

Data Architecture and Management, MS

For program contact information, please visit this website (<https://coe.northeastern.edu/academics-experiential-learning/academic-departments/mgen/ms-daam/>).

Many MS programs in the data area deal with data collection and analysis but do not however address a crucial activity that data scientists, data analysts, business analysts and many software engineers need to perform to make that data valuable, namely, data integration. That activity may also be referred to as data preparation, data curation, application integration and data engineering based on the integration of use cases and integration persona. The MS program in Data Architecture and Management focuses on these activities.

Data systems engineering occurs because data is fragmented and usually scattered across many data sources. However, even if all the data one needed were in one place, there is still an intensive need for integration. Information is data in context and the context of data as collected is different than the many ways it needs to be transformed so as to generate useful information.

Data engineering field could be thought of as a superset of business intelligence and data warehousing that brings in more elements from software engineering. This discipline also integrates specialization around the operation of so called "big data" distributed systems, along with concepts around the extended Hadoop ecosystem, stream processing, and in computation at scale.

The MS in Data Architecture and Management program offers a multitude of courses in data engineering in addition to supplementary courses that are required to deliver the data results in a meaningful way to management. We plan to cover data management, advanced data management, data warehousing and business intelligence, column data bases, data science engineering, and big-data engineering. On the software engineering side, we offer advanced big-data programming using the powerful Scala language and a course on advanced data science as well as cloud computing. Multi-thread concurrent computing is also offered as it is important for synchronizing a huge set of servers working in parallel to do large scale analytics to make things run faster by 100's fold increase in speed. Due to the high-level mathematical operations required to make these programs run, only software engineers have the capacity to work in these complicated areas. Only they can make the necessary mathematical algorithms execute quickly enough to get the finest results.

Program Requirements

Complete all courses and requirements listed below unless otherwise indicated.

Core Requirements

Code	Title	Hours
DAMG 6105	Data Science Engineering with Python	4
DAMG 6210	Data Management and Database Design	4
DAMG 7250	Big Data Architecture and Governance	4
DAMG 7370	Designing Advanced Data Architectures for Business Intelligence	4

Electives

Code	Title	Hours
Complete 16 semester hours from the following subject codes:		16
CSYE		
DAMG		
INFO		

Program Credit/GPA Requirements

32 total semester hours required

Minimum 3.000 GPA required