

# Chapter 7: “Privacy by Design” versus “Privacy Engineering”

**Lecture PETs4DS:**  
**Privacy Enhancing Technologies for Data Science**

Dr. Benjamin Heitmann and Prof. Dr. Stefan Decker  
Informatik 5  
Lehrstuhl Prof. Decker



**How to apply knowledge about privacy threats and PETs to new and existing IT systems?**

- We learned how to identify privacy threats with LINDDUN.
- We learned about several Privacy Enhancing Technologies (PETs).
- Question: Is that enough to implement IT systems with privacy in mind?

# Comparing SSL and PGP

## Two examples: SSL/TLS vs. GPG

---

Lets look at two example implementations using the same cryptographic primitives

- Lets look at two different ways to address the same privacy threat with the same cryptographic primitives:
- **Use case:** Establishing a trusted & secure channel on top of an insecure channel
- **Cryptographic primitive:** Asymmetric encryption using pairs of public/secret keys
- **Privacy threats** addressed by the two approaches to provide a secure channel:  
(Using terms from the LINDDUN threat analysis, both approaches provide a private data flow, so the following is a list of applicable threats)
  - Linkability: No
  - Identifiability: No
  - Non-repudiation: No
  - Detectability: No
  - **Disclosure of information: YES**

- Open standard, part of browsers since 1994
- Enables secure HTTP connection between browser and web server
  - Hand-shake uses public-key crypto
  - Generation and exchange of keys for symmetric encryption
  - Symmetric encryption then provides a fast and secure communication channel
- Shipped with every web browser
  - Every browser has
- Protocol and implementations are constantly updated



# SSL/TLS: Pros and Cons

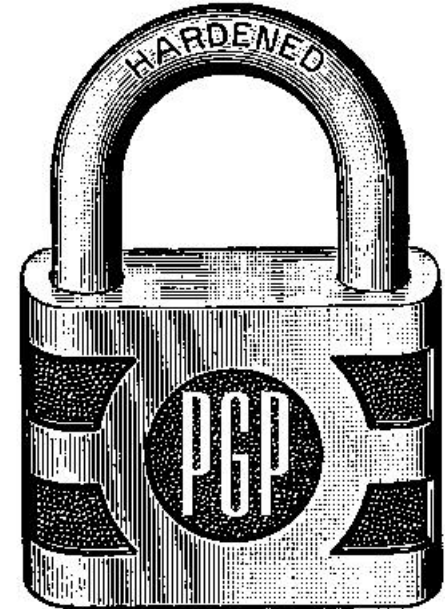
---

- Pros:
  - easy to use
  - no extra software to install
  - completely transparent to user
  - proven secure if browser is updated and current
- Cons:
  - provides high value target
  - many attacks are known (see RFC 7457, which summarises known attacks)
  - almost impossible to protect against attacks from consumer side (only vendors can address attacks)
  - sometimes insecure settings can be forced on users

## Table of Contents

<a href="#">1.</a>	<a href="#">Introduction .....</a>	
<a href="#">2.</a>	<a href="#">Attacks on TLS .....</a>	
<a href="#">2.1.</a>	<a href="#">SSL Stripping .....</a>	
<a href="#">2.2.</a>	<a href="#">STARTTLS Command Injection Attack (CVE-2011-0411) .</a>	
<a href="#">2.3.</a>	<a href="#">BEAST (CVE-2011-3389) .....</a>	
<a href="#">2.4.</a>	<a href="#">Padding Oracle Attacks .....</a>	
<a href="#">2.5.</a>	<a href="#">Attacks on RC4 .....</a>	
<a href="#">2.6.</a>	<a href="#">Compression Attacks: CRIME, TIME, and BREACH .....</a>	
<a href="#">2.7.</a>	<a href="#">Certificate and RSA-Related Attacks .....</a>	
<a href="#">2.8.</a>	<a href="#">Theft of RSA Private Keys .....</a>	
<a href="#">2.9.</a>	<a href="#">Diffie-Hellman Parameters .....</a>	
<a href="#">2.10.</a>	<a href="#">Renegotiation (CVE-2009-3555) .....</a>	
<a href="#">2.11.</a>	<a href="#">Triple Handshake (CVE-2014-1295) .....</a>	
<a href="#">2.12.</a>	<a href="#">Virtual Host Confusion .....</a>	
<a href="#">2.13.</a>	<a href="#">Denial of Service .....</a>	
<a href="#">2.14.</a>	<a href="#">Implementation Issues .....</a>	
<a href="#">2.15.</a>	<a href="#">Usability .....</a>	

- First version from 1991. Open Standard since 1998.
- Used to sign and/or encrypt emails on top of the insecure SMTP protocol
- Combines symmetric key crypto (for speed) and asymmetric key crypto (for secure key exchange)
  - Public key of recipient is used to encrypt a session key
  - Session key is used once to encrypt message
  - Encrypted message and session key are sent to recipient
- Not integrated into contemporary email clients.
- Requires users to:
  - Generate own keys.
  - Import and manage keys of communication partners.
- All other aspects of “operational security” e.g. Web of Trust” need to be managed by user.



- Pros:
  - Very secure
  - All parameters can be chosen by users
  - Users can customise workflow based on his own requirements
  - Not many well known attacks
- Cons:
  - Requires extra software
  - Makes it possible to use software in the wrong way
  - Does not enforce best practices



# Comparison between SSL and PGP

---

Why is one a failure and one a success?

- SSL is transparent for user
- SSL protects money
- All web servers which handle money support SSL
- Therefore using SSL does not require finding out if communication partner is using it.
- **Other ideas?**
- **Discuss!**
- PGP “only” protects freedom of speech
- PGP is hard to set-up
- PGP is hard to use
- Email can be used without PGP.
- Using PGP requires knowing if your communication partner is using it or not

# Privacy by Design

Based on:

Cavoukian, Ann. "Foundational Principles-Privacy by design.", 2009, published on a now defunct web site.

## How to ensure that privacy is considered from the start when designing a system?

---

- Both EU and US legislation call for implementations which respect the privacy of “data subjects”
- In both cases, “Privacy By Design” is referenced
- That means that “Privacy By Design” should be:
  - providing complete instructions.
  - directly actionable.
  - easy to implement.
- Lets check if that is the case.

# Privacy by Design: “The 7 Foundational Principles”

---


1. Proactive not Reactive; Preventative not Remedial
2. Privacy as the Default Setting
3. Privacy Embedded into Design
4. Full Functionality — Positive-Sum, not Zero-Sum
5. End-to-End Security — Full Lifecycle Protection
6. Visibility and Transparency — Keep it Open
7. Respect for User Privacy — Keep it User-Centric

Lets check if it can be argued that it fulfills all criteria ?

Open Whisper Systems


SUPPORTBLOGDEVELOPERSINSTALL

PRIVACY THAT FITS IN YOUR POCKET




“ Use anything by Open Whisper Systems.

— **Edward Snowden**, Whistleblower and privacy advocate




“ Signal is the most scalable encryption tool we have. It is free and peer reviewed. I encourage people to use it everyday.

— **Laura Poitras**, Oscar winning filmmaker and journalist



“ I am regularly impressed with the thought and care put into both the security and the usability of this app. It's my first choice for an encrypted conversation.

— **Bruce Schneier**, internationally renowned security technologist



“ After reading the code, I literally discovered a line of drool running down my face. It's really nice.

— **Matt Green**, Cryptographer, Johns Hopkins University

*“The Privacy by Design (PbD) approach is characterized by proactive rather than reactive measures. It anticipates and prevents privacy invasive events before they happen. PbD does not wait for privacy risks to materialize, nor does it offer remedies for resolving privacy infractions once they have occurred — it aims to prevent them from occurring. In short, Privacy by Design comes before-the-fact, not after.”*

**Does it apply to the Signal iOS client?**

Yes.

Explanation: Signal provides end-to-end encryption with forward security, in order to disable attacks on the ciphertext in the future.

- *“We can all be certain of one thing — the default rules! Privacy by Design seeks to deliver the maximum degree of privacy by ensuring that personal data are automatically protected in any given IT system or business practice. If an individual does nothing, their privacy still remains intact. No action is required on the part of the individual to protect their privacy — it is built into the system, by default.”*

**Does it apply to the Signal iOS client?**

Yes.

Explanation: Signal does not give users the option to not use its privacy features.

- *“Privacy by Design is embedded into the design and architecture of IT systems and business practices. It is not bolted on as an add-on, after the fact. The result is that privacy becomes an essential component of the core functionality being delivered. Privacy is integral to the system, without diminishing functionality.”*

### **Does it apply to the Signal iOS client?**

Yes.

Explanation: Signal uses a protocol specifically designed for privacy on mobile devices. In particular, its mechanism for push notifications supports encryption, and it provides encrypted group chat and encrypted voice calls. future.



- *“Privacy by Design seeks to accommodate all legitimate interests and objectives in a positive-sum “win-win” manner, not through a dated, zero-sum approach, where unnecessary trade-offs are made. Privacy by Design avoids the pretense of false dichotomies, such as privacy vs. security, demonstrating that it is possible to have both.”*

### **Does it apply to the Signal iOS client?**

Yes.

Explanation: No functionality which is part of non-private messengers is missing from signal. For instance, push notifications and group messaging.

- *“Privacy by Design, having been embedded into the system prior to the first element of information being collected, extends securely throughout the entire lifecycle of the data involved — strong security measures are essential to privacy, from start to finish. This ensures that all data are securely retained, and then securely destroyed at the end of the process, in a timely fashion. Thus, Privacy by Design ensures cradle to grave, secure lifecycle management of information, end-to-end.”*

**Does it apply to the Signal iOS client?**

Yes.

Explanation: Yes, the life cycle is quite simple, so no leakage is possible.

- *“Privacy by Design seeks to assure all stakeholders that whatever the business practice or technology involved, it is in fact, operating according to the stated promises and objectives, subject to independent verification. Its component parts and operations remain visible and transparent, to users and providers alike. Remember, trust but verify.”*

### **Does it apply to the Signal iOS client?**

Yes.

Explanation: The source code for the Signal app is available. In addition the developers engage with a community of cryptographers and security / privacy researchers whenever new bugs/threats are identified. Further, the signal developers have implemented their protocol as part of WhatsApp and FB Messenger.

- *”Above all, Privacy by Design requires architects and operators to keep the interests of the individual uppermost by offering such measures as strong privacy defaults, appropriate notice, and empowering user-friendly options. Keep it user-centric.”*

### **Does it apply to the Signal iOS client?**

Yes.

Explanation: Signal is user centric, as there is basically no value for the operators of the Signal infrastructure / back-end servers or the provision of the smart phone clients. (The business model is based on security / privacy consulting.)

## Thought experiment: Leaks of private data versus Privacy by Design

---

Is “Privacy by Design” just a label?

- Can you think of a system which follows Privacy by Design but leaks private data?

## Case studies

Based on:

Gürses, Seda, Carmela Troncoso, and Claudia Diaz. "Engineering privacy by design."  
*Computers, Privacy & Data Protection* 14.3 (2011).

- As specified by the 7 principles, PbD is very vague
- There is no documentation about how to implement the principles
- We look at two case studies which show how to implement a system following PbD in addition to the principle of data minimisation:
  1. e-Petitions
  2. Electronic Toll Pricing

- Enable users to show their support for a petition.

### Anti-Trump petition to stop UK state visit passes 1m signatures

As online campaign gathers support, Downing Street confirms PM will not be withdrawing invitation to US president





### Process:

- User provides identification, e.g. number of electronic ID.
- User authentication: signing of petition with electronic ID.
- Verification of signatures:
  - Validity of signatures is verified
  - Requirements are checked, e.g. user age, constituency
  - Duplicates are removed.

# Privacy threats for e-Petition system implementation without considering privacy

---

1. Public disclosure of signing petition
  2. Abuse of information about signing petition
- 
- In terms of LINDDUN:
  - The data storage of the e-petition system is threatened by:
    - Linkability: YES
    - Identifiability: YES
    - Non-repudiation: YES
    - Detectability: YES
    - Disclosure of information: YES

- Requirements:
  - Signature correspond to existing individuals. (e.g. not dead people)
  - Only can sign if eligible (age, constituency, ... )
  - Only sign once.
  - Number of signatures needs to be counted correctly.
- Data minimisation:
  - The goal is to count how many people support a petition, not to keep a list of people who signed a petition.

New architecture:

- **Registration authority:** hands out credentials to individuals, which encode the attributes of the individual such as age and home address.
- **E-petition web server:** allows user to prove that he is eligible to sign the petition
- **Client software of user (e.g. smart phone client):** Allows signing of petitions, and verifying that the user has been counted for petition.
  
- **User stays completely anonymous when signing petition.**
- **Which PET is enabling this?**
  - Zero knowledge proofs.
  - Allow user to proof that they have valid credentials without providing credentials to server.
  - Server can publish list of signatures on petition, which all clients can check to see if they are included.
  - But list provides repudiation, as it only lists signatures not the names of users.

- The data storage of the anonymous e-petition system is threatened by:
  - Linkability: NO
  - Identifiability: NO
  - Non-repudiation: NO
  - Detectability: NO
  - Disclosure of information: NO
- The data flow between client app and server also needs to be protected and anonymised, e.g. by using TOR.

# Electronic Toll Pricing

- Used to pay for privatised roads.
- Very popular outside of Germany, e.g. in France and Ireland.
- Some tolls require stopping and paying with coins.
- New toll systems take pictures of license plates.
- Or they use “on-board units”.



## Electronic Toll Pricing: implementation without privacy

---

- On-Board Unit (OBU) tracks car.
- Sends all details of GPS co-ordinates to the server.
- Server sends bill based on usage to user.
  
- All applicable LINDDUN privacy threats affect the centralised data storage.

- Which PET can enable privacy in this setting?
- **Cryptographic commitments.**
- Allow the OBU to send commit on having used a road at a certain time / place.
  - E.g. whenever the toll system takes a picture of the license plate.
- The OBU can prove where the car was through the commitments.



- Through the use of data minimisation, almost all threats are removed.
- In addition, the OBU needs to transmit the data only to the server when the car is at the same place.
  - E.g. always at home.

# Privacy Engineering

Based on:

- 1.) Gürses, Seda, and Jose M. del Alamo. "Privacy Engineering: Shaping an Emerging Field of Research and Practice." *IEEE Security & Privacy* 14.2 (2016): 40-46.
- 2.) Spiekermann, Sarah, and Lorrie Faith Cranor. "Engineering privacy." *IEEE Transactions on software engineering* 35.1 (2009): 67-82.

- “Privacy by Design” does not provide clear instructions on how to implement the 7 core principles
- Instead, in order to design and implement IT systems for privacy, knowledge from (at least) three areas is required:
  - Engineering
  - Policy and law
  - User experience and Human-Computer Interaction
- Privacy Engineering is an umbrella term for this knowledge

## **The three pillars of Privacy Engineering:**

- Privacy by policy
- Privacy by interaction
- Privacy by architecture

- Visible in current systems through the EULA's and privacy policy statements
- Often detached / isolated from technical implementation
  - Legal team is often not aware of potential of new PETs
- However, purely technical implementation of privacy is also not sufficient
  - Purely technical approaches might be insufficient for aligning nuanced legal policies with engineering artifacts.
- Often implemented by “Data Protection Officers” and similar roles
- Possibility to engage engineers by explaining why privacy is relevant to a system in the first place
- **Example:** We will look at the EU General Data Protection Regulation in this chapter.

- Aspects of privacy related to
  - user experience (UX) and
  - human-computer interaction (HCI)
- Privacy in the context of the social interactions enabled by the system
- Objective: the system should respect social norms
- This also involves how the system is perceived to handle:
  - ethics
  - trust
  - accountability
  - transparency
- Outside of the scope of this lecture, but an emerging research topic.

- Engineering systems with hard-coded constraints in place
- Provide guarantees against undoing these constraints.
- Involves knowledge about implementing and applying PETs.
- The majority of this lecture was about “Privacy by architecture”.

# Comparison of “privacy by policy” and “privacy by architecture”

Privacy stages	identifiability	Approach to privacy protection	Linkability of data to personal identifiers	System Characteristics
0	identified	privacy by policy (notice and choice)	linked	<ul style="list-style-type: none"><li>• unique identifiers across databases</li><li>• contact information stored with profile information</li></ul>
1	pseudonymous		linkable with reasonable & automatable effort	<ul style="list-style-type: none"><li>• no unique identifies across databases</li><li>• common attributes across databases</li><li>• contact information stored separately from profile or transaction information</li></ul>
2		privacy by architecture	not linkable with reasonable effort	<ul style="list-style-type: none"><li>• no unique identifiers across databases</li><li>• no common attributes across databases</li><li>• random identifiers</li><li>• contact information stored separately from profile or transaction information</li><li>• collection of long term person characteristics on a low level of granularity</li><li>• technically enforced deletion of profile details at regular intervals</li></ul>
3			anonymous	unlinkable



- Functional Requirements Analysis
- Data Minimisation
- Modelling Attackers, Threats and Risks
- Multilateral Security Requirements Analysis
- Implementation of the Design
- Testing of the Implementation

# Economic Benefits of Privacy Enhancing Technologies

Based on:

Economics, London. "Study on the economic benefits of privacy enhancing technologies (PETs)." *Final Report to the European Commission DG Justice, Freedom and Security*, London (2010).

# Report: Economic Benefits of Privacy Enhancing Technologies

---

- Performed by “London Economics” in 2010
- Paid for by the European Commission, DG Justice, Freedom and Security

## **Strong empirical basis:**

- Based on responses of 1337 business in 12 EU member states
- Includes details of 20 exploratory case studies and 6 commercial case studies

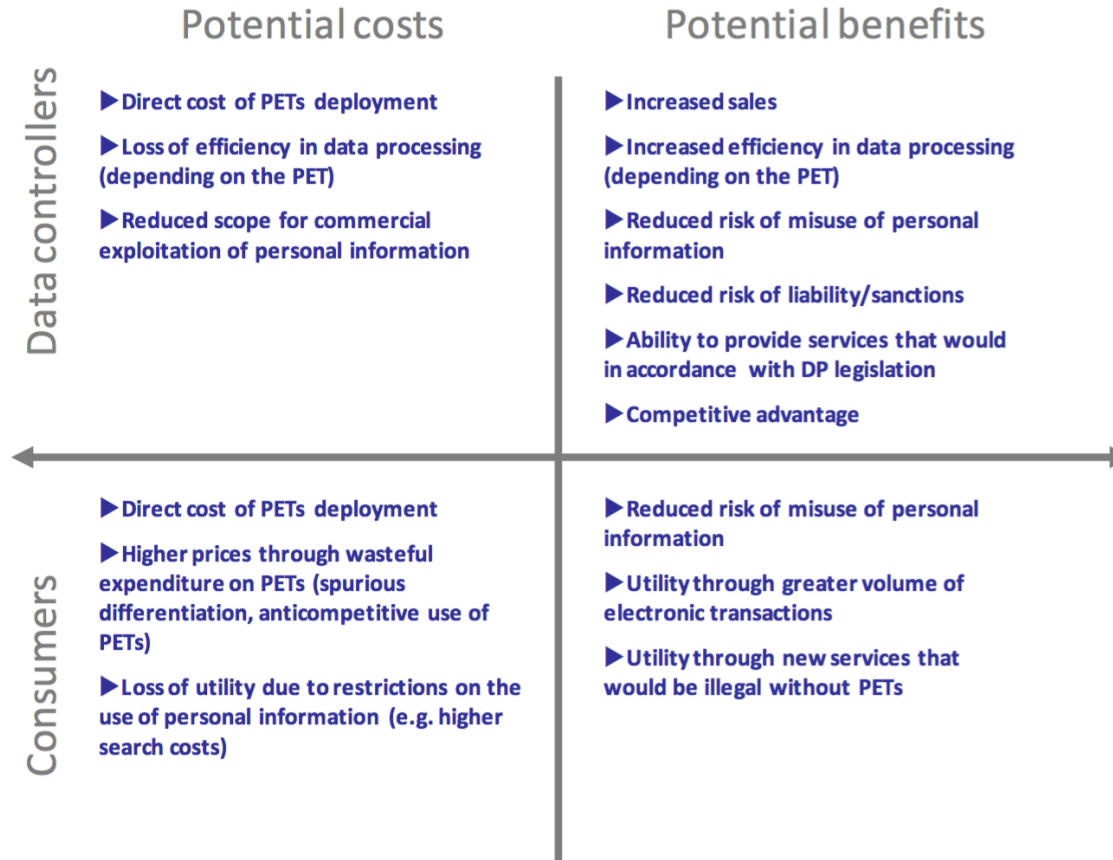
## Why are businesses deploying PETs?

---

- Most important goals for using PETs by businesses (at time of publication)
  - Data protection
  - Anonymisation
  - Data minimisation
  - Fulfilment of consent requirements
- Decision factors for deployment of PETs by businesses:
  - Required changes in business model for some PETs
  - Possible trade-off between ability to use personal data and benefits of PETs
- The net economic benefit of PETs deployment has to be assessed on a case-by-case basis

- Almost no evidence that demand by individuals driving PETs deployment
- Demand depends on:
  - risk aversion of individual user
  - risk of data loss / privacy invasion
  - efficacy of PETs reducing the risk
- However, in practice individuals are faced with:
  - Uncertainties about risk of personal data disclosure
  - Lack of knowledge about PETs
  - Behavioural biases preventing individuals from acting in accordance with their stated preference for greater privacy
- Result: widespread indifference of consumers

# Potential costs and benefits of PETs deployment



### **Potential benefits from using PETs:**

- Fear of data loss
- Competitive advantage

### **Potential benefits from NOT using PETs:**

- Efficiency of electronic processes with consumer data
- Personalising goods and services
- Exploiting personal data for new goods and services

### **Slow deployment due to:**

- unawareness of consumers
- low maturity of PETs
- requirement for training employees

Market failures / imperfections which may hold back PETs innovation:

- Asymmetric information
  - Consumers have less information about risk of privacy exposure compared to data controllers
- Externalities
- Coordination failures
- Lack of information sharing about privacy risks



# The role of legislation for incentivizing the use of PETs

---

- Data protection legislation supports data controllers by setting the framework
- Legislation has to reflect privacy demands of society at large
- Enforcement of rules helps data controllers by penalising non-compliant behaviour
  - This prevents gaining a competitive advantage by not protecting privacy!
  - But also requires strong effort of rules enforcements by governments!
- Public sector endorsements of PETs and certification programmes can also help.
- Innovation in PETs may be lacking due to “market failures” such as coordination problems.
  - These require government intervention and investment as well, e.g. through research funding.

# Upcoming EU Legislation about Privacy

**General Data Protection Regulation (GDPR),  
known as “Datenschutz-Grundverordnung” in Germany**

# General Data Protection Regulation (GDPR)

---

- Starts being valid from 25<sup>th</sup> May 2018.
- Two year transition period started on 27<sup>th</sup> April 2016.
- Replaces the “Data Protection Directive” from 1995.
- Type of law:
  - Hybrid between regulation and directive
  - Is immediately valid for all countries (like a regulation)
  - Parts of it can be customized by every country (like a directive)
- German name: “Datenschutz-Grundverordnung”

- Goals:
  - Give citizens back the control of their personal data
  - Simplify regulatory environment for international business
- Scope:
  - Applies if data control OR data processor OR data subject is based in EU.
- Single set of rules:
  - Every member state has independent Supervisory Authority (SA)
  - SA's of different countries work together.
  - The SA of a companies main establishment is responsible.

- Automated individual decision making based on purely algorithmic basis can be contested.
- “Privacy by Design” is required.
- Privacy settings must be high by default.
- Data breaches have to be reported within 72 hours to the SA of a company.
- Valid consent must be explicit for data collected and for the purposed of using the data.
  - Data controllers must be able to prove consent (opt-in)
  - Consent can also be withdrawn at any time.

- Right to erasure
  - Replaces the “right to be forgotten”
  - Can only be invoked on grounds of non-compliance
- Data portability
  - Person must be able to transfer their personal data from one electronic processing system to and into another
  - Using a structured and commonly used electronic format.
- So far it is not clear how the GDPR applies to data processors outside of the EU, for instance in the US.

- Warning in writing in cases of first and non-intentional non-compliance
  - Regular periodic data protection audits
  - A fine up to 10,000,000 EUR or up to 2% of the annual worldwide turnover of the preceding financial year in case of an enterprise, whichever is greater
  - The same kind of fine, but up to 20 million or 4%
- 
- These are much higher fines than before.

### **Open question:**

- Is this enough financial incentive for companies to deploy PETs ?

- “Privacy by Design” (PbD) is referenced in current and future regulation (EU and US)
  - However, the foundation principles of PbD are very vague
  - Implementing PbD is done on a case by case basis
- Privacy Engineering is a different way to implement privacy conscious systems:
- Privacy Engineering has three pillars:
  - Privacy by policy
  - Privacy by architecture
  - Privacy by interaction
- Economic incentives for implementing PETs are largely missing
  - Consumers are not willing to pay for PETs
  - Companies do not have pressure to implement PETs
- Economic incentives have to be provided by legislation, such as the upcoming GDPR



## Other information

## Questions about lecture content

---

- The last exercise session is on Thursday, 9.2.2017.
- If you have any remaining questions about the content of the lecture, please send them to me until Wednesday evening.
- I will go through the questions on Thursday.

## Currently open thesis topics

---

- Developing a decentralised personalisation approach for secure peer to peer environments.
  - Developing a personalisation approach using distributed processing on the blockchain with Ethereum.
  - Developing of benchmarking suite for libraries which enable algorithms to use encrypted user data.
- 
- More thesis topics in the area of PETs4DS are available upon request.