



DAMG 7374-05 KNOWLEDGE GRAPHS with GenAI/GraphDB

Course Information

Course Title: Knowledge Graphs with GenAI/GraphDB

Course Number: DAMG 7374-03

Term and Year: Fall 2025

Credit Hour: 4

CRN: 39499

Course Format: Virtual with onsite components(optional)

Instructor Information

Full Name: Kishore Aradhya

Email Address: k.aradhya@northeastern.edu

Instructor Biography

Kishore Aradhya is a senior technology executive with over twenty-five years of experience developing and scaling diverse technology organizations, spanning startups to Fortune 500 companies. Throughout his career, he has held leadership positions at organizations such as Stanley Black & Decker, Bose, Adobe, edX/MIT, Staples, and Monster Search. In his most recent role at Stanley Black & Decker, he led the AI and Enterprise Data Platform Engineering teams as a Senior Director of Advanced Analytics and Data. His work focused on bringing GenAI and LLM technologies to the forefront and advancing the use of critical Data Insights by working with key executive leadership.

His expertise is crafting production-ready Enterprise Data Platforms for Analytics, Machine Learning, and BI applications. This spans the design, development, and architecture of highly scalable enterprise SaaS cloud services, including Customer and Analytics Data Platforms, Data Engineering, Search, mobile, and e-commerce solutions. He also spearheaded an NLP and Computer-Vision-driven document extraction research initiative, introducing key innovations in Adobe Document product features.

Kishore Aradhya has an MBA from the Isenberg School, UMass Amherst, an MS in Computer Science, and multiple executive and academic certifications from MIT, Stanford, and Harvard professional programs. Outside his primary roles, he is an active contributor to the industry as a CDO Magazine Editorial Board Member, an Industry Advisory Board Member at DVSum (an AI Data insights startup), and a Product Advisory Council Member at Kensu (an Enterprise Data Observability startup). Furthermore, Kishore offers strategic advisories to various startups and founders, focusing on enterprise market fit and other technical and Go-to-market (GTM) strategies.

Teaching Assistant Information

Full Name:

Email:

Office Hours:

Course Prerequisites

- Essential Python, basic Graphs, and Graph Database familiarity using Neo4j or other Graph DB's.
- Essential understanding of data concepts like Data Quality, Data Transformation, and Data Integration.
- Basic understanding of AI concepts - Machine Learning, Deep Learning & NLP

Course Description

This seminar-style course explores the cutting-edge intersection of Large Language Models (LLMs) and Graph Databases within the context of Knowledge Graphs, emphasizing their synergistic potential in modern Data Engineering and AI. Throughout the course, students will delve into the fundamentals of LLMs, graph theory, and graph database technology. They will investigate the integration of these technologies in practical applications, including knowledge graphs, NLP, and AI-driven data analysis. The course blends theoretical learning about graphs, graph analytics, and knowledge of graphs with hands-on projects, providing deep insights into Graph Database modeling. Students will engage with current research, develop critical skills in practical graph-based application development, and complete a semester-long project that showcases their ability to leverage these technologies in building sophisticated solutions.

Course Learning Outcomes

1. Master basic and advanced graph database concepts, primarily focusing on Graph thinking.
2. Through invited speakers, learn about current trends and challenges in implementing Knowledge Graphs and Graphs with specific adoption in Graph RAG and other current AI-driven graph development.
3. Study real-world case studies to grasp practical tradeoffs, challenges, and limitations.
4. Learn to read, analyze, and present KG/Graph-based research through a practical implementation lens.
5. Build professional connections with field experts for future collaboration and growth.

Standard Learning Outcomes

Learning outcomes expected for all College of Engineering Graduate programs:

1. *An ability to identify, formulate, and solve complex engineering problems.*
2. *An ability to explain and apply engineering design principles appropriate to the program's educational objectives.*
3. *An ability to produce solutions that meet specified end-user needs with consideration of public health, safety, and welfare, as well as global, cultural, social, environmental, and economic factors.*

Required Tools:

We will be using the following services and tools as needed:

1. Neo4j Desktop Edition
2. Snowflake Cloud services (class account available)
3. AWS Cloud Services

4. Data Integration Tool: LlamaIndex, LangChain, Airbyte (free; we will use this as needed)
5. Data Transformation Tool: dbt (free, but will probably use a cloud service, dbtCloud)
6. API Services from OpenAI and/or Hugging Face.

Reference Books:

1. **Building Knowledge Graphs: A Practitioner's Guide**
Authors: Jesus Barrasa, Jim Webber
Published by O'Reilly Media, Inc.
2. **Knowledge Graphs: Fundamentals, Techniques, and Applications**
Mayank Kejriwal, Craig A. Knoblock, Pedro Szekely
Published by The MIT Press

Topics Covered:

- Graph Data Modeling
- Knowledge Graphs
- Graph Analytics
- Graph Data Science
- Semantic Modeling using Graph Thinking
- Graph RAGS using LLM
- Generative Artificial Intelligence (AI)
- Data Governance

Course Activities

1. Activity #1 (Data Problem Formulation)
Data Driven Decision Making: What Business Question are we trying to solve with Graph driven data modeling?
2. Activity #2 (Graph Data Platform Design & Architecture)
Deliver an overall data platform design, the data model, and its associated implementation architecture.
3. Activity #3 (Phased Implementation Deliverable #1)
Develop a 1st iteration of this deliverable and present this for feedback from the class.
4. Activity #4 (Phased Implementation Deliverable #2)
Develop a 2nd iteration of this deliverable and present this for feedback from the class.
5. Activity #5 (Final Project Implementation Deliverable)
Present the final project deliverable to the class for feedback and grading.
6. Group Research Paper (discussed throughout the class):
Selection, Analysis, and Presentation for grading.

Course Schedule (will change due to the dynamic nature of the class & speaker availability)

Week 1: Introduction to Large Language Models and Graph Databases

- **Topics:**
 - Overview of LLMs: Definition, evolution, and key models.
 - Graph Databases: Basic concepts, types, and how they differ from relational databases.
 - Overview of Graph Theory: Nodes, edges, properties, and types of graphs.
 - Use Cases: How LLMs and graph databases are used in real-world scenarios.
 - Future Trends: Predicting the trajectory of LLMs and graph databases in technology.

Week 2: Foundations of Large Language Models

- **Topics:**
 - Deep Dive into GPT-3 and GPT-4: Architecture, training, and capabilities.
 - Understanding Transformer Models: Self-attention mechanism.
 - Dataset Requirements: Size, diversity, and ethical considerations.
 - Limitations and Challenges: Bias, reliability, and compute requirements.
 - LLMs in Practice: Case studies of successful LLM implementations.

Week 3: Basics of Graph Databases

- **Topics:**
 - Graph Database Engines: A comparison of Neo4j, ArangoDB, and others.
 - Query Languages: Cypher, Gremlin, and AQL.
 - Graph Database Modeling: Design principles and best practices.
 - Indexing and Searching in Graph Databases.
 - Performance Considerations: Query optimization and database tuning.

Week 4: Integration of LLMs and Graph Databases

- **Topics:**
 - Synergy between LLMs and Graph Databases: Enhancing data retrieval and analysis.
 - NLP for Graph Databases: Using LLMs to interpret and query graph data.
 - Graph-based Knowledge Extraction: Techniques and tools.
 - Case Studies: Real-world examples of LLM and Graph Database integration.
 - Future Implications: Predictive analytics and intelligent decision making.

Week 5: Data Collection and Preprocessing

- **Topics:**
 - Data Sourcing: Techniques for gathering relevant data for graph databases and LLMs.
 - Data Cleaning: Standardizing, deduplicating, and preparing data.
 - Ethical Data Collection: Privacy concerns and bias minimization.
 - Data Transformation: From raw data to graph-compatible formats.

- Integration of Unstructured Data: Utilizing LLMs for processing natural language data.

Week 6: Advanced Querying in Graph Databases

- **Topics:**
 - Complex Query Design: Writing efficient and powerful graph queries.
 - Query Optimization: Execution plans and performance tuning.
 - Recursive Queries: Traversing graphs for deep insights.
 - Graph Algorithms: Pathfinding, clustering, and centrality.
 - Visualization: Tools and techniques for visualizing graph data.

Week 7: Machine Learning with Graph Databases

- **Topics:**
 - Graph-based Machine Learning Models: Overview and applications.
 - Graph Neural Networks: Concepts and practical implementations.
 - Feature Extraction: Deriving meaningful features from graph structures.
 - Predictive Modeling: Using graph data for forecasting and trends analysis.
 - LLMs for Enhanced Graph Analysis: Text analysis and data enrichment.

Week 8: Natural Language Processing (NLP) Fundamentals

- **Topics:**
 - Core NLP Techniques: Tokenization, stemming, and lemmatization.
 - Vector Space Models: Word embeddings and their significance.
 - Text Classification and Sentiment Analysis: Practical applications.
 - Named Entity Recognition (NER): Extracting entities from text for graph databases.
 - Integrating NLP with Graph Databases: Enhancing data relationships and insights.

Week 9: Midterm Project Discussion and Planning

- **Topics:**
 - Project Ideation: Brainstorming and concept development.
 - Research Methodology: Structuring an effective research project.
 - Data Requirements: Identifying data needs for projects.
 - Technical Review: Assessing the technical feasibility of project ideas.
 - Project Planning: Timeline, milestones, and deliverables.

Week 10: Knowledge Graphs and Semantic Networks

- **Topics:**
 - Building Knowledge Graphs: Principles and techniques.
 - Ontology and Taxonomy in Graphs: Structuring and classifying information.
 - Semantic Analysis: Understanding context and meaning in data.
 - Application of LLMs in Knowledge Graphs: Enhancing data connectivity and insights.
 - Case Studies: Real-world applications of knowledge graphs.

Week 11: Graph-Based Data Engineering

- **Topics:**
 - Graph Data Modeling: Advanced concepts and techniques.
 - ETL for Graph Data: Extracting, transforming, and loading processes.
 - Data Integration: Merging graph data with other data sources.
 - Graph Database Scalability: Handling large-scale graph data.
 - Monitoring and Maintenance: Ensuring database health and performance.

Week 12: Scalability and Performance

- **Topics:**
 - Scaling Techniques for Graph Databases: Partitioning, sharding, and replication.
 - High-Performance Computing (HPC) in LLMs.
 - Benchmarking and Performance Metrics.
 - Load Balancing and Resource Management.
 - Case Studies: Scaling graph databases in large enterprises.

Week 13: Security and Compliance

- **Topics:**
 - Security Best Practices for Graph Databases and LLMs.
 - Data Governance: Compliance with legal and ethical standards.
 - Encryption and Data Protection: Techniques and tools.
 - Auditing and Monitoring: Tracking data access and modifications.
 - Risk Management: Identifying and mitigating potential threats.

Week 14: Emerging Technologies and Innovations

- **Topics:**
 - Latest Developments in LLMs and Graph Databases.
 - AI and Machine Learning Breakthroughs in Graph Analysis.
 - Innovative Applications: How emerging technologies are reshaping industries.
 - Guest Lectures: Insights from industry leaders and researchers.
 - Future Predictions: Where the fields are heading.

Week 15: Final Project Presentations

- **Topics:**
 - Effective Presentation Skills: Crafting and delivering impactful presentations.
 - Peer Review: Providing constructive feedback.
 - Project Analysis: Evaluating the success and learning outcomes of projects.
 - Future Directions: Exploring further research or development opportunities.
 - Course Wrap-Up: Summarizing key learnings and next steps.

Grade Breakdown:

Class Participation: 15%

Project Completion: 40%

Project Presentation: 20%

Group Research Paper Analysis & Presentation: 25%

Please Note: If you miss more than two classes without clear, explicit permission from the Instructor/TA, you will automatically be dropped 1 grade level.

Grading Scale:	87-89.9% B+	77-79.9% C+	69.9% or below F
	84-86.9% B	74-76.9% C	
95-100% A			
90-94.9% A-	80-83.9%B-	70-73.9% C-	

Attendance/Late Work Policy.

Attendance Policy

Students are expected to complete course readings, participate in class discussions or other learning activities during the unit, and complete written assignments for each unit during the time of that unit. It is understood that there might be one week when active participation in ongoing class conversations and learning activities might be delayed. Beyond one week, if there is an absence or lateness in participation (1) faculty must be notified in advance; (2) grades will be adjusted accordingly.

Late Work Policy

Students must submit assignments by the deadline in the time zone noted in the syllabus. Students must communicate with the faculty before the deadline if they anticipate work will be submitted late. Work submitted late without prior communication with faculty will not be graded.

End-of-Course Evaluation Surveys

Your feedback regarding your educational experience in this class is critical to the College of Engineering. Your comments will make a difference in our curriculum's future planning and presentation.

At the end of this course, please take the time to complete the evaluation survey at <https://neu.evaluationkit.com>. Your survey responses are entirely **anonymous and confidential**. For courses six weeks in length or shorter, surveys will be open one week before the end of the courses; for courses greater than six weeks in length, surveys will be open for two weeks. An email will notify you when surveys are available to your Husky Mail account.

Academic Integrity

A commitment to the principles of academic integrity is essential to the mission of Northeastern University. Promoting independent and original scholarship ensures that students derive the most from their educational experience and pursuit of knowledge. Academic dishonesty violates the most fundamental values of an intellectual community and undermines the achievements of the entire University.

As academic community members, students must become familiar with their rights and responsibilities. In each course, they are responsible for knowing the requirements and restrictions regarding research and writing, examinations of whatever kind, collaborative work, using study aids, the appropriateness of assistance, and other issues. Students are responsible for learning documentation conventions and acknowledging sources in their fields. Northeastern University expects students to complete all examinations, tests, papers, creative projects, and assignments according to the highest ethical standards, as set forth either explicitly or implicitly in this Code or by the direction of instructors.

Go to <http://www.northeastern.edu/osccr/academic-integrity-policy/> to access the full academic integrity policy.

University Health and Counseling Services

As a student enrolled in this course, you are fully responsible for assignments, work, and course materials as outlined in this syllabus and in the classroom. Over the semester, if you experience any health issues, please get in touch with UHCS.

For more information, visit <https://www.northeastern.edu/uahcs>.

Student Accommodations

Northeastern University and the Disability Resource Center (DRC) are committed to providing disability services that enable students who qualify under Section 504 of the Rehabilitation Act and the Americans with Disabilities Act Amendments Act (ADAAA) to participate fully in the university's activities. To receive accommodations through the DRC, students must provide appropriate documentation demonstrating a substantially limiting disability.

For more information, visit <https://drc.sites.northeastern.edu>.

Library Services

The Northeastern University Library is at the hub of campus intellectual life. Resources include over 900,000 print volumes, 206,500 e-books, and 70,225 electronic journals.

For more information and education-specific resources, visit <https://library.northeastern.edu>.

24/7 Canvas Technical Help

For immediate technical support for Canvas, call 617-373-4357 or email help@northeastern.edu

Canvas Faculty Resources: <https://canvas.northeastern.edu/faculty-resources/>

Canvas Student Resources: <https://canvas.northeastern.edu/student-resources/> For assistance with my Northeastern e-mail, and basic technical support:

Visit ITS at <https://its.northeastern.edu>

Email: help@northeastern.edu

ITS Customer Service Desk: 617-373-4357

Diversity and Inclusion

Northeastern University is committed to equal opportunity, affirmative action, diversity, and social justice while building a climate of inclusion on and beyond campus. In the classroom, members of the University community work to cultivate an inclusive environment that denounces discrimination through innovation, collaboration, and an awareness of global perspectives on social justice.

Please visit <http://www.northeastern.edu/oidi/> for complete information on Diversity and Inclusion

Title IX

Title IX of the Education Amendments of 1972 protects individuals from sex or gender-based discrimination, including discrimination based on gender-identity, in educational programs and activities that receive federal financial assistance.

Northeastern's Title IX Policy prohibits Prohibited Offenses, which are defined as sexual harassment, sexual assault, relationship or domestic violence, and stalking. The Title IX Policy applies to the entire community, including male, female, transgender students, faculty, and staff.

In case of an emergency, please call 911.

Please visit <https://www.northeastern.edu/ouec> for a complete list of reporting options and resources both on- and off-campus.