

Chen Li, UC Irvine



Title: Thinking in DAGs: How Humans and Agents Collaborate in Data Science

Abstract: As LLM agents increasingly participate in data science, a critical question emerges: what abstraction best supports humans and agents in jointly constructing, understanding, and revising analyses? While scripts and notebooks are expressive and detailed, they often obscure data dependencies within mutable state, overwhelm agents with excessive low-level code, and hinder human inspection of intermediate steps. However, natural language, despite its ease of use, frequently suffers from inherent ambiguity and imprecision.

This talk proposes the directed acyclic graph (DAG) as a superior abstraction for human-agent collaborative data science. By modeling analyses as operators connected by explicit data dependencies, DAGs provide a compact structure for agent reasoning, a visual interface for human oversight, and a foundation for execution, optimization, and provenance. We illustrate this approach with our work called BobFlow, demonstrating how agents can iteratively construct dataflows to solve data science tasks and manage the context using dataflow-level metadata. We show how an analyst easily understands the agent's behaviors through an intuitive dataflow-based interface, how the analyst efficiently inspects the agent's past actions to provide feedback, and how the agent finishes a task with high accuracy and low cost. We also address key implementation challenges, including seamless migration between scripts and workflows, efficient caching of intermediate results, managing reproducible environments, and control blocks such as IfElse and ForLoop. We showcase our progress within Apache Texera (Incubating), demonstrating how these techniques foster transparent, efficient, and trustworthy human-agent collaboration on data science.

Bio: Prof. Chen Li is a professor in the Department of Computer Science at UC Irvine. He received his Ph.D. degree in Computer Science from Stanford University, and his M.S. and B.S. in Computer Science from Tsinghua University, China. His research interests are in the fields of data management, data science, AI/ML, databases, data-intensive computing, search, and visualization. He was a co-founder and CTO of a startup to commercialize his research. He was a recipient of an NSF CAREER award, an ACM Distinguished Member, and an IEEE fellow.

Yunyao Li, Adobe



Title: Rethinking Graphs for the Agentic Enterprise

Abstract: Knowledge graphs have long provided structured representations of the entities, relationships, and context underlying complex enterprises. Large language models and AI agents are now challenging this paradigm: how graphs are constructed and represented is changing, while agents increasingly need to reason over, update, and act through them.

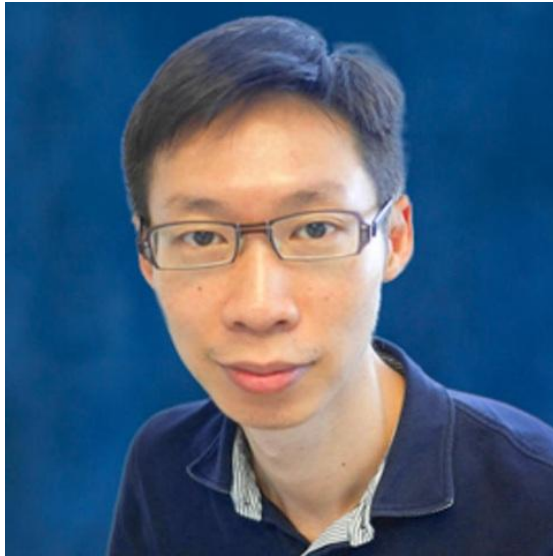
In this keynote, I will explore the emerging interplay between knowledge graphs and agents: how agents can construct and evolve graphs from heterogeneous enterprise data; how graphs can provide grounding, context, provenance, and constraints for agent reasoning and action; and how these systems can support dynamic workflows with human oversight. I will conclude with research challenges at the intersection of data, graphs, agents, and human-AI systems needed to move reliably from representation to action.

Bio: Yunyao Li is Director of Machine Learning at Adobe Experience Platform, where she leads initiatives in agentic AI for enterprise applications. Previously, she led Machine Learning for Apple Knowledge Platform and was a Distinguished Research Staff Member and Senior Research Manager at IBM Research–Almaden.

Her work spans enterprise AI, natural language processing, search, knowledge systems, and databases. She has published over 100 papers, holds 36 patents, co-authored two books, and advises universities, startups, and industry organizations on AI strategy and innovation.

Yunyao is an ACM Distinguished Member and was a member of the inaugural New Voices program of the U.S. National Academies of Sciences, Engineering, and Medicine. She received her undergraduate degrees from Tsinghua University and her M.S. and Ph.D. from the University of Michigan, Ann Arbor.

Eric Lo, The Chinese University of Hong Kong



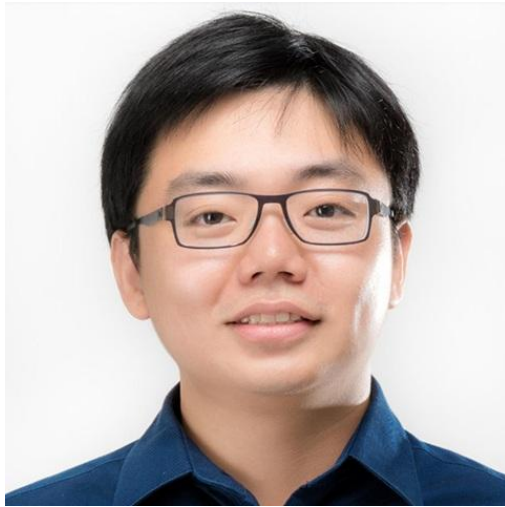
Title: Agentic Memory Through the Lens of Incremental View Updates

Abstract: Today's agent memory is largely "one-off and app-specific": each product builds its own memory logic from scratch. This leads to duplicated engineering efforts, inconsistent user experiences across products, and slow, fragmented rollout of new memory capabilities.

In this talk, I will present a fundamentally different approach: treating agent memory as multiple views of underlying base data (e.g., message history). From this perspective, various agent memories can be modeled as a graph—a lattice of cuboids—allowing us to leverage decades of efficient techniques from classical data cube research. This approach makes agent memory management dramatically more efficient, scalable, and consistent across products.

Bio: Eric Lo is currently an Associate Professor in the Department of Computer Science and Engineering at the Chinese University of Hong Kong (CUHK). He earned his PhD in Computer Science from ETH Zurich and has previously worked at both Google and Microsoft. His recent research focuses on vector databases, serving systems for AI agents, and cloud databases. He is currently the PC Chair of ACM SoCC 2026 and an Associate Editor of The VLDB Journal. His work has received recognition in the form of awards and honorable mentions at conferences such as VLDB 2005 and ICDE 2012. In 2020, he received the ACM SIGMOD Research Highlight Award.

Cheng Chen, ByteDance



Title: Graph Memory for AI Agents: Design and Practice with MemoryBase

Abstract: Agentic applications require more than retrieving relevant text. Their memory must capture domain entities, relationships, historical interactions, and evolving knowledge—and make this information reusable across sessions. This talk introduces **MemoryBase**, an enterprise-grade long-term memory system for AI agents developed at ByteDance. MemoryBase serves ByteDance products and is also offered to external customers through Volcano Engine via Mem0-compatible APIs. It provides cross-session memory, intelligent memory extraction, hybrid retrieval, and native support for **Graph Memory**.

Graph Memory in MemoryBase not only organizes memories accumulated from agent interactions and experiences into a graph of entities, events, concepts, and relationships, but also incorporates existing domain knowledge—such as schemas, taxonomies, rules, and curated knowledge graphs—to guide and constrain memory retrieval. This combination allows retrieval to move beyond semantic similarity: user requests can be mapped to relevant entities and graph neighborhoods, while domain knowledge guides graph traversal and determines which memories should be recalled. This talk presents the key components of the architecture, including memory ingestion, knowledge loading, graph construction, entity resolution, graph-based retrieval, and evidence assembly for LLM reasoning. Representative enterprise scenarios will demonstrate how structured knowledge and accumulated memories work together to produce more grounded and explainable responses.

Bio: Cheng Chen is a Tech Lead at ByteDance. His team develops graph and AI infrastructure, including agent memory, graph databases, graph computing, graph neural networks, GraphRAG, and graph foundation model platforms. These platforms support ByteDance products such as Douyin and TikTok. He received his Ph.D. in Computer Science from the National University of Singapore and has published more than 40 research papers. His work was invited to the Best of VLDB 2025 Special Issue of The VLDB Journal. He also received the Best Paper Runner-up Award at EuroSys 2024 and the Best Industry Paper Runner-up Award at VLDB 2023. He has served on the research and industry-track program committees of VLDB, SIGMOD, and ICDE.