DAMG 7370 - Designing Data Architecture Business Intelligence
Fall 2025

Course Information
Course Title: Designing Data Architecture Business Intelligence
Course Number: DAMG 7370
Term and Year: Fall 2025
Credit Hour: 4 SH
CRN: 20650
Course Format: Livecast
Time: Mondays 3pm-6:20pm PT / 6pm-9:20pm ET

Instructor Information
Full Name: Jiazhen Zhu
Email Address: jiaz.zhu@northeastern.edu
Office Hours: TBD

Instructor Biography
Jiazhen Zhu is an engineering leader, and member of the Forbes Technology Council with over 15 years of experience in data engineering, AI systems, and scalable infrastructure. He has led high-impact teams at Fortune 100 companies across industries including retail, fintech, and healthcare. In his courses, he integrates real-world use cases and emphasizes modern data architecture, SQL and Python, cloud-native engineering, data modeling, and GenAI applications—equipping students with practical skills for today's data-driven roles.

Teaching Assistant Information
Full Name: TBD
Email Address: TBD
Office Hours: TBD

Guest industry experts will join for two sessions (see Weekly Schedule). Teaching Assistant(s) and their contact info will be provided in class.

Course Prerequisites
Data Management and Database Design

Course Description
This class covers advanced data architecture patterns for supporting enterprise business intelligence (BI) and data warehouse (DW) solutions across typical structured, unstructured and semi-structured data sources. Hybrid data integration and engineering workflows, delivering analytics to all users (business casual, power, data scientist). The course has a wide array of

technology topics: relational, column, memory and NoSQL databases; data integration (ETL/ELT) pipelines; data preparation and virtualization methods; as well as analytics from descriptive/diagnostic to predictive/prescriptive. Here we'll tackle the modern cloud-based data architectures and traditional data warehouse approaches, and best practices for on-premise vs cloud deployments.

NEW for Fall 2025: the course will also include AI-native concepts – exploring how AI-powered analytics and generative AI (GenAI) can be part of data pipelines, BI deliverables, and more. They'll become versed in cloud platforms and real-time pipelines and use alongside GenAI workflows to build and support a modern data product. The course will culminate in a collaborative capstone project where students will work in teams to build a BI system that is complimented by GenAI functionality and there is an opportunity to pitch it as a startup idea. At the end of the course, students will have gained hands-on exposure to the latest tools in BI and knowledge of when to use GenAI in enterprise data architecture.

Course Learning Outcomes

At the successful conclusion of this course, students will be able to:

- Design & assess BI data architectures, comparing traditional enterprise data warehouse ones with the new generation of analytical data architecture patterns (relational, columnar, Hadoop/NoSQL, etc.).
- Design and implement organizational data models in both entity-relationship (ER) and dimensional (star and snowflake schema) modeling techniques to address and systematize the analytical query and reporting requirements.
- Build data integration workflows and ETL/ELT processes that clean, prepare data and validate data quality from different source systems applying data engineering best practices.
- Building BI & analytics solutions to convey value from data (e.g. creating insightful dashboards/visualizations such as Power BI/Tableau, and conducting descriptive, diagnostic, predictive & prescriptive analytics).
- Apply GenAI tools in BI flows to optimize data analysis and decision support – such as deploy large language models that provide insights, automate data prep tasks or use natural language to query datasets.
- Implement data governance, Security, and Ethics best practices within the domain of data architecture and BI projects to manage data compliance and integrity.

In achieving these objectives, students will develop practical full-stack capabilities to design, build, and support contemporary business intelligence data systems tied to the evolving frontier of data engineering and AI.

Required Tools and Course Textbooks.

Textbook:

- Business Intelligence Guidebook: From Data Integration to Analytics, by Rick Sherman. Published by Morgan Kaufmann, 2014.
- Data Mining Concepts and Techniques, 3rd Edition by Jiawei Han, Micheline Kamber and Jian Pei. Published by Morgan Kaufmann, 2011.

Tools:

- Cloud Platforms: Google Cloud (BigQuery, Pub/Sub), Microsoft Azure (Synapse, Data Factory), Snowflake
- Data Engineering & Pipelines: Apache Airflow, PySpark (Databricks), dbt, Talend, Kafka (streaming)

- Modeling & Architecture: ER/Studio, dbdiagram.io, Kimball methodology, Lucidchart
- BI & Visualization: Power BI, Tableau, Looker, Metabase
- Databases: PostgreSQL, MongoDB, MySQL, Azure SQL, BigQuery
- Development & Notebooks: VS Code, JupyterLab, Docker, GitHub
- GenAI & Automation Tools: ChatGPT, Claude, LangChain, ChromaDB, LlamaIndex (optional)
- Data Quality & Governance: Great Expectations, OpenLineage, DataHub
- Deployment/MLOps: Docker, FastAPI, MLflow, GitHub Actions

Course Schedule/Topics Covered.

#	Date	In Class Topic	Content	Assignment Due
1	**9/8/2025**	Industry Trends, Data Architecture & GenAI for BI	- Content Cover for this class - Role evolution (DE, BI, DS) - Modern data architecture - Google data case study - MCP intro: modular design for data pipelines - Linkedin/GitHub portfolio - Career skill map - Vibe Coding *- Resume Support*	- Hands-on Lab 0
2	**9/15/2025**	Modern Data Stack Foundations & Lifecycle Planning	- Data lifecycle: ingest → serve - GCP/Azure/AWS - Data product lifecycle planning - Terraform, IAM, infrastructure blueprints - MCP units: ingestion, transformation, storage - AI-augmented infra design	- Hands-on Lab1
3	**9/22/2025**	Source Systems, Data Ingestion, and SQL vs NoSQL Queries	- PostgreSQL vs MongoDB - ACID, BASE, CAP - Joins, windows, EXPLAIN - SQL Queries - AI: Text-to-SQL, NoSQL mapping - Claude: suggest DB choice + modular ingestion blocks (MCP)	- Assignment 1 - Hands-on Lab2
4	**9/29/2025**	Storage Systems & Abstractions	- File formats (CSV, JSON, Avro, Parquet) - Lake vs warehouse vs lakehouse *- Guest Speaker 1 - DE Life*	- Assignment 2 - Hands-on Lab3
5	**10/6/2025**	Data Modeling, Transformation, Serving, Orchestration & Monitoring	- Star vs snowflake schemas - dbt, Airflow - Modeling KPIs, fact/dim - Monitoring pipelines (Airflow, dbt tests)	

6	**10/13/2025**	Data Integration, ETL/ELT & Engineering Practices	- ETL vs ELT patterns - Talend, dbt, Airflow orchestration - Final project pitch - Vibing Coding: ideate pipeline refactors via GenAI	- Assignment 3 - Hands-on Lab4 - Project proposal (with MCP plan) + ELT design + BI
7	**10/20/2025**	Midterm Exam	- In-class written exam (covering Weeks 1–6 content)	
8	**10/27/2025**	Python, PySpark, Data Preparation, Quality & Distributed Processing	- Spark internals - PySpark transforms - Data profiling, cleaning - GPT for code conversion - Vibing Coding sprint: GenAI-assisted notebook debugging	- Hands-on Lab5
9	**11/3/2025**	Real-Time Data Pipelines & Streaming	- Kafka, Spark Streaming - Stream processing design - Event time, watermark, DAGs - MCP stream handler design	- Assignment 4 - Hands-on Lab6
10	**11/10/2025**	BI & Data Analytics Design	- Power BI, Tableau, Looker - Dashboarding best practices - Text-to-SQL, KPI summary with GenAI - Agent Coding: AI-powered KPI bot design	- Hands-on Lab7
11	**11/17/2025**	GenAI Pipelines & Prompt Engineering	- Prompt types (zero/few-shot, chain-of-thought), Agent, MCP, FunctionCalling - LangChain, RAG - Slack bot for BI - Claude vs GPT prompt tuning	- Hands-on Lab8
12	**11/24/2025**	Governance, Observability, DataOps & MLOps	- Data contracts & SLAs - Great Expectations, OpenLineage - *Guest Speaker 2 – AI/BI Innovation & Startups*	- Hands-on Lab9
13	**12/1/2025**	Data System Design	- End-to-end architecture (batch vs stream trade-offs) - System design critique session - GPT: contract gen + alerts	- Hands-on Lab10
14	**12/8/2025**	Final Presentation Demos	- Capstone team demo - Peer feedback & wrap-up - *Resume Support*	- GitHub - Deck

Assignment Grading

- Attendance – 10 %
- Assignment/Quiz 1 – 5%
- Assignment/Quiz 2 – 5%
- Assignment/Quiz 3 – 5%
- Assignment/Quiz 4 – 5%
- In Class Hands-on Labs (1-10) - 10%

- Midterm Exam – 15%
- Final Project – 25%
- Presentation Demo – 20%

Grading Scale

Percentage Range	Letter Grade	Grade Point Equivalent
95.0–100.0%	A	4.000
90.0–94.9%	A-	3.667
87.0–89.9%	B+	3.333
84.0–86.9%	B	3.000
80.0–83.9%	B-	2.667
77.0–79.9%	C+	2.333
74.0–76.9%	C	2.000
70.0–73.9%	C-	1.667
69.9% and Below	F	0.000

Incomplete Grades

An incomplete grade may be reported by the instructor when a student has failed to complete a major component of a required course, such as homework, a quiz or final examination, a term paper, or a laboratory project. Students may make up an incomplete grade by satisfying the requirements of the instructor. Be aware that instructors' policies on the granting of incomplete grades may vary and that the final decision on an incomplete grade is up to the instructor. **Instructors may deny requests for an incomplete grade.** If the missing assignment(s) have not been submitted to the instructor within 30 days from the end of the term in which the course was offered, or the agreed upon due date, the grade entered will reflect the student's grade in the course for the work completed and the missing assignments receiving no credit toward the final grade.

Attendance/Late Work Policy
Attendance Policy

In each term, students enrolled in on-ground sections are expected to be on campus and attending class beginning with the first day of classes. Students in online sections are expected to log in and participate in class beginning with the first day of classes.

Students who join a class after the first day of class during the university add period, or who are approved for late registration by the instructor and the Graduate School of Engineering, are responsible for all coursework missed prior to enrolling. In the interest of students' success, the college does not support the arrival of students to class after the university add deadline. **Enrolled students who do not attend class during the first week of a semester risk being dropped from the course.**

In cases where an enrolled student cannot arrive to campus by the first day of class due to circumstances beyond their control, it is the student's responsibility to contact the instructor for approval and notify the Graduate School of Engineering.

Students registered in MGEN courses (INFO, CSYE, and DAMG) are allowed **a maximum of 2 absences per course,** with 3 or more absences resulting in an 'F' for that course. Course instructors are not expected to make accommodations and students are expected to inform their instructors of any absences in advance of the class. Should a student anticipate being unable to attend 3 or more classes, they should discuss their situation with their Academic Advisor to explore other types of leave and accommodations in accordance with the University's academic and global entry expectations. Students may be asked to share communications about class absences with their Academic Advisor. If a student is sick long-term or experiences a medical issue that prevents class attendance, it is strongly encouraged that they speak with their Academic Advisor (coe-gradadvising@northeastern.edu) to learn more about the Medical Leave of Absence. International students should review the Office of Global Services webpage to understand their visa compliance requirements.

Teaching Assistants (TAs) or Instructional Assistants (IAs) will be present at each class to collect student attendance.

Late Work Policy

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

Course Evaluations
Student feedback on their learning experience is valuable and helps improve future courses. We encourage all students to complete the course evaluation surveys when they become available.

Surveys are distributed at both the midterm mark and the end of the term via email and are completely anonymous and confidential. Any questions about the surveys can be directed to mgen-programs@coe.northeastern.edu

MGEN Student Feedback
Students who would like to provide the MGEN unit with anonymous feedback on this particular course, Teaching Assistants, Instructional Assistants, professors, or to provide general feedback regarding their program, may do so using this survey:
https://neu.co1.qualtrics.com/jfe/form/SV_cTIAbH7ZRaaw0Ki

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

As members of the academic community, 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, the use of study aids, the appropriateness of assistance, and other issues. Students are responsible for learning the conventions of documentation and acknowledgment of sources in their fields. Northeastern University expects students to complete all examinations,

tests, papers, creative projects, and assignments of any kind according to the highest ethical standards, as set forth either explicitly or implicitly in this Code or by the direction of instructors.

The following is a broad overview, but not an all-encompassing definition, of what constitutes a violation of academic integrity:

*Cheating***:** The University defines cheating as using or attempting to use unauthorized materials, information, or study aids in any academic exercise. When completing any academic assignment, a student shall rely on their own mastery of the subject.

*Fabrication***:** The University defines fabrication as falsification, misrepresentation, or invention of any information, data, or citation in an academic exercise.

*Plagiarism***:** The University defines plagiarism as using as one's own the words, ideas, data, code, or other original academic material of another without providing proper citation or attribution. Plagiarism can apply to any assignment, either final or drafted copies, and it can occur either accidentally or deliberately. Claiming that one has "forgotten" to document ideas or material taken from another source does not exempt one from plagiarizing.

Unauthorized Collaboration: The University defines unauthorized collaboration as instances when students submit individual academic works that are substantially similar to one another. While several students may have the same source material, any analysis, interpretation, or reporting of data required by an assignment must be each individual's independent work unless the instructor has explicitly granted permission for group work.

Participation in Academically Dishonest Activities: The University defines participation in academically dishonest activities as any action taken by a student with the intention of gaining an unfair advantage over other students.

Facilitating Academic Dishonesty: The University defines facilitating academic dishonesty as intentionally or knowingly helping or contributing to the violation of any provision of this policy.

Please visit https://osccr.sites.northeastern.edu/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 course of the semester if you experience any health issues, please contact UHCS.

For more information, visit https://www.northeastern.edu/uhcs.

Student Accommodations/Disability Access Services (DAS)

Northeastern University and Disability Access Services (DAS) 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 activities of the university. To receive accommodations through DAS, students must provide documentation of a disability that demonstrates a current substantial limitation. Accommodations are approved based on a review of the information that is submitted and reviews are done on a case-by-case basis.

If the course is conducted in an on-ground (in-person) format, students are expected to attend class physically as scheduled. Professors are **not required to provide virtual attendance links** unless a student has documented accommodation approved by the **Disability Access Services (DAS) office** and their **Academic Advisor**. If a student requires accommodation for remote participation, they must submit a formal request through the **Disability Office** and coordinate with their **Academic Advisor** prior to the course start date.

For more information, visit https://disabilityaccessservices.sites.northeastern.edu/

Office of Global Services
As an F-1, J-1, or Study Permit student, you must meet certain obligations in order to maintain lawful nonimmigrant status. Maintaining status is necessary in order to retain eligibility for the benefits of F-1 or J-1 status, such as employment authorization and program extension, and can be crucial to a successful application for a change or adjustment of nonimmigrant status in the future. Failure to maintain your nonimmigrant status can result in serious problems with immigration and *could lead to deportation from the U.S. or Canada.*

Students must maintain on-ground presence throughout the academic term. At Northeastern, there are four different defined instructional methods: Traditional, Hybrid, Live Cast, and Online. Traditional, Hybrid, and Live Cast courses meet the Visas' on-ground presence requirements. **Online courses do not meet the Visas' on-ground presence requirements.**

Students enrolled in Summer courses should adhere to OGS guidelines on maintaining status during the Summer term.

For more information please visit,
https://international.northeastern.edu/ogs/current-students/understanding-visa-requirements/guidelines-on-maintaining-status/

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 for education specific resources, visit https://library.northeastern.edu
Network Campus Library Services: Northeastern University Library Global Campus Portals

24/7 Canvas Technical Help
For immediate technical support for Canvas, call 617-373-4357 or email help@northeastern.edu

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

Outreach, Engagement, Belonging

Northeastern University is committed to fostering a community of belonging, which is essential to the advancement of Northeastern University's mission of teaching and research. Our university is stronger as a result of the varied backgrounds, experiences, and perspectives that all members of our global community bring to the pursuit of knowledge.
Embracing this pluralism is not the work of one office, department, or academic unit. It is a shared responsibility that spans disciplines and boundaries. By harnessing the power of our differences, we will continue to light the path to bold new ideas and life-changing discoveries.

It is my intention that students from all backgrounds and perspectives will be well served by this course, and that the diverse experiences that students bring to this class will be viewed as an asset. I welcome individuals of all ages, backgrounds, beliefs, ethnicities, genders, gender identities, gender expressions, national origins, religious affiliations, sexual orientations, socioeconomic background, family education level, ability – and other visible and nonvisible differences. All members of this class are expected to contribute to a respectful, welcoming and belonging environment for every other member of the class. Your suggestions are encouraged and appreciated.

Please visit Belonging at Northeastern – Northeastern Provost for complete information.

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.

The Office for University Equity and Compliance (OUEC) leads Northeastern University's efforts in maintaining compliance with all federal, state, and provincial civil rights laws and prohibits discrimination within any of its programs, activities, and services. Please visit https://ouec.northeastern.edu/ for more information and for the link to file a report.