VirginiaTech *Invent the Future®* **Master of Information Technology**

ECE 5984 Data Engineering Project

This is an elective course for the Masters of Information Technology program. Fundamentals of data engineering. The role of data engineer. Data engineering lifecycle. Data quality and valuation. Data provenance. Data generation, ingestion, transformation, storage, serving Artificial Intelligence and Machine Learning (AI/ML), visualization and business analytics. Automation and task orchestration. Data systems. E(xtract)T(ransform)L(oad) data. Build, test, maintain data pipelines. Data lakes. Real-world problems with emphasis on end-to-end engineering solutions. Cloud services and open-source data engines/platforms. Engineering portfolio. Master of Information Technology (MIT) students only.

Master of Information Technology (MIT) students only. The delivery format is online and asynchronous.

To begin, visit **Getting Started. (https://canvas.vt.edu/courses/176740/pages/getting-started)**

🔗 Quick Links

Please click on the following quick links to get more information about the course:

- **Course Schedule (https://canvas.vt.edu/courses/176740/pages/course-schedule)**
- **Course Syllabus (https://canvas.vt.edu/courses/176740/assignments/syllabus)**
- **Learner Support (https://canvas.vt.edu/courses/176740/pages/learner-support)**
- **Privacy (https://canvas.vt.edu/courses/176740/pages/privacy)**
- **Accessibility (https://canvas.vt.edu/courses/176740/pages/accessibility)**

✔️ Course Objectives

Having successfully completed this course, the student should be able to:

- Identify the component of data engineering lifecycle
- Illustrate data quality issues
- Describe provenance issues
- Apply data validation methods
- Design data pipeline orchestration environments
- Evaluate data storage environments

- Analyze the role of data lakes in data engineering
- Propose end-to-end engineering solutions using cloud services and open-source data engines/platforms
- Construct an engineering deliverable portfolio

📚 Course Modules

Please click on the following course modules to have quick acce

Aug 21 - 27	Aug 28 - Sept 10
quick link image for module 1 (https://canvas.vt.edu/courses/176740/pages/module-1-fundamentals-of-data-engineering)	quick link image for module 2 (https://canvas.vt.edu/courses/176740/pages/module-2-data-pipelines)

Sept 25 - Oct 8	Oct 9 - 15
quick link image for module 4 (https://canvas.vt.edu/courses/176740/pages/module-4-data-engineering-project)	quick link image for module 5 (https://canvas.vt.edu/courses/176740/pages/module-5-data-storages-and-data-provenance)

Oct 30 - Nov 12	Nov 13 - 26
quick link image for module 7 (https://canvas.vt.edu/courses/176740/pages/module-7-data-visualization-and-business-analytics-pipeline)	quick link image for module 8 (https://canvas.vt.edu/courses/176740/pages/module-8-privacy-and-security-issues-in-data-pipelines)

---

icon image for contacting me section **Synchronous Session**

- Zoom Date/time: Wednesday 7-8 pm Eastern Time
- Zoom URL: **https://virginiatech.zoom.us/j/5718583104 (https://virginiatech.zoom.us/j/5718583104)**

👤 Contacting me

Please feel free to reach out to me for any questions or concerns. You can also post your questions in the Discussions section.

- Course Instructor: Dr. Nektaria Tryfona
- Email: **tryfona@vt.edu (mailto:tryfona@vt.edu)**

- Details of Contact: I check email frequently M-F and generally respond the same day. On weekends, I check email once per day and will respond within 36 hours.

TA: Ayush Dhar

- Email: adhar@vt.edu
- Office Hours:
- Zoom:
- Details of Contact: TA checks the discussion forum and email at least once a day and generally responds within 24 hours. On weekends, he will respond within 36 hours.