Sample syllabus - students receive the detailed syllabus at the beginning of the semester enrolled in the class.

CS 598 Data Curation

Course Description

Welcome to CS 598 Data Curation! Data curation is the active and on-going management of data through its lifecycle of interest and usefulness to scholarship, science, and education; curation activities and policies enable data discovery and retrieval, maintain data quality and add value, and provide for re-use over time. This course provides an overview of a broad range of theoretical and practical problems in the emerging field, examining issues related to appraisal and selection, long-lived data collections, research lifecycles, workflows, metadata, and legal and intellectual property issues.

Course Goals and Objectives

Upon successful completion of this course, you will be able to:

- Describe the significance of abstraction in data management and the relationships among the common key data abstraction strategies
- Understand the nature of representation hierarchies and strategies for data transformation and transcoding
- Explain the process of data derivation and the importance of provenance documentation
- Compare and contrast various data preservation strategies
- Understand the importance of dataset identifiers and citation
- Describe management of heterogeneity, including schema matching techniques
- Explain the role metadata plays in data management and identify a variety of metadata schemes
- Describe common data behaviors of managers, programmers, scientists, and other users
- Summarize the role institutions, agencies, policies, and laws play in data curation

Textbook and Readings

There is no required textbook for this course, but there are weekly required readings that can be found in each weekly overview page.

Course Outline

This 4-credit hour course is 16 weeks long. You should invest 10-12 hours every week in this course.

Week		Topics
1		Orientation, Introduction to Data Curation
2		Data Models: Relational Model
3		Trees, Text and Documents
4		Data Models: Ontologies; Schemas; Abstractions; Conceptual Modeling
5		Data Cleaning and Integration; Managing, Processing, and Policy Heterogeneity; Schema Integration
6		Data Concepts; Identity Problems; Ontology for Data Concepts
7		Metadata
8		Preservation
9		Identifiers
10		Standards
11		Workflow, Provenance, and Reproducibility
12		Communication

13		Practices
14		Policy, Law, and Ethics
15		Organization and Governance
16		Review

Assignment Deadlines

For all assignment deadlines, please refer to the **Course Assignment Deadlines, Late Policy, and Academic Calendar** page.

Elements of This Course

The course is comprised of the following elements:

- **Lecture Videos.** In each week, the concepts you need to know will be presented through a collection of short video lectures. You may stream these videos for playback within the browser by clicking on their titles or download the videos. You may also download the slides that go along with the videos. **The videos usually total 1.5 to 2 hours each week.** You generally should spend at least the same amount of time digesting content in the video. The actual amount of time needed to digest the content will vary based on your background.
- **Orientation Quiz.** The purpose of the orientation quiz is to ensure that you have gone through the orientation module and acquired the necessary information about the course before you start it. The orientation quiz is a required activity, but it's not part of the course grading. You have unlimited attempts on the orientation quiz. You need to answer all questions correctly in order to pass the orientation quiz.
- **Weekly Quizzes**. Each week concludes with an ungraded quiz to help ensure you understood that week's content. You will be allowed unlimited attempts for each quiz, and there is no time limit on how long you take to complete each attempt at the quiz.
- **Exercises.** There are three exercises for you to complete in this course, each of which will account for 20% of your final grade. You will submit this assignment for peer review to get feedback from your classmates. You will then incorporate the feedback you receive and submit a final version of your exercise to the instructor and TAs for grading. You will be allowed one submission attempt for

these exercises. Though you are encouraged to discuss these assignments with your classmates, everyone must submit their own work.

- **Final Project.** The course concludes with a final project in lieu of a final exam. It will account for 40% of your final grade. You will also submit your final project for peer review, incorporate that feedback, and submit your final project to the instructor and TAs for grading. For more information about the final project, please read the About the Final Project page in the course orientation.

Grading Distribution and Scale

Grading Distribution

Assignment	Percent of the Final Grade
Monthly Exercises	60% (20% each)
Final Project	40%

Grading Scale

Letter Grade	Percent Needed	Letter Grade	Percent Needed	Letter Grade	Percent Needed
A+	**95%**	**B+**	**85%**	**C**	**70%**
A	**90%**	**B**	**80%**	**D**	**60%**
A-	**88%**	**B-**	**78%**	**F**	**Below 58%**

Student Code and Policies

A student at the University of Illinois at the Urbana-Champaign campus is a member of a University community of which all members have at least the rights and responsibilities common to all citizens, free from institutional censorship; affiliation with the University as a student does not diminish the rights or responsibilities held by a student or any other community member as a citizen of larger communities of the state, the nation, and the world. See the University of Illinois Student Code for more information.

Academic Integrity

All students are expected to abide by the campus regulations on academic integrity found in the Student Code of Conduct. These standards will be enforced and infractions of these rules will not be tolerated in this course. Sharing, copying, or providing any part of a homework solution or code is an infraction of the University's rules on academic integrity. We will be actively looking for violations of this policy in homework and project submissions. Any violation will be punished as severely as possible with sanctions and penalties typically ranging from a failing grade on this assignment up to a failing grade in the course, including a letter of the offending infraction kept in the student's permanent university record.

Again, a good rule of thumb: *Keep every typed word and piece of code your own*. If you think you are operating in a gray area, you probably are. If you would like clarification on specifics, please contact the course staff.

Disability Accommodations

Students with learning, physical, or other disabilities requiring assistance should contact the instructor as soon as possible. If you're unsure if this applies to you or think it may, please contact the instructor and Disability Resources and Educational Services (DRES) as soon as possible. You can contact DRES at 1207 S. Oak Street, Champaign, via phone at (217) 333-1970, or via email at disability@illinois.edu.

Assignment Deadlines

Assignment	Release Date	Hard Deadline
Assignment 1	First day of class	Sunday of Week 4
Assignment 2	First day of class	Sunday of Week 8
Assignment 3	First day of class	Sunday of Week 12
Final Project	First day of class	Sunday of Week 16

Late Policy

- Unless otherwise specified, all assignments are due at **11:59 PM US Central Time** on the due date. (Time Zone Converter)
- No late assignments will be accepted without instructor approval prior to the assignment due date.

Academic Calendar

- The Graduate College at the University of Illinois maintains a Graduate College Calendar. The calendar includes important dates such as final exam dates, course registration and cancellation, and holidays.
- There is also a campus-wide calendar available.
- The CS Department also sends reminders about upcoming deadlines. You will also receive the Graduate College newsletter in your Exchange email account.

*Syllabus is subject to change