ATENEO DE MANILA UNIVERSITY
Loyola Schools

SYLLABUS FOR GRADUATE COURSES

A. COURSE INFORMATION

COURSE NUMBER	CSCI 173 (undergraduate students) CSCI 273 (graduate students)			**NO. OF UNITS**	3
COURSE TITLE	Big Data Processing				
PREREQUISITE/S	None				
DEPARTMENT/ PROGRAM	DISCS			**SCHOOL**	Science and Engineering
SCHOOL YEAR	SY 2025-2026			**SEMESTER**	Second
INSTRUCTOR/S	William Emmanuel Yu, Ph.D.				
VENUE	CTC 112 (ST2) Online (ZZZ)	**SECTION**	ST2, ZZZ	**SCHEDULE**	SAT 8-11

B. COURSE DESCRIPTION

This course tackles principles, best practices, and technologies involved in large-scale data processing. Among the topics covered are distributed databases and programming, cloud computing, and unstructured databases.

Updated to include recent developments in the fast paced big data (particularly the Hadoop ecosystem) and NoSQL.

C. *PROGRAM LEARNING OUTCOMES

**This section does not apply to a student who is not a major of the program under which this course is administered*

PROGRAM LEARNING OUTCOMES
PLO1: Model and analyze data from a variety of domains ranging from Internet social networks
PLO2: Apply a variety of data mining techniques for modeling datasets, handling incomplete and erroneous data, and automatically predicting and classifying elements in a dataset.
PLO3: Select and apply big data processing and cloud

Alignment of the Course to the Program Learning Outcomes

	PLO1	PLO2	PLO3
CLO1		X	
CLO2		X	X
CLO3		X	X
CLO4		X	X
CLO5		X	X
CLO6		X	X
CLO7		X	X
CLO8		X	X

D. COURSE LEARNING OUTCOMES

By the end of this course, students should be able to:

COURSE LEARNING OUTCOMES
CLO1: WILL be able to understand the history and development of massively parallel processing using commodity computing technology.
CLO2: WILL be familiar with distributed computing and parallel programming concepts.
CLO3: WILL be able to understand the basics of how these were developed and how they work.
CLO4: WILL be able to install and configure a working big data system (using Hadoop) on both their own computing environments and in the cloud.
CLO5: WILL be able to develop applications to utilize this big data systems using parallel processing concepts.
CLO6: WILL be able to configure and develop database like querying environments with Hive.
CLO7: WILL be able to learn and develop using the Map-Reduce development paradigm using Hadoop Streaming.
CLO8: WILL be able to develop additional parallel applications using the Spark framework.

E. COURSE OUTLINE and LEARNING HOURS

Course Outline	CLOs	Estimated Contact or Learning Hours

History of Commodity High Performance Computing	CLO 1	2 weeks
Introduction to Massively Distributed Systems	CLO 2, 3	2 weeks
Introduction to Hadoop, Hadoop On-Premise/Cloud	CLO 4, 5	3 weeks
Hive	CLO 3, 5, 6	2 weeks
Hadoop Streaming	CLO 3, 5, 7	2 weeks
Spark, SparkQL, Spark ML	CLO 3, 5, 8	2 weeks
Visualization and Optional Tooling	CLO 3, 5	2 weeks

Classes are ONSITE (CTC 112), however please take note that the certain days may be ASYNCH (not onsite) and there will be activities posted in the LMS for action.

F. ASSESSMENTS AND RUBRICS

Assessment Tasks	Assessment Weight	CLOs
Quizzes, Seatwork and Exams	70%	All CLOs
Laboratory Assignments	30%	All CLOs

G. TEACHING and LEARNING METHODS

TEACHING & LEARNING METHODS & ACTIVITIES	CLOs
Lectures	All CLOs
Online Instructional Videos	All CLOs
Interactive Forums and Chat	All CLOs
Hands-on Walkthrough	All CLOs

H. REQUIRED READINGS

- *Refer to materials in Learning Management System*
- *Hadoop: The Definitive Guide 4th ed, by White, T.; O'Reilly (2015)*

I. SUGGESTED READINGS

- *Large-scale and Big Data: Processing and Management; Taylor and Francis (2014)*
- *Principles of Data Integration, by Doan, Halevy, and Ives; Elsevier (2012)*
- *MapReduce Design Patterns: Building Effective Algorithms and Analytics for Hadoop and Other Systems, by Miner and Shook; O'Reilly (2014)*
- *Spark: The Definitive Guide by Bill Chambers and Matei Zaharia.; O'Reilly (2018)*

J. GRADING SYSTEM

Numerical Grade	Letter Grade (Undergrads)	Letter Grade (Grad Students)
93 - 100	A	A
87 - 92	B+	A-
81 - 86	B	B+
75 - 80	C+	B
69 - 74	C	C (unsatisfactory, no credit)
60 - 68	D	F (fail)
< 60	F	F (fail)

K. CLASS POLICIES

- Quizzes, assignments and activities are always available electronically on schedule.
- In case you cut, it will be your responsibility to know the material covered for the day. There will be no make-up lectures or sessions. However, the LMS does have material for self-paced learning. For purely online learning, you are responsible for following the course via the LMS and perusing the materials at the appropriate pace
- No make-up quizzes / tests will be given unless you have a very valid reason, and proof supporting it.
- Cheating will not be tolerated. Cheating in any requirement will result in a minimum penalty of having a grade of 0 for that requirement, and will be reported to the appropriate authorities, as provided for by the Student Handbook. Duplicate work will merit penalties for both the student who copied and the student from whom the work was copied.
- Students are expected to comply with the DISCS Academic Integrity Policy. With each submission, students must include a certification that their work is substantially their own and not copied from others. In addition, students must clearly acknowledge and specify any help from outside sources such as other classmates, the Web, books, etc., that they received while doing their projects. Failure to acknowledge such may be interpreted as intellectual dishonesty. Consult the course website for details on these policies.
- No eating in the classroom or during class.
- Announcements for class scheduling and online quizzes will be provided in the LMS. It is a student's responsibility to ensure that he/she regularly checks the course website for new activities.
- Please familiarize yourself with our Gender Policy <http://2012.ateneo.edu/ls/ls-gender-policy> and Code Decorum<https://2012.ateneo.edu/policies/code-decorum-investigation-sexual-harassment> documents.
- Additional policies, with due consultation with the students, may be implemented by the instructor to adapt to the class environment.
- Undergraduate students should refer to the policies from the Undergraduate Academic Policies memo sent by the ADAA to the community, dated January 10, 2023
- Graduate students should refer to the policies from the LS Graduate Level Academic Policies memo sent to the community, dated January 12, 2023

L. CONSULTATION HOURS

NAME OF FACULTY	EMAIL	DAY/S	TIME
William Emmanuel Yu	wyu@ateneo.edu (online only)	M-F	8:00am - 5:00pm