65) 9658 2486 | [dumingzhe@gmail.com](mailto:dumingzhe@gmail.com) | <https://mingzhe.space>  
Innovation 4.0, 3 Research Link, Singapore, 117602

## RESEARCH INTERESTS

---

Code Generation / Preference Alignment / Bias Evaluation and Mitigation

## EDUCATION

---

### Nanyang Technological University

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

- Advisor: Prof. Anh Tuan Luu

Singapore

*Aug. 2022 – Present*

### The University of Melbourne

*Master of Information Technology, Faculty of Engineering*

- Advisor: Prof. Richard O. Sinnott
- Specialization: Distributed Computing (Honorable Dissertation)

Melbourne, Australia

*Oct. 2017 – Dec. 2019*

### Central South University

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

- Excellent Scholarships (GPA: 3.7/4.0)

Changsha, China

*Sep. 2013 – Jun. 2017*

## PROFESSIONAL EXPERIENCE

---

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

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

*Aug. 2021 – Present*

### ByteDance, *TikTok Search*

- Software Architect (Full Time)
- Architected a scalable, general-purpose search engine (*Search Eternal*) that now powers core search services for major ByteDance platforms, including *Douyin*, *TikTok* and *TouTiao*.

*Dec. 2019 – Aug. 2021*

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

- Research Intern (Mentor: Wei Liu)
- Designed and implemented a Transformer-based model for text-to-image synthesis, generating visually coherent images from textual descriptions.

*Nov. 2018 – Mar. 2019*

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

- Software Engineer Intern
- Developed components for a core system in Alipay's payment clearing flow, enabling the aggregation of millions of orders at minute-level latency for data exchange with banking partners.

*Jul. 2016 – Oct. 2016*

### Tencent, *Social Network Group*

- Software Engineer Intern
- Built and deployed a URL analysis platform to process and detect suspicious content in high-volume public message traffic, significantly improving the filtering of malicious information to enhance content security.

*Mar. 2016 – Jun. 2016*

## ACADEMIC SERVICES

---

- **Conference Reviewer:** *NeurIPS (24,25), ICLR (24,25,26), ICML (25,26), AISTATS (24,26), AAAI'26, AISI'26, CIKM'25, ECAI'25, NAACL'24, EMNLP'24, ECAI'24, CLEF'23.*
- **Journal Reviewer:** *ACM Computing Survey, Frontiers in Artificial Intelligence.*
- **Volunteer:** *NeurIPS'25, ACL'25, WWW'24, EMNLP'24, AAAI'23.*
- **STEM Mentor:** *The University of Melbourne (2019-Present).*

## SELECTED PUBLICATIONS

---

[[NeurIPS'25](#)] [Afterburner: Reinforcement Learning Facilitates Self-Improving Code Efficiency Optimization](#).  
Mingzhe Du, Luu Anh Tuan, Yue Liu, Yuhao Qing, Dong Huang, Xinyi He, Qian Liu, Zejun Ma, See-kiong Ng

[[ACL'25](#)] [CodeArena: A collective evaluation platform for LLM code generation](#).

Mingzhe Du, Luu Anh Tuan, Bin Ji, Xiaobao Wu, Dong Huang, Terry Yue Zhuo, Qian Liu, See-Kiong Ng

[[NeurIPS'24](#)] [Mercury: A Code Efficiency Benchmark for Code Large Language Models](#).

Mingzhe Du, Luu Anh Tuan, Bin Ji, Qian Liu, See-Kiong Ng

## PUBLICATIONS

---

[[Preprint](#)] [Nexus: Execution-Grounded Multi-Agent Test Oracle Synthesis](#)

Dong Huang, Mingzhe Du<sup>†</sup>, Jie M. Zhang, Zheng Lin, Meng Luo, Qianru Zhang, See-kiong Ng.

[[Preprint](#)] [SWE-Perf: Can Language Models Optimize Code Performance on Real-World Repositories?](#)

Xinyi He, Qian Liu, Mingzhe Du, Lin Yan, Zhijie Fan, Yiming Huang, Zejian Yuan, Zejun Ma.

[[Preprint](#)] [Secure Code Generation via Online Reinforcement Learning with Vulnerability Reward Model](#).

Tianyi Wu, Mingzhe Du, Yue Liu, Chengran Yang, Terry Yue Zhuo, Jiaheng Zhang, See-Kiong Ng.

[[Preprint](#)] [Scaling Code LLM Training and Test-Time Inference via Execution-Free Reward Models](#).

Xiao Zhu, Xinyu Zhou, Boyu Zhu, Hanxu Hu, Mingzhe Du, Haotian Zhang, Huiming Wang, Zhijiang Guo.

[[Preprint](#)] [Paper Espresso: From Paper Overload to Research Insight](#).

Mingzhe Du, Luu Anh Tuan, Dong Huang, See-kiong Ng.

[[ACL'26](#)] [Semantics-Aligned, Curriculum-Driven, and Reasoning-Enhanced Vulnerability Repair Framework](#)

Chengran Yang, Ting Zhang, Jinfeng Jiang, Xin Zhou, Haoye Tian, Mingzhe Du, Jieke Shi, Junkai Chen, Yikun Li, Eng Lieh Ouh, Lwin Khin Shar, David Lo.

[[TOSEM'26](#)] [Benchmarking LLMs for Unit Test Generation from Real-World Functions](#).

Dong Huang, Jie M. Zhang, Mark Harman, Qianru Zhang, Mingzhe Du<sup>†</sup>, See-Kiong Ng.

[[EACL'26](#)] [Pro-QuEST: Prompt-chaining Quiz Engine for testing Specialized Technical Product Knowledge](#)

Sujatha Das Gollapalli, Mouad Hakam, Mingzhe Du, See-Kiong Ng, Mohammed Hamzeh.

[[AMIYA'26](#)] [Improving Arabic Dialectness in LLMs with Reinforcement Learning](#)

Sujatha Das Gollapalli, Mouad Hakam, Mingzhe Du, See-Kiong Ng, Mohammed Hamzeh.

[[ICLR'26](#)] [Beyond Prompt-Induced Lies: Investigating LLM Deception on Benign Prompts](#). (Oral)

Zhaomin Wu, Mingzhe Du, See-Kiong Ng, Bingsheng He.

[[EMNLP'25](#)] [On Assigning Product and Software Codes to Service Requests with Large Language Models](#).

Sujatha Das Gollapalli, Mouad Hakam, Mingzhe Du, See-Kiong Ng, Mohammed Hamzeh.

[[NeurIPS'25](#)] [EffiBench-X: A Multi-Language Benchmark for Measuring Efficiency of LLM-Generated Code](#).

Yuhao Qing\*, Boyu Zhu\*, Mingzhe Du\*, Zhijiang Guo, Terry Yue Zhuo, Qianru Zhang, Jie M. Zhang, Heming Cui, Siu-Ming Yiu, Dong Huang, See-Kiong Ng, Luu Anh Tuan.

[[NeurIPS'25](#)] [Guardreasoner-VL: Safeguarding VLMs via Reinforced Reasoning](#).

Yue Liu, Shengfang Zhai, Mingzhe Du, Yulin Chen, Tri Cao, Hongcheng Gao, Cheng Wang, Xinfeng Li, Kun Wang, Junfeng Fang, Jiaheng Zhang, Bryan Hooi.

[[ICSE'25](#)] [Measuring the Influence of Incorrect Code on Test Generation](#).

Dong Huang, Jie M. Zhang, Mark Harman, Mingzhe Du<sup>†</sup>, and Heming Cui.

[[ACL'25](#)] [AntiLeakBench: Preventing Data Contamination by Automatically Constructing Benchmarks with Updated Real-World Knowledge](#). (Oral and SAC Highlight Award)

Xiaobao Wu, Liangming Pan, Yuxi Xie, Ruiwen Zhou, Shuai Zhao, Yubo Ma, **Mingzhe Du**, Rui Mao, Anh Tuan Luu, and William Yang Wang.

[[ICML'25](#)] [Position: Current Model Licensing Practices are Dragging Us into a Quagmire of Legal Noncompliance.](#) (Oral)

Moming Duan\*, **Mingzhe Du**\*, Rui Zhao, Mengying Wang, Yinghui Wu, Nigel Shadbolt, and Bingsheng He.

[[AAAI'25](#)] [Towards Verifiable Text Generation with Generative Agent.](#) (Oral)

Ji, Bin, Huijun Liu, **Mingzhe Du**, Shasha Li, Xiaodong Liu, Jun Ma, Jie Yu, and See-Kiong Ng.

[[SAC'25](#)] [Curriculum Demonstration Selection for In-Context Learning.](#)

Duc Anh Vu, Cong-Duy Nguyen, Xiaobao Wu, Nhat Hoang, **Mingzhe Du**, Thong Nguyen, and Anh Tuan Luu.

[[JMIR'25](#)] [Unraveling Online Mental Health Through the Lens of Early Maladaptive Schemas: AI-Enabled Content Analysis of Online Mental Health Communities.](#)

Beng Heng Ang, Sujatha Das Gollapalli, **Mingzhe Du**, and See-Kiong Ng.

[[AAAI'24](#)] [From Static to Dynamic: Knowledge Metabolism for Large Language Models.](#)

**Mingzhe Du**, Anh Tuan Luu, Bin Ji, and See-Kiong Ng.

[[AAAI'24](#)] [Chain-of-Thought Improves Text Generation with Citations in Large Language Models.](#)

Bin Ji, Huijun Liu, **Mingzhe Du**, and See-Kiong Ng.

[[ECAI'24](#)] [Counseling Responses for Mental Health Forum Questions with Early Maladaptive Schema Prediction.](#)

Sujatha Das Gollapalli, Beng Heng Ang, **Mingzhe Du**, and See-Kiong Ng.

[[TLDK'23](#)] [Constituency-Informed and Constituency-Constrained Extractive Question Answering with Heterogeneous Graph Transformer.](#)

**Mingzhe Du**, Mouad Hakam, See-Kiong Ng, and Stéphane Bressan.

[[WWW'23](#)] [Identifying Checkworthy Cure Claims on Twitter.](#)

Sujatha Das Gollapalli, **Mingzhe Du**, and See-Kiong Ng.

[[AAAI'23](#)] [Generating Reflective Questions for Engaging Gallery Visitors in ArtMuse.](#)

Sujatha Das Gollapalli, **Mingzhe Du**, and See-Kiong Ng.

[[CLEF'22](#)] [NUS-IDS at CheckThat! 2022: Identifying Check-worthiness of Tweets using CheckthaT5.](#)

**Mingzhe Du**, Sujatha Das Gollapalli, and See-Kiong Ng.

† *Corresponding Author*, \* *Equal Contribution*.

## AWARDS & HONORS

---

- **NVIDIA Academic Grant:** *Predicting and Optimizing LLVM Compiler Pass Order in a Unified Model.*
- **Thinking Machines Lab Research Grant:** *Scaling Reinforcement Learning for Efficient Code Generation.*
- **Staff Research Award** at *Institute of Data Science, National University of Singapore, 2025.*
- **Google Academic Research Award:** *SafeCodeX: Security-Aware Code Generation with LLMs, 2025.*
- **First Runner-Up Award** at the **Singapore Healthcare AI Datathon:** *Identifying ICU Patient with Central Lines at Risk of CLABSI, 2024.*
- **First Prize** at **Catalyst:** *A Path Recommendation System based on Data Aggregation and Analysis, 2019.*
- **First Prize** at **Codebrew:** *After Dark: Finding Safe and Fast Routes for Pedestrians at Night, 2019.*
- **Technology Pioneer Award** at the *Cybersecurity Talent Development and Entrepreneurship Forum, 2015.*
- **First Prize** at the *National Internet of Things Competition, 2015.*
- **First Prize** at the *Mathematical Contest in Modeling: Study on the Wreckage Position of Malaysia Airlines Flight MH370, 2014.*

- **Second Prize** at the *8th National Information Security Competition: An Efficient Express Label System Based on Hierarchical Encryption, 2014.*

#### MISCELLANEOUS

---

- **LinkedIn:** <https://www.linkedin.com/in/dumingzhe>    **WeChat:** *spirit\_song*
- **GitHub:** <https://www.github.com/elfsong>    **Language:** *Mandarin (native) | English (professional)*