General overview of projects:

Carried out by 4-5 students (depending on the number of students)

Tutoring by me and the TA

If you take this course for credit, **be proactive** on your choice of project!

You can propose your own subject and we will discuss its appropriateness; it can be related to your ongoing research -> please email me ASAP

I will also provide potential project topics **(refer to Moodle)**

Ideally, a successful project can lead to a publication

Novelty and **effort** on the project are important

Deliverables include final report, midterm presentation, and final presentation

Project content:

Project can either be (i) research or (ii) implementation based.

(i) Research:

- Focus on a particular topic
- Do a literature survey
  - NDSS, ACM CCS, IEEE S&P, Usenix Security, PETS ->relevant conference proceedings are a good start!
- Analyze the existing work and criticize (determine weaknesses and potential improvements)
- Make suggestions, propose your improvements, come up with a systematic solutions including pseudocodes, system figures etc.

Examples:

- Privacy in social networks
- Privacy-enhanced access control, authentication, and identity management
- Privacy-preserving, secure systems (e.g., distributed machine learning, federated learning)
- De-anonymization
- ….

(ii) Implementation

- Focus on a particular application or dataset
- Decide on the architecture and system model
- Determine the privacy requirements

- Implement PETs for your application or dataset
- Examples:
  - Linkage attacks
  - CryptDB for genomic data
  - Inference/reconstruction attacks on machine learning models
  - Defenses for machine learning systems
  - Applications of PETs
  - Attacks/defenses for healthcare systems/genomic data
  - ….

Examples from previous semesters (Instructor: Prof. Erman Ayday):

- Effect of Indirect Information Sharing on Privacy
- Kin Genomic Privacy: Inference Attacks
- De-anonymizing Unstructured Online Social Networks
- De-anonymizing Private Instagram Profiles via Twitter
- De-anonymization of Social Network Data
- Privacy Preserving Active Learning with Secure Multiparty Computation
- Microsoft Malware Classification Challenge
- Data Protection Legislation in Turkey, EU and USA
- Privacy-Preserving Community Detection
- Practical Differential Privacy via Grouping and Smoothing
- Side Channel Attacks: A Historical Survey
- Privacy Preserving Genome Wide Association Studies (GWAS) Using Hadoop
- De-anonymizing Call Records
- Data Sharing and Privacy in Genomics
- De-anonymizing medical databases
- Privacy Preserving Dynamic Time Warping
- Privacy preserving of RIMARC algorithm
- Detecting Fake Accounts on Social Networks

- Bioinformatic Data Sharing
- De-anonymizing Online Social Networks

Project grading:

Overall, project is the 60% of the course:

- Midterm presentation: 10%

- Final presentation: 25%

- Final report: 25%

Deadlines for the deliverables:

By the end of Week 2: Finalize project group formation by the end of the week

By the end of Week 3: Finalize topic selection by the end of the week

By the end of Week 5: Flash talks about the projects 3 minute per group due by the end of the week

Week 7: Midterm Presentations (live)

Week 15 and 16: Final Presentations (live)

By the end of Week 16: Final reports (and source codes, if any) are due by the end of the week

Flash talks:

They are not graded

The goal is for me to give you some initial feedback about your projects

No slides are required (unless you want to)

You will just, in 3-4 minutes, talk about your project idea

I will reserve office hours or end of the lecture for this

Project reports (final):

- Submit to Moodle: ProjectAcronym.pdf (along with the source codes, if any)

- Format: ACM 2-column format (https://www.acm.org/publications/proceedings-template)

- No page limitation

Tentative outline includes:

- Problem statement and literature search

- Planned methods (for midterm presentations)

- Next steps (for midterm presentations)

- Expected output (for midterm presentations)

- Proposed solution (for final reports/presentations)

- Future work and conclusion (for final reports/presentations)

Grading of the written reports will be based on:

–Organization (sectioning, graphs, figures, citations, references, etc.)

–Grammar

–Readability

–Novelty and difficulty of the problem

–Publishability

Presentations (final):

- On the morning of your presentation, please email me and our TA your presentation slides in **PDF format by 9am (EST).**

Duration:

For final presentations: 20 minutes + 5 minutes Q&A

Grading of the presentations will be based on:

– Timing

– Quality of slides (graphics, text, figures)

– Quality of the presentation

– Novelty and difficulty of the problem

– Publishability