

CS 432/532, Fall 2021 Syllabus

Course Overview

The Web has fundamentally changed how we learn, play, communicate, and work. Its influence has become so monumental that it has given birth to a new science: Web Science, or the science of decentralized information structures. Although Web Science is interdisciplinary by nature, this course will be focusing mainly on the computing aspects of the Web.

Catalog Description: Provides an overview of the World Wide Web and associated decentralized information structures, focusing mainly on the computing aspects of the Web: how it works, how it is used, and how it can be analyzed. Students will examine a number of topics including: web architecture, web characterization and analysis, web archiving, Web 2.0, social networks, collective intelligence, search engines, web mining, information diffusion on the web, and the Semantic Web.

Prerequisites: A grade of C or better in CS 361 and CS 330.

Course Objectives

After completing this course, you should be able to do the following:

- Explain the interdisciplinary nature of web science.
- Explain the main components of web architecture.
- Explain how search engines index and rank webpages given a query.
- Write Python code to interact with various web APIs and services, including web servers, web archives, Memento aggregators, and the Twitter API.
- Explain the impact of social networks on how information (and disinformation) is spread on the web.
- Explain the differences between classification and clustering algorithms.
- Identify appropriate machine learning algorithms to apply given a problem.

Course Delivery Method

All sections of this course will be delivered fully online, asynchronously (no meeting time requirement). Course materials, including lecture videos, will be distributed via Blackboard. All deadlines are based on the local timezone in Norfolk, VA.

Sections

CS 432 (undergraduate) sections:

- CRN 20629 - WEB2 (in Hampton Roads)
- CRN 20630 - WEB5 (in Virginia, but outside of Hampton Roads)
- CRN 20631 - WEB7 (in the US, but outside of Virginia)
- CRN 20636 - WEB8 (outside of the US)

CS 532 (graduate) sections:

- CRN 20632 - WEB2 (in Hampton Roads or in Virginia)

- CRN 20633 - WEB5 (in Virginia, but outside of Hampton Roads)
- CRN 20634 - WEB7 (in the US, but outside of Virginia)
- CRN 20635 - WEB8 (outside of the US)

Instructor Contact and Office Hours

Dr. Michele Weigle: mweigle at cs.odu.edu, <https://www.cs.odu.edu/~mweigle/>

My office hours will be *TBA*, or by appointment. All office hours will be held via Zoom (see Blackboard for the link to the Zoom meeting room). Students will be placed into the waiting room if I am already meeting with another student. If you cannot attend during regular office hours, please contact me to set up an alternate appointment time.

Requirements

Prerequisites

Prerequisites: Standing as an undergraduate senior, graduate student, or approval from the instructor. Undergraduates should have a grade of C or better in CS 361 and CS 330.

We will be using Python this semester. You are not required to know Python ahead of time, but since you are required to have previous programming experience, I expect you to be able to pick up the syntax quickly. If you are unfamiliar with Python, I strongly suggest that you spend a bit of time with it before the semester begins. Here are some good Intro to Python resources:

Add links

- [Introduction to Python](#)
- [CS 1110: Introduction to Computing Using Python](#) (Cornell University)
- [A Gentle Introduction to Programming Using Python](#) (MIT)
- [Python Track: Introduction](#) (Cal Tech)
- [Python in One Easy Lesson](#), Nick Parlante

Textbook

There is no required textbook, but here is a list of recommended books:

- [Python in a Nutshell](#) - 3rd edition available [via ODU](#)
 - mentioned in Module 2
- [Learning Python](#) - 5th edition available [via ODU](#)
 - mentioned in Module 2
- [An Introduction to Search Engines and Web Navigation](#)
 - used in Module 6
- [Search Engines Information Retrieval in Practice](#) - pdf available
 - used in Module 6
- [Introduction to Information Retrieval](#) - pdf available
 - used in Module 6
- [Networks, Crowds, and Markets: Reasoning About a Highly Connected World](#) - preprint available
 - used in Modules 7, 8
- [Programming Collective Intelligence](#) - available [via ODU](#)

- used in Modules 11-14
- [Speech and Language Processing](#) - pdf available
 - mentioned in Modules 2, 6, 13

Grading

There will be a total of 100 points available for the semester.

Assignment Types

Your grade in this class will be based on the following components:

Learning Checks - 5 points

- Preparation for the week's topic
- Learning Checks released via Blackboard on Saturdays and due on Fridays
- I will not grade your submissions for correctness, but correct answers with references will be provided after the due date.
- Full credit will be awarded only if all Learning Checks are complete and have been submitted on time.

Online Discussion - 5 points

- Class discussion based on the week's topic or current events.
- Discussion Question released via Piazza on Saturdays. Your initial response is due on Fridays and a follow-up to at least one classmate's posting is due on Sundays.
- Full credit will be awarded only if all Discussion Questions have been answered on time and follow-ups have been posted on time in more than half of the weeks (so 8 follow-ups if we have 14 weeks of discussion questions).

Homework (HW) - 90 points

- Implement concepts from class along with a report
- 9 homework assignments (HW1-HW9), each worth 10 points
- Released on Saturdays and usually due 2 weeks later on Sunday
- *Don't wait until the last minute -- these are hard and time-consuming!*
- Grading scale: 0-10
 - 2 points are reserved for the required report
 - many homework assignments will have extra credit opportunities
- All work must be your own. You may use resources on the Internet for reference, but you must not copy large sections of code. If you use online resources, you must cite your sources (including URL). Group work on HW assignments is not acceptable.

There is no mid-term or final exam

Grading Scale

The grading scale for undergraduate students (CS 432) is as follows:

points	letter	points	letter
100-92	A	77-72	C

points	letter	points	letter
91-90	A-	71-70	C-
89-87	B+	69-67	D+
86-82	B	66-62	D
81-80	B-	61-60	D-
79-77	C+	59-00	F

The grading scale for graduate students (CS 532) is as follows:

points	letter	points	letter
100-94	A	79-78	C+
93-90	A-	77-74	C
89-88	B+	73-70	C-
87-84	B	69-0	F
83-80	B-		

Late Assignments

Any assignment submitted after its deadline is considered late. Late assignments lose 1 point for every 24 hours they are late. Submissions over 72 hours late are not accepted. This time limit includes weekends -- they are counted just like weekdays. I reserve the right to specify that late submissions will not be accepted for particular assignments.

- 0-24 hours late: -1 point
- 25-48 hours late: -2 points
- 49-72 hours late: -3 points
- over 72 hours late: not accepted

Summary Schedule

Note: This is a tentative schedule and may change during the semester. The complete schedule with assignments and due dates is posted on Blackboard.

Typical weekly schedule:

- Saturday - new module course materials, Learning Checks (LC), Discussion Question (DQ), and HW assignment released
- Friday - LC due, initial DQ response due
- Sunday - followup DQ response due, HW assignment due (usually from a previous week)

[ODU Fall 2021 academic schedule](#)

Week	Week Start Date	Topic
------	-----------------	-------

Week	Week Start Date	Topic
1	Aug 28	Introduction to Web Science and Web Architecture
2	Sep 4	Introduction to Python
3	Sep 11	Introduction to Info Vis with R, Python
4	Sep 18	Measuring the Web
5	Sep 25	Archiving the Web
6	Oct 2	Searching the Web
	Oct 9	NO CLASS - Fall Break
7	Oct 16	Social Networks
8	Oct 23	Selection and Social Influence
9	Oct 30	Visualizing Social Networks
10	Nov 6	Disinformation
11	Nov 13	Collective Intelligence and Recommender Systems
12	Nov 20	Clustering Algorithms Nov 24-28 - Thanksgiving Break
13	Mon, Nov 29	Document Filtering (Classification)
14	Dec 3	kNN and Algorithm Summary
	Fri, Dec 10	<i>last day of classes</i>

Related Courses

- All things HTTP - [CS 531 - Web Server Design](#)
- Information retrieval, metadata - [CS 834 - Intro to Information Retrieval](#)
- Visualization, Analytics - [CS 625 - Data Visualization](#), [CS 725/825 - Information Visualization](#)
- Web programming, LAMP - [CS 418/518 - Web Programming](#)
- Data science - [CS 620 - Intro to Data Science](#)

Course History

This course at ODU was originally developed by [Dr. Michael Nelson](#), based on Dr. Frank McCown (ODU alum)'s Intro to Web Science course at Harding University. In previous semesters, it has also been taught by Dr. Alexander Nwala (former PhD student at ODU).

Course Policies

Email/Piazza

Each student must check the class Piazza site and email daily. You should use Piazza to ask and answer general course-related questions. I will also use this to notify you about important updates (assignment

deadline changes, office hours cancellations, etc.).

Participation

Since this is an online class, it is essential that you regularly stay involved in class activities. This includes checking the class website and Piazza for announcements, submitting assignments on time, and responding to discussion posts in a timely manner.

Make-up Work

Make-ups for graded activities are possible only with a valid written medical or university excuse. It is the student's responsibility to give the instructor the written excuse and to arrange for any makeup work to be done.

Disability Services

In compliance with PL94-142 and more recent federal legislation affirming the rights of disabled individuals, provisions will be made for students with special needs on an individual basis. The student must have been identified as special needs by the university and an appropriate letter must be provided to the course instructor. Provision will be made based upon written guidelines from the University's [Office of Educational Accessibility](#). All students are expected to fulfill all course requirements.

Students are encouraged to self-disclose disabilities that have been verified by the Office of Educational Accessibility by providing Accommodation Letters to their instructors early in the semester in order to start receiving accommodations. Accommodations will not be made until the Accommodation Letters are provided to instructors each semester.

Seeking Help

The course Blackboard site should be your first reference for questions about the class. If you have questions about course requirements or materials, post questions using the class Piazza site. For extra help, attend office hours.

Academic Integrity

Old Dominion University is committed to students' personal and academic success. In order to achieve this vision, students, faculty, and staff work together to create an environment that provides the best opportunity for academic inquiry and learning. All students must be honest and forthright in their academic studies. Your work in this course and classroom behavior must align with the expectations outlined in the Code of Student Conduct, which can be found at <https://www.odu.edu/oscai>.

The following behaviors along with classroom disruptions violate this policy, corrupt the educational process, and will not be tolerated.

- Cheating: Using unauthorized assistance, materials, study aids, or other information in any academic exercise.
- Plagiarism: Using someone else's language, ideas, or other original material without acknowledging its source in any academic exercise.
- Fabrication: Inventing, altering or falsifying any data, citation or information in any academic exercise.

- Facilitation: Helping another student commit, or attempt to commit, any Academic Integrity violation, or failure to report suspected Academic Integrity violations to a faculty member.

In particular, submitting anything that is not your own work without proper attribution (giving credit to the original author) is plagiarism and is considered to be an academic integrity violation. It is not acceptable to copy source code or written work from any other source (including other students, online resources), unless explicitly allowed in the assignment statement. In cases where using resources such as the Internet is allowed, proper attribution must be given.

Any evidence of an academic integrity violation (cheating) will result in a 0 grade for the assignment/exam, and the incident will be submitted to the Department of Computer Science for further review. Note that academic integrity violations can result in a permanent notation being placed on the student's transcript or even expulsion from the University. Evidence of cheating may include a student being unable to satisfactorily answer questions asked by the instructor about a submitted solution. Cheating includes not only receiving unauthorized assistance, but also giving unauthorized assistance. For class files kept in Unix space, students are expected to use Unix file permission protections (chmod) to keep other students from accessing the files. Failure to adequately protect files may result in a student being held responsible for giving unauthorized assistance, even if not directly aware of it.

Students may still provide legitimate assistance to one another. You are encouraged to form study groups to discuss course topics. Students should avoid discussions of solutions to ongoing assignments and should not, under any circumstances, show or share code solutions for an ongoing assignment.

All students are responsible for knowing the rules. If you are unclear about whether a certain activity is allowed or not, please contact the instructor.

More information on academic integrity is available on the ODU-CS [academic integrity page](#).