CS/IS 7492: Digital Privacy: Technology and Ethics
Cornell Tech - CS/IS 7492
Credits: 3 hours
Fall 2020
Helen Nissenbaum, Information Science: Cornell Tech
Vitaly Shmatikov, Computer Science: Cornell Tech

COURSE OVERVIEW

This course will analyze privacy from the perspective of contextual integrity and other philosophical, legal, and social approaches, providing students with an understanding of why certain activities, technologies, and systems are considered to be privacy threats and/or violations, while others are seen as helpful in protecting and promoting privacy. Concurrently, this course will examine contemporary research on computer systems and technologies that both undermine and protect digital privacy. Students will be expected to apply themselves to both the technical and nontechnical material with equal energy and enthusiasm. The course is open to Ph.D. students only.

LEARNING OUTCOMES

The class is oriented around the impact of digital technologies on privacy. Although the field is already much larger than any single semester-long class can cover, students enrolled in this class will gain insight into a range of the most significant contemporary digital systems that threaten privacy as well as those that have developed in response to such threats, aiming to protect and promote privacy. While instructors anticipate that the academic backgrounds and interests of enrolled students will differ, the course aims to provide a strong technical working grasp of the respective technologies it will cover.

Through assigned readings and in-class discussions students will gain insight into privacy as a philosophical concept and an ethical value. They will appreciate why privacy is important to individuals, relationships, and societies and deserves protecting. The skills imparted in the philosophical and ethical dimensions of the course will enable students to identify practices and (socio-technical) systems that do and may compromise privacy, to identify circumstances where privacy conflicts with other rights and values, and how to reason about them. Assuming that students give full attention to the material in the philosophical and ethical aspects of the class, they will learn that reasoning well about these aspects demands both rigor and precision. In moving back and forth between the technical and philosophical, they will develop a rare ability to map respective concepts onto one another and achieve a grasp of meaningful privacy.

SYLLABUS

- Web and mobile tracking: How it happens; why it happens; can anything be done?
- Ethics and philosophy of privacy: Defining privacy as a **meaningful** concept; explaining why it is worth protecting.
- Privacy as contextual integrity: A theory of privacy as appropriate flow of information.
- Communications privacy: Digital networks, digital platforms, and digital services claim to offer “anonymous”, “private”, and/or “censorship-resistant” communication, but do they really? Technical solutions and tools, and the properties they do (and do not) provide.
- Data analytics, including statistical databases and large-scale data collection (threat), and differential privacy (solution?)
- Machine learning: Privacy threats from applications of machine learning such as face recognition; leakage of training data and other privacy issues in from ML models; federated and distributed learning.
- Location tracking: How it happens and why it deserves special attention and investigation, technical mechanisms for location tracking in mobile apps.
- Genetic privacy: Is *your* data really yours? Implications for science and the meaning of privacy

COURSE MATERIALS

Course readings, required and recommended will be available in two ways: 1) as a download from the CANVAS website, or 2) through links given in the Course Schedule.

ASSIGNMENTS AND GRADING CRITERIA

We expect to attend all classes and, beforehand, to have complete reading assignments thoroughly and with a critical eye. We encourage and will reward attendance and active participation during classes. Special note regarding Zoom: we expect students to attend classes with their video feed activated. If your circumstances do not allow for this on a given occasion, please reach out to instructors and we will evaluate case-by case. Over the course of the semester, we will assign 5-6 homework assignments, which include technical or nontechnical elements, or both.

Homework assignments (5): 50%
Attendance and participation: 10%
Final project: 40% -- due December 15, 2020

The final project will be a substantial research project building from course topics and material covered. Project ideas, which can be related to students' existing thesis work, should be discussed with Instructors and settled approximately by mid-semester. Whichever of the disciplinary perspectives students choose to adopt (i.e. technical or philosophical), the project must include some reference to the other. Here, too, students are invited to consult with instructors to locate any additional reference materials.

ACADEMIC INTEGRITY

We expect you to abide by Cornell's Code of Academic Integrity at all times. Please note that the Code specifically states that a "Cornell student's submission of work for academic credit indicates that the work is the student's own. All outside assistance should be acknowledged, and the student's academic position truthfully reported at all times." Please contact us if you have any questions or concerns about appropriately acknowledging others' work in your submitted assignments. You should expect that we will rigorously enforce the Code.

WEEKLY SCHEDULE

Sep 3 **Introduction to the course and the students**

Overview of topics and readings
Introducing the idea of "meaningful privacy"
Course requirements
About the final project (some sample project topics)

Online Tracking

Sep 8 **Tracking: Overview of Technology and Policy**

Readings:
Mayer and Mitchell. "Third-Party Web Tracking: Policy and Technology".
Hoofnagle, Chris Jay, et al. "Behavioral advertising: The offer you can't refuse." *Harv. L. & Pol'y Rev.* 6 (2012): 273.
United States Department of Health, Education & Welfare. 1973. *Records, Computers, and the Rights of Citizens*. Report of the Secretary's Advisory Committee on Automated Personal Data Systems, Excerpt.

Sep 10 **Web Tracking**

Readings:
Englehardt and Narayanan. "Online Tracking: A 1-million-site Measurement and Analysis".
Optional:
Acar et al. "The Web Never Forgets: Persistent Tracking Mechanisms in the Wild".

Sep 15 **Mobile Tracking**

Readings:
Reardon et al. "50 Ways to Leak Your Data: An Exploration of Apps' Circumvention of the Android Permissions System".

Sep 17 **Online Advertising**

Readings:
Bashir and Wilson. "Diffusion of User Tracking Data in the Online Advertising Ecosystem".
Faizullabhoy and Korolova. "Facebook's Advertising Platform: New Attack Vectors and the Need for Interventions".

Sep 22 **"The Price of Free"**

Readings:
Hoofnagle, Chris Jay, and Jan Whittington. "Free: accounting for the costs of the internet's most popular price." *UCLA L. Rev.* 61 (2013): 606.
Bamberger et al. "Can You Pay For Privacy? Consumer Expectations and the Behavior of Free and Paid Apps". 35 Berkeley Technology Law Journal 327 (2020).

Homework #1 (due Oct 2)

Ethical and Philosophical Approaches to Privacy

This section will address questions: What is privacy? And why do we care? We will follow the search for a meaningful conception.We will ask why privacy deserves to be understood and protected. To this end, the section introduces the landscape of philosophical definitions of privacy and theories of its ethical importance for individuals and societies.

Nissenbaum, Helen. *Privacy in context: Technology, policy, and the integrity of social life*. Stanford University Press, 2009, will serve as primary text for this topic. It will be referenced as [PIC], below. Book chapters will be supplemented with selected articles/book chapters as primary readings and "classic" texts as Optional readings for those who would like to familiarize themselves with first hand renderings.

Sep 24 Ethical and Philosophical Approaches

Readings:
PIC Chapter 4
Optional:
Cohen, Julie E. "Examined lives: Informational privacy and the subject as object." *Stan. L. Rev.* 52 (1999): 1373.
Gavison, Ruth. "Privacy and the Limits of Law." *The Yale Law Journal* 89.3 (1980): 421.
Prosser, W. "Privacy: A Legal Analysis'(1960)." *California Law Review* 48: 383.
Solove, Daniel J. "The meaning and value of privacy." *Social dimensions of privacy: Interdisciplinary perspectives* 71 (2015): 71.
Reiman, Jeffrey H. "Driving to the panopticon: A philosophical exploration of the risks to privacy posed by the highway technology of the future." *Santa Clara Computer & High Tech. LJ* 11 (1995): 27.

Sep 29 **Ethical and Philosophical Approaches**

Readings:
PIC Chapter 5
Griffin et al. "Hails: Protecting Data Privacy in Untrusted Web Applications"

Oct 1 **Fundamental Challenges to Ethical and Philosophical Approaches**

Readings:
Myers and Liskov. "Protecting Privacy using the Decentralized Label Model"
Acquisti, Alessandro, Laura Brandimarte, and George Loewenstein. "Privacy and human behavior in the age of information." *Science* 347.6221 (2015): 509.
Optional:
Allen, Anita L. "Coercing privacy." *Wm. & Mary L. Rev.* 40 (1998): 723.

Oct 6 **Fundamental Challenges to Ethical and Philosophical Approaches**

Readings:
PIC Chapter 6
Acquisti, Alessandro, Laura Brandimarte, and George Loewenstein. "Privacy and human behavior in the age of information." *Science* 347.6221 (2015): 509.
Optional:
Allen, Anita L. "Coercing privacy." *Wm. & Mary L. Rev.* 40 (1998): 723.

Communication Privacy

Oct 8 **Communication Privacy**

Readings:
Dingledine et al. "Tor: The Second-Generation Onion Router".

Oct 13 **Communication Privacy**

Readings:
Corrigan-Gibbs et al. "Riposte: An Anonymous Messaging System Handling Millions of Users".
Kift and Nissenbaum, "Metadata in Context"

Contextual Integrity

Oct 15 **Contextual Integrity**

Readings:
PIC Chapter 7

Oct 20 **Contextual Integrity**

Readings:
PIC Chapter 8
Optional:
Barth et al. "Privacy and Contextual Integrity: Framework and Applications"

Homeworks #2 and #3 (due Oct 29)

Oct 22 **Applications of Contextual Integrity**

Readings:
Shvartzshnaider et al. "VACCINE: Using Contextual Integrity For Data Leakage Detection"

Data Analytics and Differential Privacy

Oct 27 **Anonymization**

Readings:
Narayanan and Shmatikov: "Robust De-Anonymization of Large Sparse Datasets"

Oct 29 **Anonymization**

Readings:
Ohm, Paul. "Broken promises of privacy: Responding to the surprising failure of anonymization." *UCLA l. Rev.* 57 (2009): 1701.

Nov 3 **Impossibility of Privacy in Statistical Databases**

Readings:
Dinur and Nissim. "Revealing Information while Preserving Privacy".

Nov 5 **Differential Privacy**

Readings:
Dwork et al. "Calibrating Noise to Sensitivity in Private Data Analysis".
McSherry. "Privacy Integrated Queries" (CACM).
Optional:
McSherry. "Privacy Integrated Queries" (SIGMOD).

Nov 10 **The Extent and Limits of Tech Solutions**

Readings:
Barocas, Solon, and Helen Nissenbaum. "Big data's end-run around anonymity and consent." In *Privacy, big data, and the public good: Frameworks for engagement,* Eds. J. Lane, V. Stodden, S. Bender, H. Nissenbaum, Cambridge: Cambridge University Press.(2014): 44-75.
J. Shattuck. "Computer Matching is a Serious Threat to Individual Rights." Communications of the ACM 27(6), June 1984, pp. 538-541.
R. Kuserow. "The Government Needs Computer Matching to Root Out Waste and Fraud." Communications of the ACM 27(6), June 1984, pp. 542-545.
Optional:
Cohen, Julie E. "Right to Read Anonymously: A Closer Look at Copyright Management in Cyberspace, A." Conn. L. Rev. 28 (1995): 981.

Nov 12 **Differential Privacy in Data Analytics Systems**

Readings:
Bittau et al. "PROCHLO: Strong Privacy for Analytics in the Crowd".

Homework #4 (due Nov 20)

Privacy in Emerging Domains

Nov 17 Location Privacy

Nov 19 Location Privacy

Readings:
Troncoso et al. "Decentralized Privacy-Preserving Proximity Tracing".
Carpenter v. United States (EPIC amicus brief)
Martin, Kirsten E., and Helen Nissenbaum. "What Is It About Location?" Berkeley Technology Law Journal (2020): 103.

Nov 24 **Genomic Privacy**

Readings:
Gymrek et al. "Identifying Personal Genomes by Surname Inference".
Alondra Nelson: Watch this lecture (start at minute 13)
https://www.youtube.com/watch?v=giiXWtBdVFc&feature=youtu.be&t=780
Optional:
Barocas, Solon, and Karen Levy. "Privacy Dependencies." *Washington Law Review*

Nov 26 No class (Thanksgiving)

Dec 1 **End-to-end Secure Messaging**

Readings:
Abelson et al. "Keys under doormats: Mandating insecurity by requiring government access to all data and communications"
Levy S. "Battle of the Clipper Chip," *The New York Times*, June 12, 1994
https://www.forbes.com/sites/zakdoffman/2020/03/14/new-warning-issued-for-all-whatsapp-and-imessage-users-major-threat-to-encryption/?sh=278febe153f5

Dec 3 **Machine Learning and Privacy**

Readings:
McMahan et al. "Communication-Efficient Learning of Deep Networks from Decentralized Data".
Shokri et al. "Membership Inference Attacks Against Machine Learning Models".

Homework #5 (due Dec 10)

Dec 8 **Face Recognition and Privacy**

Readings:
Shan et al. "Fawkes: Protecting Privacy against Unauthorized Deep Learning Models".
See articles and opinions in Canvas by: Kashmir Hill, Luke Stark, Hartog and Selinger

Dec 10 Project Presentations

Leftover readings

- Zuboff, Shoshana. "Big other: surveillance capitalism and the prospects of an information civilization." *Journal of Information Technology* 30.1 (2015): 75.
- Regan, Priscilla M. Legislating privacy: Technology, social values, and public policy. Univ of North Carolina Press, 2000.

Sample project ideas

- Investigate location-tracking SDKs in mobile apps.
- Do “privacy-preserving” cryptographic protocols for computing on genomic data actually preserve privacy?
- Design and prototype a privacy-preserving image recognition service.
- Design and prototype a privacy-preserving communication service that hides traffic in some gaming protocol.
- Investigate the connection between robustness to privacy and robustness to adversarial examples in machine learning models.
- Analyze privacy properties, with reference to Contextual Integrity, of AutoML, Google CloudML, Amazon SageMaker, Azure Machine Learning Studio, or similar platforms for building custom ML models.
- Visualizing data flows, dynamically, in multiple dimensions