Overview

Course:	CS 1: Great Ideas in Computer Science
Course Level:	Introductory undergraduate
Course Description:	"An introduction to the most important discoveries and intellectual paradigms in computer science, designed for students with little or no previous background. Explores problem-solving using high and low-level programming languages; presents an integrated view of computer systems, from switching circuits up through compilers and GUI design. Examines theoretical and practical limitations related to unsolvable and intractable computational problems, and the social and ethical dilemmas presented by such issues as software unreliability and invasions of privacy."[1]
Module Topic:	Electronic Privacy
Module Author:	Lyndal Grant
Semesters Taught:	Spring 2019-2020, Spring 2020-2021
Tags:	social networks (CS), anonymization (CS), privacy (both), consent (phil), public interest (phil), electronic privacy (phil), informational privacy [phil]
Module Overview:	This module introduces the concept of electronic privacy and considers why electronic privacy matters. We begin by introducing students to what Solon Barocas and Helen Nissenbaum call "the traditional privacy protection paradigm," according to which tech companies can adequately protect users' privacy through a combination of informed consent and data anonymization. We explore a number of challenges to this paradigm by considering questions like the following: Are lengthy and complicated terms of service agreements an effective way to secure informed consent? How does the fact that machine learning allows highly sensitive personal information to be inferred from seemingly mundane user data complicate questions about consent? Is de-identification an effective way to ensure data remains anonymous, given that a small number of datapoints are often adequate to re-identify particular users (as Latanya Sweeney has shown[2])? We also explore how robust electronic privacy protections might benefit both individuals and society at large.
Connection to Course Material:	This module follows up on class sessions in which students learn simple data encryption techniques. Since encrypting user data is one way to protect user privacy, this sets us up for a discussion of how user privacy has traditionally been understood in the tech industry, as well as recent challenges to that understanding. CS 1 is an introduction to computer science for students who do not have a background in the discipline. This being the case, it makes sense to focus on a relatively nontechnical topic that

[1] http://sites.fas.harvard.edu/~cs1/no1.cs1.2016.pdf

[2] Yoo, Ji Su, Alexandra Thaler, Latanya Sweeney, and Jinyan Zang. "Risks to Patient Privacy: A Re-identification of Patients in Maine and Vermont Statewide Hospital Data." *Technology Science* (October 2018).

		will resonate with a broad audience.

Goals

Module Goals:	1. Encourage and help students to reflect on the question: Why is electronic privacy valuable? 2. Have students recognize and identify reasons to care about electronic privacy that are based in (1) self-interest and/or (2) public interest. 3. Have students think critically about whether systems put in place by social media companies, online retailers, and the like genuinely protect their electronic privacy. 4. Discuss requirements for meaningful informed consent when it comes to the collection and use of private information.	
Key Philosophical Questions:	1. Why is electronic privacy valuable? 2. Does a person who has "nothing to hide" nonetheless have self-interested reasons to care about electronic privacy? 3. Are strong electronic privacy protections in the public interest? 4. Does data inference (the fact that additional information can be inferred from information you have knowingly provided) complicate questions about informed consent?	Prior to the module, many students consider electronic privacy to be primarily valuable insofar as it protects one from criminal activity. The key philosophical questions are designed to help students challenge that view.

Materials

Key Philosophical Concepts:	● Privacy ● Electronic Privacy ● Informational Privacy ● Self-interest ● Public interest ● Informed consent	The distinction between self-interest and public interest helps students think more critically about a common view concerning state or corporate surveillance on which only those who are engaging in illegal or disreputable activities have any reason to be concerned about such surveillance (the "nothing to hide, nothing to fear" view). This view presupposes a merely self-interested concern with privacy, stemming from the desire to keep unfavorable things hidden. Students are encouraged to consider other self-interested reasons we might want to be free from surveillance, as well as

		reasons that are based in public interest.
Assigned Readings:	● Salon Barocas and Helen Nissenbaum, "Big Data's End Run Around Procedural Privacy Protections" (*Communications of the ACM*). ● Carissa Veliz, "Privacy and Digital Ethics After the Pandemic" (*Nature Electronics*).	Barocas and Nissenbaum's article introduces students to what they call "the traditional privacy protection paradigm," according to which informed consent and anonymization are sufficient to protect user's electronic privacy. The authors argue that informed consent and data anonymization are not merely difficult to achieve, but, in the age of big data, unsuited to the job of protecting privacy. The authors draw a parallel to the role of informed consent in biomedicine to argue that companies' approaches to protecting user privacy ought to be informed by review processes that assess "the substantive values at stake in these informational practices" (p.33). While the article is dense, it is accessible to a general audience (which makes it a good fit here). Veliz's article surveys the new (and not-so-new) ways in which electronic privacy has come under threat since the beginning of the Coronavirus pandemic. It discusses the potentially concerning ways in which access to data can shift the power dynamics between individuals, corporations, and governments. It therefore highlights both the personal and geopolitical risks involved in lax privacy regulations, and calls into question whether we should think of personal data as something that should be bought and sold.

Class Agenda:	1. Class activity: present and discuss a hypothetical COVID tracing app, "COVID-TRACE PLUS."

	2. What is privacy? Defining informational and electronic privacy. 3. The traditional privacy protection paradigm: informed consent + anonymization. 4. Explore challenges to the traditional paradigm. 5. Case studies: Amazon Kindle's terms of service agreement, Strava fitness app and US military bases, and Samaritan's RADAR suicide prevention program. 6. Do we have self-interested reasons to care about privacy if we have nothing to hide? Are there public interest reasons to care about privacy?	
Sample Class Activity:	At the beginning of the class, the Embedded EthiCS TA presents a hypothetical Covid tracing app, "COVID-TRACE PLUS." The app has the following features: ● Auto-installed on all smart phones; users are required to update their health status daily. ● Keeps a log of nearby users. ● For positive COVID test, automatically notifies all users who have come within 15ft radius of infected individual; shares time/location of potential contact and infected person's full name. ● All data is stored centrally and shared with local/national health authorities. In small groups of 3-4, students briefly discuss the following questions, followed by a full-class debrief: ● What privacy concerns would COVID-TRACE PLUS raise? ● Given the stakes, should an app like COVID-TRACE PLUS be used?	Having students consider and articulate their reactions to this hypothetical COVID-tracing app at the start of the module (before the introduction of key philosophical concepts and arguments) has several benefits. First, it helps to set the tone for the class: students are expected to consider their own intuitions about cases, talk with their classmates, and consider others' points of view (as opposed to merely absorbing information delivered by the TA). Second, it inevitably leads students to generate ideas and arguments that will come up again when the TA introduces philosophical concepts and views later in the module. This helps students to see the value of those theoretical tools. Third, it helps the instructor get a better feel for what students already understand and how they are thinking about the ethical issues under discussion.
Module Assignment:	In the follow-up assignment, students are asked to write a short essay in response to an excerpt from a New York Times article about the sale of user data by telecommunications firms and mobile apps. The excerpt and essay prompt appear below. *Excerpt:* "Telecommunications firms and mobile-based apps make billions of dollars per year by selling customer location data to marketers and other businesses, offering	It is helpful to use a concrete case study as the basis for homework questions, as this gives students practice applying the theoretical material from the module to a real-world scenario.

a vast window into the whereabouts of cell phone and app users, often without their knowledge.

That practice, which has come under increasing scrutiny and criticism in recent years, is now the subject of a proposed ban in New York City […]

The Bill would restrict cell phone companies and mobile apps from sharing location data to situations where they were "providing a service explicitly requested" by the customer. The language is designed to challenge the vague agreements customers click on when signing up for an app or a cellular service. The legislation would also exclude the collection of location data in 'exchange for products or services.'"

- From a July 23, 2019 article in the New York Times, "New York City to Consider Banning Sale of Cellphone Location Data" (by Jeffrey C. Mays).

Essay prompt: Do you think New York City should pass the legislation? Please support your answer using ideas from the lecture and/or readings, and make sure to consider at least one reason based in public interests. (Your response should be approximately two to three paragraphs in length.)

Lessons Learned: Student reaction to the module was very positive.

- In previous iterations of this module, only a small subset of students initially participated in class discussion. We found that starting the module with small-group discussion about a concrete case helped to set expectations about class participation early.
- Student feedback in the post-class survey indicated that students found the discussion of data-inference (and the associated example involving Strava's activity heat-map) particularly interesting.
- Student responses to the homework questions were generally thoughtful, though they indicated that students may not have fully grasped the distinction between reasons based in self-interest vs. those based in the public interest. This is something to keep in mind for future iterations of this module.