

MING SHAN HEE

✉ E-mail: mingshan_hee@mymail.sutd.edu.sg

🌐 LinkedIn ◊ 🎓 Google Scholar ◊ 🌐 www.mingshanhee.com

EDUCATION

SINGAPORE UNIVERSITY OF TECHNOLOGY AND DESIGN (SUTD) SINGAPORE
PH.D IN INFORMATION SYSTEM AND TECHNOLOGY DESIGN JAN 2021 - AUG 2025 (EXPECTED)

SINGAPORE MANAGEMENT UNIVERSITY (SMU) SINGAPORE
B.SCI. IN INFORMATION SYSTEMS AUG 2017 - DEC 2020

- Summa Cum Laude, cGPA 3.81/4.00 (Dean's List AY 2019 - 2020)

RESEARCH EXPERIENCE

MOHAMED BIN ZAYED UNIVERSITY OF ARTIFICIAL INTELLIGENCE ABU DHABI
VISITING RESEARCHER, ADVISOR: PROF. PRES LAV NAKOV NOV 2024 - APR 2025

- Misinformation Fact-Checking
- Media Bias/Fact-Check
- Visual Question Answering

PUBLICATIONS

- [1] **Ming Shan Hee** and Roy Ka-Wei Lee. "Demystifying Hateful Content: Leveraging Large Multimodal Models for Hateful Meme Detection with Explainable Decisions". In: *Proceedings of the International AAAI Conference on Web and Social Media (ICWSM)*. 2025.
- [2] Yuhao Wu, **Ming Shan Hee**, Zhiqing Hu, and Roy Ka-Wei Lee. "LongGenBench: Benchmarking Long-Form Generation in Long Context LLMs". In: *The Thirteenth International Conference on Learning Representations (ICLR)*.
- [3] **Ming Shan Hee**, Aditi Kumaresan, and Roy Ka-Wei Lee. "Bridging Modalities: Enhancing Cross-Modality Hate Speech Detection with Few-Shot In-Context Learning". In: *Proceedings of the Association for Computational Linguistics: EMNLP (2024)*.
- [4] **Ming Shan Hee**, Shivam Sharma, Rui Cao, Palash Nandi, Preslav Nakov, Tanmoy Chakraborty, and Roy Lee. "Recent Advances in Online Hate Speech Moderation: Multimodality and the Role of Large Models". In: *Findings of the Association for Computational Linguistics: EMNLP (2024)*.
- [5] Yuhao Wu, **Ming Shan Hee**, Zhiqing Hu, and Roy Ka-Wei Lee. "LongGenBench: Benchmarking Long-Form Generation in Long Context LLMs". In: *arXiv preprint arXiv:2409.02076 (2024)*.
- [6] Ri Chi Ng, Nirmalendu Prakash, **Ming Shan Hee**, Kenny Tsu Wei Choo, and Roy Ka-Wei Lee. "SGHateCheck: Functional Tests for Detecting Hate Speech in Low-Resource Languages of Singapore". In: *Proceedings of the 8th Workshop on Online Abuse and Harms (WOAH)*. 2024.
- [7] Ying Ying Lim, **Ming Shan Hee**, Xun Wei Yee, Weng Kuan Yau, Xinming Sim, Wesley Tay, Wee Siong Ng, See-Kiong Ng, and Roy Ka-Wei Lee. "AISG's Online Safety Prize Challenge: Detecting Harmful Social Bias in Multimodal Memes". In: *Companion Proceedings of the ACM Web Conference*. 2024.
- [8] **Ming Shan Hee**, Karandeep Singh, Charlotte Ng Si Min, Kenny Tsu Wei Choo, and Roy Ka-Wei Lee. "Brinjal: A Web-Plugin for Collaborative Hate Speech Detection". In: *Companion Proceedings of the ACM on Web Conference*. 2024.
- [9] Zhiqiang Hu, Thao Thanh Nguyen, Yujia Hu, Chia-Yu Hung, **Ming Shan Hee**, Chun Wei Seah, and Roy Ka-Wei Lee. "Contrastive Disentanglement for Authorship Attribution". In: *Companion Proceedings of the ACM on Web Conference*. 2024.

- [10] Han Wang, Nirmalendu Prakash, Nguyen Khoi Hoang, **Ming Shan Hee**, Usman Naseem, and Roy Ka-Wei Lee. “Prompting large language models for topic modeling”. In: *2023 IEEE International Conference on Big Data (BigData)*. IEEE. 2023.
- [11] Rui Cao, **Ming Shan Hee**, Adriel Kuek, Wen-Haw Chong, Roy Ka-Wei Lee, and Jing Jiang. “Pro-cap: Leveraging a frozen vision-language model for hateful meme detection”. In: *Proceedings of the 31st ACM International Conference on Multimedia*. 2023.
- [12] Nirmalendu Prakash, Han Wang, Nguyen Khoi Hoang, **Ming Shan Hee**, and Roy Ka-Wei Lee. “Promttopic: Unsupervised multimodal topic modeling of memes using large language models”. In: *Proceedings of the 31st ACM International Conference on Multimedia*. 2023, pp. 621–631.
- [13] **Ming Shan Hee**, Aditi Kumaresan, Nguyen Khoi Hoang, Nirmalendu Prakash, Rui Cao, and Roy Ka-Wei Lee. “Matk: The meme analytical tool kit”. In: *Proceedings of the 31st ACM International Conference on Multimedia*. 2023.
- [14] **Ming Shan Hee**, Wen-Haw Chong, and Roy Ka-Wei Lee. “Decoding the underlying meaning of multimodal hateful memes”. In: *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence*. 2023.
- [15] Han Wang, **Ming Shan Hee**, Md Rabiul Awal, Kenny Tsu Wei Choo, and Roy Ka-Wei Lee. “Evaluating gpt-3 generated explanations for hateful content moderation”. In: *Proceedings of the Thirty-Second International Joint Conference on Artificial Intelligence*. 2023.
- [16] Nirmalendu Prakash, **Ming Shan Hee**, and Roy Ka-Wei Lee. “Totaldefmeme: A multi-attribute meme dataset on total defence in singapore”. In: *Proceedings of the 14th Conference on ACM Multimedia Systems*. 2023.
- [17] **Ming Shan Hee**, Roy Ka-Wei Lee, and Wen-Haw Chong. “On explaining multimodal hateful meme detection models”. In: *Proceedings of the ACM Web Conference*. 2022.

AWARDS & FELLOWSHIPS

SINGAPORE DATA SCIENCE CONSORTIUM DISSERTATION RESEARCH FELLOWSHIP	JAN 2023
SUTD PRESIDENT’S GRADUATE FELLOWSHIP	JAN 2021

INVITED TALKS

DECODING HATE: HARNESSING AI FOR MULTIMODAL CONTENT MODERATION

- DSO National Laboratories, Singapore AUG 2024
- AI Wednesdays (Organised by GovTech Singapore & MDDI Singapore), Singapore AUG 2024

SEEING THROUGH HATE: EXPLAINABLE AI IN MULTIMODAL CONTENT MODERATION

- SUTD X Google Lightning Presentations, Singapore MAY 2024

UNDERSTANDING (DARK) HUMOUR WITH INTERNET MEME ANALYSIS

- [Tutorial] ACM Web Conference 2024 (WWW 2024) MAY 2024

ACADEMIC SERVICES

PROGRAM COMMITTEE (REVIEWER)

- EMNLP, 2023
- WW4SG Workshop, 2024
- ASONAM, 2024
- DSAA - Machine Learning Journal Track, 2024
- ACL Rolling Review (ARR), 2024
- TheWebConf (WWW) - Web4Good, 2025
- TheWebConf (WWW), 2024 - 2025
- ACM Multimedia (MM), 2024 - 2025

- ICLR Workshop on Reasoning and Planning for Large Language Models, 2025

CONFERENCE ORGANISING COMMITTEE

- Singapore Symposium on Natural Language Processing 2024 (SSNLP 2024), Local Chair
- Singapore Symposium on Natural Language Processing 2025 (SSNLP 2025), Local Chair

TEACHING ACTIVITIES

SINGAPORE DATA SCIENCE CONSORTIUM (SDSC)

NOV 2024

CO-INSTRUCTOR, 1-DAY TECHNICAL COURSE

- “Online Trust and Safety Series: Leveraging Retrieval-Augmented Generation (RAG) Systems for Content Moderation”

SINGAPORE UNIVERSITY OF TECHNOLOGY AND DESIGN

MAY 2022 - AUG 2022

TEACHING ASSISTANT, 50.040 NATURAL LANGUAGE PROCESSING

- Prepared and graded assignments focusing on Part-Of-Speech (POS) tagging for 30 junior and senior level undergraduates

SINGAPORE POLYTECHNIC

OCT 2019 - DEC 2019

ADJUNCT LECTURER, WEB APPLICATION DEVELOPMENT

- Prepared lectures and assignments focusing on Web application development (HTML, CSS, C#, .NET Core 2.1) for 42 second-year students

SINGAPORE POLYTECHNIC

OCT 2019 - DEC 2019

ADJUNCT LECTURER, APPLICATION DEVELOPMENT

- Prepared lectures and assignments focusing on native Window application development (Windows Form and Windows Presentation Foundation) for 20 first-year students