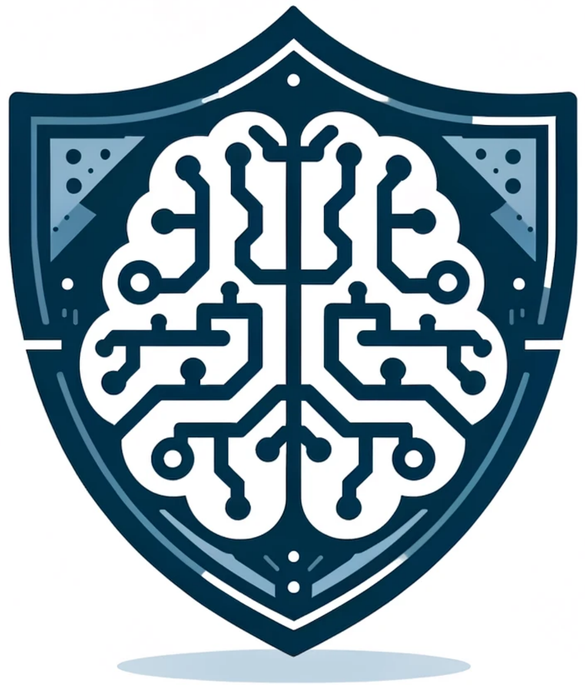
PPAI 2025

6-th AAAI Workshop on

Privacy-Preserving AI

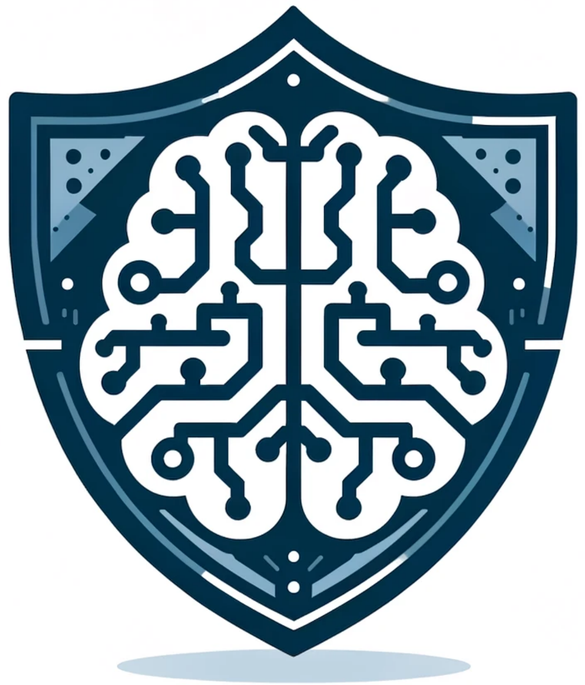


PPAI 2025

6-th AAAI Workshop on

Privacy-Preserving AI

Posters:
In front of the
Registration desk



Welcome

Welcome

- **Why PPAI ?**

  Provide a multi-disciplinary platform for researchers, AI practitioners, and policymakers to focus on both theory and practice of Privacy-preserving AI systems.

Welcome

- **Why PPAI ?**

  Provide a multi-disciplinary platform for researchers, AI practitioners, and policymakers to focus on both theory and practice of Privacy-preserving AI systems.

- **Scope and Topics**
  - Algorithmic approaches to protect data privacy in the context of learning, optimization, and decision making.
  - Social issues related to tracking, tracing, and surveillance programs.
  - **Policy considerations and legal frameworks for privacy; Broader implications of privacy in LLMs; and The societal impact of privacy within AI.**

Welcome

- **Why PPAI ?**

  Provide a multi-disciplinary platform for researchers, AI practitioners, and policymakers to focus on both theory and practice of Privacy-preserving AI systems.

- **Scope and Topics**
  - Algorithmic approaches to protect data privacy in the context of learning, optimization, and decision making.
  - Social issues related to tracking, tracing, and surveillance programs.
  - **Policy considerations and legal frameworks for privacy; Broader implications of privacy in LLMs; and The societal impact of privacy within AI.**

- Submissions
  - 54 submissions: accepted — 8 orals + 36 posters

Welcome

Workshop Chairs

Ferdinando Fioretto
University of Virginia

fioretto@virginia.edu

Juba Ziani
Georgia Institute of Technology

juba.ziani@isye.gatech.edu

Wanrong Zhang
TikTok/University of British Columbia

wanrongzhang@fas.harvard.edu

Jeremy Seeman
Urban Institute/University of Michigan

JSeeman@urban.org

Student Organization Committee

Diptangshu Sen
Georgia Institute of Technology

dsen30@gatech.edu

Saswat Das
University of Virginia

saswatdas@email.virginia.edu

Program

9:00	**Invited Talk** by *Aaron Roth*
9:30	**Invited Talk** by *Alexis Shore Ingber*
10:00	**Contributed Talks**
	Talk 1: *Understanding and Mitigating the Impacts of Differentially Private Census Data on State Level Redistricting*
	Talk 2: *Fairness Issues and Mitigations in (Private) Socio-demographic Data Processes*
	Talk 3: *Privacy-Preserving Retrieval Augmented Generation with Differential Privacy*
	Talk 4: *Hacking the CRC Archive: Evaluating empirical privacy metrics on deidentified data*
10:30	**Break**
11:00	**Contributed Talks**
	Talk 5: *LLM on the wall, who *now*, is the appropriate one of all?": Contextual Integrity Evaluation of LLMs*
	Talk 6: *Understanding Memorization In Generative Models Through A Geometric Framework*
	Talk 7: *Streaming Private Continual Counting via Binning*
	Talk 8: *Laplace Transform Interpretation of Differential Privacy*
11:30	**Tutorial** by *Eugene Bagdasarian*
12:15	**Poster Session** (by the registration desk)
13:30	**Lunch** (on your own)
14:45	**Invited Talk** by *Amy Cyphert*
15:15	**Panel Discussion**
15:45	**Break**
16:15	**Invited Talk** by *Rachel Cummings*
16:45	**Concluding Remarks**

Program

9:00	**Invited Talk** by *Aaron Roth*
9:30	**Invited Talk** by *Alexis Shore Ingber*
10:00	**Contributed Talks**
	Talk 1: *Understanding and Mitigating the Impacts of Differentially Private Census Data on State Level Redistricting*
	Talk 2: *Fairness Issues and Mitigations in (Private) Socio-demographic Data Processes*
	Talk 3: *Privacy-Preserving Retrieval Augmented Generation with Differential Privacy*
	Talk 4: *Hacking the CRC Archive: Evaluating empirical privacy metrics on deidentified data*
10:30	**Break**
11:00	**Contributed Talks**
	Talk 5: *LLM on the wall, who *now*, is the appropriate one of all?": Contextual Integrity Evaluation of LLMs*
	Talk 6: *Understanding Memorization In Generative Models Through A Geometric Framework*
	Talk 7: *Streaming Private Continual Counting via Binning*
	Talk 8: *Laplace Transform Interpretation of Differential Privacy*
11:30	**Tutorial** by *Eugene Bagdasarian*
12:15	**Poster Session** (by the registration desk)
13:30	**Lunch** (on your own)
14:45	**Invited Talk** by *Amy Cyphert*
15:15	**Panel Discussion**
15:45	**Break**
16:15	**Invited Talk** by *Rachel Cummings*
16:45	**Concluding Remarks**

Aaron Roth
University of Pennsylvania

Program

9:00	**Invited Talk** by *Aaron Roth*
9:30	**Invited Talk** by *Alexis Shore Ingber*
10:00	**Contributed Talks**
	Talk 1: *Understanding and Mitigating the Impacts of Differentially Private Census Data on State Level Redistricting*
	Talk 2: *Fairness Issues and Mitigations in (Private) Socio-demographic Data Processes*
	Talk 3: *Privacy-Preserving Retrieval Augmented Generation with Differential Privacy*
	Talk 4: *Hacking the CRC Archive: Evaluating empirical privacy metrics on deidentified data*
10:30	**Break**
11:00	**Contributed Talks**
	Talk 5: *LLM on the wall, who *now*, is the appropriate one of all?": Contextual Integrity Evaluation of LLMs*
	Talk 6: *Understanding Memorization In Generative Models Through A Geometric Framework*
	Talk 7: *Streaming Private Continual Counting via Binning*
	Talk 8: *Laplace Transform Interpretation of Differential Privacy*
11:30	**Tutorial** by *Eugene Bagdasarian*
12:15	**Poster Session** (by the registration desk)
13:30	**Lunch** (on your own)
14:45	**Invited Talk** by *Amy Cyphert*
15:15	**Panel Discussion**
15:45	**Break**
16:15	**Invited Talk** by *Rachel Cummings*
16:45	**Concluding Remarks**

Aaron Roth
University of Pennsylvania

Alexis Shore Ingber
University of Michigan

Program

9:00	**Invited Talk** by *Aaron Roth*
9:30	**Invited Talk** by *Alexis Shore Ingber*
10:00	**Contributed Talks**
	Talk 1: *Understanding and Mitigating the Impacts of Differentially Private Census Data on State Level Redistricting*
	Talk 2: *Fairness Issues and Mitigations in (Private) Socio-demographic Data Processes*
	Talk 3: *Privacy-Preserving Retrieval Augmented Generation with Differential Privacy*
	Talk 4: *Hacking the CRC Archive: Evaluating empirical privacy metrics on deidentified data*
10:30	**Break**
11:00	**Contributed Talks**
	Talk 5: *LLM on the wall, who *now*, is the appropriate one of all?": Contextual Integrity Evaluation of LLMs*
	Talk 6: *Understanding Memorization In Generative Models Through A Geometric Framework*
	Talk 7: *Streaming Private Continual Counting via Binning*
	Talk 8: *Laplace Transform Interpretation of Differential Privacy*
11:30	**Tutorial** by *Eugene Bagdasarian*
12:15	**Poster Session** (by the registration desk)
13:30	**Lunch** (on your own)
14:45	**Invited Talk** by *Amy Cyphert*
15:15	**Panel Discussion**
15:45	**Break**
16:15	**Invited Talk** by *Rachel Cummings*
16:45	**Concluding Remarks**

Aaron Roth
University of Pennsylvania

Alexis Shore Ingber
University of Michigan

Eugene Bagdasarian
University of Massachusetts, Amherst

Posters
In front of the
Registration desk

Program

9:00	**Invited Talk** by *Aaron Roth*
9:30	**Invited Talk** by *Alexis Shore Ingber*
10:00	**Contributed Talks**
	Talk 1: *Understanding and Mitigating the Impacts of Differentially Private Census Data on State Level Redistricting*
	Talk 2: *Fairness Issues and Mitigations in (Private) Socio-demographic Data Processes*
	Talk 3: *Privacy-Preserving Retrieval Augmented Generation with Differential Privacy*
	Talk 4: *Hacking the CRC Archive: Evaluating empirical privacy metrics on deidentified data*
10:30	**Break**
11:00	**Contributed Talks**
	Talk 5: *LLM on the wall, who *now*, is the appropriate one of all?": Contextual Integrity Evaluation of LLMs*
	Talk 6: *Understanding Memorization In Generative Models Through A Geometric Framework*
	Talk 7: *Streaming Private Continual Counting via Binning*
	Talk 8: *Laplace Transform Interpretation of Differential Privacy*
11:30	**Tutorial** by *Eugene Bagdasarian*
12:15	**Poster Session** (by the registration desk)
13:30	**Lunch** (on your own)
14:45	**Invited Talk** by *Amy Cyphert*
15:15	**Panel Discussion**
15:45	**Break**
16:15	**Invited Talk** by *Rachel Cummings*
16:45	**Concluding Remarks**

Aaron Roth
University of Pennsylvania

Alexis Shore Ingber
University of Michigan

Eugene Bagdasarian
University of Massachusetts, Amherst

Posters
In front of the
Registration desk

Program

9:00	**Invited Talk** by *Aaron Roth*
9:30	**Invited Talk** by *Alexis Shore Ingber*
10:00	**Contributed Talks**
	Talk 1: *Understanding and Mitigating the Impacts of Differentially Private Census Data on State Level Redistricting*
	Talk 2: *Fairness Issues and Mitigations in (Private) Socio-demographic Data Processes*
	Talk 3: *Privacy-Preserving Retrieval Augmented Generation with Differential Privacy*
	Talk 4: *Hacking the CRC Archive: Evaluating empirical privacy metrics on deidentified data*
10:30	**Break**
11:00	**Contributed Talks**
	Talk 5: *LLM on the wall, who *now*, is the appropriate one of all?": Contextual Integrity Evaluation of LLMs*
	Talk 6: *Understanding Memorization In Generative Models Through A Geometric Framework*
	Talk 7: *Streaming Private Continual Counting via Binning*
	Talk 8: *Laplace Transform Interpretation of Differential Privacy*
11:30	**Tutorial** by *Eugene Bagdasarian*
12:15	**Poster Session** (by the registration desk)
13:30	**Lunch** (on your own)
14:45	**Invited Talk** by *Amy Cyphert*
15:15	**Panel Discussion**
15:45	**Break**
16:15	**Invited Talk** by *Rachel Cummings*
16:45	**Concluding Remarks**

Aaron Roth
University of Pennsylvania

Alexis Shore Ingber
University of Michigan

Eugene Bagdasarian
University of Massachusetts, Amherst

Amy Cyphert
West Virginia University

Posters
In front of the
Registration desk

Program

9:00	**Invited Talk** by *Aaron Roth*
9:30	**Invited Talk** by *Alexis Shore Ingber*
10:00	**Contributed Talks**
	Talk 1: *Understanding and Mitigating the Impacts of Differentially Private Census Data on State Level Redistricting*
	Talk 2: *Fairness Issues and Mitigations in (Private) Socio-demographic Data Processes*
	Talk 3: *Privacy-Preserving Retrieval Augmented Generation with Differential Privacy*
	Talk 4: *Hacking the CRC Archive: Evaluating empirical privacy metrics on deidentified data*
10:30	**Break**
11:00	**Contributed Talks**
	Talk 5: *LLM on the wall, who *now*, is the appropriate one of all?": Contextual Integrity Evaluation of LLMs*
	Talk 6: *Understanding Memorization In Generative Models Through A Geometric Framework*
	Talk 7: *Streaming Private Continual Counting via Binning*
	Talk 8: *Laplace Transform Interpretation of Differential Privacy*
11:30	**Tutorial** by *Eugene Bagdasarian*
12:15	**Poster Session** (by the registration desk)
13:30	**Lunch** (on your own)
14:45	**Invited Talk** by *Amy Cyphert*
15:15	**Panel Discussion**
15:45	**Break**
16:15	**Invited Talk** by *Rachel Cummings*
16:45	**Concluding Remarks**

Aaron Roth
University of Pennsylvania

Alexis Shore Ingber
University of Michigan

Eugene Bagdasarian
University of Massachusetts, Amherst

Amy Cyphert
West Virginia University

Rachel Cummings
Columbia University

Program 3:15PM

PPAI-25 Panel:

“Understanding and regulating the privacy risks in an embodied agents world”

Annette Zimmermann
University of Wisconsin-Madison

Lauryn P. Gouldin
Syracuse University

Eugene Bagdasarian
University of Massachusetts, Amherst

Aloni Cohen
University of Chicago

A bit of shameless ad

Contributing Authors

James Anderson
Columbia University

Kallista Bonawitz
Google Research

Konstantinos Chatzikokolakis
University of Athens

Giovanni Cherubin
Microsoft Research

Graham Cormode
University of Warwick

Rachel Cummings
Columbia University

Damien Desfontaines
Tumult Labs

Liyue Fan
University of North Carolina at Charlotte

Ferdinando Fioretto
University of Virginia

Marco Gaboardi
Boston University

Marzyeh Ghassemi
Massachusetts Institute of Technology

Bryant Gipson
Google

Anna Goldenberg
University of Toronto

Michael Hay
Tumult Labs and Colgate University

Peter Kairouz
Google Research

Steven H. Low
California Institute of Technology

Ashwin Machanavajjhala
Tumult Labs and Duke University

Brendan McMahan
Google Research

Catuscia Palamidessi
Inria and École Polytechnique

Nicolas Papernot
University of Toronto and Vector Institute

David Pujol
Tumult Labs

Reza Shokri
National University of Singapore

Jeremy Seeman
Urban Institute

Thomas Steinke
Google DeepMind

Vinith M. Suriyakumar
Massachusetts Institute of Technology

Yurii Sushko
Goolge

Yuchao Tao
Snap Inc

Christine Task
Knexus Research

Andreas Terzis
Google

Abhradeep Thakurta
Google Deepmind

Salil Vadhan
Harvard University

Pascal Van Hentenryck
Georgia Institute of Technology

Jiayuan Ye
National University of Singapore

Fengyu Zhou
California Institute of Technology

Juba Ziani
Georgia Institute of Technology

now
the essence of knowledge

Student sponsorship

We will be able to provide several student scholarships!

Diamond

Google

Deloitte.

Gold

Silver

OpenDP

Thank you!!

Student sponsorship

We will be able to provide several student scholarships!

Diamond

Gold

Career Opportunities:

If you'd like to hear about available opportunities from our sponsors, please fill out the form on the PPAI website https://ppai-workshop.github.io/

OpenDP

Thank you!!

Thank You!

Program committee

Alyssa	Columbus	Johns Hopkins University
Amin	Rahimian	University of Pittsburgh
Anshuman	Suri	University of Virginia
Antti	Koskela	Nokia Bell Labs
Aur√©lien	Bellet	INRIA
Bishnu	Bhusal	University of Missouri
Catuscia	Palamidessi	Laboratoire d'informatique de l'√âcole polytechnique
Difang	Huang	University of Hong Kong
Diptangshu	Sen	Georgia Institute of Technology
Edo	Roth	Google
Ejaz	Ahmed	School of Public Policy, Georgia Tech
Fan	Mo	Imperial College London
Ferdinando	Fioretto	University of Virginia
Gharib	Gharibi	TripleBlind.ai
Graham	Cormode	University of Warwick
Hongyan	Chang	NUS
James	Flemings	University of Southern California
Jeremy	Seeman	Urban Institute
Joonas	J√§lk√∂	University of Helsinki
Juba	Ziani	Georgia Tech
Kobbi	Nissim	Georgetown University
Krishna	Acharya	Georgia Institute of Technology
Krystal	Maughan	University of Vermont
Ludmila	Glinskih	Google
Manel	Slokom	CWI
Marco	Romanelli	Hofstra University
Moshe	Shenfeld	Hebrew University
Pierre	Tholoniat	Columbia University
Prajwal	Panzade	Old Dominion University
Pratiksha	Thaker	Carnegie Mellon University
Rakshit	Naidu	Georgia Institute of Technology
Ranya	Aloufi	Imperial College London
Rojin	Rezvan	University of Wisconsin-Madison
Rola	Alseidi	Philadelphia University
Rui-Jie	Yew	Brown University
Sandeep	Silwal	University of Wisconsin-Madison
Sandro	Radovanovic	University of Belgrade
Sankarshan	Damle	EPFL
Saswat	Das	University of Virginia
Shahnewaz Karim	Sakib	The University of Tennessee at Chattanooga
Stacey	Truex	Denison University
Tianhao	Li	Duke University
Tianhao	Wang	University of Virginia
Vahid	Behzadan	University of New Haven
Vasanta	Chaganti	Swarthmore College
Xi	He	University of Waterloo
Yeojoon	Youn	Georgia Institute of Technology
Yidi	Xu	University of Pennsylvania
Yifan	Liu	Georgia Tech

Thank You!

Invited Speakers

Aaron Roth
University of Pennsylvania

Rachel Cummings
Columbia University

Amy Cyphert
West Virginia University

Alexis Shore Ingber
University of Michigan

Eugene Bagdasarian
University of Massachusetts, Amherst

Panelists

Annette Zimmermann
University of Wisconsin-Madison

Lauryn P. Gouldin
Syracuse University

Eugene Bagdasarian
University of Massachusetts, Amherst

Aloni Cohen
University of Chicago

Big Thanks to all the workshop participants!

PPAI-25 Panel

“Understanding and regulating the privacy risks in an embodied agents world”

Annette Zimmermann
University of Wisconsin-Madison

Lauryn P. Gouldin
Syracuse University

Eugene Bagdasarian
University of Massachusetts, Amherst

Aloni Cohen
University of Chicago

Scenario

Consider a robotic agents deployed to "observe" individuals (e.g., skilled workers at work, children at play, or by-standing) to collect data such as personal behaviors, location data, and "skills", which are later pooled to improve machine learning models.

Similar scenarios can be pictured for personal assistant agents.

- **home assistant agents** (either embodied or not) that are designed to learn user habits to improve service quality but inadvertently record sensitive conversations and personal routines.
- **medicine** and real-life symptom tracking and monitoring + automated treatment and symptom management + personalized health recommendations.

Scenario

Consider a robotic agents deployed to "observe" individuals (e.g., skilled workers at work, children at play, or by-standing) to collect data such as personal behaviors, location data, and "skills", which are later pooled to improve machine learning models.

Similar scenarios can be pictured for personal assistant agents.

- **home assistant agents** (either embodied or not) that are designed to learn user habits to improve service quality but inadvertently record sensitive conversations and personal routines.
- **medicine** and real-life symptom tracking and monitoring + automated treatment and symptom management + personalized health recommendations.

Issue: This data might is not only sensitive on its own, but could even be shared with third-party developers for training commercial models.

Technical focus

1. - Current privacy techniques are likely insufficient today for such large scale, dynamic, and multi-modal data. What is missing and how do we get there?

2. - What metrics and auditing techniques can we develop to quantify the privacy risk of such data collection?

3. - How do we even measure the privacy risk in these settings (e.g., multimodal data)?

Policy / Legal focus

1. Current regulatory frameworks are likely not be sufficient for these privacy concerns. What are we missing and how do we get there?

2. What do you think is the role of consent when it comes to scrapping and subtle data collection?

3. Assume a privacy break arises (an LLM leaks some sensitive user information). Where does the responsibility lie?

4. How do we establish clear guidelines for data retention and post-violation remedies?

Closing remarks

1. Panel Discussion
    1. A
    2. B
    3. C

The 6th AAAI Workshop on Privacy-Preserving Artificial Intelligence (PPAI 2025)

PPAI will resume at 2:45 PM

(Poster session now!)
By the registration desk