Final Project Guidelines

CS2080: Applied Privacy for Data Science, Spring 2025

The final projects in the course are meant to give you the opportunity to further explore an aspect of differential privacy that interests you, give you the experience of formulating, carrying out, and presenting an interesting, short-term independent project that is similar to an experience you might have in a career as an applied data scientist confronting privacy issues and/or as a data privacy researcher.[1]

Projects will be done in groups of three or four. (We cannot support solo or pair projects due to the size of the class.)

Many different types of projects are possible:

- Implement and experimentally evaluate differentially private algorithms or attacks on real-life datasets. For example, identify a dataset that resembles a sensitive data use case and a type of statistical analysis that would be useful on such a dataset, implement and tune a differentially private algorithm for that analysis and evaluate the privacy-utility tradeoff.
- Develop and implement a methodology for evaluating downstream utility (going beyond accuracy) of a data release under differential privacy
- Critically evaluate an existing system for privacy protection, identify potential vulnerabilities and propose or demonstrate improvements using techniques from this class.
- Explore how the “noisy” results from differential privacy can be properly incorporated into usable scientific or consumer data products in a specific use case.
- Explore how differential privacy might be incorporated into a larger system design (with some particular application domain in mind).
- Seek new theoretical results on some aspect of differential privacy or related topics.
- Connect differential privacy to some other area of interest to you (whether in CS or outside).
- Analyze how differential privacy relates and doesn’t relate to other theories of privacy (beyond contextual integrity, which we cover in class)
- Design, prototype, and evaluate a human-centered tool (e.g., an interactive interface) that supports data analysts or other users in using differentially-private algorithms
- Develop and evaluate tools that help express differential privacy’s guarantees or other relevant attributes to parties without CS expertise (e.g., the general public, policymakers)
- Model some new problem that might not be captured by the current literature
- Synthesize and write an exposition of several papers in the differential privacy research literature (beyond those covered in detail in class)
- Contribute to the OpenDP project; see here for a current wishlist.
- And more... be creative!

Refer to the course syllabus, reading list, and other suggestions we provide as sources of inspiration.

[1] We don't expect you to produce publishable results during the time you have this semester, but we do hope that some of the projects in the class can subsequently be developed into publishable research.

The steps for various aspects of the project are as follows:

- **Topic Ideas (Friday 2/28, revision on Friday 3/21):** As part of PS4 (due 2/28), submit about a paragraph of discussion about 1-3 ideas you have for potential project topics. For each topic, you should include the general kind of problem or use case that you'd like to address in your project and the general methodology (e.g. is it a theory project or an experimental project or...). The point of these topic ideas is both to get you thinking about the project early on and to enable us to give you early feedback and suggestions, which we will provide approximately one week. After that, we will set up a forum for you all to seek project partners with similar interests to yours, and your group will submit a revised, more detailed set of topic ideas, including an initial list of relevant papers, after Spring Break.
- **Project Description (Friday 4/11):** Submit a couple of pages giving a detailed description of what your final project will look like. You should be able to clearly state your research questions, briefly articulate how your project relates to what has been done in the past, describe the approach you are taking, give your timeline for completing various aspects of the project, and *discuss your fallback plan in case you don't obtain the results that you're hoping to obtain.*
- **Written paper (Weds 4/30, revision on Thursday 5/8):** Submit a paper (approx. 10 pages) describing your completed project. The paper should motivate your research questions and results, explain how the project fits into the context of previous work (with proper scholarly citations), justify the methodology, and present and interpret the results in a convincing manner. It should also include a brief description of what the different team members contributed to the project. (Naturally, the form will differ depending on the type of project.) You will submit a revised version of the paper based on the feedback from your draft.
- **Poster Presentation (date):** We will hold a poster presentation event. Your poster should motivate the problem you worked on, describe your approach, and present and interpret the results. As in your paper, be sure to give proper credit to previous work and clearly distinguish what you’ve done from what’s been done before. You should develop a short (2min) presentation of core findings that you can give to other attendees who visit your poster. We will provide a course code to pay for poster printing.

Some online resources for preparing talks and research papers:

- “How to write a technical paper” by Michael Ernst.
- “How to give a technical presentation” by Michael Ernst.
- “How to give a great research talk” by Simon Peyton-Jones.
- “How to write a great research paper: Simon's seven easy steps.” by Stephanie Weirich.
- “How to give a good research talk.” by Stephanie Weirich.
- “How to Give Bad Talk” by John Ousterhout, Tom Anderson, Dave Patterson (channeled by Mike Dahlin).

We encourage you to discuss your ideas with us throughout the process, at office hours or by email. We hope you enjoy the experience!