Ao Jiao

+1 (216)507-0653 | contact@joeyjiao.dev | https://github.com/Joey-Jiao

About Me

I'm a second-year M.S. student in Computer Science at Case Western Reserve University, advised by Prof. Yinghui Wu. My research interests lie at the intersection of Graph Machine Learning, Data Management, and Knowledge Discovery. I am primarily focused on developing principled frameworks for Explainable AI, with a specific application to Graph Neural Networks. I also possess proficient full-stack webapp development skills, specializing in architecting complex, customized and interactive data visualization systems.

Education

Case Western Reserve University — Cleveland, OH
M.S. in Computer Science — *Sept. 2024 – Present*

Southern University of Science and Technology — Shenzhen, China
B.E. in Computer Science and Technology — *Sept. 2020 – Jun. 2024*

Publications

- M. Wang*, C. Ma*, **A. Jiao***, T. Liang, P. Lu, S. Hegde, Y. Yin, E. Gurkan-Cavusoglu, Y. Wu, "**Assessing LLMs for Serendipity Discovery in Knowledge Graphs: A Case for Drug Repurposing**", *AAAI Conference on Artificial Intelligence (AAAI)*, 2026. (* equal contributions)
- Y. Huang*, Z. Zhang*, **A. Jiao**, Y. Ma, R. Cheng, "**A Comparative Visual Analytics Framework for Evaluating Evolutionary Processes in Multi-objective Optimization**", *IEEE Transactions on Visualization and Computer Graphics (TVCG)*, 2024. (* equal contributions)

Projects

Explaining the Explainers: A Model-Agnostic Framework for GNN Explainability *Feb. 2025 – Sept. 2025*

- Designed a model-agnostic GNN explainer using a Straight-Through Estimator (STE) to enable discrete subgraph selection without gradient access, compatible with RGCN, HAN, and HGT architectures.
- Developed a relation-aware optimization and visualization system that tracks mask evolution across epochs and decomposes importance by relation type.
- Identified the epoch parameter as a dynamic optimization variable rather than a fixed hyperparameter, providing a novel early-stopping diagnostic that could improves reproducibility and efficiency.
- Introduced a complete metrics quartet for multi-objective evaluation, revealing architecture-specific attention concentration patterns, e.g., HAN's dominance vs. RGCN's balanced spread.
- Provided a critical analysis of STE's gradient mismatch and scalability limits, proposing Gumbel-Softmax and trade-off balance as future extensions toward more interpretable, architecture-agnostic XAI systems.

NL2Graph: A Dual-Method Framework for Natural Language to Graph Query Translation *Sept. 2025 – Present*

- Built a research framework comparing fine-tuned seq2seq (BART) and prompt-based LLM approaches for translating natural language into graph database queries across Cypher, SPARQL, and Gremlin.
- Designed a unified evaluation pipeline (generate → execute → evaluate) with protocol-based polymorphism, enabling both methods to share the same execution, scoring, and reporting infrastructure.
- Engineered thread-safe parallel processing with idempotent staged persistence, handling 400K+ records with timeout protection and graceful degradation.

– Achieved 85.0% accuracy and 90.9% F1 on MetaQA Cypher with DeepSeek-Reasoner, demonstrating that reasoning-capable LLMs with schema-aware prompting can rival fine-tuned seq2seq models on multi-hop knowledge graph QA.

A From-Scratch Interactive Visualization Architecture for Python Profiling *Nov. 2023 – Jun. 2024*

– Engineered a from-scratch visualization framework with a layered rendering and interaction architecture, decoupling logical nodes from visual instances to enable independent control of layout, style, and user interaction states.

– Developed a mid-layer coordination system for layout caching, spatial synchronization, and state propagation across interdependent components, ensuring consistent behavior under continuous interaction.

– Implemented an incremental computation strategy that reuses cached geometric states to eliminate redundant layout evaluations, maintaining real-time responsiveness on large, deeply nested call-stack structures.

– Optimized the rendering pipeline to the browser import-size threshold, achieving a fluid, high-density visualization experience within strict resource constraints.

Kraken: A Distributed Orchestration System for LLM-Driven Knowledge Mining *Jan. 2025 – Jul. 2025*

– Engineered resilient remote execution abstractions with declarative configuration, idempotent execution logic and graceful degradation, validated through production deployment processing large-scale LLM inference workloads across distributed clusters.

– Designed a checksum-based incremental synchronization protocol for code and dependency distribution, enabling rapid iterative deployments across heterogeneous nodes while maintaining consistency under concurrent modification.

– Implemented an index-based workload partitioning strategy coordinating 70-way parallel execution per node, with centralized state management through persistent queues to ensure exactly-once task semantics across long-running workflows.

Services

Case Western Reserve University Cleveland, OH

Teaching Assistant / Lecturer, CSDS 237: Python Programming *Jan. 2025 – May. 2025*

– Delivered 6 full lectures during the latter half of the semester, including data persistence, exception and control flow, basic GUI design and introduction to ML.

– Graded assignments and conducted office hours, providing academic support to students.

Case Western Reserve University Cleveland, OH

Judge, Intersection Poster Symposium *Nov. 2025*

– Evaluated undergraduate research posters and provided feedback at the Intersection Poster Symposium.