CS208: Applied Privacy for Data Science
Spring 2022 Syllabus

Course Website: http://seas.harvard.edu/~salil/cs208/
Time: Tuesdays & Thursdays 11:15-12:30, starting January 25
Place: Harvard Science & Engineering Complex (SEC) LL2.211, 150 Western Ave, Allston, MA

Course Staff

Instructors:
Salil Vadhan (he/him/his; they/them/theirs)
http://salil.seas.harvard.edu/, salil_vadhan@harvard.edu, SEC 3.327

James Honaker (he/him/his)
http://hona.kr/, james@hona.kr, SEC 4.442 / 4.444

Wanrong Zhang (she/her/hers)
https://wanrongz.github.io, wanrongzhang@fas.harvard.edu, SEC 3.330

Teaching Fellows:
Daniel Alabi (he/him/his)
alabid@g.harvard.edu, SEC 3.334

Jayshree Sarathy (she/her/hers)
jsarathy@g.harvard.edu, SEC 3.331

Michael Shoemate [half time] (he/him/his)
shoematem@g.harvard.edu, SEC 4.447

Connor Wagaman [half time] (he/him/his)
wagaman@college.harvard.edu

Faculty Coordinator:
Allison Choat (she/her/hers; they/them/theirs), achoat@seas.harvard.edu

Whenever possible, please post questions for the course staff on Ed (privately if needed) rather than emailing us, so that we can all see the question and responses.

Overview

Data scientists, including industry analysts, scientific researchers and data-driven policy makers, often want to analyze data that contains sensitive personal information that must remain private. However, common techniques for data sharing that attempt to preserve privacy either bring great privacy risks or great loss of information. Moreover, the increasing ability of big data, ubiquitous sensors, and social media to record lives in detail brings new ethical responsibilities to safeguard privacy.

The traditional approach to protecting privacy when sharing data is to remove "personally identifiable information," but it is now known that this approach does not work, because seemingly innocuous information is often sufficient to uniquely identify individuals. A long literature has shown that anonymization techniques for data releases are generally open to re-identification attacks. Indeed, there have been many high-profile examples in which individuals in supposedly anonymized datasets were re-identified by linking the remaining fields with other, publicly available datasets. Aggregated information can reduce but not prevent this risk, while also reducing the utility of the data to researchers.

This class will provide an overview of the risks of private data leakage in data science applications and a firm foundation in how to measure and protect against these risks using the framework of differential privacy, together with a hands-on examination of how to build algorithms and software to preserve privacy, including a review of the deployed solutions in industry and government.

Differential privacy, deriving from roots in cryptography, is a formal, mathematical conception of privacy preservation. It guarantees that any released statistical result does not reveal information about any single individual. That is, the distribution of answers one would get with differentially private algorithms from a dataset that does not include myself must be indistinguishable from the distribution of answers where I have added my own information.

Using differential privacy enables us to provide wide access to statistical information from a privacy sensitive dataset without worries of individual-level information being leaked inadvertently or due to an adversarial attack. There is now both a rich theoretical literature on differential privacy and numerous efforts to bring differential privacy closer to practice, including large-scale deployments by Google, Apple, Microsoft and the US Census Bureau, as well as the OpenDP open-source software project that was founded here at Harvard. This course will set out a foundation in the underlying theory of differential privacy, and then consider the practical elements of implementing and deploying privacy-preserving techniques for data analysis.

Format and Goals

The class will have a mix of lecture/discussion meetings, which will focus on learning the fundamentals of the underlying theory and discussion of important issues, and practicum

sessions where there will be some lecture, some demonstration code, and some hands-on computer work. Homeworks will typically involve some analytical/mathematical work to learn techniques, and increasingly as the term progresses, hands-on data-immersive coding tasks to test and experiment with approaches to privacy preservation within the context of real datasets and data science questions.

The main components of the course are as follows:

- **Class Participation:** Attendance is mandatory, as our meetings will be highly interactive, especially our practicum meetings. For most class meetings, we will provide material or videos for you to read or watch and comment on in advance. Participation also includes your engagement in section and office hours and on Ed.
- **Problem Sets:** There will be problem sets due approximately once per week, until the last third of the semester, when the frequency will be reduced to give you more time to work on your projects. These will be progressive, and require reuse of previous solutions, so it is important both to keep up on the problems, review feedback to submissions, and organize and document previous submitted code so that it can be reused. We will drop your lowest problem set score when determining final grades.
- **Final Project:** You will do a final project on a topic of your choosing. Projects can be done individually or in pairs, with groups of three allowed for ambitious projects. You can do a project that is experimental, or involves system-building, or is theoretical. The project should provide good opportunities to connect the course material to your other interests and get some exposure to the frontier of research in differential privacy. The project will involve submitting topic ideas for feedback (due approx 4 weeks into the semester, with a revision a few weeks later), a detailed project proposal (due approx 3 weeks before the end of the semester), a written paper (draft due in reading period, final version in exam period), and a project presentation (in exam period). We will post more details about the final project, including some directions to look for topics, early in the course.

We anticipate placing roughly equal weight on each of the above three elements in determining final grades.

Late Days: You will have 6 late days for the semester, at most 3 of which can be used on any one problem set or final project milestone. These late days are meant to offer flexibility for minor disruptions, like mild illness or other deadlines you may have. Extensions beyond the late day policy require a note from your resident dean (for undergraduates) or advisor (for graduate students).

Learning Outcomes

By the end of the course, we hope that you will all be able to:

- Identify and demonstrate risks to privacy in data science settings,
- Correctly match differential privacy technology with an application,
- Safely implement privacy solutions, and experimentally validate the performance and utility of algorithms,

- Understand differential privacy at a level sufficient to engage in research about best practices in implementation, apply the material in practice, and/or connect it to other areas,
- Analyze the ethical and policy implications of differential privacy deployments,
- Formulate and carry out an interesting, short-term independent research project, and present the work in both written and oral form.

Prerequisites

The minimum prerequisites are basic probability, algorithms, and Python or R programming at the level of CS109/AC209. STAT110 and CS120 or CS124 should also be sufficient preparation.

Diversity and Inclusion[1]

We would like to create a learning environment in our class that supports a diversity of thoughts, perspectives and experiences, and honors your identities (including race, gender, class, sexuality, socioeconomic status, religion, ability, etc.). We (like many people) are still in the process of learning about diverse perspectives and identities. If something was said in class (by anyone) that made you feel uncomfortable, please talk to us about it. If you feel like your performance in the class is being impacted by your experiences outside of class, please don't hesitate to come and talk with us. As a participant in course discussions, you should also strive to be open-minded and respectful of your classmates.

Health Accommodations[2]

If you have a physical or mental health condition that affects your learning or classroom experience, please let us know as soon as possible so that we can do our best to support your learning (at minimum, providing all of the accommodations listed in your AEO letter if you have one).

Support Structures[3]

Everyone can benefit from support during challenging times. If you experience significant stress or worry, changes in mood, or problems eating or sleeping this semester, whether because of CS208 or other courses or factors, please do not hesitate to reach out to Salil or other members of the course staff. Not only are we happy to listen and discuss how we can help you cope in CS208, we can also refer you to additional support structures on campus, including, but not limited to, the below.

- Academic Resource Center
- InTouch

---

[1] Based on text by Dr. Monica Linden at Brown University.
[2] Based on text by Prof. Krzysztof Gajos at Harvard University.
[3] Based on text in the Harvard CS50 Syllabus.

- Counseling and Mental Health Services, 617-495-2042
- Let's Talk
- Room 13, 617-495-4969

COVID Policies

We are planning CS208 as an interactive in-person experience, and look forward to engaging with you all in the classroom together. However, we are mindful that the pandemic may cause disruptions. To minimize such risks, it is important that everyone is diligent about following Harvard's COVID policies, including masking in the classroom (please use a high-quality one, and wear it properly and consistently), staying home when ill, and testing according to your assigned cadence.

For students who are unable to come to class due to COVID isolation or symptoms, we will have a livestream of every class meeting on Zoom, and will also be recording lectures via Panopto. The livestream and recordings can be accessed through the class Canvas page. If you have questions or comments while watching the livestream, post them in the Ed discussion thread for that lecture and the staff or a fellow student can respond. We also encourage all of you to identify a "study buddies" who can take notes and gather information for you from lecture, section, and office hours in case you are not able to attend. While you are ultimately responsible for keeping up with the class while in isolation, do let us know if this occurs and we will do our best to help support you.

Auditor Policies

We are happy to have auditors attend the class, subject to following Harvard's latest COVID policies, including being fully vaccinated and boosted, masking, and at least weekly PCR testing. For auditors from outside Harvard who are not officially cross-registered in the class, email Susan Welby <swelby@seas.harvard.edu> and Allison Choat <achoat@seas.harvard.edu> so that we can apply to get you registered as a "Person of Interest (POI)" and thus eligible to participate in Harvard's COVID testing system.

Collaboration Policy

Students are encouraged to discuss the course material and the homework problems with each other in small groups (2-3 people). Discussion of homework problems may include brainstorming and talking through possible solutions, but should not include one person telling the others how to solve the problem. In addition, each person must write up their solutions independently, and these write-ups should not be checked against each other or passed around. While working on your problem sets, you should not refer to existing solutions, whether from other students, past offerings of this course, materials available on the internet, or elsewhere. All sources of ideas, including the names of any collaborators, must be listed on your submitted homework.

In general, we expect all students to abide by the Harvard College Honor Code. We view us all (teaching staff and students) as engaged in a *shared mission* of learning and discovery, not an adversarial process. The assignments we give and the rules we set for them (such as the collaboration policy) are designed with the aim of maximizing what you take away from the course. We trust that you will follow these rules, as doing so will maximize your own learning (and thus performance on exams) and will maintain a positive educational environment for everyone in the class. We welcome and will solicit feedback from you about what more we can do to support your learning.

Topics to be Covered

- Privacy attacks on "de-identified" data and statistical data releases
    - Reidentification attacks
    - Reconstruction attacks
    - Membership attacks
    - Interpretation and debates about the meaning of attacks
- Foundations of differential privacy
    - Definition, interpretation, and variants
    - Basic mechanisms (Laplace, Gaussian, histograms, exponential)
    - Composition of differential privacy
    - Survey of known algorithms and experimental validation
- Implementing (centralized) differential privacy
    - Statistical releases by the US Census Bureau, Opportunity Insights, and others
    - Privacy budgeting
    - Differentially private machine learning and deployments by Google and Facebook
    - Interactive query interfaces
    - Differentially private programming platforms such as OpenDP
- Distributed models of differential privacy
    - The local model: basic theory and mechanisms
    - Randomized response, histograms, SGD
    - Comparison with the centralized model
    - Deployments by Google and Apple
    - Intermediate models: shuffle, federated learning, DP+crypto
- Non-technical facets of privacy
    - Ethical considerations
    - Privacy law and policy
    - Power dynamics in sociotechnical systems
- Other possible topics (depending on time and interest)
    - Differential privacy for graph and social network data
    - Statistical inference under differential privacy
    - Side-channel & randomness attacks on implementations