CS2080: Applied Privacy for Data Science
Spring 2025 Syllabus

Course Website: https://opendp.github.io/cs208/spring2025/
Time: Mondays & Wednesdays 11:15am-12:30pm, starting January 27
Place: 114 Western Ave, Room 2.112, Allston, MA

Course Staff

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

James Honaker (he/they)
http://hona.kr/, james@hona.kr, SEC 4.447

Priyanka Nanayakkara (she/her/hers)
https://priyakalot.github.io/, priyankan@g.harvard.edu, SEC 2.101

Teaching Fellows:
Zachary Ratliff (he/him)
https://zacharyratliff.org/, zacharyratliff@g.harvard.edu, SEC 3.331

Christian Aagnes (he/him/his)
christianaagnes@g.harvard.edu

Sahil Kuchlous (he/him/his)
sahilkuchlous@college.harvard.edu

Jason Tang (he/him/his)
jasontang@college.harvard.edu

Yanis Vandecasteele (he/him)
yanis_vandecasteele@g.harvard.edu

Henry Wu (he/him/his)
hhwu@college.harvard.edu

Faculty Coordinator:
Allison Choat (she/they), 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 with names. 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 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, Meta, Microsoft, Apple, Mozilla, 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:** For every class, there will be pre-readings assigned, which you are expected to comment on using Perusall (more detailed instructions provided separately). In class, we will typically have group discussions about the reading, after which each of you will submit an individual response. Thus, attendance is mandatory, except in cases of absences excused for illness or other circumstances approved 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, typically Friday 5pm, 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 and a group of classmates will do a final project on a topic of your choosing. 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 poster 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 1:2:2 weighting on each of the above three elements, respectively, 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 program/research advisor (for graduate students), and otherwise will be penalized by a drop in modified letter grade per day (e.g. an A- would become a B after two excess late days).

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 at the level of STAT 110, and algorithms and Python or R programming at the level of CS1090A/AC209A or CS1200.

Diversity and Inclusion[1]

We would like to create a learning environment in our class that supports a diversity of thoughts, perspectives and experiences, honors your identities (including race, gender, class, sexuality, socioeconomic status, religion, ability, etc.), and gives us all a chance to learn from each other, even when we disagree. We (like many people) are still in the process of learning about diverse perspectives and identities. If something is said in class (by anyone) that feels inappropriate, 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).

In cases where you cannot come to class due to illness, you should inform the instructors so that your absence can be excused. If you are well enough, you can attend the Zoom livestream or watch the recording in Panopto. The livestream and recordings can be accessed through the

[1] Based on text by Dr. Monica Linden at Brown University.
[2] Based on text by Prof. Krzysztof Gajos at Harvard University.

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.

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 CS2080 or other courses or factors, please do not hesitate to reach out to one of us. Not only are we happy to listen and discuss how we can help you cope in CS2080, we can also refer you to additional support structures on campus, including, but not limited to, the below:

- Academic Resource Center
- InTouch
- Counseling and Mental Health Services, 617-495-2042
- Room 13, 617-495-4969

Auditor Policies

We are happy to have auditors attend the class, provided that you inform us (by filling out the first-class survey) and respect Harvard policies, including those governing access to student and course information.

Collaboration & AI 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, whether human collaborators, AI tools, or internet websites, must be cited on your submitted homework, with an explanation of how you used those resources.

ChatGPT and other AI tools should be treated similarly to collaboration with your peers in the class. You may use these tools to help you understand the material and as part of your brainstorming process, but you should not be asking the tools to solve the homework problems for you. If you do use such tools, you must cite them and submit a document listing all of the prompts you entered and responses obtained.

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

[3] Based on text in the Harvard CS50 Syllabus.

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 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.

Textbook and Readings

The recommended textbook for the class is *Hands-On Differential Privacy*, by Ethan Cowan, Michael Shoemate, and Mayana Pereira. We will also assign readings from many other sources throughout the semester; you can read these in the Perusall e-reader (which you will also use for commenting on the readings) and for most (but not Hands-On Differential Privacy) you will be able to download pdfs by following links from the course schedule (in some cases using Harvard's institutional journal subscriptions). We will also provide a continually updated annotated course bibliography to help you navigate the literature as you pursue your course projects and other interests in the field.

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 & other measures of privacy
    - Survey of known algorithms and experimental validation
- Implementing (centralized) differential privacy
    - Deployments by the US Census Bureau, Microsoft, Wikimedia, and others
    - Synthetic data releases and statistical releases
    - DP machine learning and deployments by Google and Meta
    - Programming platforms such as OpenDP
    - Interfaces & usability
    - Evaluating downstream utility
- Distributed models of differential privacy
    - Local vs. federated vs. centralized DP
    - Basic theory & mechanisms: randomized response, histograms, SGD
    - Combining DP with other PETS (e.g. secure multiparty computation)
    - Deployments by Google, Apple, Meta, Mozilla.
- Social perspectives on DP
    - Differential privacy in relation to other (non-CS) privacy philosophies
    - Communicating differential privacy guarantees to affected parties

    - 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

Other Courses that Cover DP

- CS 1260 “Fairness & Privacy: Perspectives of Law & Probability” (ugrad, F24, S26)
    - Tackles 3 concepts: Algorithmic Fairness, Differential Privacy and Cryptography - so broader and less time on DP (about 1mo)
    - Tighter integration with Law (alternates CS & Law lectures)
    - Less emphasis on deployments and implementation of DP (no implementation)
    - Limited enrollment
    - Final exam rather than project
    - Prerequisites: Experience writing proofs and familiarity with statistics basics is required; a theoretical computer science course such as CS 1210 (formerly CS 121) or CS 1240 (formerly CS 124) is recommended.
- Stat 188 “Variations, Information and Privacy” (ugrad, F24, F25?)
    - More depth on statistical aspects of DP, so less broad coverage of other facets
    - Final exam rather than project
    - Covers deployments like US 2020 Census, but not so much implementation
    - Prerequisites are Stat 110 & 111
- AC 221 “Critical Thinking in Data Science (grad, S25)
    - Brief coverage of differential privacy
- CS 2260 “Topics in Theory for Society: Differential Privacy” (grad, F25)
    - Focuses more on DP theory, less on DP practice
    - Has a final project
- Boston U. “Privacy in Statistic and Machine Learning” (grad, S25, TuTh 2pm-3:15pm)
    - Focuses more on DP theory, less on DP practice
    - The lecture notes are an excellent resource for DP theory
    - Has a final project