



## Advances in Large Language Models and Data-Centric AI: Insights from HIT Shenzhen



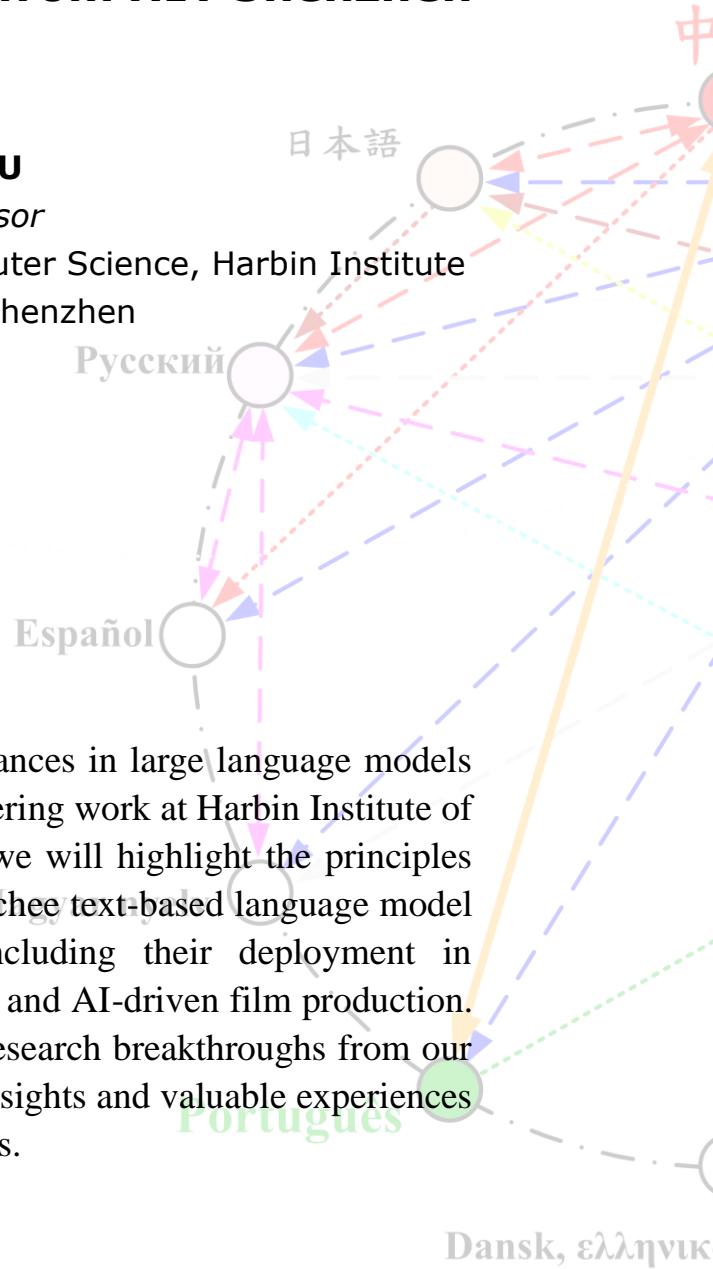
**Prof. Xuebo LIU**

*Assistant Professor*

School of Computer Science, Harbin Institute  
of Technology, Shenzhen

### Abstract

This presentation will explore recent advances in large language models and data-centric AI, drawing on the pioneering work at Harbin Institute of Technology, Shenzhen. In the first part, we will highlight the principles and applications of our self-developed Lychee text-based language model and the Jiutian multimodal model, including their deployment in innovative fields such as cooking robotics and AI-driven film production. The second part will focus on the latest research breakthroughs from our lab in data-centric AI, offering practical insights and valuable experiences for natural language processing researchers.



## Biography

Xuebo Liu is an Assistant Professor at the School of Computer Science, Harbin Institute of Technology, Shenzhen. In 2021, he earned his Ph.D. from the Faculty of Science and Technology at the University of Macau. His research focuses on natural language processing, Data-Centric AI, domain-specific large language models, and machine translation. He has published over 40 papers in top-tier conferences such as ACL, EMNLP, ICLR, and NeurIPS, with more than 30 as the first or corresponding author. He serves as Area Chair for ACL 2022/2024, EMNLP 2022-2024, and NAACL 2024. His awards include the Second Prize of the 2022 Macao Technological Invention Award (second contributor), the 2022 Macao Graduate Science and Technology Innovation Award, and a nomination for the 2022 Outstanding Dissertation Award from the Chinese Information Processing Society.

**Venue: N1-1006**  
**Date: 10-Oct-2024, Thursday**  
**Time: 15:00~16:00**  
**Language: English**

**Contacts:** Prof. Derek F. Wong  
**Telephone:** (853) 8822-4478  
**Website:** <http://nlp2ct.cis.um.edu.mo/>

**Address:**  
E11-2023, University of Macau,  
Avenida da Universidade,  
Taipa, Macau, China