IT-590 Fall 2017
Advanced Data Management

Dates: 08/22-12/15/2017	Time: Online, instructor-led	Location: http://blackboard.valpo.edu
Instructor: Sonja Streuber	Office Hours: MTWRF 11:00-12:00 on Google Hangouts and in 145H Meier	Contact: sonja.streuber@valpo.edu

Introduction

Welcome to IT-590, Advanced Data Management! Advanced data management has always been at the core of efficient database and information systems. Recent trends like big data and cloud computing have aggravated the need for sophisticated and flexible data storage and processing solutions. This course provides a comprehensive overview of the principles of data management developed in the last decades with a focus on data structures and query languages. It treats a wealth of different data models and surveys the foundations of structuring, processing, storing and querying data according to these models.

Starting off with the topic of database design, it further discusses weaknesses of the relational data model, and then proceeds to convey the basics of graph data, tree-structured XML data, key-value pairs and nested, semi-structured JSON data, columnar and record-oriented data as well as object-oriented data. The course culminates in an analysis of fragmentation, replication and consistency strategies for data management in distributed databases as well as recommendations for handling polyglot persistence in multi-model databases and multi-database architectures.

NOTE that this course builds on the content of IT-533 Data Mining and IT-603 Information Management.

Learning Objectives

Students who successfully complete this course will gain an advanced architectural understanding of which Data Management Systems are appropriate for which kinds of data, under which performance conditions, and for which analytical purpose. Students will be able to:

- Analyze large sets of data and use cases to determine whether relational or various non-relational or object-based management systems are appropriate
- Design and implement data management architectures across various data distribution models using tools such as the Unified Modeling Language (UML)
- Manage data and analytical approaches in a number of applications and tools, both, locally and in the cloud
- Apply the learned techniques to many common data management and storage problems, including content management systems
- Re-engineer data management systems for increased efficiency and speed

Course Format

This is **not** a self-paced MOOC, but an intense project-based lab course. Your weekly work schedule is:

- **Monday&Tuesday:** Read and watch the materials for the week. Start the upcoming lab.
- **Wednesday**: Post your initial response to the discussion question by 11:59 pm CST.
- **Thursday**:   Review more materials for the week; finish lab and start testing your code
- **Friday**: Respond to your peer's discussion post AND post lab (if due) by 11:59 pm CST.

Textbooks & Materials

- **Tools**:
    - A laptop computer (Windows, Mac, or Linux) with at least 4GB RAM and a webcam
    - Microsoft Visio, free for Valparaiso University students through Imagine.
    - Oracle VIrtualBox, available at http://www.oracle.com/technetwork/server-storage/virtualbox/downloads/index.html . This course uses a number of Virtual Machines, which will be distributed through Blackboard links for the weeks in which they will be used.
    - Screen-cast-o-matic for your team's final case study demo video.
- **Books etc.**
    - Wiese, Lena (2015), *Advanced Data Management For SQL, NoSQL, Cloud and Distributed Databases*. De Gruyter. https://www.degruyter.com/view/product/460529?format=B
    - Lynda.com at Valparaiso University. https://apps.valpo.edu/authentication/lynda/

Workload and Grading

This 2-credit course requires individual and team-based work and may involve a few hours of independent research a week. You will be completing the following:

1. **Weekly Discussions (15*(5 points for post; 5 points for response) =150 points)**: Every Wednesday, you will answer the discussion question for the week. This is usually a rather substantial assignment requiring at least 1 hour of work. Every Friday, you will respond to two of your peers and evaluate their solutions. Speed, correctness, and thoughtful responses that help your peer increase their understanding of the topic will earn you full credit.
2. **Labs (4*10 points each=40 points)**:   The course contains 4 larger team-based lab assignments about a dataset you will receive in week 2. Solutions must be posted on Blackboard by 11:59 pm CST on Friday. **NO EMAIL SUBMISSIONS ACCEPTED.**
3. **Final Lab (10 points)**: Your final project will be a Big Data case study together with a dataset description, an evaluation of the optimal management solution, and an associated architecture. Using Screencast-o-matic, you will produce a video presentation of your case study and post it to YouTube for inclusion in your e-Portfolio.

You can earn up to 200 points in this course. No extra credit assignments will be given.

Letter Grade Conversion:

A	>93%	A-	90 to 93%	B+	87 to <90%	B	83 to <87%	B-	80 to <83%
C+	77 to <80%	C	73 to < 77%	C-	70 to <73%	F	<70%		

Assignment Submission, Late Work, and Academic Honesty

- **Assignment Submission:** All Assignments must be submitted on Blackboard. **No emailed Assignments will be accepted.**
- **Late Work:** Labs are considered late if not posted to Blackboard by Sunday 11:59 pm CST of the week in which they are due. **Late work will lose 50% of the grade. Work submitted more than 1 week late will receive 1 point only.**
- **Academic Honesty:** All work you submit for this course must be your own. You may NOT use anyone else's words (from blogs, webpages, purchased solutions, etc.) without giving a clear source citation. If you are unsure, consult http://www.plagiarism.org/ or the Writing Center. In addition, you must write and sign with your name the following statement on all course work:

  I have neither given nor received, nor have I tolerated others' use of unauthorized aid.

  For more information about Valparaiso University's Academic Honor Code, case review cycles, and potential penalties, please refer to http://www.valpo.edu/student/honorcouncil/index.php

  Any work suspected in noncompliance the Valparaiso University Honor Code will receive 0 points and be referred to the Graduate School for adjudication.

Diversity and Inclusion

Valparaiso University aspires to create and maintain a welcoming environment built on participation, mutual respect, freedom, faith, competency, positive regard, and inclusion. This course will not tolerate language or behavior that demeans members of our learning community based on age, ethnicity, race, color, religion, sexual orientation, gender identity, biological sex, disabilities (visible and invisible), socio-economic status, or national origin. The success of this class relies on all students' contribution to an anti-discriminatory environment where everyone feels safe, welcome, and encouraged to engage, to explore, and ultimately, "to embark on a rewarding personal and professional journey" (Pres. Heckler).

Title IX

Valparaiso University strives to provide an environment free of discrimination, harassment, and sexual misconduct (sexual harassment, sexual violence, dating violence, domestic violence, and stalking). If you have been the victim of sexual misconduct, we encourage you to report the incident. If you report the incident to a University faculty member or instructor, she or he must notify the University's Title IX Coordinator about the basic facts of the incident. Disclosures to University faculty or instructors of sexual misconduct incidents are not confidential under Title IX. Confidential support services available on campus include: Sexual Assault Awareness & Facilitative Education Office "SAAFE" (219-464-6789), Counseling Center (219-464-5002), University Pastors (219-464-5093), and Student Health Center (219-464-5060). For more information, visit http://www.valpo.edu/titleix/ .

Disability Support Services

As part of its mission for its students, Valparaiso University has developed a nondiscrimination policy which identifies its intention to provide a safe and tolerant environment for all, including those with disabilities. Please contact the Director of Disability Support Services, at 6496, if you believe you have a disability that might require a reasonable accommodation in order for you to perform as expected in this class. See also http://www.valpo.edu/disabilityss/

Academic Support

To get help, use the Academic Success Center (ASC) online directory (valpo.edu/academicsuccess) or contact the ASC (academic.success@valpo.edu) to help point you in the right direction for academic support resources for this course. Valpo's learning centers offer a variety of programs and services that provide group and individual learning assistance for many subject areas. These learning centers include:

- Graduate Tutoring Lab: Serves the academic needs of Graduate students – tutors offer suggestions on organization of papers, assist in research and citations, and help in understanding difficult assignments. Additional one on one tutoring is also available.
- Writing Center: Primarily serves the needs of undergraduate students, but is also available for Graduate students. Writing Consultants provide proofreading and editing assistance for papers and assignments.

Class Cancellations

Notifications of class cancellations will be made through Blackboard with as much advance notice as possible. It will be both posted on Blackboard and sent to your Valpo e-mail address. If you don't check your Valpo e-mail account regularly or have it set-up to be forwarded to your preferred e-mail account, you may not get the message. Please check Blackboard and your Valpo e-mail (or the e-mail address it forwards to) before coming to class.

Schedule

Week	Week Starting	Topic	Readings and Videos (read and watch before Monday)	Due by 11:59 pm CST W = Wednesday F = Friday
1	08/22/2017	Introduction and Background: Database Design and Modeling, Unified Modeling Language	Wiese Chapter 1 Lynda.com course on Learning Relational Databases, Chapter 1-4 Other materials as posted on Blackboard	W: Discussion F: Response
2	08/28/2017	Relational Database Management Systems and Query Processing	Wiese Chapter 2 Lynda.com course on Learning Relational Databases, Chapter 5-End Other materials as posted on Blackboard	W: Discussion F: Response
3	09/04/2017	NoSQL and the Cloud	Wiese Chapter 3 Lynda.com course on NoSQL for SQL Professionals Other materials as posted on Blackboard	W: Discussion F: LAB_01, Response
4	09/11/2017	Graph Databases	Wiese Chapter 4 Other materials as posted on Blackboard	W: Discussion F: Response
5	09/18/2017	XML Databases	Wiese Chapter 5 Other materials as posted on Blackboard	W: Discussion F: Response
6	09/25/2017	Key-Value Stores and Document Databases	Wiese Chapter 6 Other materials as posted on Blackboard	W: Discussion F: LAB_02, Response
7	10/02/2017	Column Stores	Wiese Chapter 7 Other materials as posted on Blackboard	W: Discussion F: Response

8	10/09/2017	Extensible Record Stores	Wiese Chapter 8 Other materials as posted on Blackboard	W: Discussion F: Response
9	10/16/2017	Object Databases	Wiese Chapter 9 Other materials as posted on Blackboard	W: Discussion F: LAB_03, Response
10	10/23/2017	Distributed Database Systems	Wiese Chapter 10 Other materials as posted on Blackboard	W: Discussion F: Response
11	10/30/2017	Data Fragmentation	Wiese Chapter 11 Other materials as posted on Blackboard	W: Discussion F: Response
12	11/06/2017	Replication and Synchronization	Wiese Chapter 12 Other materials as posted on Blackboard	W: Discussion F: Response
13	11/13/2017	Consistency	Wiese Chapter 13 Other materials as posted on Blackboard	W: Discussion F: LAB_04, Response
THANKSGIVING BREAK				
14	11/27/2017	More Database Technologies	Wiese Chapter 14 Other materials as posted on Blackboard	W: Discussion F: Response
15	12/04/2017	Database Engineering and Polyglot Architectures	Wiese Chapter 15 Other materials as posted on Blackboard	W: Discussion F: LAB_05 (Final Lab) (Video Presentation on YouTube), Response

APPENDIX

Student Learning Objectives—Graduate School
1. Students will understand and practice methods of inquiry and strategies of interpretation within the student's field of study.
2. Students will master the knowledge and skills pertinent to the student's field of study.
3. Students will effectively articulate the ideas, concepts, and methods through written and oral presentation.
4. Students will understand the connection between their knowledge and skills on the one hand and their professional identity, responsibilities, and demands on the other.
5. Students will integrate knowledge and methods of their study with cognates and other disciplines.
6. Students will study, reflect upon, and practice ethical behavior and cultural sensitivity as they relate to professional and personal responsibility.

Student Learning Objectives—Information Technology Program
1. To understand and practice methods of inquiry and strategies of interpretation within the student's field of study.
    1A. Students will master several programming environments.
    1B. Students will learn to identify and isolate problems.
2. To master the knowledge and skills pertinent to the student's field of study.
    2A. Students will acquire an extensive technology related vocabulary.
    2B. Students will become comfortable using a wide range of technology environments.
3. To effectively articulate the ideas, concepts, and methods through written and oral presentation.
    3A. Students will be taught how to make formal, oral presentations and be required to give 6 such presentations during their program.
    3B. Students will write numerous, thorough papers requiring extensive research. They will be required to use the   services on the writing center.
4. To understand the connection between their knowledge and skills on one hand and their professional identity, responsibilities, and demands on the other.
    4A. Students will understand the implications of legal and professional regulations as they relate to information technology.
    4B. Students will study how technology can be made available to people that are traditionally less advantaged.
5. To integrate knowledge and methods of their study with cognates and other disciplines.
    5A. Students will learn techniques of modeling data from other disciplines.
    5B. Students will study human factors in IT.
6. To practice ethical and cultural sensitivity as it relates to professional and personal responsibility.
    6A. Students will examine a wide range of ethical issues related to technology and the potential effects on people and the environment.
    6B. Students will explore the relationship between IT and ethnic and cultural diversity.