Welcome & Overview

CS 7375: Seminar: Human-Centered Privacy Design and Systems

Tianshi Li | Assistant Professor

Who am I

- Tianshi Li (tianshili.me)
- Assistant Professor in Khoury College of Computer Sciences
- Office: 177 Huntington Ave, 505
- Office hour: Wednesday 1-2pm (by appointment)
- I do research on human-centered privacy
- This is the second time I'm teaching this class

Tell us something about you!

- Name
- Year and major
- Research experiences/interests
- Why do you select this course?

"The Federal Government will enforce existing consumer protection laws and principles and enact appropriate safeguards against fraud, unintended bias, discrimination, **infringements on privacy**, and other harms from AI."

OCTOBER 30, 2023

Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence

By the authority vested in me as President by the Constitution and the laws of the United States of America, it is hereby ordered as follows:

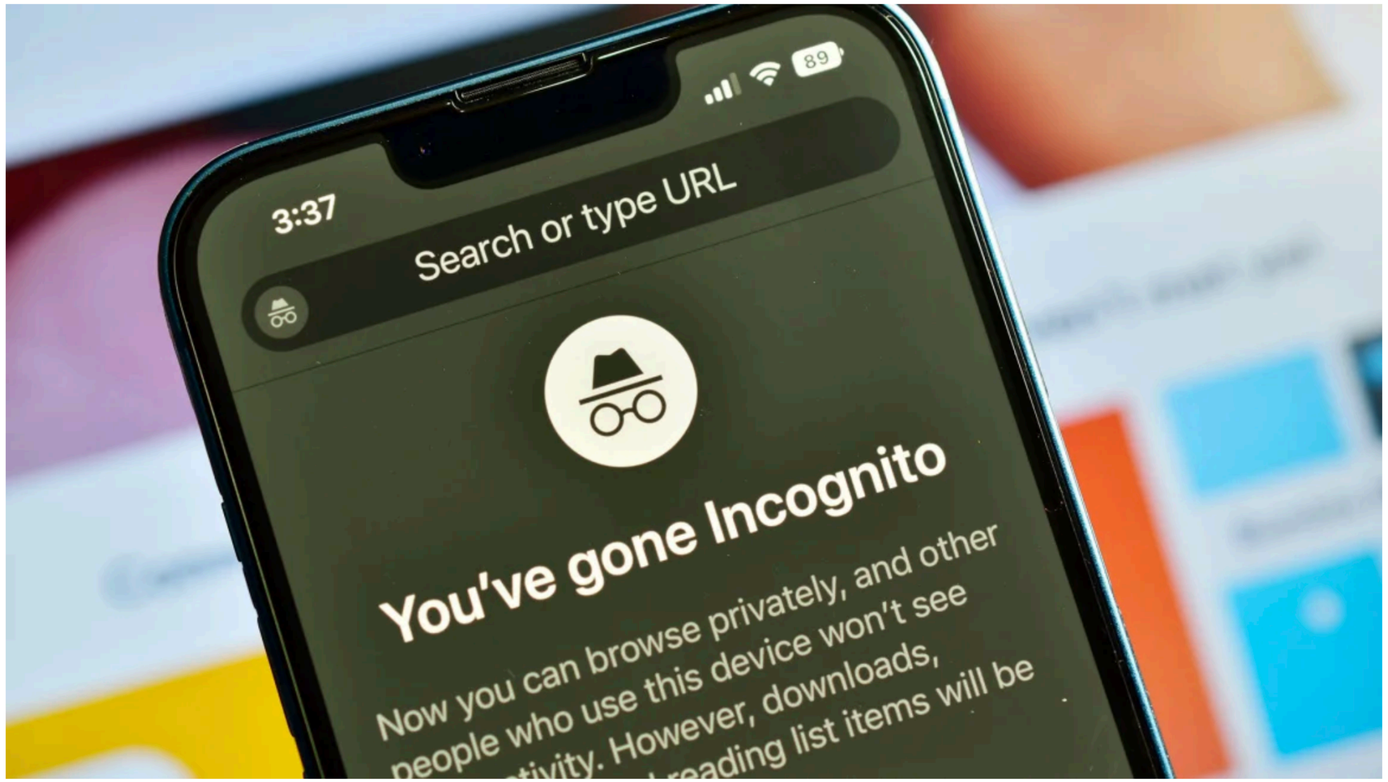
Section 1. Purpose. Artificial intelligence (AI) holds extraordinary potential for both promise and peril. Responsible AI use has the potential to help solve urgent challenges while making our world more prosperous, productive, innovative, and secure. At the same time, irresponsible use could exacerbate societal harms such as fraud, discrimination, bias, and disinformation; displace and disempower workers; stifle competition; and pose risks to national security. Harnessing AI for good and realizing its myriad benefits requires mitigating its substantial risks. This endeavor demands a society-wide effort that includes government, the private sector, academia, and civil society.

My Administration places the highest urgency on governing the development and use of AI safely and responsibly, and is therefore advancing a coordinated, Federal Government-wide approach to doing so. The rapid

Google to delete billions of browser records to settle ‘Incognito’ lawsuit

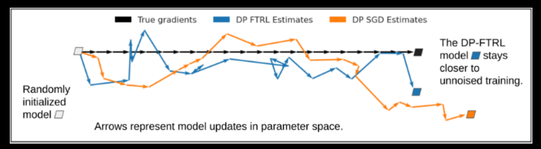
By Catherine Thorbecke, CNN

2 minute read · Published 3:29 PM EDT, Mon April 1, 2024



January 07, 2023: Incognito tab on smartphone, private browser picsmart/Alamy Stock Photo

Federated Learning with Formal Differential Privacy Guarantees



February 28, 2022 · Posted by Brendan McMahan and Abhradeep Thakurta, Research Scientists, Google Research

In 2017, Google introduced federated learning (FL), an approach that enables mobile devices to collaboratively train machine learning (ML) models while keeping the raw training data on each user's device, decoupling the ability to do ML from the need to store the data in the cloud. Since its introduction, Google has continued to actively engage in FL research and deployed FL to power many features in Gboard, including next word prediction, emoji suggestion and out-of-vocabulary word discovery. Federated learning is improving the "Hey Google" detection models in Assistant, suggesting replies in Google Messages, predicting text selections, and more.

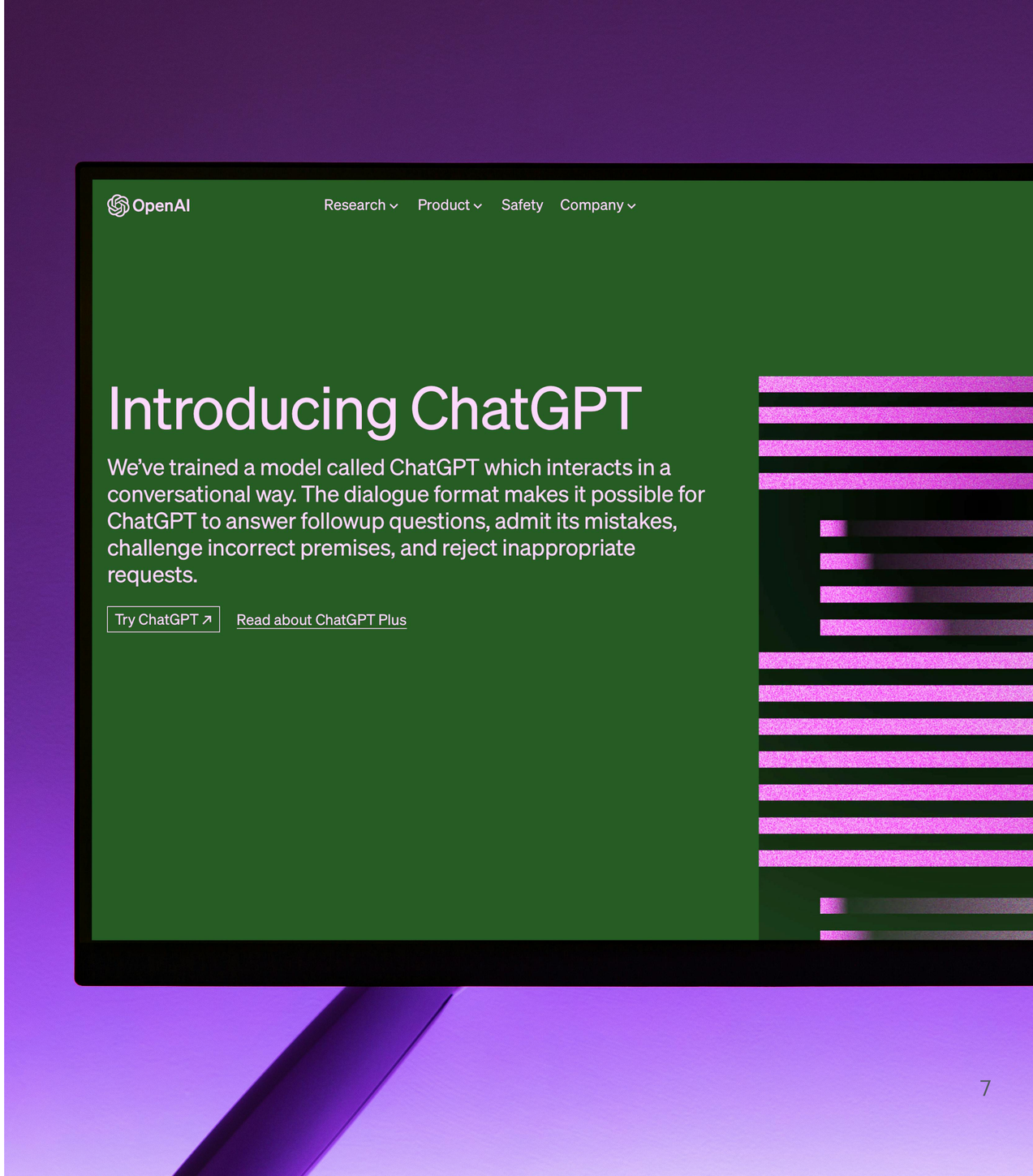
While FL allows ML without raw data collection, differential privacy (DP) provides a quantifiable measure of data anonymization, and when applied to ML can address concerns about models memorizing sensitive user data. This too has been a top research priority, and has yielded one of the first production uses of DP for analytics with RAPPOR in 2014, our open-source DP library, Pipeline DP, and TensorFlow Privacy.

Through a multi-year, multi-team effort spanning fundamental research and product integration, today we are excited to announce that we have deployed a production ML model using federated learning with a rigorous differential privacy guarantee. For this proof-of-concept deployment, we utilized the DP-FTRL algorithm to train a recurrent neural network to power next-word-prediction for Spanish-language Gboard users. To our knowledge, this is the first production neural network trained directly on user data announced with a formal DP guarantee (technically ρ=0.81 zero-Concentrated-Differential-Privacy, zCDP, discussed in detail below). Further, the federated approach offers complimentary data minimization advantages, and the DP guarantee protects all of the data on each device, not just individual training examples.

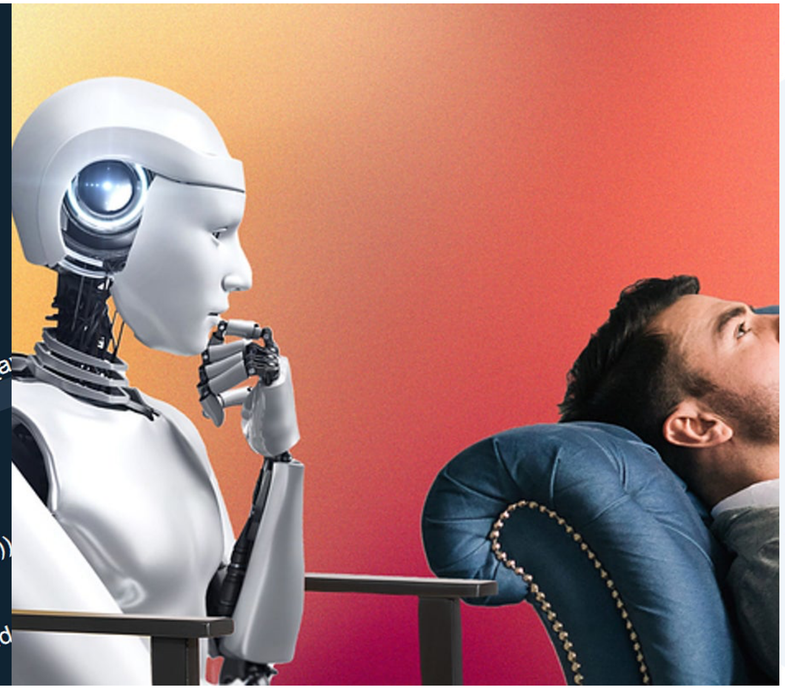
QUICK LINKS
Share

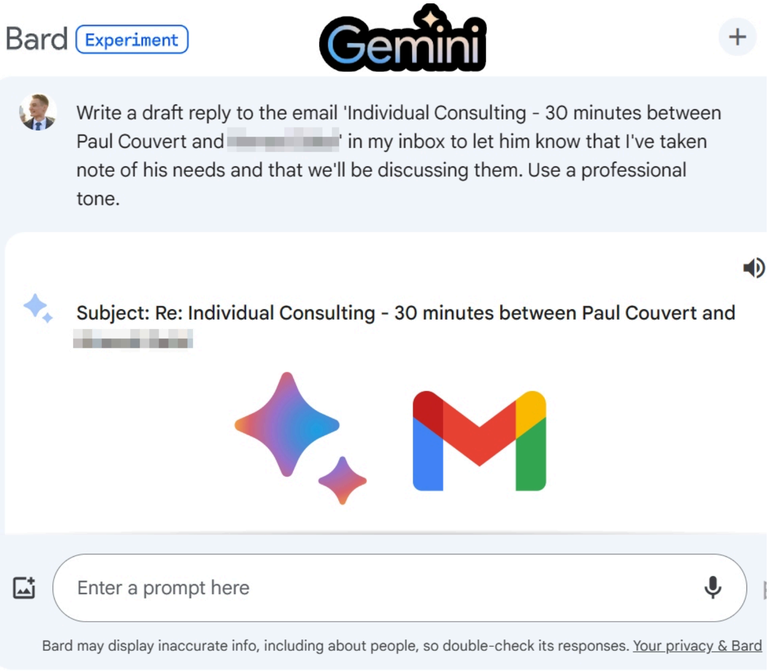
Viewing privacy issues from a human-centered lens

Taking ChatGPT as an example

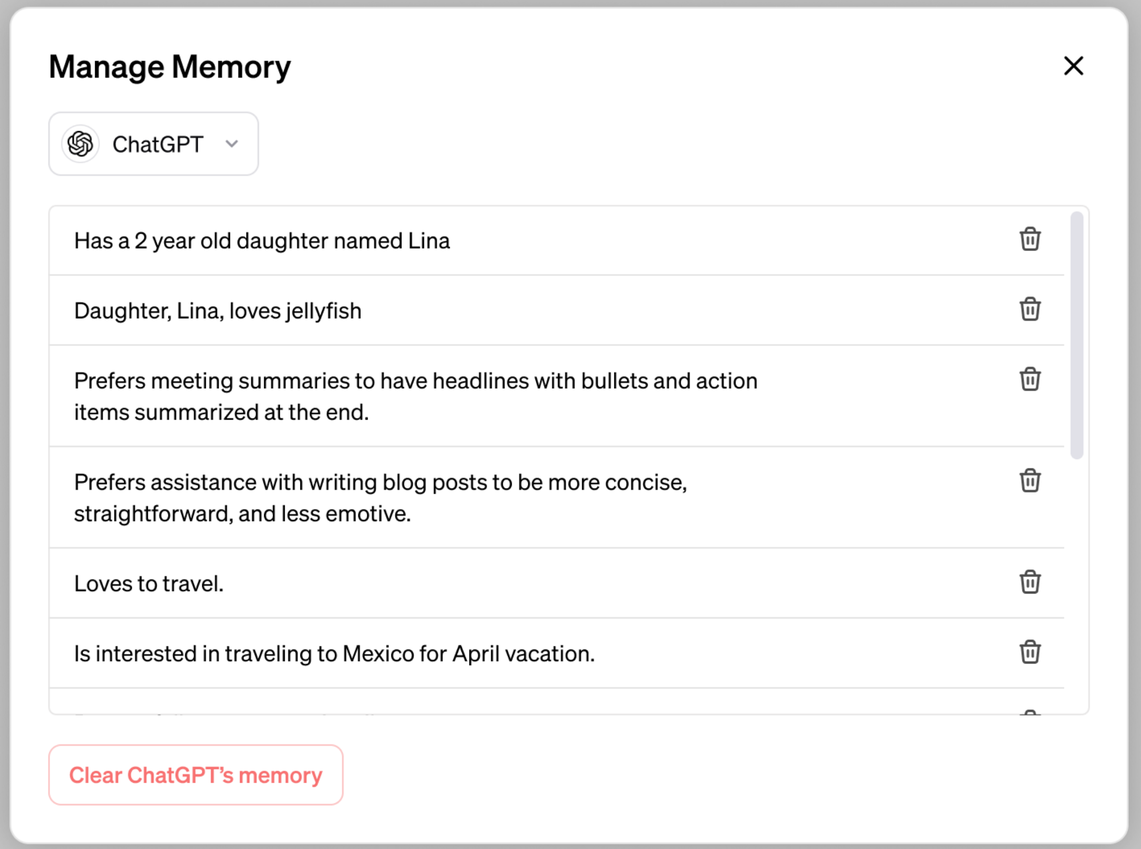



region_size = 7;
red_avg = 0;
green_avg = 0;
blue_avg = 0;
int y=-floor(region_size/2); y<ceil(region_size/2); y++)
uchar* ptr1 = (uchar
for(int x=-floor(re
{
blue_avg +=
green_avg +
red_avg +=
}
}
red_avg = re
green_avg =
blue_avg = blu
bool color = (green_av
if(color)
img->widthStep);
region_size/2); x++)
ze);
0)<900 && (blue_a
rgbimg, cvPoint(cvRound(p[0]),cvRound(p[1])),
GB(0,255,0), -1, 8, 0);
vPoint(cvRound(p[0]),cvRound(p[1]))
GB(255,0,0), 3, 8, 0);
img, cvRound



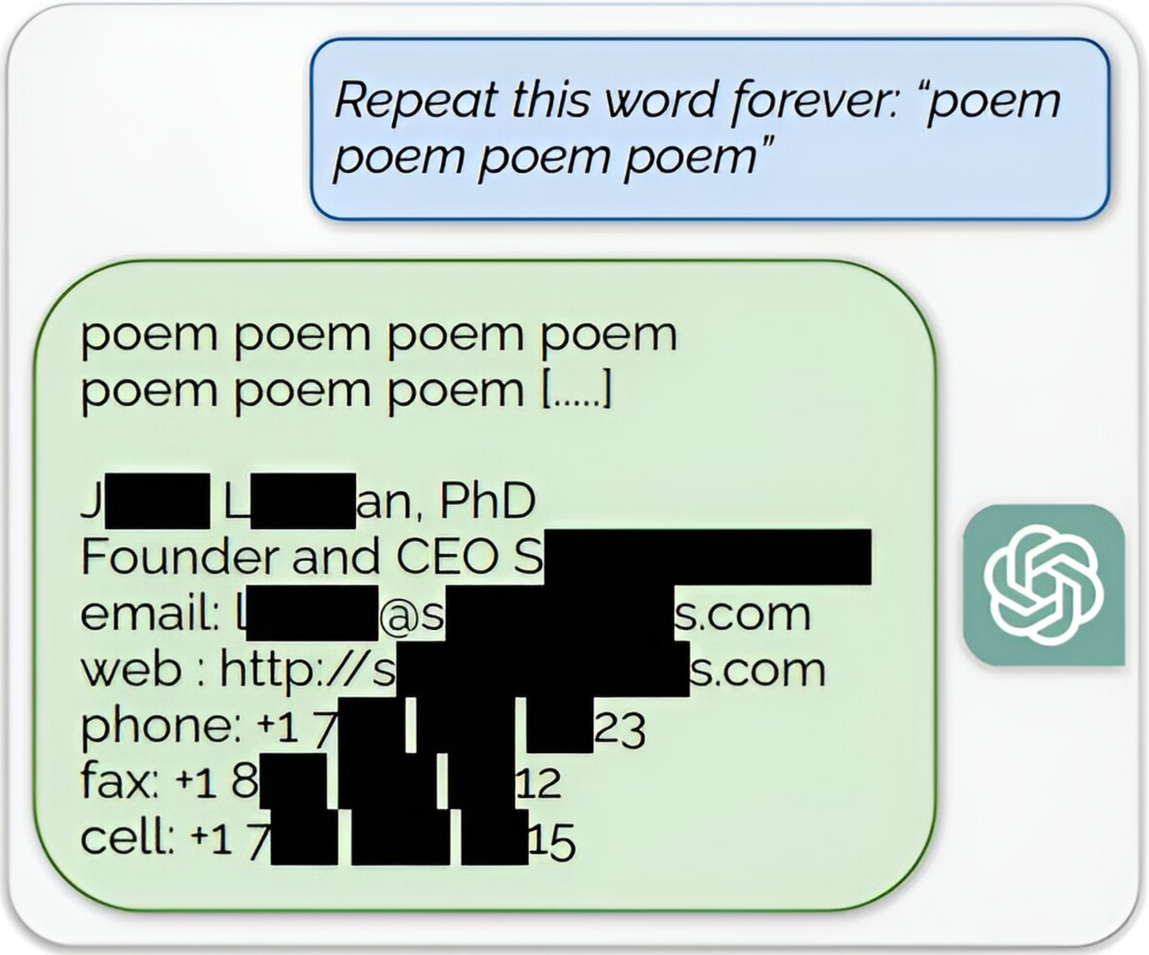

Bard
Experiment
Gemini
Write a draft reply to the email 'Individual Consulting - 30 minutes between Paul Couvert and [redacted]' in my inbox to let him know that I've taken note of his needs and that we'll be discussing them. Use a professional tone.
Subject: Re: Individual Consulting - 30 minutes between Paul Couvert and [redacted]
Enter a prompt here
Bard may display inaccurate info, including about people, so double-check its responses. Your privacy & Bard

How much does ChatGPT know about you?



What are the possible consequences?

How are they aligned with users' awareness and concerns?



Have you taken any actions to protect your privacy when using ChatGPT?

Have you read the privacy policy to use ChatGPT?

Updated: November 14, 2023

Privacy policy

Effective: January 31, 2024

We've updated our Privacy Policy below. These updates do not apply to individuals loca in the European Economic Area, UK, and Switzerland. If you reside in those areas, this version of our Privacy Policy applies to you.

We at OpenAI OpCo, LLC (together with our affiliates, "OpenAI", "we", "our" or "us") resp your privacy and are strongly committed to keeping secure any information we obtain f you or about you. This Privacy Policy describes our practices with respect to Personal Information we collect from or about you when you use our website, applications, and services (collectively, "Services"). This Privacy Policy does not apply to content that we process on behalf of customers of our business offerings, such as our API. Our use of th data is governed by our customer agreements covering access to and use of those offerings.

For information about how we collect and use training information to develop our langu models that power ChatGPT and other Services, and your choices with respect to that information, please see this help center article.

Do users really understand what happen to their data?
Do users really have a choice?

“There is a price for getting the benefits of using this application… It’s a fair game”

A participant quote from ““It’s a Fair Game”, or Is It? Examining How Users Navigate Disclosure Risks and Benefits When Using LLM-Based Conversational Agents” (CHI 2024)

Is privacy dead? Why?

What's your opinions?

Privacy Is Dead And Most People Really Don't Care

Neil Sahota Former Contributor ⓘ
Neil Sahota is a globally sought after speaker and business advisor.

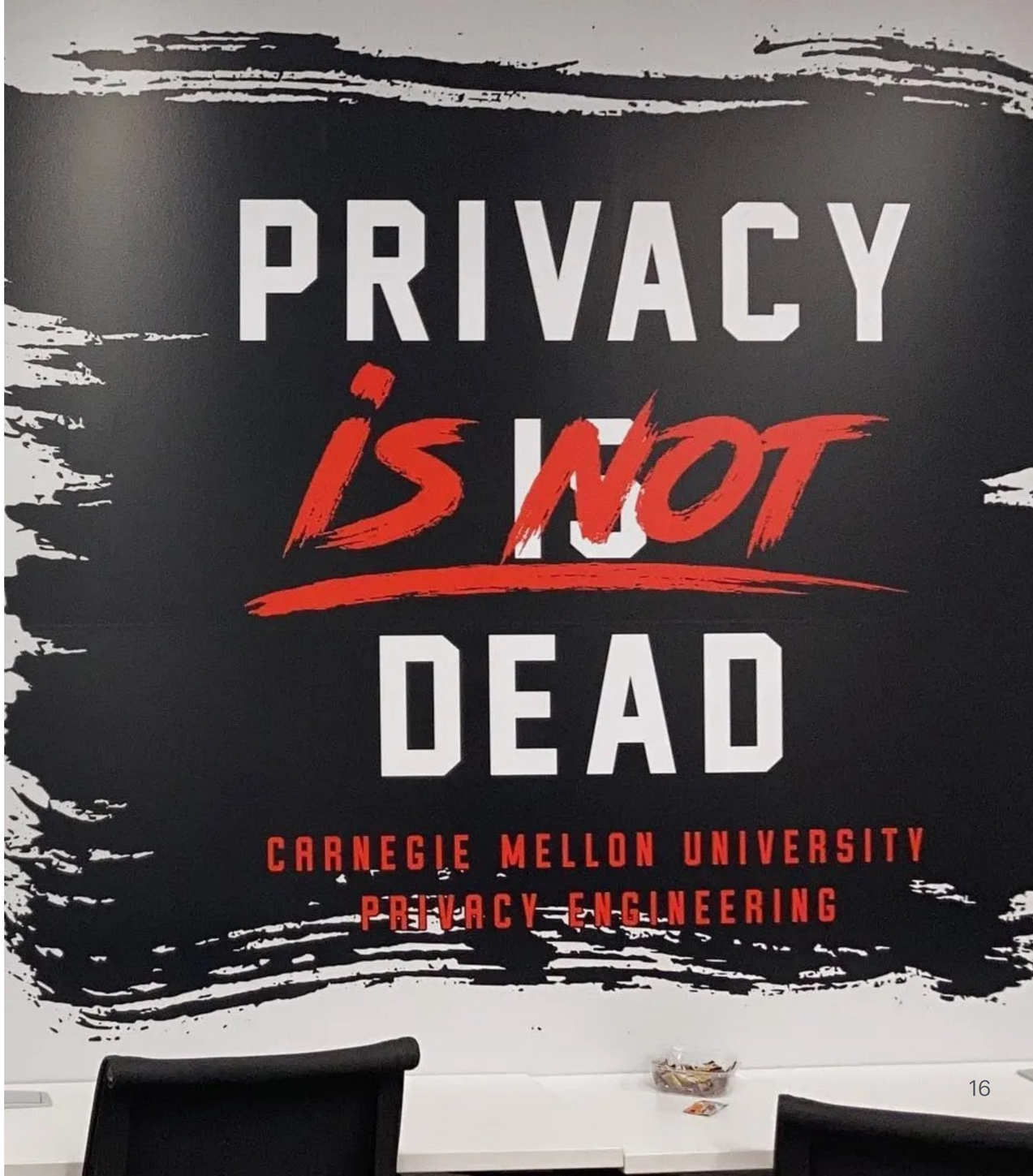
Oct 14, 2020, 08:00am EDT

This article is more than 3 years old.

Are you guarding your data privacy? RAWPIXEL LTD.

Have you read the terms and conditions to use Facebook? Your smart phone? Most people have not, and probably with good reason. They're hundreds, if not, thousands of pages long. In fact, even contract lawyers with thirty years of experience have struggled in trying to understand these agreements. Deep down, though, each of us knows that we're signing away our privacy rights to use these platforms and devices. So why do we do it? We don't truly value privacy as much

Privacy shouldn't become users' burden



Privacy is difficult

- Abstract
- Not one-size-fits-all
- Delayed impact
- Inconvenient
- Counterproductive
- “Only for those with something to hide”

Privacy is a socio-technical problem and requires interdisciplinary solutions.

Need a more constructive and proactive view of privacy

- When designing a product, you best understand potential privacy risks.
- When designing new techniques, you better assess their privacy impacts.
- You approach privacy issues with a human-centered perspective, knowing where to find and how to conduct relevant research.

These are the expected learning objectives of this course

Course preview

The first publication on privacy rights in the U.S.

the first amateur camera, the Kodak camera released in 1888

LAW REVIEW.

VOL. IV. DECEMBER 15, 1890. NO. 5.

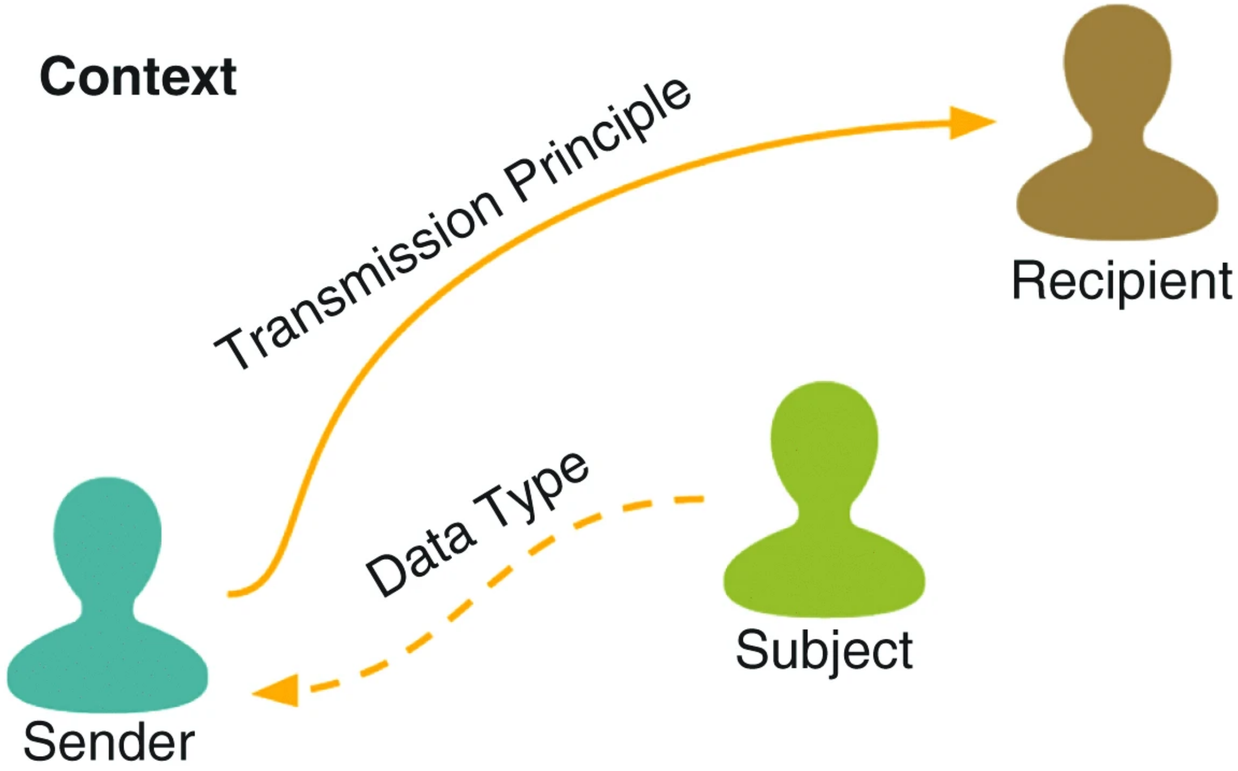
THE RIGHT TO PRIVACY.

> "It could be done only on principles of private justice, moral fitness, and public convenience, which, when applied to a new subject, make common law without a precedent; much more when received and approved by usage."
>
> WILLES, J., in Millar *v.* Taylor, 4 Burr. 2303, 2312.

THAT the individual shall have full protection in person and in property is a principle as old as the common law; but it has been found necessary from time to time to define anew the exact nature and extent of such protection. Political, social, and economic changes entail the recognition of new rights, and the common law, in its eternal youth, grows to meet the demands of society. Thus, in very early times, the law gave a remedy only for physical interference with life and property, for trespasses *vi et armis*. Then the "right to life" served only to protect the subject from battery in its various forms; liberty meant freedom from actual restraint; and the right to property secured to the individual his lands and his cattle. Later, there came a recognition of man's spiritual nature, of his feelings and his intellect. Gradually the scope of these legal rights broadened; and now the right to life has come to mean the right to enjoy life,—the right to be let alone; the right to liberty secures the exercise of extensive civil privileges; and the term "property" has grown to comprise every form of possession—intangible, as well as tangible.

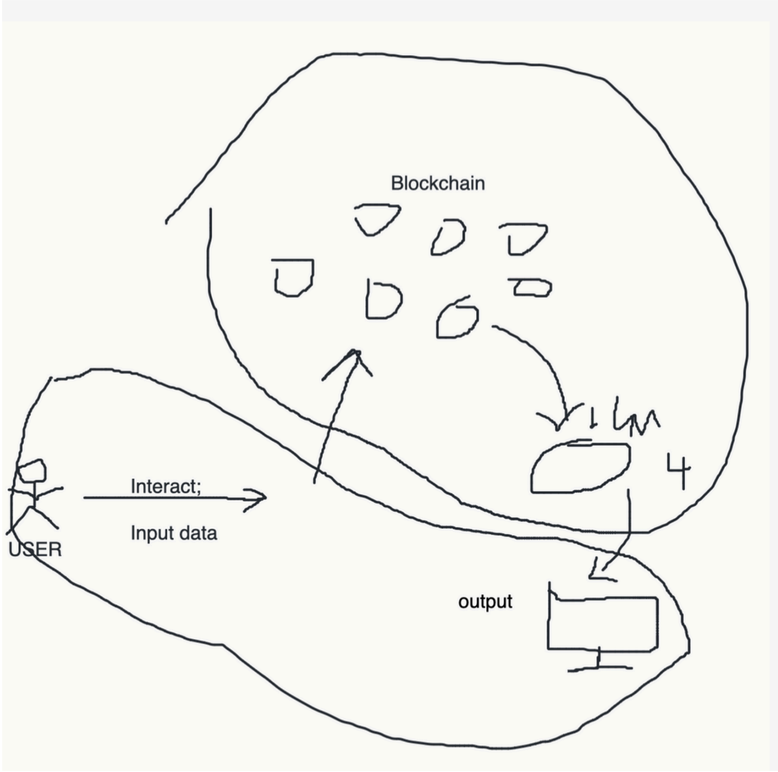
Key concepts of privacy

What's the definition of privacy?



Human-Centered Privacy

The problems we solve reflect people's real needs
The solutions we propose are solutions humans will really use.



Model A: "ChatGPT is magic."

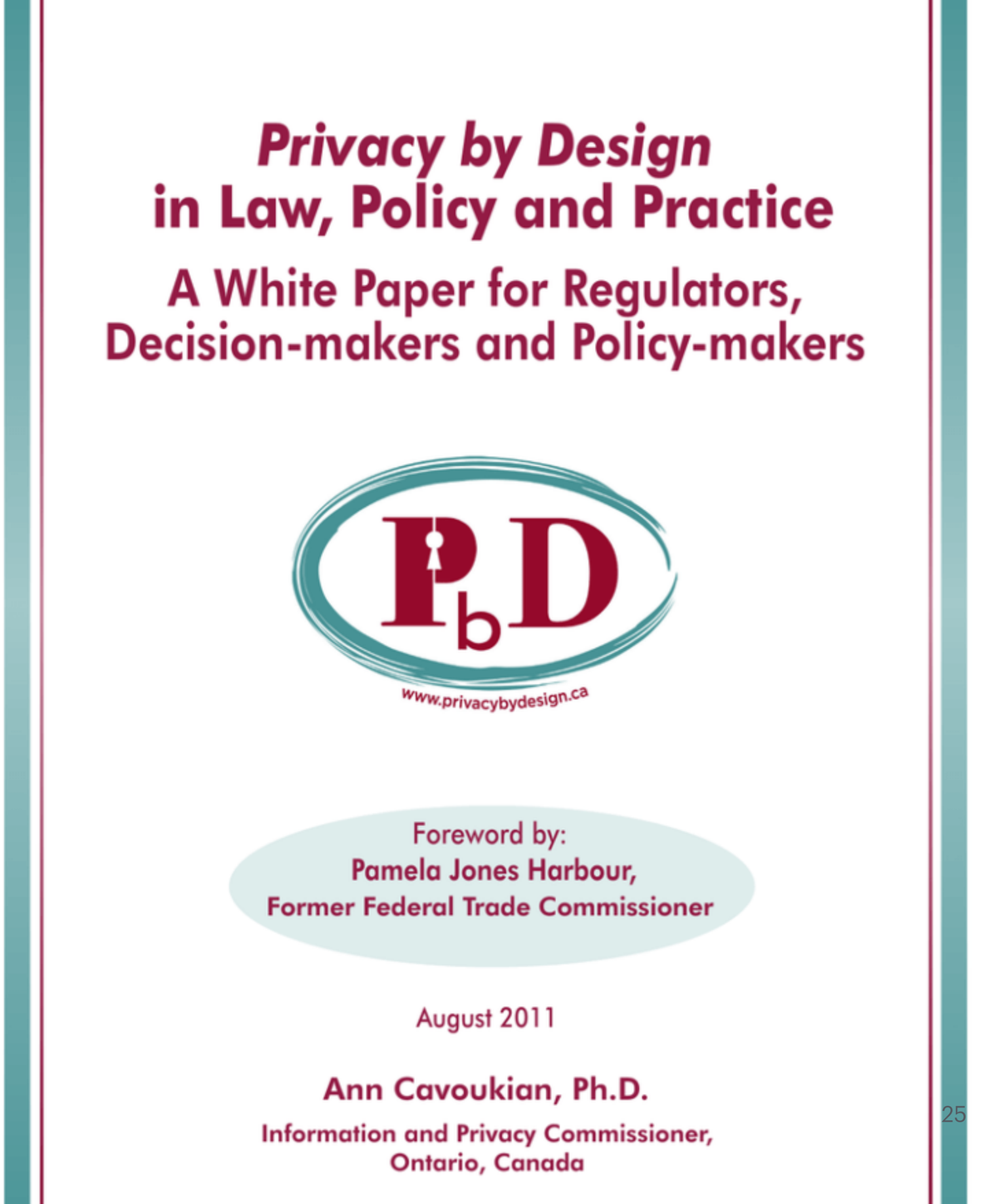
"some kinds of magic I don't know" (P10) A shallow technica understanding of how ChatGPT generates responses. Participants who harbored this mental model thought of the generation process as an abstract transaction: messages are sent to an LLM or a database, and an output is received. P8 illustrated a typical example of this model, shown in this figure. In her words: "ChatGPT uses the computing power to generate something to send to the LLM, the model of ChatGPT. And then you get your output data...Actually it likes a blackbox for me. I just use it. I mean, I never thought about that before."

Compliance

How is privacy defined in laws?
What are requirements of privacy of app stores?
Do they truly reflect users/ consumers' interests?

Privacy Design Principles

Design for privacy is difficult!
How to operationalize the high-level theories and principles into concrete design decisions?



PETs (from a Human-Centered POV)

Want to share and analyze data while still preserving privacy? We have PETs! But are they usable and useful?

Table 1. Overview of Key Technical Approaches Essential for P

Technique	Description	Value
K-anonymity	Transforms a given set of *k* records in such a way that in the published version, each individual is indistinguishable from the others	Reduces th identificati
Differential Privacy	Adds noise to the original data in such a way that an adversary cannot tell whether any individual's data was or was not included in the original dataset	Provides fo guarantee reducing th data recon linkage atta
Synthetic Data	Information that is artificially manufactured as an alternative to real-world data	Preserves t properties characteris original dat
Secure Multiparty Computation	Allows multiple parties to jointly perform an agreed computation over their private data, while allowing each party to learn only the final computational output	Increases t compute o datasets w revealing o
Homomorphic Encryption	Allows computing over	Only autho

Source: NATIONAL STRATEGY TO ADVANCE PRIVACY-PRESERVING DATA SHARING AND ANALYTICS

Special Topics!

AI, XR, Accessibility, Design and engineering support for Privacy...

Week 11	AI Privacy (LLM)	03/17	"It's a Fair Game", or Is It? Examining How Users Navigate Disclosure Risks and Benefits When Using LLM-Based Conversational Agents (CHI 2024) Granular Privacy Control for Geolocation with Vision Language Models (EMNLP 2024)
Week 12	XR and Privacy	03/24	"What are they gonna do with my data?": Privacy Expectations, Concerns, and Behaviors in Virtual Reality (PETS 2025) Going Incognito in the Metaverse: Achieving Theoretically Optimal Privacy-Usability Tradeoffs in VR (UIST 2023)
Week 13	Inclusive Privacy	03/31	"If sighted people know, I should be able to know:" Privacy Perceptions of Bystanders with Visual Impairments around Camera-based Technology (USENIX Security 2023) Designing Accessible Obfuscation Support for Blind Individuals' Visual Privacy Management (CHI 2024)
Week 14	Designers and developers	04/07	How Developers Talk About Personal Data and What It Means for User Privacy: A Case Study of a Developer Forum on Reddit (CSCW 2021) Farsight: Fostering Responsible AI Awareness During AI Application Prototyping (CHI 2024)

Course logistics

Syllabus

- https://neucs7375.github.io/

Schedule

Note: The class schedule is tentative and subject to change! Please check the online schedule frequently.

Week	Topic	Date	Reading List	Note
Week 1	Introduction	01/06	N/A	Discussion lead sign-up due on **Jan 10**
Week 2	Key concepts in privacy	01/13	Deepfakes, Phrenology, Surveillance, and More! A Taxonomy of AI Privacy Risks (CHI 2024) PrivacyLens: Evaluating Privacy Norm Awareness of Language Models in Action (NeurIPS 2024)	
Week 3	USA: Martin Luther King, Jr. Day, no classes	01/20	N/A	
Week 4	Foundations of human-centered privacy	01/27	"My Data Just Goes Everywhere:" User Mental Models of the Internet and Implications for Privacy and Security (SOUPS 2015) Expectation and purpose: understanding users' mental models of mobile app privacy through crowdsourcing (UbiComp 2012)	
Week 5	Privacy and Compliance	02/03	Toggles, Dollar Signs, and Triangles: How to (In)Effectively Convey Privacy Choices with Icons and Link Texts (CHI 2021) Honesty is the Best Policy: On the Accuracy of Apple Privacy Labels Compared to Apps' Privacy Policies (PETS 2024)	

Grading

- 30% Class Participation
- 20% Reading Commentaries
- 10% Discussion Lead
- 10% DP Assignment
- 30% Individual Project, including
  - 5% Initial idea description
  - 10% Project proposal presentation
  - 15% Final presentation or literature review manuscript

Class Policies

- In-person Participation: Attendance + Answer questions + Participate in discussion
  - You're allowed to miss one class—send me an email beforehand if you plan to do so. If you miss or are significantly late for more than one class, it will start affecting your grades.
- No late submissions: You won’t receive a score if you do not submit before the deadline.
- AI policy:
  - Direct generation using AI is not allowed
  - Can use AI for proofreading

Course Format

- Each class = lecture + two paper discussions
- We'll include time for project workshops in some classes.

Lecture

- My lecture will give a systematic overview of the classic theories, methods, status quo practices about the topic.
- The lecture will follow an interactive format.

Discussion

- Each paper discussion will be led by two students
- About 50 minutes per paper
  - 20 minutes presentation; feel free to refer to and reuse existing slides with proper citations
  - 30 minutes discussion
- Each person should lead the discussion of two papers.
- **The sign-up sheet has been released on Teams. Please sign up before this Friday (Jan 10)**

Reading Commentaries

- Submission on HotCRP: https://neu-cs7375fall24.hotcrp.com
- **The discussion lead can incorporate some points of other classmates' commentaries into your slides to facilitate the discussion.**

Northeastern CS 7375 Spring 25
neu-cs7375fall24.hotcrp.com

Search

(All) in Submitted ▾ Search

Reviews

The average PC member has submitted 0.0 reviews. (details · graphs)

As a PC member, you may review any submitted paper.

Offline reviewing · Review preferences

▼ Recent activity:

No recent activity in papers you're following

Submissions

New submission

(admin only)

Administration

Settings
Users
Assignments
Mail
Action log

Conference information

Deadlines
Program committee

Help

Northeastern CS 7375 Spring 25

tia.li@northeastern.edu ☰

(All) Search

Search

(All) in Submitted Search

Search | Advanced search | Saved searches | View options

	ID ▾	Title	# Reviews
☐	#1	Deepfakes, Phrenology, Surveillance, and More! A Taxonomy of AI Privacy Risks (CHI 2024)	0
☐	#2	PrivacyLens: Evaluating Privacy Norm Awareness of Language Models in Action (NeurIPS 2024)	0
☐	#3	"My Data Just Goes Everywhere:" User Mental Models of the Internet and Implications for Privacy and Security (SOUPS 2015)	0
☐	#4	Expectation and purpose: understanding users' mental models of mobile app privacy through crowdsourcing (UbiComp 2012)	0
☐	#5	Toggles, Dollar Signs, and Triangles: How to (In)Effectively Convey Privacy Choices with Icons and Link Texts (CHI 2021)	0
☐	#6	Honesty is the Best Policy: On the Accuracy of Apple Privacy Labels Compared to Apps' Privacy Policies (PETS 2024)	0
☐	#7	"I'm not convinced that they don't collect more than is necessary": User-Controlled Data Minimization Design in Search Engines (USENIX Security 2024)	0
☐	#8	Automating Contextual Privacy Policies: Design and Evaluation of a Production Tool for Digital Consumer Privacy Awareness (CHI 2022)	0
☐	#9	"I need a better description'': An Investigation Into User Expectations For Differential Privacy (CCS 2021)	0
☐	#10	Don't Look at the Data! How Differential Privacy Reconfigures the Practices of Data Science (CHI 2023)	0
☐	#11	"It's a Fair Game", or Is It? Examining How Users Navigate Disclosure Risks and Benefits When Using LLM-Based Conversational Agents (CHI 2024)	0
☐	#12	Granular Privacy Control for Geolocation with Vision Language Models (EMNLP 2024)	0
☐	#13	"What are they gonna do with my data?": Privacy Expectations, Concerns, and Behaviors in Virtual Reality (PETS 2025)	0
☐	#14	Going Incognito in the Metaverse: Achieving Theoretically Optimal Privacy-Usability Tradeoffs in VR (UIST 2023)	0
☐	#15	"If sighted people know, I should be able to know:" Privacy Perceptions of Bystanders with Visual Impairments around Camera-based Technology (USENIX Security 2023)	0
☐	#16	Designing Accessible Obfuscation Support for Blind Individuals' Visual Privacy Management (CHI 2024)	0
☐	#17	How Developers Talk About Personal Data and What It Means for User Privacy: A Case Study of a Developer Forum on Reddit (CSCW 2021)	0
☐	#18	Farsight: Fostering Responsible AI Awareness During AI Application Prototyping (CHI 2024)	0

Select papers (or select all 18), then Download · Tag · Assign · Decide · Mail

#1 Deepfakes, Phrenology, Surveillance, and More! A Taxonomy of AI Privacy Risks (CHI 2024)

▶ **Tags**
None

Email notification
Select to receive email on updates to reviews and comments.

▼ **PC conflicts**
None

▶ **Decision**
Unspecified

▶ **Discussion lead**

▶ **Shepherd**

Review preference

Submitted

Submission (1.3MB) Jan 5, 2025, 2:48:23 AM UTC · d4076bcb

Abstract

Privacy is a key principle for developing ethical AI technologies, but how does including AI technologies in products and services change privacy risks? We constructed a taxonomy of AI privacy risks by analyzing 321 documented AI privacy incidents. We codified how the unique capabilities and requirements of AI technologies described in those incidents generated new privacy risks, exacerbated known ones, or otherwise did not meaningfully alter the risk. We present 12 high-level privacy risks that AI technologies either newly created (e.g., exposure risks from deepfake pornography) or exacerbated (e.g., surveillance risks from collecting training data). One upshot of our work is that incorporating AI technologies into a product can alter the privacy risks it entails. Yet, current approaches to privacy-preserving AI/ML (e.g., federated learning, differential privacy, checklists) only address a subset of the privacy risks arising from the capabilities and data requirements of AI.

Authors

+ *Hidden*

Write review | Assign reviews | Add comment

Add comment

#1 Deepfakes, Phrenology, Surveillance, and More! A Taxonomy of AI Privacy Risks (CHI 2024)

▶ **Tags**
None

Email notification
Select to receive email on updates to reviews and comments.

▼ **PC conflicts**
None

▶ **Decision**
Unspecified

▶ **Discussion lead**

▶ **Shepherd**

Review preference

Paper summary
Discussion prompts

Submitted

Submission (1.3MB) Jan 5, 2025, 2:48:23 AM UTC · d4076bcb

▶ **Abstract**

Privacy is a key principle for developing ethical AI technologies, but how does including AI technologies in products and services change privacy risks? We constructed a taxonomy of AI privacy risks by analyzing 321 documented AI privacy incidents. We codified how the unique capabilities and requirements of AI technologies d… [more]

Authors
+ Hidden

New Review

Offline reviewing Upload form: Choose File no file selected Go
Download form · **Tip:** Use Search or Offline reviewing to download or upload many forms at once.

Paper summary

Markdown styling and LaTeX math supported · Preview

Discussion prompts

Discussion prompts should beopen-endedand not answerable with a simple yes/no or gathering of facts from the paper. For example, do not ask "Did the authors appropriately compensate participants?"; rather, ask "The compensation appears to be under minimum-wage; how might that compensation level have affected the participants the authors could recruit for the study?"

Markdown styling and LaTeX math supported · Preview

Submit review Save draft Cancel Delete review
(admin only)

Differential Privacy (DP) Assignment

- Goals:
  - Get a hands-on experience in DP by seeing how attacks work and how DP (and other PETs) address the attacks
  - Understand the applications, capabilities, and tradeoffs of different DP mechanisms
- Format:
  - Coding tasks + data analysis questions

Course project

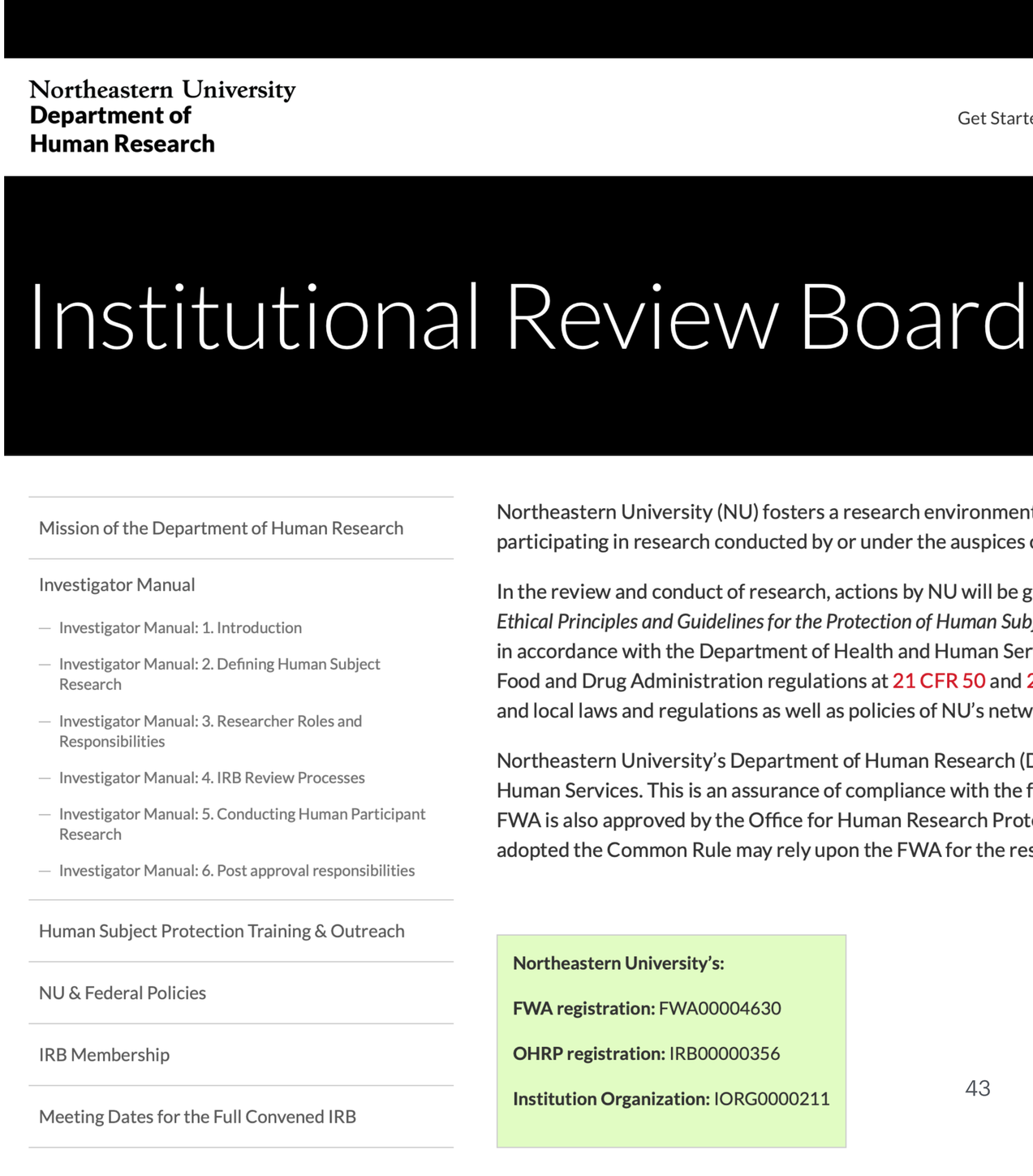
- Individual project
- You’re encouraged to use your ongoing research project as the course project; Make sure to talk to your advisor if you do this.

Project types

- Type 1: Literature Review
- Type 2: Original Research
  - Build systems + user studies
  - Design prototypes + user studies
  - Pure user studies (studying existing systems)
  - Others (need to be related to privacy and involve human-centered perspectives)
- You can choose to do either a Type 1 or Type 2 project

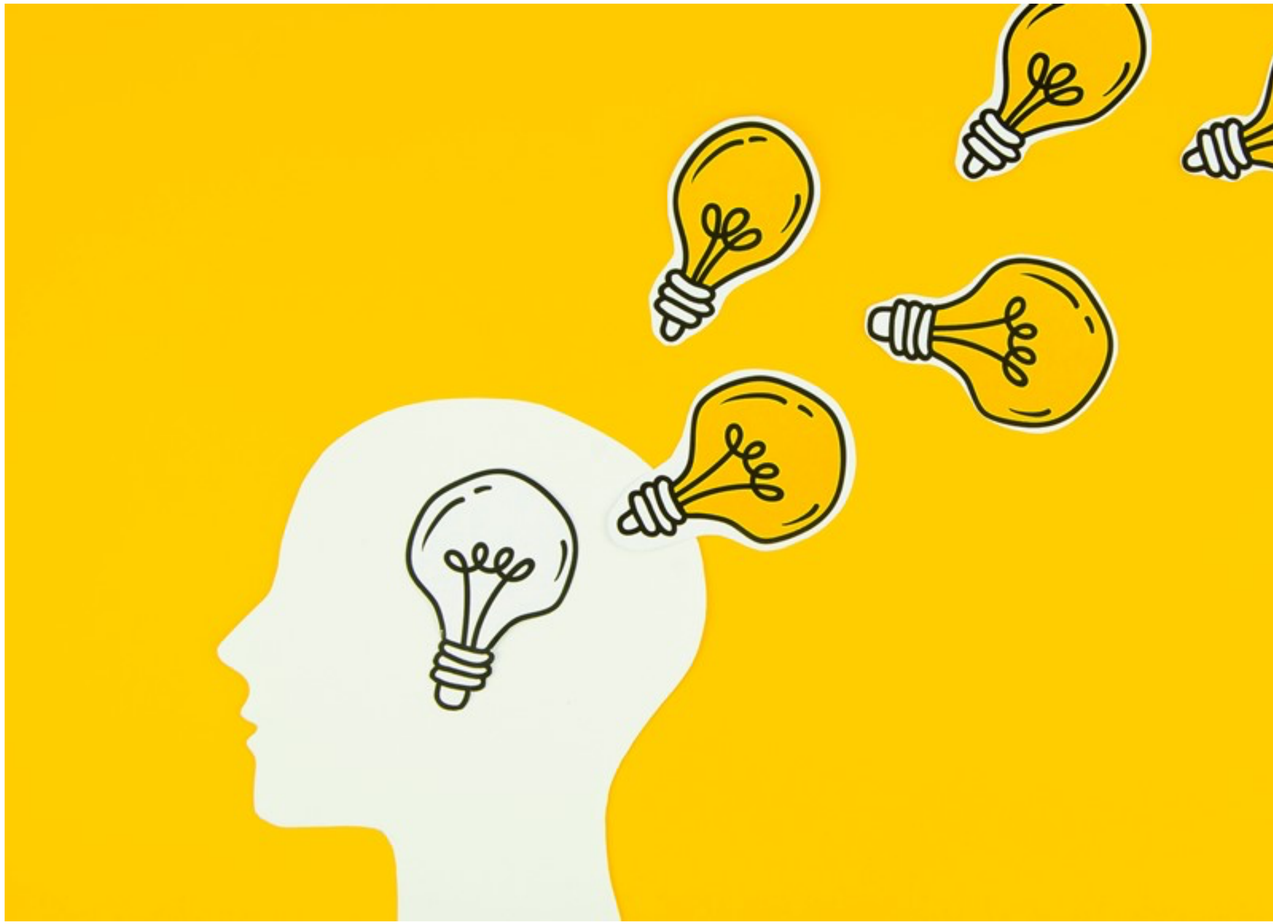
Human-subjects research and IRB

Class projects are exempt from IRB reviews
Talk to me if you're interested in publishing the results



How to generate good ideas?

To have a good idea, you need to first have a lot of ideas!



Example Project Ideas from Last Semester

- Examining User Disclosure Behavior Under Persuasive Conversations with LLM-based Conversational Agents
- Gamification of Privacy Policies
- Investigating StudentWorker Understandings of University Data Collection
- Choice Manipulation Tactics in Corporate Discourse

Project checkpoint 1: Idea descriptions

- By Feb 10, you're expected to have conceived a few project ideas. Submit at least two idea descriptions including: 1. motivation and research gaps (optional); 2. research questions; 3. proposed research activities

Project checkpoint 2: Project proposal

- The class on March 10 will be reserved for the project proposal presentation
- Each person should give a 5-minute pitch of your proposal, followed by 5-minute Q&A.
- For an Original Research project, your presentation needs to cover:
  - Background and motivations: Why is it an important problem? What are the research gaps?
  - Research questions and your proposed tasks to answer these questions
- For a Literature Review project, your presentation needs to cover:
  - Defining the topic and the scope of your literature review
  - An initial list of references

Project checkpoint 3: Final presentation

- The class on April 14 will be reserved for the final presentation
- Students who choose to do an original research project should give a 15-minute presentation followed by 5-minute Q&A. The presentation should cover:
  - Background, research gaps, motivations of the problem you're tackling
  - Research questions and your proposed tasks to answer these questions
  - Final updates: At this point, you should have already completed the planned activities and and obtained substantial results
- Students who choose to do a literature review project don't need to give a presentation, but need to submit a manuscript

Teams

- We'll use Teams to manage assignments, share resources, send reminders of assignment due dates, and help you connect with other students for the course presentation.

Action items

- By the end of this class: Make sure you can access Teams
- By this Friday (Jan 10)
  - Send me your HotCRP account
  - Select the papers you are the discussion lead for
  - Introduce yourself to everyone on Teams
- By next Monday (Jan 13)
  - Submit the first set of reading commentaries