

## Overview

<b>Course:</b>	CS 165: Data Systems	
<b>Course Level:</b>	Upper-level undergraduate	
<b>Course Description:</b>	<p>“We are in the big data era and data systems sit in the critical path of everything we do. We are going through major transformations in businesses, sciences, as well as everyday life - collecting and analyzing data changes everything and data systems provide the means to store and analyze a massive amount of data. This course is a comprehensive introduction to modern data systems. The primary focus of the course is on the modern trends that are shaping the data management industry right now: column-store and hybrid systems, shared nothing architectures, cache conscious algorithms, hardware/software co-design, main-memory systems, adaptive indexing, stream processing, scientific data management, and key-value stores. We also study the history of data systems, traditional and seminal concepts and ideas such as the relational model, row-store database systems, optimization, indexing, concurrency control, recovery and SQL. In this way, we discuss both how and why data systems evolved over the years, as well as how these concepts apply today and how data systems might evolve in the future. We focus on understanding concepts and trends rather than specific techniques that will soon be outdated - as such the class relies largely on recent research material and on a semi-flipped class model with a lot of hands-on interaction in each class.”<sup>1</sup></p>	
<b>Module Topic:</b>	Data Privacy	
<b>Module Author:</b>	Sophie Gibert	
<b>Semesters Taught:</b>	Fall 2021	
<b>Tags:</b>	Privacy [phil], big data [CS], data systems [CS], data sharing [CS], intrinsic value [phil], extrinsic value [phil], liberty [phil], power [phil], harm [phil], deletion [CS], columnar storage [CS], GDPR [CS]	
<b>Module Overview:</b>	<p>This module focuses on the questions of what privacy is, why it is valuable, and how data system design choices can promote or hinder its valuable functions. We begin with an overview of four philosophical accounts of privacy. Then, we introduce a distinction between intrinsic and extrinsic value, discuss four ways in which privacy is extrinsically valuable, and explore arguments for the claim that privacy is also intrinsically valuable. Lastly, we take a closer look at Google’s data practices and privacy policy. Students discuss the extent to which Google’s privacy policy protects privacy and the extent to which Google’s current data practices limit the valuable functions of privacy. Special attention is paid to Google’s deletion practices and the trade-offs that exist between row- and column-based data storage systems as it pertains to deletion.</p>	<p>Google uses a combination of row-oriented and columnar data storage systems. In general, columnar databases allow for more efficient queries of data but less efficient insertion and deletion. In columnar systems, deletes are always logical, or “soft,” meaning that relevant data is marked as “deleted” but is not actually deleted, at least not for a period of time. Columnar databases thus may reduce the control that users have over their personal data.</p>
<b>Connection to Course Material:</b>	<p>Students in this course learn about modern systems for data management and storage. They also design a columnar database in an ongoing assignment. This module gives them tools for thinking about how data practices can affect privacy and how data system</p>	<p>The topic was chosen because of its direct connection to the technical material covered in the course. The topic is also timely: data privacy incidents regularly appear in the news, and new data</p>

<sup>1</sup> [Link](#).

design choices can promote or hinder privacy's valuable functions.

privacy regulations have gone into effect recently in the EU and in California.

### Goals

**Module Goals:** By the end of the module, students should be able to:

1. Describe four philosophical accounts of what privacy is and explain why two of them are more plausible.
2. Distinguish between intrinsic and extrinsic value.
3. Identify four valuable functions of privacy.
4. Explain why one might think that privacy is also intrinsically valuable.
5. Analyze a company's privacy policy. Make judgments about:
  - a. Whether the policy protects privacy, and
  - b. Whether the company's data practices limit any of the valuable functions of privacy.

**Key Philosophical Questions:**

1. What is privacy?
2. Why is privacy valuable? What are its important functions?
3. How much privacy is too much?
4. How can data system design choices protect or fail to protect privacy?
5. How can data system design choices limit the valuable functions of privacy?

Another version of this module might replace Question 1 with the following question: What is the *right* to privacy, and what constitutes an infringement of this right? There are benefits and drawbacks to this alternative approach. On the one hand, philosophical accounts of the right to privacy can be easier for students to grasp than accounts of privacy itself, and conceptual analysis of the notion of privacy may not be the best way to begin addressing normative questions about privacy. On the other hand, the alternative approach requires introducing additional material about moral rights.

### Materials

- Key Philosophical Concepts:**
- Privacy
  - Intrinsic value
  - Extrinsic value
  - Liberty of action
  - Power

The concepts of intrinsic and extrinsic value are used to structure the discussion of why privacy is valuable. Students are introduced to four ways in which privacy is extrinsically valuable, as well as arguments for the claim that privacy is intrinsically valuable.

<p><b>Assigned Readings:</b></p>	<ul style="list-style-type: none"> <li>● Warzel and Ngu. "<a href="#">Google's 4,000-word privacy policy is a secret history of the internet.</a>" <i>The New York Times</i>.</li> <li>● Excerpts from Google's current <a href="#">privacy policy</a>: "Information Google collects," "Why Google collects data," "Your privacy controls," "Sharing your information," "Exporting &amp; deleting your information, and "Retaining your information."</li> </ul>	<p>Warzel and Ngu walk through the evolution of Google's privacy policy from 1999 to today and discuss what it reveals about the internet's transformation. This article makes Google's privacy policy come to life and contextualizes our choice to focus on Google's policies and practices in this module.</p> <p>Excerpts from Google's privacy policy prepare students to engage in the in-class activity/discussion.</p>
----------------------------------	--	--

<b>Implementation</b>		
<p><b>Class Agenda:</b></p>	<ol style="list-style-type: none"> <li>1. Introduction: Why data are useful; Why there's cause for concern.</li> <li>2. Philosophical accounts of what privacy is: Solitude, secrecy, limited access, control.</li> <li>3. Distinction between intrinsic and extrinsic value.</li> <li>4. Four valuable functions of privacy: Promotes liberty of action, protects from harm and power, allows for intimate and trusting relationships, promotes democratic ideals. <ol style="list-style-type: none"> <li>a. Discussion of feminist critiques of privacy.</li> </ol> </li> <li>5. Privacy in practice: Brief overview of Google's data practices and privacy policy.</li> <li>6. Activity/discussion.</li> </ol>	
<p><b>Sample Class Activity:</b></p>	<p>In small groups (4-5 students), students engage in a three-stage analysis of Google's privacy policy and data practices. They are asked to discuss the following questions and subsequently to share their thoughts with the class:</p> <ol style="list-style-type: none"> <li>1. To what extent does Google's privacy policy protect privacy, understood as limited access and as control over personal information?</li> <li>2. To what extent do Google's data practices limit people's liberty of action, subject them to harm and power, limit their ability to have different kinds of relationships, including intimate and trusting relationships, and hinder democratic ideals?</li> <li>3. How could Google better protect privacy? Who would bear the costs?</li> </ol>	<p>This class is made up of approximately 25 in-person students. The classroom is flat, and students are seated at round tables that seat up to 5. This makes small-group discussion natural and feasible.</p> <p>The activity focuses on Google because Google's data practices are representative of other companies' practices and because Google collects and stores more personal data than most other entities. Google's privacy policy is also readable and clear.</p>
<p><b>Module Assignment:</b></p>	<p>Students receive a fact sheet on the General Data Protection Regulation (GDPR). They are asked to answer the following questions:</p>	<p>During the module, students are introduced to philosophical accounts of privacy and to four valuable functions of privacy.</p>

1. Based on your understanding of the GDPR, to what extent does it protect EU citizens' privacy? In your response, appeal to either the limited access theory or the control theory.
2. As far as you can tell, to what extent does the GDPR protect EU citizens from the following consequences?
  - a. Having their liberty of action limited
  - b. Being subjected to harm or power
  - c. Not being able to engage in different types of relationships, including intimate and trusting relationships
  - d. Not being able to express dissenting political opinions

During the in-class activity, students apply their new knowledge to Google's privacy policy and data practices. The module assignment is intended to *reinforce* learning by asking students to apply the same knowledge in a new context—this time, to a public policy that aims to protect privacy.

The assignment also introduces students to an important piece of legislation that shapes the current privacy landscape worldwide.

**Lessons Learned:** Students found the distinction between the four valuable functions of privacy to be particularly useful in thinking about the Google case. They also appreciated the different philosophical accounts of privacy.

Pedagogical lessons learned:

- Modules are more engaging and memorable to the extent that they connect with students' own experiences. In discussing the valuable functions of privacy, it is useful to have students think about how privacy functions in their own lives—for example, what activities they would not engage in if they had no privacy, what powers they are subjected to due to our current data practices, etc.
- Assignments can reinforce learning by having students apply concepts and processes taught in class to new but related contexts.