CSYE 7200: Big Data Systems Engineering with Scala
Fall 2025

Course Information
Course Title: Big Data Systems Engineering with Scala
Course Number: CSYE 7200
Term and Year: Fall 2025
Credit Hour: 4
CRN: 17551
Course Format: Classroom

Instructor Information
Full Name: Robin Hillyard
Email Address: r.hillyard@neu.edu
Office Hours: Tuesday 2pm—6pm; Friday all day (online/Teams only)

Instructor Biography
Associate Teaching Professor Robin Hillyard has taught at Northeastern University since 2015. Before that, he enjoyed a long and successful career in software development and consulting. He has served in every capacity from Chairman/President on down—in startups and more established companies—and in many industries: financial, healthcare, eCommerce, and computer-aided design. Throughout this time, he was an innovator, especially in the areas of unit testing and agile development. He wrote his first program in 1968, and it worked first time (an accomplishment in those days!). His undergraduate degree was in Engineering Science at the University of Oxford (1st class honours) with a specialty in control engineering. Switching to the "other" place, he received his Ph.D. in Computer Science in 1978 from the University of Cambridge, with a dissertation on dimensions and tolerances in computer-aided design. Always fascinated by the expressiveness of computer languages for solving problems, Dr. Hillyard has become proficient in a wide range of languages, from assembly language, BCPL, Algol68, Fortran, C, Perl, Java, to Scala and many others. A long-time proponent of unit testing, he wrote his first unit test suite (for assembly language) in 1972. Having developed perhaps the world's first object-relational database, he has previously taught relational databases at the Khoury College of Computer Sciences. He has a patent concerning the use of reactive programming for healthcare information systems. Most recently, he has been involved with "big data" on a Hadoop platform using Spark, GraphX, Scala, Elastic Search, and Zeppelin for the recognition of gaps in (health) care. Other classes taught: INFO 6205 (Program Structure and Algorithms), CSYE 7374 (Cryptography).

Teaching Assistant Information
Full Names: TBDClick or tap here to enter text.
Email Address:
Office Hours: See Canvas

Course Prerequisites
None.
Course Description
This course demonstrates the importance of functional programming, particularly Scala, to the high-performance processing of Big Data with Apache Spark. Equally important is Scala's use in reactive (concurrent) programming. Please note that the emphasis is on Scala and functional programming, more than Spark and Big Data, although both aspects are well-covered. A significant part of the class is devoted to a major team project, chosen by the team in either machine learning or reactive programming. Above all, this is a *practical* class: you will learn many aspects of programming and software engineering that are useful whatever language you are using.

Course Learning Outcomes

1) Students will understand why functional programming is essential in a high-performance big data scenario.
2) Students will learn how to use the various modules of Apache Spark, including—and especially—the machine learning library.
3) Students will learn how to build reactive (concurrent) systems using actors.
4) Students will learn how to parse (deserialize) information into features for machine learning or other purposes.
5) Students will learn about referential transparency and its role in proving programs to be correct.
6) Students will gain skills in a programming language (Scala) that is valued highly in the job marketplace.

Required Tools and Course Textbooks.

1. **(strongly recommended)** *Functional Programming in Scala*—Pilquist, Bjarnason, and Chiusano, **Manning** (2nd edition)
2. **(recommended)** *Programming in Scala*—Odersky, Spoon, and Venners, **Artima** (4th edition).

Course Schedule/Topics Covered.

Week	Date	In Class Topic	Assignment Due
1	**09/01**	**Introduction to Scala, Spark, Big Data**	
2	**09/08**	**A language for Big Data**	
3	**09/15**	**Coding in Scala: functional programming**	
4	**09/22**	**Fundamental aspects of Scala**	
5	**09/29**	**Exceptional Conditions; containers; methods; intro to for-comprehensions**	
6	**10/06**	**Functional Composition; Asynchronous Programming**	
7	**10/13**	**(de)-Serialization; implicits; monads; actors**	
8	**10/20**	**Spark introduction**	
9	**10/27**	**Spark detail and machine learning**	
10	**11/03**	**Advanced topics and Major Team Project**	
11	**11/10**	**Advanced topics and Major Team Project**	
12	**11/17**	**Advanced topics and Major Team Project**	
13	**11/24**	**Advanced topics and Major Team Project**	**Thanksgiving**

14	**12/01**	**Advanced topics and Major Team Project**	
	12/08	**Final Exams (date/time TBD)**	
	12/16	**Final Grades**	

Assignment Grading

- Attendance – 4 %
- Assignments – 20%
- Quizzes – 12%
- Team Project – 23 %
- Midterm Exam – 19%
- Final Exam – 22%

Grading Scale

There is no fixed scale for determining final grades. As a rough guide, 95% or above will generally earn an A. 80% or more should get you a B-. But I repeat: there is no fixed scale for grading.

Attendance/Late Work Policy

Attendance Policy

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 automatic 'F' for that course.** Students are expected to inform their instructors of any absences in advance of the class; 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-mgen-gradadvising@northeastern.edu) to learn more about the Medical Leave of Absence. 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 in accordance with the University's academic and global entry expectations. 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.

End-of-Course Evaluation Surveys

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

At the end of this course, please take the time to complete the evaluation survey at **https://neu.evaluationkit.com**. Your survey responses are **completely anonymous and confidential**. For courses 6 weeks in length or shorter, surveys will be open one week prior to the end of the courses; for

courses greater than 6 weeks in length, surveys will be open for two weeks. An email will be sent to your Northeastern University Mail account notifying you when surveys are available.

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.

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

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_cTlAbH7ZRaaw0Ki

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
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 activities of the university. To receive accommodations through the DRC, students must provide appropriate documentation that demonstrates a current 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 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

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*.