Privacy & Identity Lab
Radboud University
Tilburg University
University of Groningen

Introduction

Jaap-Henk Hoepman
Privacy & Identity Lab
Radboud University
Tilburg University
University of Groningen

@jhh.cs.ru.nl // www.cs.ru.nl/~jhh // blog.cs.ru.nl // @xotoxot

Agenda

• Course overview
• Privacy: an introduction
Organisation

- Teachers
  - Jaap-Henk Hoepman (jhh@cs.ru.nl; room 19.12)
- Blackboard not used
  - Website: https://www.cs.ru.nl/~jhh/secsem.html
  - Wiki: http://wiki.science.ru.nl/privacy/

Seminar

- Seminar
  - Student lecture
  - Student paper
  - Student opposition
- Grade = weighted average
  - But only if all grades at least 5.5
  - If not, lowest grade is final grade!
- Working in groups
  - 2 or 3 people
- Attendance required
- Lecture rooms
  - From February 6 to March 19 in room HFML 0220. From April 16 to June 1 in HG 00.310.
## Topics (first come first serve)

- **Privacy in databases**
  - How to provide (controlled) access to personal data stored in databases, without immediately threatening the privacy of the people involved, using mechanisms like differential privacy or statistical disclosure control.

- **Privacy friendly search**
  - How to hide the query (i.e. what is searched for) from the party hosting the database.

- **Searching in encrypted databases**
  - How to also hide the underlying data in the database from the party hosting the database.

- **Privacy in machine learning**
  - How to ensure that individual data used to train a machine learning model is not leaked when using the model.

- **Polymorphic encryption**
  - How to protect privacy in e.g. health care where data must be made conditionally accessible to certain care providers while staying encrypted in general.

- **Privacy friendly identity management**
  - How to use e.g. attribute based credentials or other claims based approaches to make identity management more privacy friendly.

- **Privacy friendly revocation of credentials**
  - How to (efficiently) revoke anonymous credentials. I.e. how to revoke a particular credential, even though individual credentials cannot be traced by definition.

- **Revocable privacy**
  - How to guarantee privacy while also guaranteeing that all users of a system abide by some predetermined rules, i.e. how to design systems that are both privacy friendly and secure.

- **Privacy friendly location based services**
  - How to provide a service that depends on the user's current location, without revealing the actual, exact location?

- **Privacy in asynchronous messaging**
  - How to establish contact anonymously, and how to subsequently exchange messages in an unlinkable fashion that prevents the service provider to learn who is communicating with who.

- **Anonymous cryptocurrencies**
  - How to make Bitcoin like cryptocurrencies privacy friendly.

- **Secure multiparty computation**
  - How to jointly compute the output of a function (e.g. some aggregate statistic) without revealing the individual inputs.

## Research

- **analyse a particular practical case**
  - what are the privacy issues (from a societal and legal perspective) and how are they dealt with

- **give a precise and concise problem description**
  - in technical terms: define your model; your assumptions

- **investigate possible PETs that apply**
  - summarise your analysis

- **pick one and solve the problem (involves a protocol)**
  - describe this in sufficient detail!

- **(informally) prove or argue correctness**
**Student lecture**

- **Goal of lecture**
  - to inform other students about your research

- **Important**
  - make lecture interactive
  - add additional material

- **Discuss draft**
  - thursday 13:00-13:15 the week before, in my office
  - mail slides etc. at least two day before

---

**Student lecture: grading**

<table>
<thead>
<tr>
<th>Content</th>
<th>Form and performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Argumentation and Depth</td>
<td></td>
</tr>
<tr>
<td>Is your lecture provide a solid basis and backing of all statements and claims made?</td>
<td>Structure: Logical ordering of your lecture, the relationship between the topics.</td>
</tr>
<tr>
<td>Is intelligibility</td>
<td></td>
</tr>
<tr>
<td>Whether the message comes across, whether your arguments and discussion are clear and understandable</td>
<td>Attractiveness: Whether your lecture captivates the audience, your use of supporting materials (e.g. powerpoint).</td>
</tr>
<tr>
<td>Comprehensiveness</td>
<td></td>
</tr>
<tr>
<td>Whether your lecture covers all important aspects, and clearly separates important issues from secondary details</td>
<td>Delivery: How well your argumentation and content connect with the audience, your presence in front of the class, the liveliness and tone of your lecture.</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Student paper**

- **Goal**
  - Report on research
  - Express own perspective and opinion on PETs

- **Format**
  - Roughly 12 pages (excluding references)
  - A4, reasonable margins, 10-11 pt font

- **Beware**
  - Collect your own literature as well
  - Use input obtained during presentation in class
**Student paper**

- Typical structure
  - Context
  - Problem description
    - Including legal/social analysis
  - Proposed solution
  - Technical analysis
  - Conclusions

**Student paper: planning**

- Average timespan
  - Literature study: 2 weeks
  - Perform research: 2 weeks
  - Write skeleton: 1 week
  - Write final paper: 3 weeks

- Deadlines
  - April 23: Skeleton
  - June 11: Final paper

- So start as soon as you can!

**Student paper: grading**

<table>
<thead>
<tr>
<th>Content</th>
<th>Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>Technical quality</td>
<td>Clarity of writing, objectiveness, linguistic quality (in terms of spelling and grammar)</td>
</tr>
<tr>
<td>Analysis</td>
<td>Logical structure of the paper, helping the reader understand the paper and giving the paper a natural flow</td>
</tr>
<tr>
<td>Quality of references</td>
<td>Accuracy of endnotes, addressing relevant references properly</td>
</tr>
<tr>
<td>Own opinion</td>
<td>Whether the paper clearly expresses and argues over one's opinions on the subject matter</td>
</tr>
</tbody>
</table>
**Working in groups**

- Everyone responsible for all output
- Review each others work!
- Work together, not separately
- Plan your work
- Equally divide work
  - And make sure everyone delivers
  - If not: notify me before everything escalates….

---

**Remaining points**

- Contribute to the wiki!
Contents

- Privacy under threat
  - Government
  - Business
  - People
- What is privacy?
  - The value of privacy
    - Individual liberty
    - Social value
- How the law protects privacy

Government surveillance

Fraud detection, policing
They know things before you yourself do!

They track you even in real shops

People...
- Online 24 hours/day
- Do many things over the Internet
  - Social networking
  - Communications
  - Reading
  - Video
  - Finance
  - Maps
  - Platforms (Airbnb, Uber)
Privacy

what is privacy according to you?

Privacy typology (Koops et. al. 2017)

bodily privacy
spatial privacy
communicational privacy
proprietary privacy
intellectual privacy
decisional privacy
associational privacy
behavioral privacy

7 types of privacy

- the (physical) person,
- behaviour and action,
- personal communication,
- data and image,
- thoughts and feelings,
- location and space, and
- association (including group privacy).

Finn, R.L., Wright, D., and Friedewald, M.: Seven types of privacy. CPDP 2012
Clarke, R.: Introduction to Cybercrime and Information Privacy, and Definitions of Privacy, 2007
Different definitions

- The right to be let alone
  - [Warren & Brandeis, 1890]
- Informational self-determination: The right to determine for yourself when, how and to what extent information about you is communicated to others
  - [Westin, 1967]
- The freedom from unreasonable constraints on the construction of one’s identity
  - [Agre, 1994]
- Contextual integrity: the right to prevent information to flow from one context to another
  - [Nissenbaum, 2004]

Contextual integrity

Don’t confuse these concepts!
Privacy invasions

- Collect
- Process
- Disseminate


Privacy computing (1950 -)
- searching becomes efficient
- data kept forever

Networking (1980 -)
- datasharing becomes easy
- data accessible on-line

"network effect"

Different types of data/information

- Volunteered
  - What you reveal explicitly when asked

- Observed
  - What you reveal implicitly by your behaviour

- Inferred
  - What is derived from other data about you

[World Economic Forum Report Personal Data: The Emergence of a New Asset Class]
Data vs Metadata

- Metadata (= Behavioural data)
  - Condensed information (rich, easy to process)
  - More "true" (judge a man not on what he says but on what he does)

Why is privacy important

“Privacy is essential for freedom, democracy, psychological well-being, individuality and creativity”

Moral basis for data protection

- prevention of information-based harm
  - Like guns, information may kill people
- prevention of informational inequality
  - The "market" of information
  - Non-discrimination
- prevention of informational injustice
  - Spheres of privacy must be protected
- respect for moral autonomy.
  - People change

Searching for the right metaphor

- Orwell / Big Brother
- Chandler / Little Sister
- Kafka / The Trial

Image: The Matrix
I have nothing to hide....

Wrong assumption

- The point is not that there is data that
- is a priori "wrong" or illegal

(assuming by the sender)

- The point is that "innocent" data can
- (later) be used wrongly

(by the current receiver)

Everybody

- Assumes
- that even "innocent" data as is...?

Freedom of thought

- That job
- That was legal (moral) vs ...

No distinction between illegal (moral) vs ...

What is the function of data?
Applies when you **process personal data?**

1. **Personal data** means any information relating to an identified or identifiable natural person (the data subject); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person.

   - **So...**
     - Name
     - Social security number
     - Email address
   - **But also...**
     - License plate
     - IP Address
     - Likes
     - Tweets
     - Search terms

2. **Processing** means any operation or set of operations which is performed on personal data, whether or not by automated means, such as collecting, recording, organizing, storing, maintaining, adapting or altering, retrieving, using, broadcasting, publishing, disclosing or otherwise making available, alignment or combination, restriction, erasure or destruction.
Subject / controller / processor

- consent
- necessity

Data Subject Rights
- notification
- access
- rectification
- object to profiling

Data Protection Principles
- purpose limitation
- data minimisation
- duration of retention
- accuracy of the data

Accountability
- risk based-approach
- transparency of processing
- data protection by design
- data protection impact assessment
Resources

- **Websites**
  - [https://wiki.science.ru.nl/privacy/](https://wiki.science.ru.nl/privacy/)
  - [https://www.eff.org/](https://www.eff.org/)
  - [https://www.bof.nl](https://www.bof.nl)

- **Books**
  - Ilija Trojanov, Juli Żeh “Aanslag op de vrijheid”, de Geus, 2010