

Final Project

cs199r – Spring 2007

Important dates

- Thursday, 3/22 – Final project (FP) handout distributed.
- Tuesday, 4/3 – Open class period to discuss FP ideas.
- Friday, 4/13 – FP proposals due before 5 PM.
- Tuesday-Thursday, 5/8-10 – FP presentations.
- Monday, 5/14 – FP write-up due before 5 PM.

Goal

The goal of the final project is to allow you (or you and a partner) to explore an aspect of the topic of this course in greater detail. Every project should require research into an idea or issue beyond that directly discussed during the semester. The primary output of your research may be a position paper, experimental analysis, or engineering prototype. A later section in this document lists some possible FP ideas. Students are also encouraged to think of their own project ideas, with a focus on better understanding the interactions between privacy and information technology.

Before Friday, April 13th, you must propose a project and have it ok'd by the teaching staff. To help you in proposing an interesting and appropriate project, we have set aside a class period early in April where the teaching staff will be available to discuss your ideas. Between that class period and April 13th, you should do a little research and get started on your proposed FP so that you know it will be feasible to complete.

The only check point before the FP deadline is a presentation of your work to the teaching staff and the other students in the course during the first week of reading period. Each presentation will be 15 minutes in length (20 minutes for two-person FPs). During this short presentation, you will not be able to present all of the details of your work. Rather, you should present a concise problem statement, summarize your main findings or conclusions, and clearly answer questions posed by the course staff.

Particulars of the final write-up will vary from project to project and should be discussed with the course staff.

Project ideas

1. Prepare a briefing book for a high-level interdisciplinary task force discussing some new technology and its privacy implications. Explain the technology and key aspects of its implementation and potential use, and present a detailed list of critical discussion questions for this task force. The briefing should contain an annotated list of readings, discuss historical analogs, and recommend one or two potential experts to interview.

2. In discussion of surveillance, one proposed explanation for the fact that so much private information can be found on Facebook.com is that 17-year olds are historically bad at thinking about the consequences of their actions. Is this statement refutable? One might compare the type and amount of personally identifiable information (PII) released on Facebook.com vs. something like LinkedIn.com. In theory, there's a significant difference in age and purpose for the two sites. Is there some difference in the quantity of PII revealed? Which provides better information for identity theft? Which provides better data for other privacy-invasive purposes?
3. Learn about Harvard's initiative on Personally Controlled Health Records (PCHR) and develop a detailed threat model.
4. How many UPC codes that we carry around are required to identify us uniquely or for successfully tracking us through a fairly crowded area? Does the surveillance threat of RFID already exist, if one could read those UPC codes from some distance? How does the movement from product category identifiers to unique identifiers change this? Page 8 of the Juels article hints at this question.
5. Is there a reasonable way to reduce the amount of information that is disclosed during an electronic transaction? In particular, can one design a functional payment system where there is no link between payment and individuals (sort of like electronic cash, but in the context of credit card)? The least one hopes for is a system where credit card numbers are never revealed by the customers (related systems have been designed in the late 90's but haven't been deployed — it would be interesting to find out why not). More ambitiously, can we get rid of any disclosure of identifiers during a transaction (anonymous cash)? What are the implications in terms of enforcement?
6. Find out what scenarios are the most relevant/pressing/challenging in the context of protecting privacy in databases. What are the relevant difficulties? Look for existing solutions and list the respective advantages/disadvantages.
7. Long-term longitudinal studies (like the Framingham study) have been used to find surprising and valuable results in medicine. The push to electronic medical records offers a much larger data set for such research. But the privacy regulations from the HIPPA act restrict the use of "personally identifiable information" in research without the informed consent of those who are identifiable. What are the technical and policy considerations and tradeoffs between the search for medical data and personal privacy?
8. DNA testing provides lots of information, not only about the person tested, but about those who are closely related to that person. What data is being revealed by DNA testing, and what influence does this have in discussions about control of personal information?
9. Traditionally, the balance between the needs of law enforcement (for local or national security) and individual privacy has rested on the separation of powers between the legislature (which decides what can be done), the executive (which does the things), and the

judiciary (which reviews and interprets). Has the recent acceleration of technological change upset this balance?

10. While privacy may have some universal underpinnings, there are also considerable variations in cultural attitudes towards privacy. How can the global nature of technology be made to co-exist with local notions of privacy?
11. Should we have a national ID card? What are the advantages and disadvantages? Which identifiers in our wallets or sitting somewhere at home would we eliminate? What stateholders would profit? What technologies are most important to make such a card successful?
12. Are there privacy concerns with smart homes or smart vehicles?
13. Many of the online social systems feature a tension between the desire to share information locally and a global privacy concern. Identify one such problem on a system such as Second Life or Facebook, and design a technical solution.
14. Technology can make things that were already possible cheaper, and sometimes more expensive. Identify a privacy harm, and present a contemporary threat model analysis, as well as a threat model in an earlier era. Provide empirical data about the costs of surveillance and data processing in the two periods, and discuss how economic realities alter the threat model.
15. Is it possible to recover your identity after it has been stolen? If not today, is it possible to develop technologies and policies that would enable a victim of identity theft to quickly and cheaply recover from the crime?