Course Syllabus

Edit

Instructor:

Dr. Nicholas Weber (please call me Nic)

email: nmweber@uw.edu;

phone: 206-221-6973

Office hours: Friday 1pm on Google Hangouts   (or by appointment)

TA:

An Yan (please call her An)

email: yanan15@uw.edu

Course Description

Data curation is the active and ongoing management of data throughout a lifecycle of use. This course provides a practical introduction to the field. It is meant to give you survey of topics for libraries, public institutions, and private corporations that are engaged in stewarding data for reuse. Throughout the quarter we will read, discuss, and engage in activities that are designed to give you a competency in data curation topics. Leaving this course you will be able to employ the correct vocabulary in describing data curation challenges, identify and select from a range of potential approaches to solving those challenges, and finally, take some initial steps to implement an appropriate solution.

This course **does not** offer an in-depth or even cursory introduction to managing databases, IT systems, or data warehouses. (However, there are a number of other courses in the INFX tract that are designed to do exactly this. I am happy to direct you to those if you have registered for the wrong course.)

Course Overview

 This course is designed around three modules: (1) Data, (2) Data Systems, and (3) Privacy, Policy and Ethics. At the end of each module you will be given a short quiz that covers important concepts or ideas that you should retain (more on quizzes below). Each module is composed of three related topics**.** Each topic has a set of readings and materials that you should engage with over the course of one week. (3 modules X 3 topics = 9 weeks of content)

The modules and topics are outlined in the following table. Each topic is a link to the page that has your readings, lecure slides, homework, and other instructions.

Data	Data Systems	Policy, Privacy, & Ethics
Types & Roles	Repositories	Policy
Documentation	Preservation	Privacy
Standardization	Cost Models	Ethics

Course Schedule

Lectures: On Monday of each week I will post a lecture that outlines important concepts that you will engage with in readings. These lectures are meant to introduce you to important topics or ideas in data curation.

Labs / Demonstrations: On Wednesday of each week I will post a lecture, demonstration, or more practically oriented lecture. These sessions are meant to introduce you to some of the more procedural aspects of doing data curation.

Chat: On Friday at 1pm (PT) I will be online for open office hours. You can drop by these hours to ask questions, get feedback on assignments, or talk about anything related to data curation. Use this Hangout group for Discussion: https://hangouts.google.com/call/qiw74jrhE-p_za6zlSBLAAEE

Discussion: You are not required to use the discussion forum in this class. However, it makes for a much more interesting and dynamic course if you do. I highly encourage you to post thoughts about readings, interesting links, or anything else you find relevant to the discussion forums.

 I have created a repository for our course on Github.   I will ocassionaly need to link to content there (data, software, etc). It may help to create a Github account if you do not already have one. Instructions here

Grades
Your grade has four components, describe in detail below.

1. Homework - 30% (6 homework assignments x 5% each)
2. Quizzes - 30% (3 quizzes x 10% each)
3. Data Profile - 20%
4. Curation Protocol - 20%

Homework (30% of total grade)
Homework assignments are meant to give you practice in applying concepts discussed in lectures and readings.
I will provide a short set of questions, or instructions and ask you to submit short answers (2-3 paragraphs). Homework should take you no more than 1 hour per week to complete. If you give a substantial amount of effort in your homework you will receive full credit.

There will be 6 total homework assignments.

Homework is due on Sunday evening at 11:59PM. For a full schedule of due dates, please consult the course calendar.

Quizzes (30% of total grade)

Each module will conclude with a short 10 question quiz. Your answers will consist of short written responses (e.g. 2-4 sentences).
Quizzes are meant to act as a test of how well you retained the essential content of the module. If you keep up with readings and lectures these quizzes will not be difficult.
A quiz will be made available on a Friday, and you must submit a response by the following Sunday evening (11:59pm). If you need to reschedule a quiz you must contact me in advance.

Dates for each quiz will be as follows:

- Module 1 Quiz - 01.26-01.28
- Module 2 Quiz - 02.16-02.18
- Module 3 Quiz - 03.09-03.11

Dataset Profile (20% of total grade)
You will write one paper that reviews a dataset publicly available online. The directions for this assignment clearly describe the requirements, and the rubric I will use to determine your mark.

There are two due dates associated with this assignment:
- January 17th - post to a discussion forum
(https://canvas.uw.edu/courses/1128207/discussion_topics/4102619) the dataset you have selected.
- January 31st - submit your profile to Canvas.

Curation Protocol (20% of your total grade)
You will be given a source of data, and a candidate use-case.

Your responsibility for this assignment will be to curate this data to satisfy the use-case. In doing so, you will document your curation in the form of a repeatable protocol.
The directions for this assignment clearly describe the requirements, and the rubric I will use to determine your mark.

The due date for your protocol will be March 7th.

Technology and Software Requirements

The iSchool has a set of technology requirements for both online and residential students. We highly recommend that students adhere to these standards which are updated annually and located at https://canvas.uw.edu/courses/721562/pages/technology-requirements. Students who do not meet these standards may experience technology problems throughout the course.

In this course you may be required to access a large number of databases through the Internet. Several of these databases are publicly available, but some are proprietary and access requires authentication through the UW Libraries   . Information about logging in to use these databases is available on the Connecting to the Libraries page   .

Academic Conduct

Please review the iSchool Academic Policies   which cover: Students with Disabilities, Grading Criteria, Academic Conduct, Academic Integrity, Copyright, Privacy, Student Code of Conduct, Evaluation of Student Work.

Course Summary:

Date	Details	
Sun Jan 28, 2018	Data Quiz	due by 11:59pm
Wed Jan 31, 2018	Data Profile	due by 11:59pm
Sun Feb 18, 2018	Data Systems Quiz	due by 11:59pm
Wed Mar 7, 2018	Curation Protocol	due by 11:59pm
Sun Mar 11, 2018	Policy, Privacy, & Ethics Quiz	due by 11:59pm