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:

Sep 10  Web Tracking
Readings:
Englehardt and Narayanan. "Online Tracking: A 1-million-site Measurement and Analysis".
Optional:

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:

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:
Ethical and Philosophical Approaches

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

 Fundamental Challenges to Ethical and Philosophical Approaches

Readings:
Myers and Liskov. “Protecting Privacy using the Decentralized Label Model”
Optional:

Communication Privacy

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

Communication Privacy

Readings:
Kift and Nissenbaum, "Metadata in Context"

Contextual Integrity

Readings:
PIC Chapter 7

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:

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:
Optional:

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)  

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:  

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”  
message-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

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