

DEPARTMENT OF COMPUTER SCIENCE

SEMINAR

2025 SERIES

Towards Neural Graph Databases: Data, Algorithms, and Applications

DATE & TIME

26 AUG 2025 (TUE) 10:30 - 11:30 AM

VENUE

JC3 UG04, JOCKEY CLUB CAMPUS OF CREATIVITY



DR. JIAXIN BAI

Postdoctoral Researcher
Hong Kong University of Science and Technology

ABSTRACT

The rapid expansion of AI has exposed critical limitations in how systems process and reason over structured data. As large language models struggle with complex structured knowledge and traditional databases remain disconnected from neural approaches, a fundamental question emerges: What data and reasoning paradigms will power the next generation of more reliable and capable AI systems? This requires reimagining how we represent, query, and reason over structured knowledge at scale. In this talk, I will present my research on building Neural Graph Databases (NGDBs) that bridge this gap through a principled approach to structured data reasoning. First, I will introduce ALTAS, the world's largest LLM-constructed knowledge graph containing billions of facts from diverse sources, which serves as crucial infrastructure for NGDB research while simultaneously functioning as the most extensive GraphRAG dataset available. Second, I will present novel reasoning algorithms that transform how we query structured data, including approaches for expressive multi-modal distribution query embeddings, numerical inference within knowledge graphs, and event-based reasoning with temporal constraints. Finally, I will demonstrate how these advances translate to real-world impact through their application to Amazon Search products, proving their value beyond academic settings. Together, these contributions—constructing comprehensive knowledge infrastructures, developing expressive reasoning algorithms, and implementing secure, practical systems—have defined the emerging field of Neural Graph Databases. As AI continues to evolve toward more factual and reasoning-capable systems, this work establishes a foundation for next-generation AI that can effectively leverage the exponentially growing structured data across domains while maintaining privacy, scalability, and reasoning capabilities.



**SPEAKER'S
BIOGRAPHY**



REGISTER NOW