

Dr Maciej Besta, ETH Zürich

Graphs in the LLM Era: From Effective LLM Ecosystems to Next-Generation Graph Analytics

"Graphs are foundational across domains but remain hard to use without deep expertise. Simultaneously, the emergence of large language models (LLMs) and their most recent reasoning variants such as DeepSeek-R1 brings the promise of novel methods for processing and mining complex graph datasets. In this talk, we first overview and map the emerging fascinating landscape at the overlap of graph processing and LLMs. We then observe that - while LLMs promise accessible natural language (NL) graph analytics - they yet fail to process industry-scale property graphs effectively and efficiently: such datasets are large, highly heterogeneous, structurally complex, and evolve dynamically. To address this, we devise a novel abstraction for complex multi-query analytics over such graphs. Its key idea is to replace brittle generation of graph queries directly from NL with planning over a Semantic Catalog that describes both the graph schema and the graph operations. Concretely, this induces a clean separation between a Semantic Plane for LLM planning and broader reasoning, and an Execution Plane for deterministic, database-grade query execution over the full dataset and tool implementations. This design yields substantial gains in both token efficiency and task effectiveness even with small-context LLMs. We use this abstraction as the basis of the first LLM-enhanced graph analytics framework called GraphSeek. GraphSeek achieves substantially higher success rates (e.g., 86% over enhanced LangChain) and points toward the next generation of affordable and accessible graph analytics that unify LLM reasoning with database-grade execution over large and complex property graphs."