Data Protection By Design

An Overview of Obligations and Technical Approaches

Seda Gürses
seda@esat.kuleuven.be
COSIC, University of Leuven

22. February 2017
SecAppDev

objective

fear of a GDPR planet!!?

not a course on how to be compliant!

objective

GDPR is an invitation to develop a vision

assess implications of your system on people's rights and freedoms

bring that assessment to system design

bring that assessment to legal requirements

confront legal reality with technical research

GDPR

GDPR

many ways of summarizing!

GDPR

lawfulness:
legal ground for processing (e.g., consent, contract, balancing)

purpose limitation & data minimization:
processing only for limited, specific, explicit purpose

sensitive data:
strict rules for personal data revealing sensitive attributes

transparency:
with respect to processing and purposes towards data subject

data subject access rights:
access, correction, object, erasure, portability, profiling

GDPR

storage limitation:
kept in identifiable form for no longer than necessary

accuracy:
accurate and kept up to date

security:
processed in a manner that ensures appropriate integrity and confidentiality

accountability:
ability to demonstrate compliance, risk management and DPbD

somebody said the magic words

Article 25:
Data Protection by Design and by Default

Taking into account the state of the art, the cost of implementation and the nature, scope, context and purposes of processing as well as the risks of varying likelihood and severity for rights and freedoms of natural persons posed by the processing, the controller shall, both at the time of the determination of the means for processing and at the time of the processing itself, implement appropriate technical and organisational measures, such as pseudonymisation, which are designed **to implement data-protection principles, such as data minimisation**, in an effective manner and to integrate the necessary safeguards into the processing in order to meet the requirements of this Regulation and protect the rights of data subjects.

diversity of issues

do you have policies on your data flows? how long does data remain in your system?

depending on the legal ground that you have, you may or may not need consent. That completely determines how you design your system (free and informed consent is not a trivial task).

do you have mechanisms for deletion/stopping processing/profiling? do you have conflicting requirements for data deletion and data retention?

are there situations in which your data minimization and unsinkability requirements are contradicting your data subject access rights requirements?

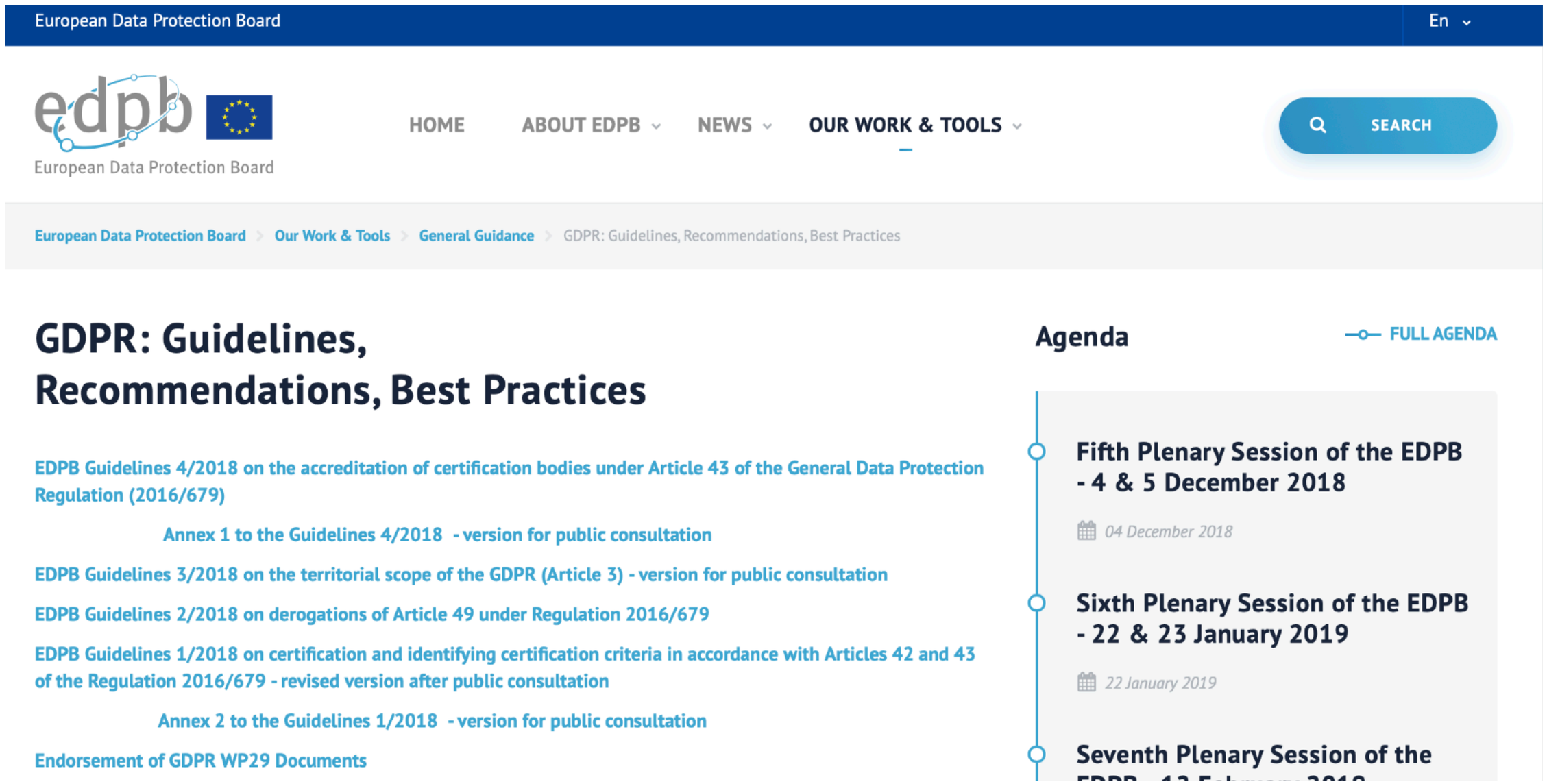
why would a developer be asked to do DPbyD?

somebody walked into your office and asked you to make it all compliant!!?

mitigate a specific risk: the system you are developing has clearly defined privacy risks, somebody ask you to mitigate those risks?

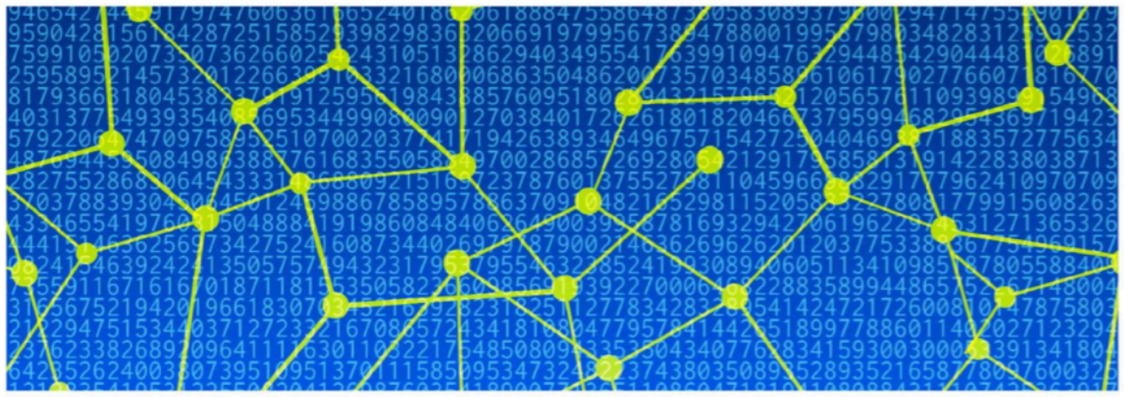
you may be tasked to implement a specific legal requirement: e.g., informed consent, data portability

European Data Protection Board
https://edpb.europa.eu/edpb_en



European Data Protection Supervisor

https://edps.europa.eu



EUROPEAN DATA PROTECTION SUPERVISOR

Opinion 5/2018

Preliminary Opinion on privacy by design

ENISA
https://www.enisa.europa.eu/publications

Privacy and Data Protection by Design

This report contributes to bridging the gap between the legal framework and the available technolog-ical implementation measures by providing an inventory of existing approaches, privacy design strat-egies, and technical building blocks of various degrees of maturity from research and development. Starting from the privacy principles of the legislation, important elements are presented as a first step towards a design process for privacy-friendly systems and services.

Published January 12, 2015
Language English

Privacy by design in big data

The extensive collection and further processing of personal information in the context of big data analytics has given rise to serious privacy concerns, especially relating to wide scale electronic surveillance, profiling, and disclosure of private data. In order to allow for all the benefits of analytics without invading individuals' private sphere, it is of utmost importance to draw the limits of big data processing and integrate the appropriate data protection safeguards in the core of the analytics value chain. ENISA, with the current report, aims at supporting this approach, taking the position that, with respect to the underlying legal obligations, the challenges of technology (for big data) should be addressed by the opportunities of technology (for privacy).

Published December 17, 2015
Language English

Norwegian Data Protection Authority
https://www.datatilsynet.no

Datatilsynet

Guide

Software development with Data Protection by Design and by Default

The Norwegian Data Protection Authority has developed these guidelines to help organisations understand and comply with the requirement of data protection by design and by default in article 25 of the General Data Protection Regulation. We have cooperated with security professionals and software developers in public and private sector among others.

Print guide

Unabhängiges Landeszentrum für Datenschutz
https://www.datenschutzzentrum.de/sdm/

ULD

Unabhängiges Landeszentrum für Datenschutz
Schleswig-Holstein

Suche

Drucken | Impressum | Datenschutzerklärung

ULD
- Wir über uns
- Meldungen an das ULD

Themen
- Privatwirtschaft
- Medizin und Soziales
- Öffentliche Sicherheit und Justiz
- Öffentliche Verwaltung
- Informationsfreiheit

Das Standard-Datenschutzmodell (SDM)

» Standard-Datenschutzmodell

Als "Standard-Datenschutzmodell" (SDM) bezeichnen die deutschen Datenschutzaufsichtsbehörden eine Methode, mit der für den Bereich des operativen Datenschutzes sichergestellt ist, dass eine einheitliche Datenschutz-Beratungs- und Prüfpraxis in Bezug insbesondere zu den technisch-organisatorischen Maßnahmen der DS-GVO erreicht werden kann.

- **SDM-Methodik-Handbuch, V1.1 (Deutsch)**
- **SDM-Methodology, V1.0 (English)**

- Vorangegangene Versionen

Federal Trade Commission
https://www.ftc.gov/tips-advice/business-center/privacy-and-security/tech

Mobile Health App Developers: FTC Best Practices

TAGS: Advertising and Marketing | Health Claims | Privacy and Security | Consumer Privacy | Data Security | Tech | Health Care

When developing a health app, sound privacy and security practices are key to consumer confidence. Here are some best practices to help you build privacy and security into your app. These practices also can help you comply with the FTC Act.

Start with Security: A Guide for Business offers tips for any business wanting to implement sound data security. For health app developers, here's tailored advice and additional questions to ask.

- Minimize data.
- Limit access and permissions.
- Keep authentication in mind.
- Consider the mobile ecosystem.
- Implement security by design.
- Don't reinvent the wheel.
- Innovate how you communicate with users.
- Don't forget about other applicable laws.

National Institute of Standards and Technology (NIST)
https://www.nist.gov/itl/applied-cybersecurity/privacy-engineering

PRIVACY ENGINEERING PROGRAM

About +
Collaboration Space +
Resources
Events
Get Involved

Resources

NIST Internal Report (NISTIR) 8062: *An Introduction to Privacy Engineering and Risk Management in Federal Systems*

NISTIR 8062 introduces the concept of applying systems engineering practices to privacy and provides a new model for conducting privacy risk assessments on federal systems.

PDF

More content coming soon!

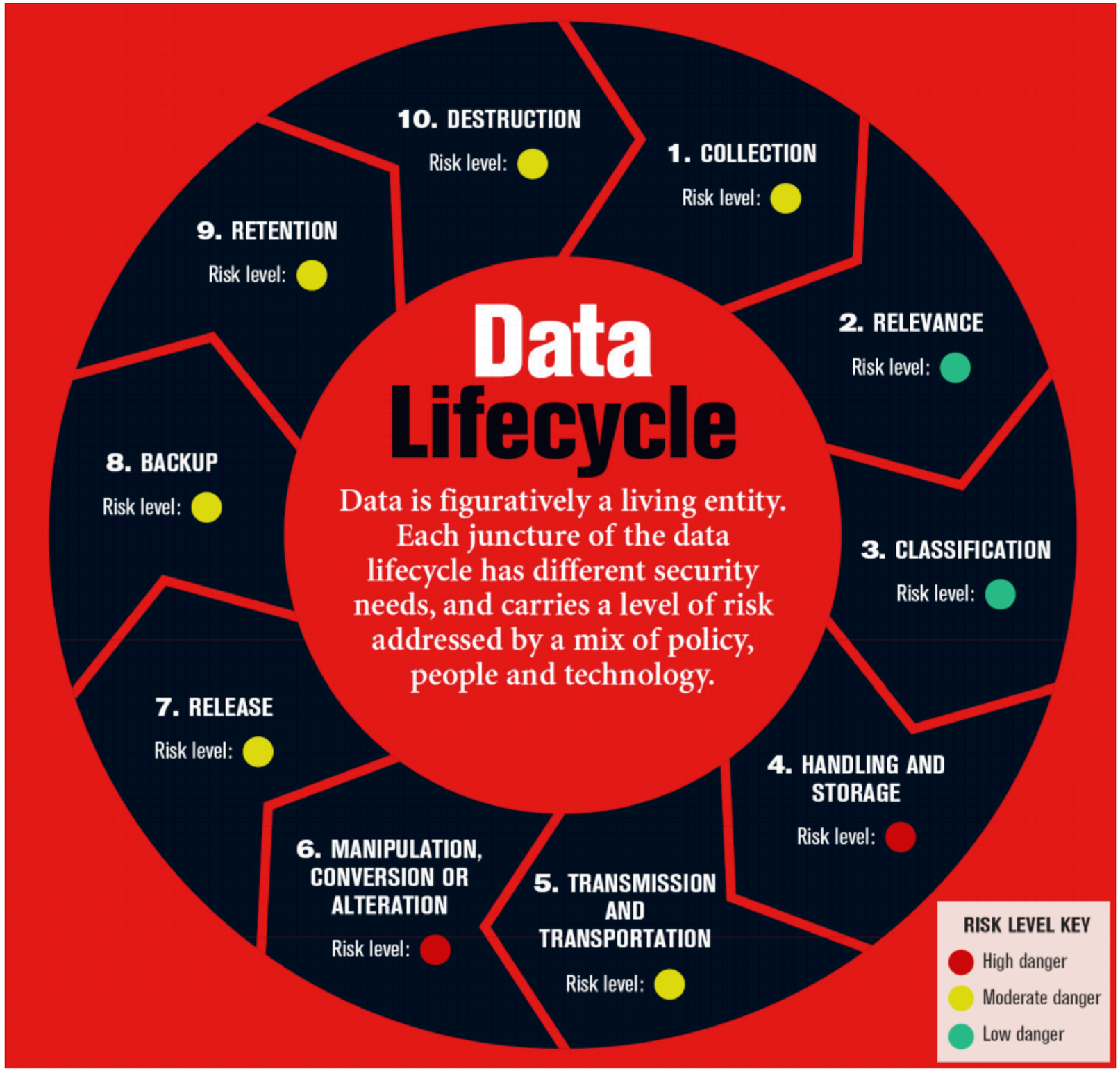
Data Protection as a Service

Accountability Life Cycle Activities

The table below lists the phased activities that support the Accountability Life Cycle.

Phase	Activity
PHASE I: Prepare	Activity A: Obtain the buy-in of key business stakeholders Activity B: Establish your GDPR readiness program team Activity C: Identify and assess relevant business functions Activity D: Identify and assess in-scope Third Party Processing activities Activity E: Establish a central Personal Data register Activity F: Distribute updated Data Protection policies and Privacy Notices Activity G: Educate internal Personal Data Handlers and external Data Processors
PHASE II: Operate	Activity H: Disseminate and maintain external Privacy Notices Activity I: Justify and record lawful Processing mechanisms Activity J: Process and record Data Subject rights requests Activity K: Validate and record Third Country data transfers Activity L: Report and manage Personal Data Breach incidents
PHASE III: Maintain	Activity M: Evidence understanding of Data Protection policies

Data Protection Laws are Data Centric



http://cdn.ttgtmedia.com/informationsecurity/images/vol4iss7/ism_v4i7_f4_DataLifecycle.pdf

privacy engineering

the field of research and practice that designs, implements, adapts and evaluates theories, methods, techniques, and tools to systematically capture and address privacy issues when developing socio-technical systems.

Gurses and Del Alamo, Privacy Engineering: Shaping an emerging field of research and practice, IEEE S&P, 2016.
http://vous-etes-ici.net/wp-content/uploads/2016/04/IEEESP_Pre.pdf

privacy theory

methods

techniques

tools

socio-technical systems

standalone privacy technology

Tor/PreTP

privacy enhancement of system or function

privacy policy languages

research into privacy violations

web census

methods:
approaches for systematically capturing and addressing privacy issues during information system development, management and maintenance

IEEE TRANSACTIONS ON SOFTWARE ENGINEERING, VOL. 35, NO. 1, JANUARY/FEBRUARY 2009

Engineering Privacy

Sarah Spiekermann and Lorrie Faith Cranor, *Senior Member, IEEE*

Privacy stages	identifiability	Approach to privacy protection	Linkability of data to personal identifiers	System Characteristics
0	identified	privacy by policy (notice and choice)	linked	• unique identifiers across databases • contact information stored with profile information
1	pseudonymous		linkable with reasonable & automatable effort	• no unique identifies across databases • common attributes across databases • contact information stored separately from profile or transaction information
2		privacy by architecture	not linkable with reasonable effort	• no unique identifiers across databases • no common attributes across databases • random identifiers • contact information stored separately from profile or transaction information • collection of long term person characteristics on a low level of granularity • technically enforced deletion of profile details at regular intervals
3	anonymous		unlinkable	• no collection of contact information • no collection of long term person characteristics • *k*-anonymity with large value of *k*

PRIPARE

PRIPARE

PReparing Industry to Privacy-by-design by supporting its Application in REsearch

Home | Partners | Factsheet | Deliverables | Events | Outreach | Resources | Advisory Board | Contact | News

Home

PRIPARE Handbook: Methodological Tools to Implement Privacy and Foster Compliance with the GDPR

http://pripareproject.eu

PRIPARE

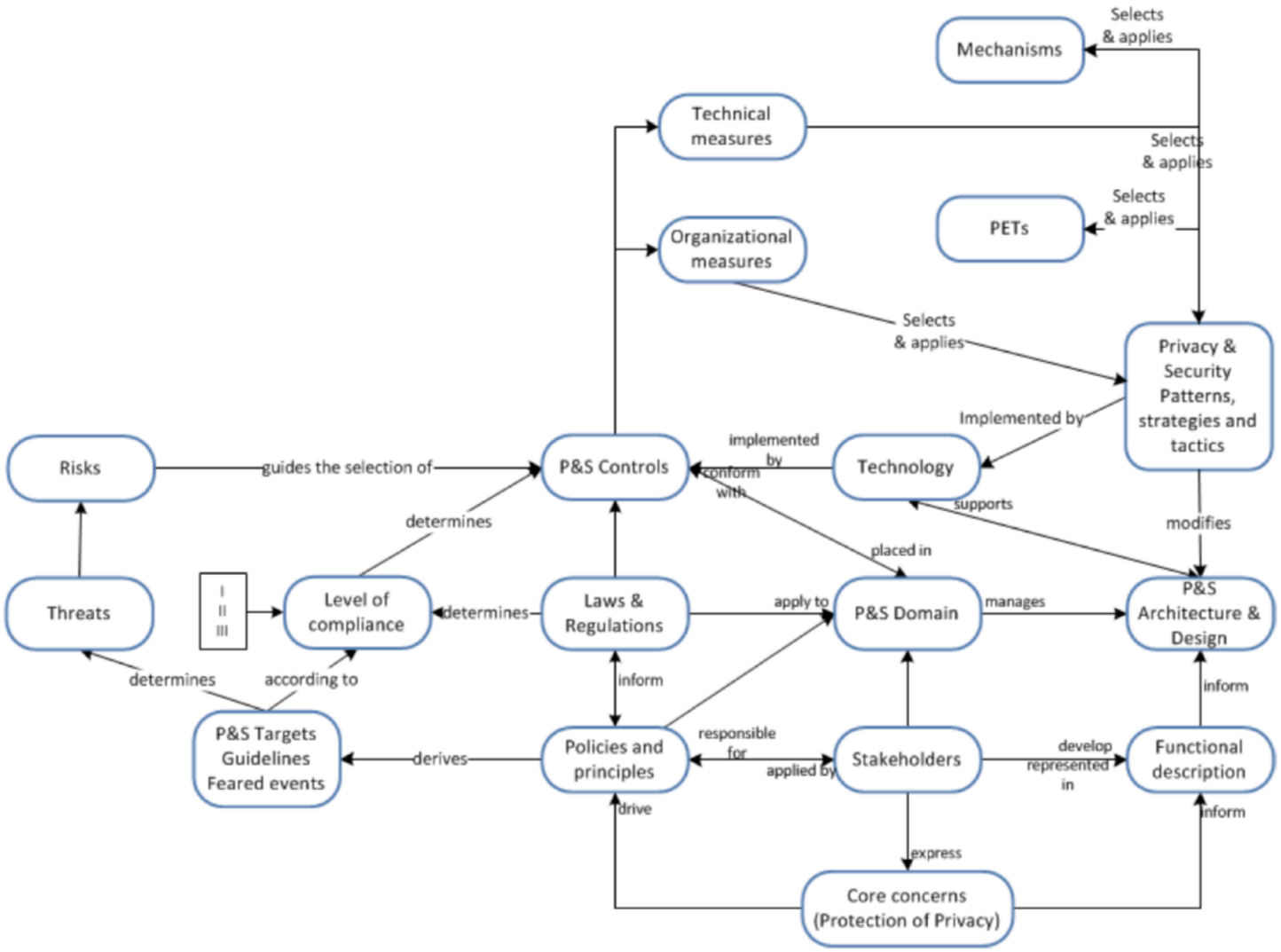
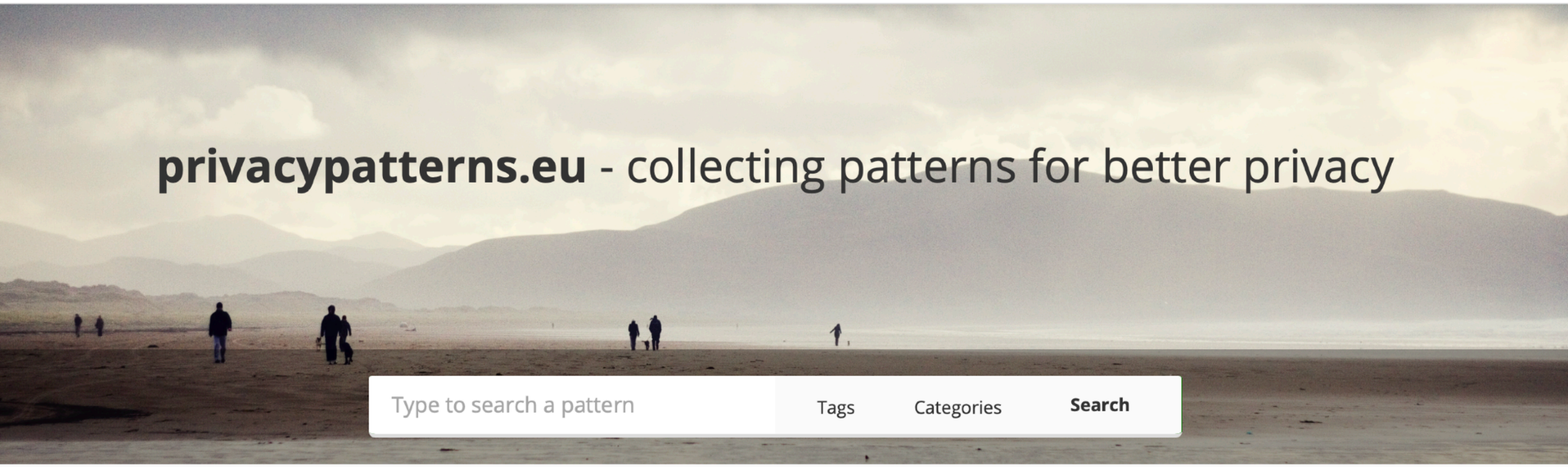


Figure 1: PRIPARE's methodology reference model

Prepare Methodology Handbook: http://pripareproject.eu/wp-content/uploads/2013/11/PRIPARE-Methodology-Handbook-Final-Feb-24-2016.pdf

privacypatterns.eu

Home About Log In / Sign Up



Anonymity Set

In a system with different users we have the problem that we can often distinguish between them. This enables location tracking, analyzing the behaviour of the users or other privacy-infringing practices.

1 Comments 3 Upvotes 0 Downvotes

Strip Metadata

There are multiple types of metadata. There is user-generated metadata data like exif-data. Exif is a format for storing metadata in pictures. There is also metadata which exists to ensure the functionality of some services like headers in email or http, or timestamps in files. Often the user is not aware of this additional data that is attached to the content. When publishing data, this could lead to a

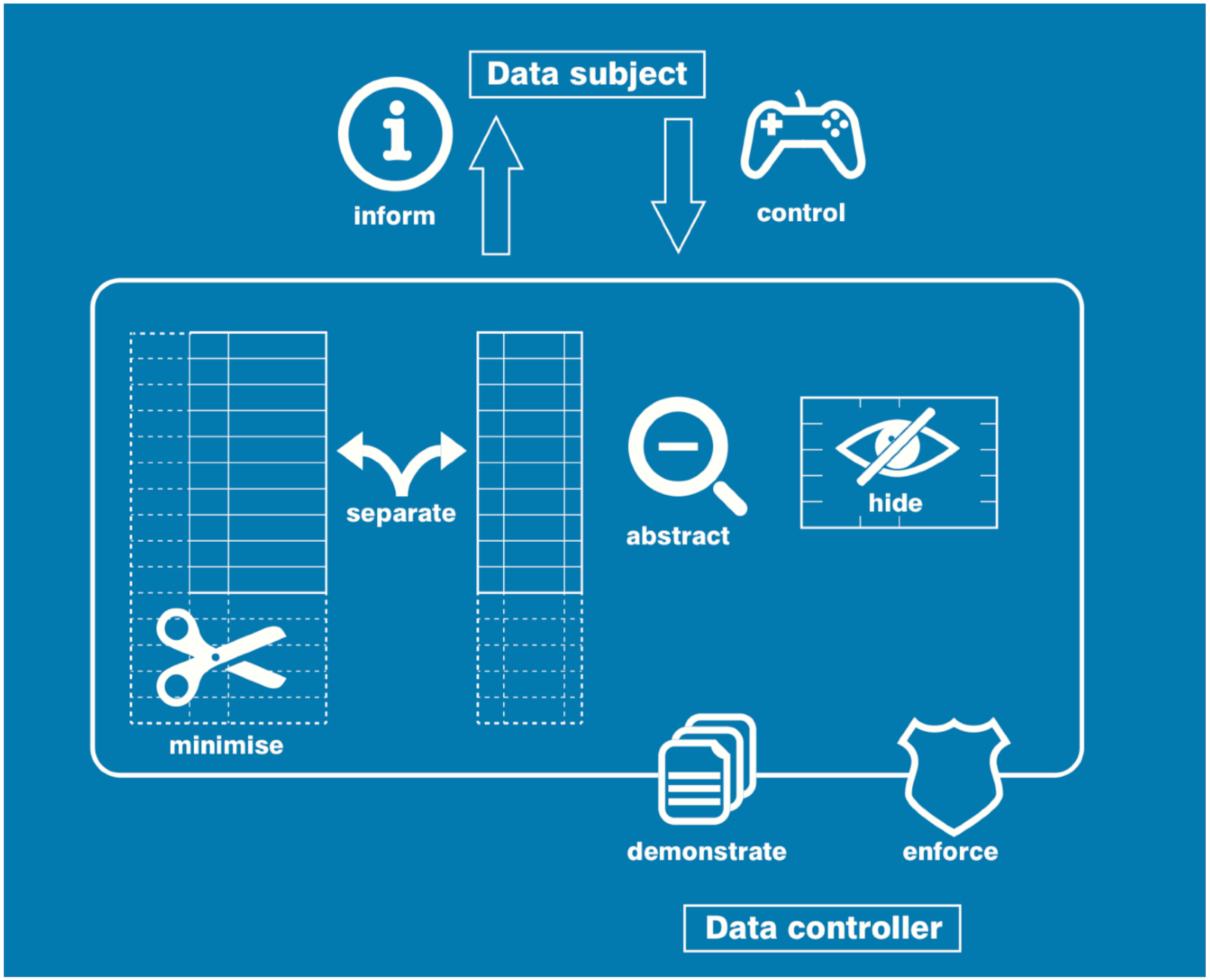
3 Comments 4 Upvotes 0 Downvotes

Pseudonymous Identity

Many kinds of sensitive informations are released through web interactions, email, data sharing or location-based systems, which can contain the name of a user or header information in packets. Another problem could be to interact anonymously in a forum. However too much interaction in a forum with an anonymous identity can be dangerous in the sense that the relation between original

3 Comments 0 Upvotes 0 Downvotes

Privacy Design Strategies (Hoepman et al.)



http://www.cs.ru.nl/~jhh/publications/pds-booklet.pdf

7 Inform

Inform data subjects about the processing of their personal data in a timely and adequate manner.

Transparency about which personal data is being processed, how they are processed and for which purpose, is an essential prerequisite for better privacy protection. It allows users to take informed decisions about using a system and agreeing to the processing of their personal data (see also the control strategy). Moreover it allows society at large to verify whether organisations are processing our personal data responsibly. ("Sunlight is said to be the best of disinfectants.")

7.1 Tactics

Transparency can be achieved following these tactics.

Supply Supply information about *which* personal data is processed, *how* they are processed, and *why*. Clearly specify how long personal data is retained, and how it is deleted. List all third parties with which you share this personal data, be clear about the conditions that cover each third party data exchange, and specify how these conditions are enforced. Put a link to your privacy policy on your homepage, and in your app. Clearly indicate how people can get in touch with questions about their privacy.

Explain Explain which personal data you process, and why. Argue why this is necessary. Do this in a clear and easy to understand manner, even for a layperson. Target this information to different user groups: novices, experts, the authorities. Consider using a layered approach: first provide an overview, and provide links to more detailed information.

Notify Notify users (in real time) the moment you process their personal data, share it with third parties, or as soon as you become aware of a data leak. Prepare clear procedures for this. Make notifications short but informative. Be sure not to notify too often. Allow users to control for which events they wish to receive a notification.

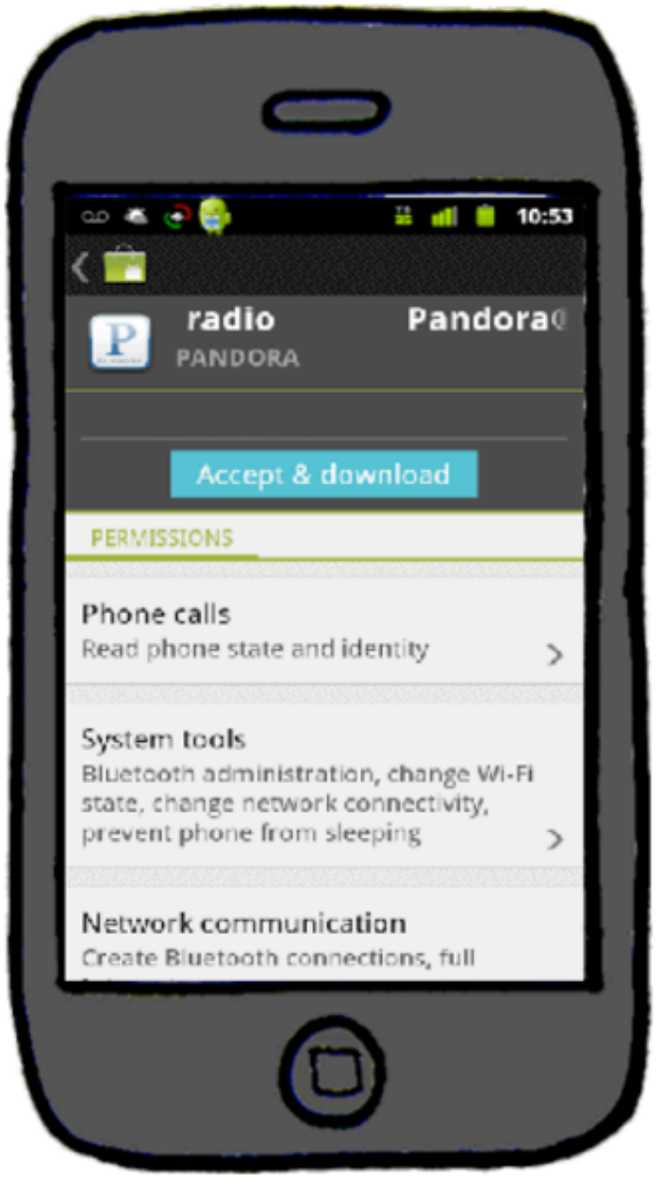
Android Permissions: User attention, comprehension, and Behavior (Felt et al., 2012)

Permission	*n*	Options	Responses	
`INTERNET` Category: Network communication Label: Full Internet access	109	Send information to the application's server	45	41.3%
		Load advertisements	30	27.5%
		None of these	16	14.7%
		Read your text messages	13	11.9%
		Read your list of phone contacts	11	10.1%
		I don't know	36	33.0%
`READ_PHONE_STATE` Category: Phone calls Label: Read phone state and identity	85	Read your phone number	41	47.7%
		See who you have called	37	43.0%
		Track you across applications	20	23.3%
		Load advertisements	11	12.8%
		None of these	10	11.6%
		I don't know	15	17.4%
`CALL_PHONE` Category: Services that cost you money Label: Directly call phone numbers	83	Place phone calls	30	35.3%
		Charge purchases to your credit card	27	31.8%
		None of these	16	18.8%
		See who you have made calls to	14	16.5%
		Send text messages	11	12.9%
		I don't know	16	18.8%
`WRITE_EXTERNAL_STORAGE` Category: Storage Label: Modify/delete SD card contents	92	Read other applications' files on the SD card	41	44.6%
		Change other applications' files on the SD card	39	42.4%
		None of these	16	17.4%
		See who you have made phone calls to	15	16.3%
		Send text messages	11	12.0%
		I don't know	15	16.3%
`WAKE_LOCK` Category: System tools Label: Prevent phone from sleeping	81	Keep your phone's screen on all the time	49	60.5%
		Drain your phone's battery	37	45.7%
		None of these	7	8.6%
		Send text messages	4	4.9%
		Delete your list of contacts	4	4.9%
		I don't know	13	16.0%
`CHANGE_NETWORK_STATE`		Turn your WiFi on or off	36	52.9%
		Send information to the application's server	13	19.1%

Android Permissions: User attention, comprehension, and Behavior (Felt et al., 2012)

Permission	*n*	Options	Responses	
`INTERNET` Category: Network communication Label: Full Internet access	109	✔ Send information to the application's server ✔ Load advertisements ✗ None of these ✗ Read your text messages ✗ Read your list of phone contacts *I don't know*	45 30 16 13 11 36	41.3% 27.5% 14.7% 11.9% 10.1% 33.0%
`READ_PHONE_STATE` Category: Phone calls Label: Read phone state and identity	85	✔ Read your phone number ✗ See who you have called ✔ Track you across applications ✗ Load advertisements ✗ None of these *I don't know*	41 37 20 11 10 15	47.7% 43.0% 23.3% 12.8% 11.6% 17.4%
`CALL_PHONE` Category: Services that cost you money Label: Directly call phone numbers	83	✔ Place phone calls ✗ Charge purchases to your credit card ✗ None of these ✗ See who you have made calls to ✗ Send text messages *I don't know*	30 27 16 14 11 16	35.3% 31.8% 18.8% 16.5% 12.9% 18.8%
`WRITE_EXTERNAL_STORAGE` Category: Storage Label: Modify/delete SD card contents	92	✔ Read other applications' files on the SD card ✔ Change other applications' files on the SD card ✗ None of these ✗ See who you have made phone calls to ✗ Send text messages *I don't know*	41 39 16 15 11 15	44.6% 42.4% 17.4% 16.3% 12.0% 16.3%
`WAKE_LOCK` Category: System tools Label: Prevent phone from sleeping	81	✔ Keep your phone's screen on all the time ✔ Drain your phone's battery ✗ None of these ✗ Send text messages ✗ Delete your list of contacts *I don't know*	49 37 7 4 4 13	60.5% 45.7% 8.6% 4.9% 4.9% 16.0%
`CHANGE_NETWORK_STATE`		✔ Turn your WiFi on or off ✗ Send information to the application's server	36 13	52.9% 19.1%

How to ask for permission? (Felt et al., 2012)



PRO

Applicable to all permissions, even advance approval

CON

Interruptive, looks like EULAs, habit-forming

INSTALL-TIME WARNINGS

How to ask for permission? (Felt et al., 2012)

PRO

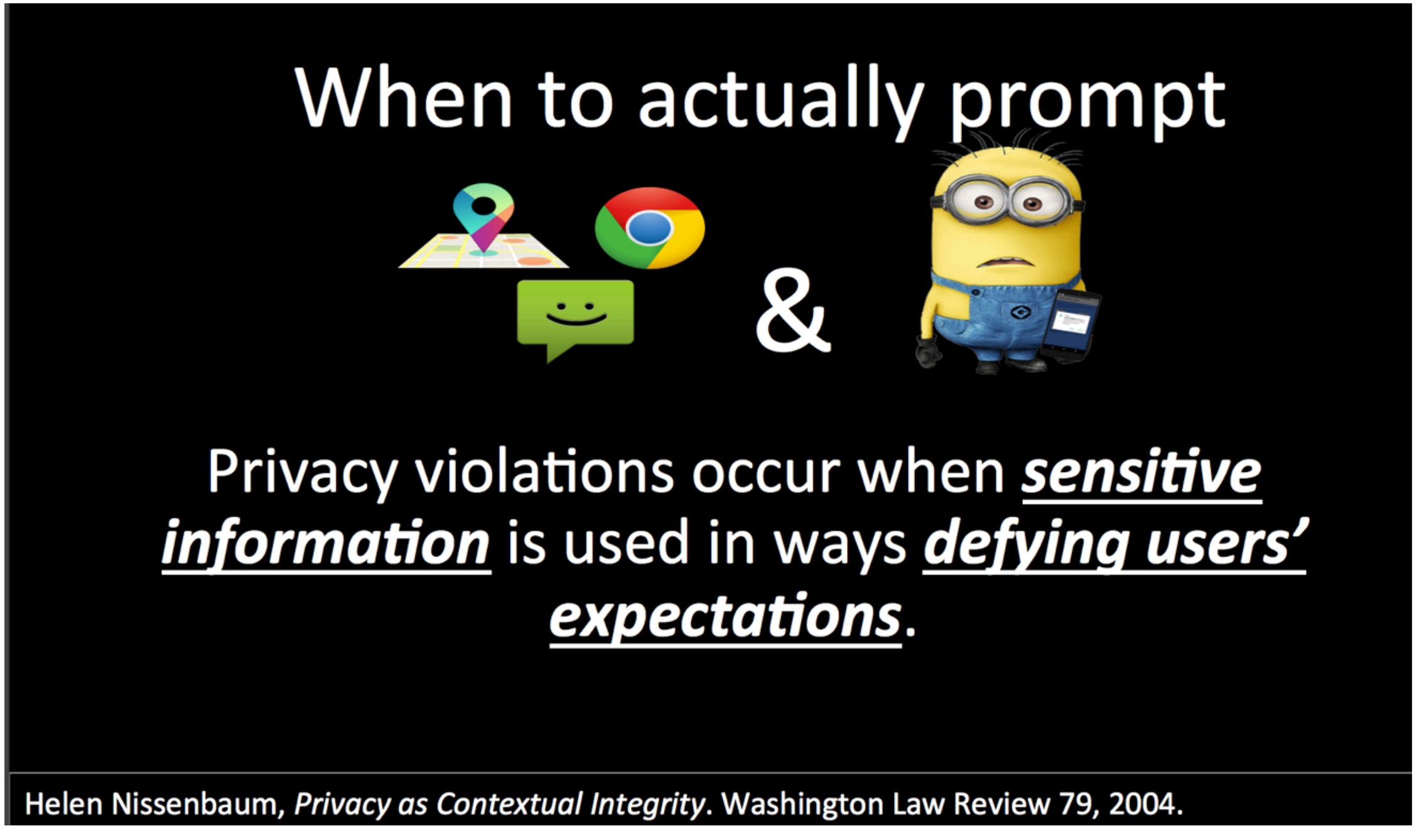
Applicable to almost all permissions

CON

Interruptive, habit-forming, not useful for advance approval

RUNTIME CONSENT DIALOGS

Android Permissions Remystified: A field study of Contextual Integrity (Wijesekera et al. 2015)



Android Permissions Remystified: A field study of Contextual Integrity (Wijesekera et al.)

The experiment

36 Android smartphone users

6,048 hours of real-world use

27 million permission requests

Android Permissions Remystified: A field study of Contextual Integrity (Wijesekera et al.)

Users want a choice

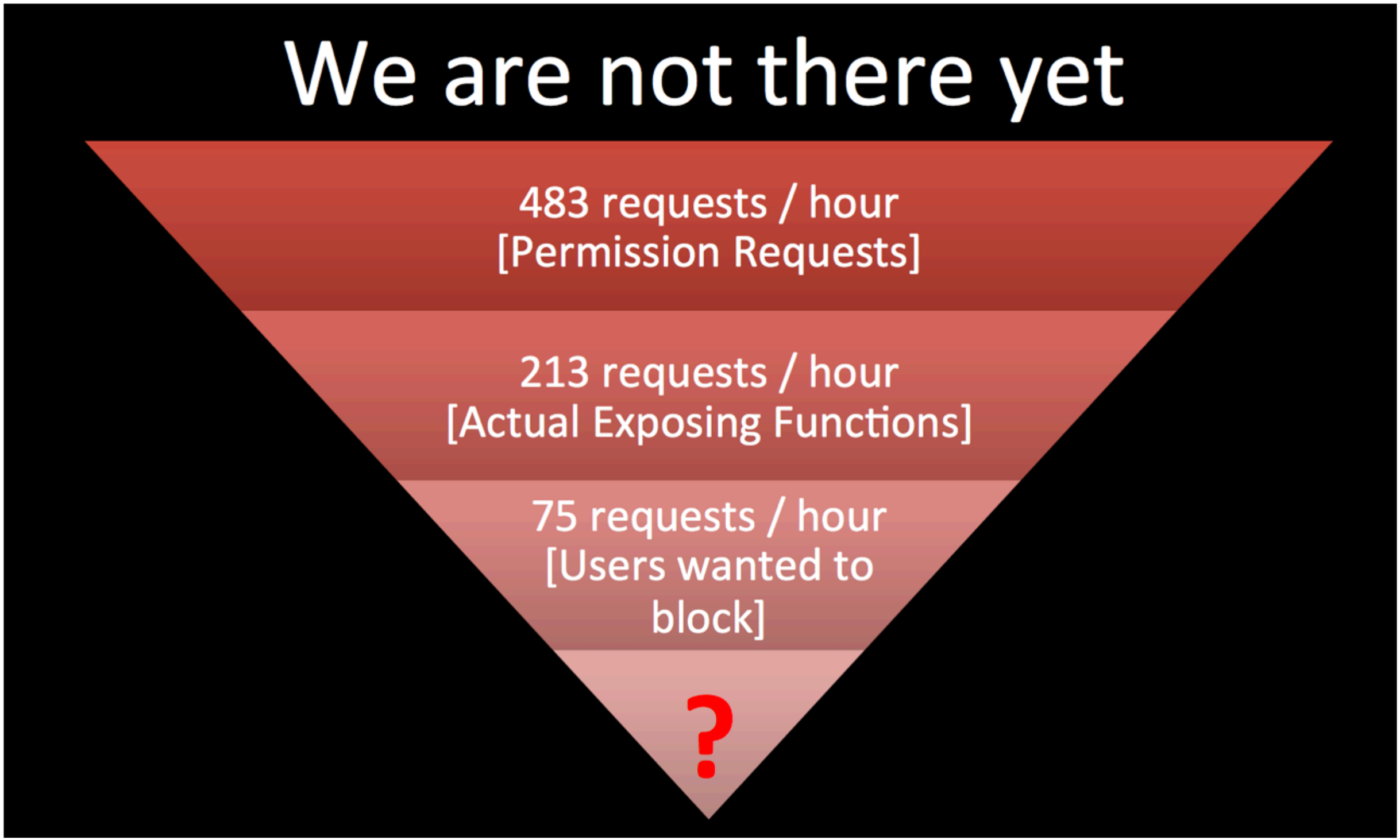
80% of users

would block at least one permission request.

35% of all requests

were deemed inappropriate.

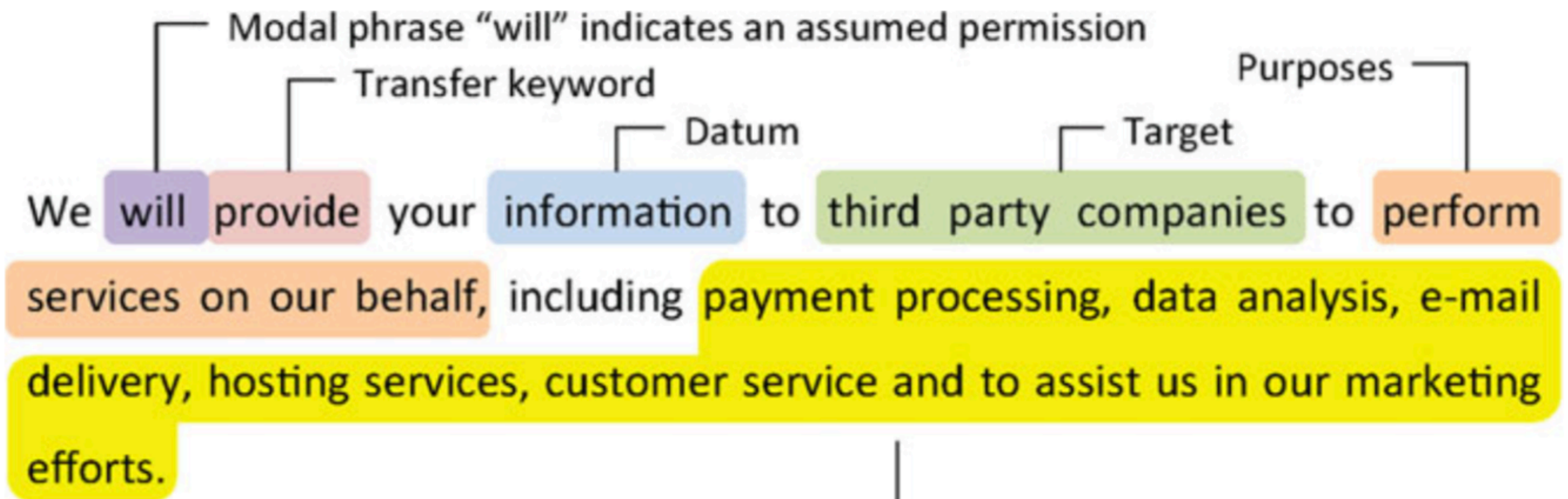
Android Permissions Remystified: A field study of Contextual Integrity (Wijesekera et al.)



techniques:
procedures, possibly with a prescribed language or notation, to accomplish privacy-engineering tasks or activities

Eddy, a formal language for specifying and analyzing data flow specifications for conflicting privacy requirements

Travis D. Breaux · Hanan Hibshi · Ashwini Rao



LINDDUN (Wuyts, Scandariato, Joosen)

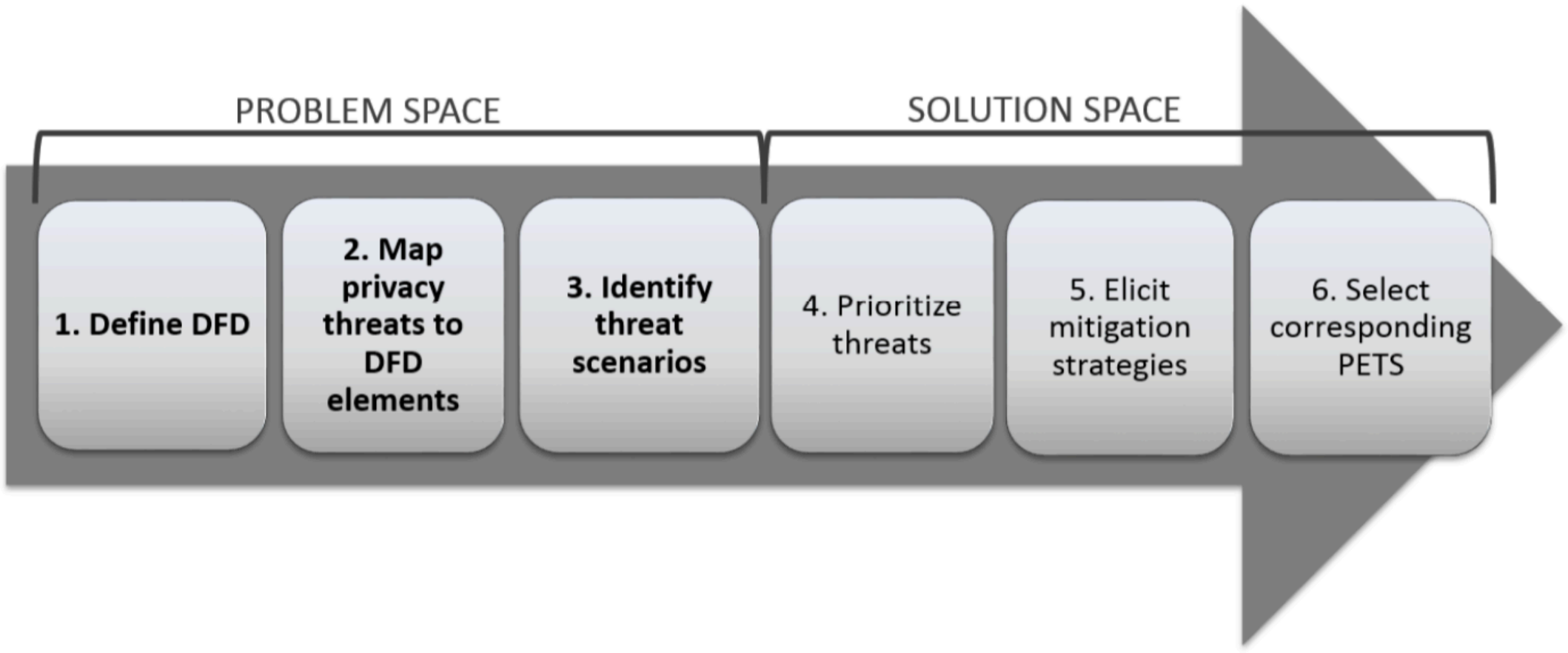
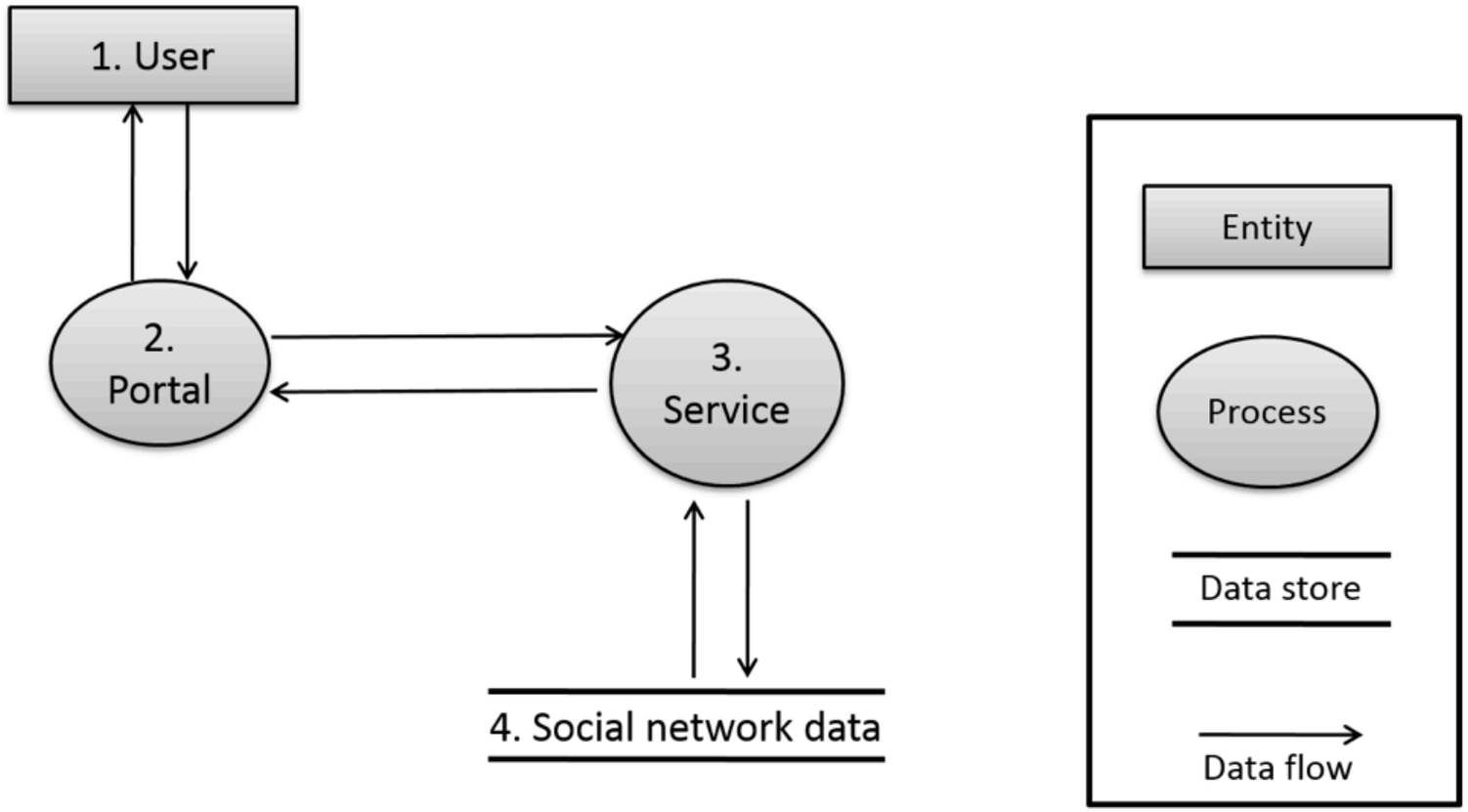


FIGURE 1: LINDDUN METHODOLOGY STEPS

https://distrinet.cs.kuleuven.be/software/linddun/index.php

LINDDUN (Wuyts, Scandariato, Joosen)



The data flow diagram (DFD) of the Social Network application

Linkability (L) occurs when one can sufficiently distinguish whether 2 items of interest (IOI, such as requests from a user) are related

Identifiability (I) occurs when it is possible to pinpoint the identity of a subject (e.g., a user)

Non-repudiation (Nr) occurs when it is possible to gather evidence so that a party cannot deny having performed an action

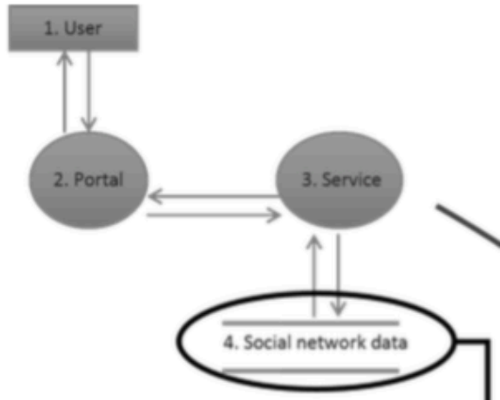
Detectability (D) occurs when one can sufficiently distinguish whether an IOI exists, e.g., in a system

Disclosure of information (Di) is the exposure of information to individuals who are not supposed to have access to it

Unawareness (U) occurs when the user is unaware of the information he is supplying to the system and the consequences of his/her act of sharing

Non-compliance (Nc) occurs when the system is not compliant with the (data protection) legislation, its advertised policies and the existing user consents

1. Model DFD



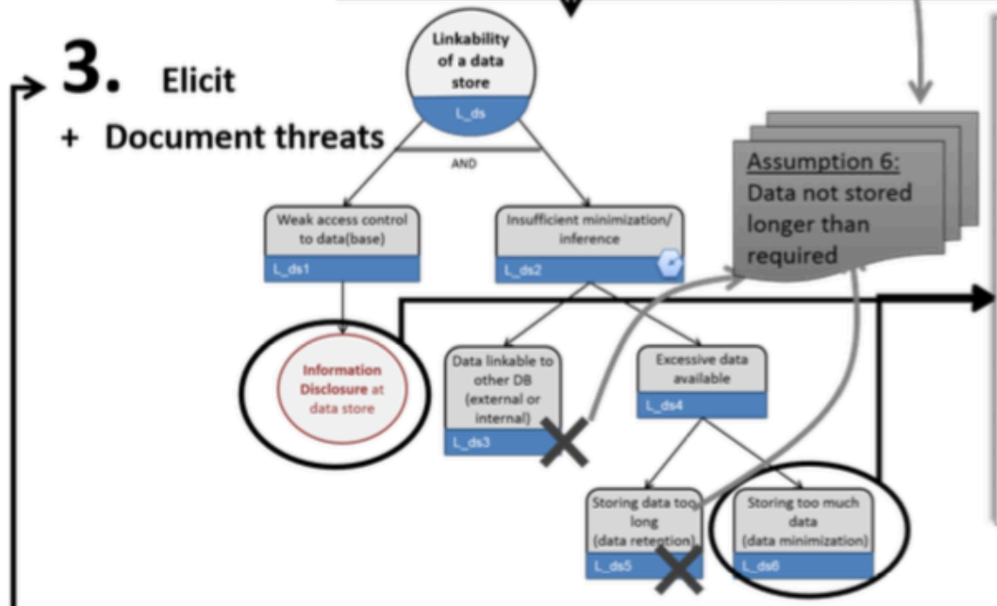
2. Map threats to DFD

	Threat target	L	I	N	D	D	U	N
Data store	Social network db	X	X			X		X*
Data flow	User data stream (user-portal)	X	X			X		X*
	Service data stream (portal-service)							X*
	DB data stream (service – DB)							X*
Process	Portal							X*
	Social network service							X*
Entity	User	X	X				X	

LINDDUN	L	I	N	D	D	U	N
Data store	X	X	X	X	X		X
Data flow	X	X	X	X	X		X
Process	X	X	X	X	X		X
Entity	X	X				X	

assumptions

3. Elicit + Document threats



MUC 01: linking data

Summary: Data entries can be linked to the same person

Primary mis-actor: skilled insider / skilled outsider

Basic path:
bf1. The misactor gains access to the database
bf2. Because too much data are stored, information can be infered ...

Consequence: By combining the entries, the misactor has access to more information about the data subject than anticipated

Reference to threat tree node(s)*: L_ds1, L_ds6

Parent threat tree(s)*: L_ds

DFD element(s)*: 4. social network data

Remarks*: data not stored longer than required (assumpt. 6) + ...

mapping table completed ?

no

yes

Move to **LINDDUN** *solution space*

SOK: Secure Messaging (Unger et al.)

well-defined goal	(interoperable/federated) secure messaging
	trust establishment
	conversation security
	transport privacy
privacy requirements	confidentiality + perfect forward/backward secrecy
	message/participation deniability
	anonymity ...
threat model (adversary)	local/global/ISP...
other quality requirements	usability and adoption

CONVERSATION SECURITY PROTOCOLS AND THEIR USABILITY AND ADOPTION IMPLICATIONS. NO APPROACH REQUIRES ADDITIONAL USER E

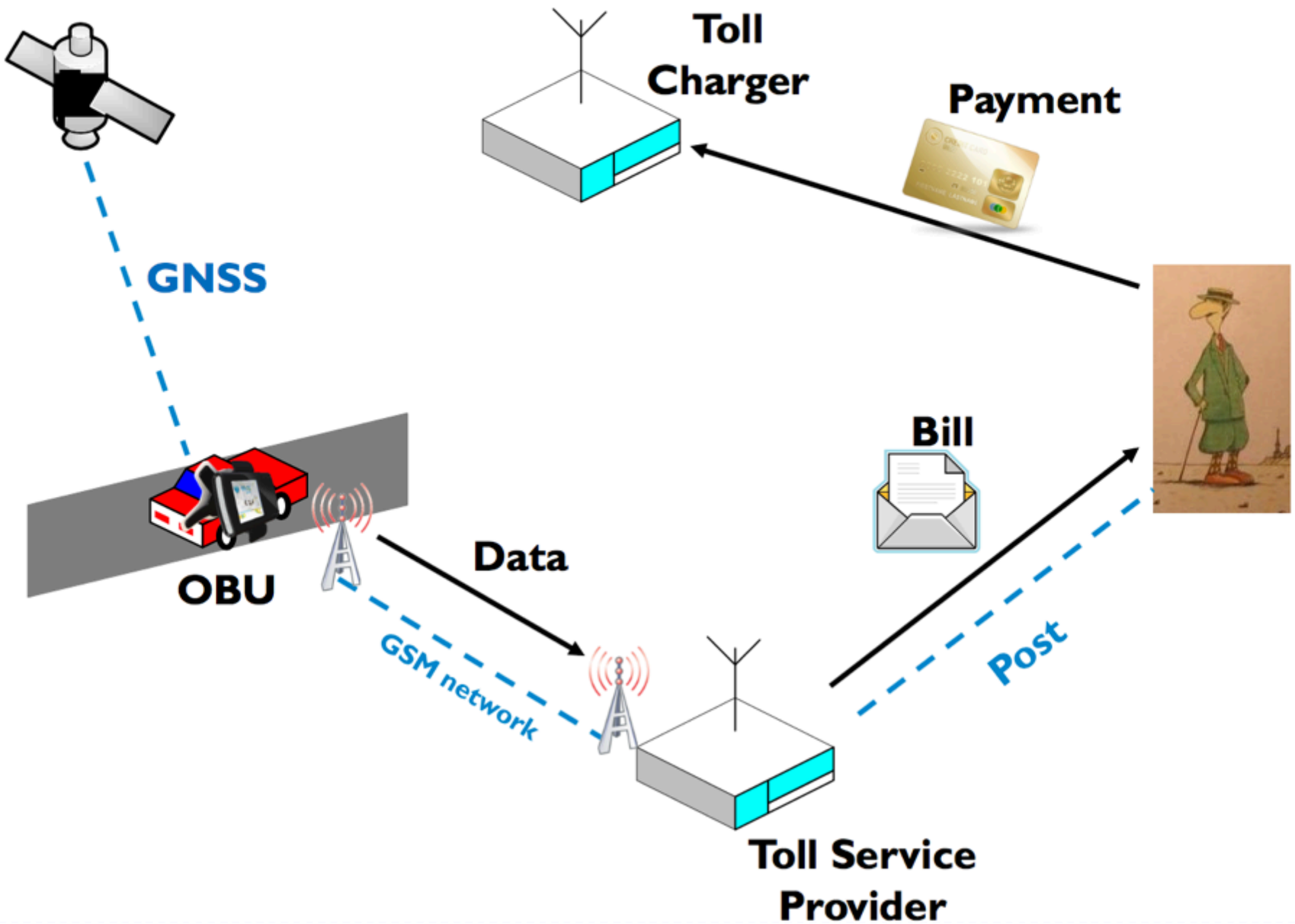
Scheme	Example	Security and Privacy													Adoption					Group			
		Confidentiality	Integrity	Authentication	Participant Consistency	Destination Validation	Forward Secrecy	Backward Secrecy	Anonymity Preserving	Speaker Consistency	Causality Preserving	Global Transcript	Message Unlinkability	Message Repudiation	Particip. Repudiation	Out-of-Order Resilient	Dropped Message Resilient	Asynchronicity	Multi-Device Support	No Additional Service	Computational E	Trust Equal	Subgro	C
TLS+Trusted Server†*	Skype	-	-	-	-	-	-	-	-	-	-	-	●	●	●	●	●	●	●	-	●	●	●	●
Static Asymmetric Crypto†*	OpenPGP, S/MIME	●	●	●	-	-	-	-	●	-	-	-	-	-	-	●	●	●	●	●				
+IBE†	Wang et al.	-	●	●	-	-	-	-	●	-	-	-	-	-	-	●	●	●	●	-				
+Short Lifetime Keys	OpenPGP Draft	●	●	●	-	-	◐	◐	●	-	-	-	-	-	-	●	●	●	●	-				
+Non-Interactive IBE†	Canetti et al.	●	●	●	-	-	●	-	●	-	-	-	-	-	-	◐	●	●	●	●				
+Puncturable Encryption†	Green and Miers	●	●	●	-	-	●	-	●	-	-	-	-	-	-	●	●	●	●	●				
Key Directory+Short Lifetime Keys†	IMKE	●	●	◐	-	●	◐	◐	-	-	-	-	●	●	●	●	●	-	-	-				
+Long-Term Keys†	SIMPP	●	●	◐	-	●	◐	◐	-	-	-	-	●	●	-	●	●	-	-	-				
Authenticated DH†*	TLS-EDH-MA	●	●	●	●	●	◐	◐	●	-	-	-	●	●	◐	●	●	-	-	●				
+Naïve KDF Ratchet*	SCIMP	●	●	●	●	●	●	◐	●	◐	-	-	●	●	◐	◐	◐	-	-	●				
+DH Ratchet†*	OTR	●	●	●	●	●	◐	●	●	◐	◐	-	●	●	◐	◐	◐	-	-	●				
+Double Ratchet†*	Axolotl	●	●	●	●	●	●	●	●	◐	◐	-	●	●	◐	●	◐	-	-	●				
+Double Ratchet+3DH AKE†*	-	●	●	●	●	●	●	◐	◐	◐	◐	-	●	●	●	●	◐	-	-	●				
+Double Ratchet+3DH AKE+Prekeys†*	TextSecure	●	●	●	●	●	●	●	-	◐	●	-	●	●	●	◐	◐	●	-	-				
Key Directory+Static DH+Key Transport†	Kikuchi et al.	●	●	-	-	●	◐	◐	-	-	-	-	●	●	-	●	●	●	-	-	-	-	-	●
+Authenticated EDH+Group MAC†	GROK	●	●	◐	-	●	◐	◐	●	-	-	-	●	●	-	●	●	●	-	-	-	-	-	●
GKA+Signed Messages+Parent IDs†	OldBlue	●	●	●	●	●	◐	◐	●	●	●	-	-	-	-	●	●	◐	-	●	●	●	-	-
Authenticated MP DH+Causal Blocks†*	KleeQ	●	●	◐	◐	◐	●	●	●	◐	◐	●	-	●	●	●	●	◐	-	●	●	●	-	-
OTR Network+Star Topology†	GOTR (2007)	●	●	-	-	-	◐	●	-	-	-	-	●	●	●	◐	●	◐	-	●	-	-	-	●
+Pairwise Topology†		●	●	●	●	●	◐	●	●	-	-	-	●	●	●	◐	●	◐	-	●	●	●	●	●
+Pairwise Axolotl+Multicast Encryption*	TextSecure	●	●	●	-	●	●	●	-	●	●	-	●	●	●	●	●	●	-	-	●	●	●	●
DGKE+Shutdown Consistency Check†	mpOTR	●	●	●	●	●	◐	◐	●	◐	-	-	-	●	●	●	●	-	-	●	●	●	-	-
Circle Keys+Message Consistency Check†	GOTR (2013)	●	●	●	●	●	◐	◐	●	●	●	●	●	●	●	●	-	-	-	◐	●	●	-	●

Engineering Privacy by Design

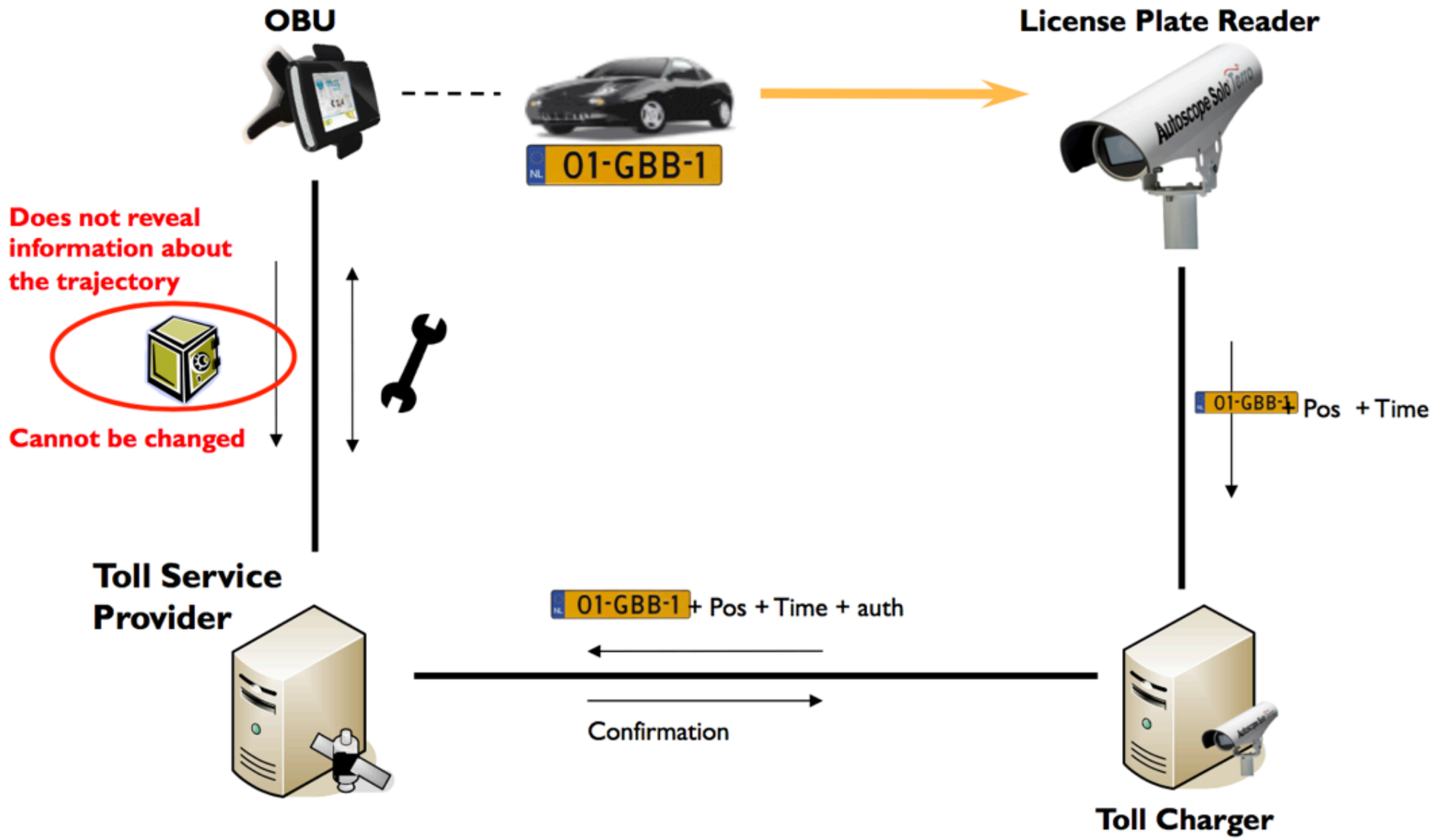
Is it possible to fulfill the desired functionality in a privacy preserving way?

Is it possible to fulfill the desired functionality while practicing data minimization?

EETS straightforward implementation



How does it work?



it is not that “data” is minimize (in the system as a whole)

the “amount” of data is the same as in the straightforward approach

it is kept in user devices, sent encrypted to a server, distributed over multiple servers

maybe data minimization is not the right metaphor?

Unpacking Data Minimization: Privacy By Design Strategies

Overarching goal

minimizing privacy risks and trust assumptions placed on other entities

Seda Gurses, Carmela Troncoso, Claudia Diaz. Engineering Privacy by Design Reloaded. Amsterdam Privacy Conference. 2015

Unpacking Data Minimization: Privacy By Design Strategies

Overarching goal

minimizing privacy risks and trust assumptions placed on other entities

Seda Gurses, Carmela Troncoso, Claudia Diaz. Engineering Privacy by Design Reloaded. Amsterdam Privacy Conference. 2015
Seda Gurses and Claudia Diaz. "Two tales of privacy in online social networks." IEEE Security & Privacy Magazine. 2013

Unpacking Data Minimization: Privacy By Design Strategies

Overarching goal

minimizing privacy risks and trust assumptions placed on other entities

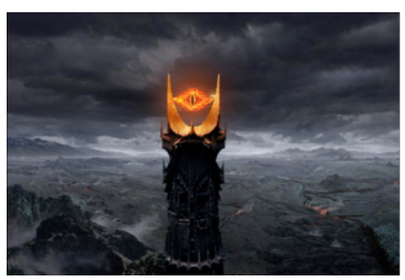
Other users
Third parties

+

semi-trusted
service provider

privacy as practice

privacy as control



The Adversary

EVERYONE

privacy as confidentiality

Seda Gurses, Carmela Troncoso, Claudia Diaz. Engineering Privacy by Design Reloaded. Amsterdam Privacy Conference. 2015
Seda Gurses and Claudia Diaz. "Two tales of privacy in online social networks." IEEE Security & Privacy Magazine. 2013

Unpacking Data Minimization: Privacy By Design Strategies

Overarching goal

minimizing privacy risks and trust assumptions placed on other entities

strategies

- Minimize Collection
- Minimize Disclosure
- Minimize Linkability
- Minimize Centralization
- Minimize Replication
- Minimize Retention

Seda Gurses, Carmela Troncoso, Claudia Diaz. Engineering Privacy by Design Reloaded. Amsterdam Privacy Conference. 2015

Unpacking Data Minimization: Privacy By Design Strategies

Overarching goal

minimizing privacy risks and trust assumptions placed on other entities

strategies

- Minimize Collection
- Minimize Disclosure
- Minimize Linkability
- Minimize Centralization
- Minimize Replication
- Minimize Retention

Great! but... how do we use these strategies?

We make explicit the activities and reasoning in **privacy engineering design** process

Seda Gurses, Carmela Troncoso, Claudia Diaz. Engineering Privacy by Design Reloaded. Amsterdam Privacy Conference. 2015

Case study: Electronic Toll Pricing

Motivation: European Electronic Toll Service (EETS)

Toll collection on European Roads trough On Board Equipment

Two approaches: Satellite Technology / DSRC

Starting assumptions

1) Well defined functionality

Charge depending on driving

2) Security, privacy & service integrity requirements

Users location should be private

No cheating clients

3) Initial reference system

Commission Decision of 6 October 2009 on the definition of the European Electronic Toll Service and its technical elements
http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32009D0750

Case study: Electronic Toll Pricing

Motivation: European Electronic Toll Service (EETS)

Toll collection on European Roads trough On Board Equipment

Two approaches: Satellite Technology / DSRC

Starting assumptions

1) Well defined functionality

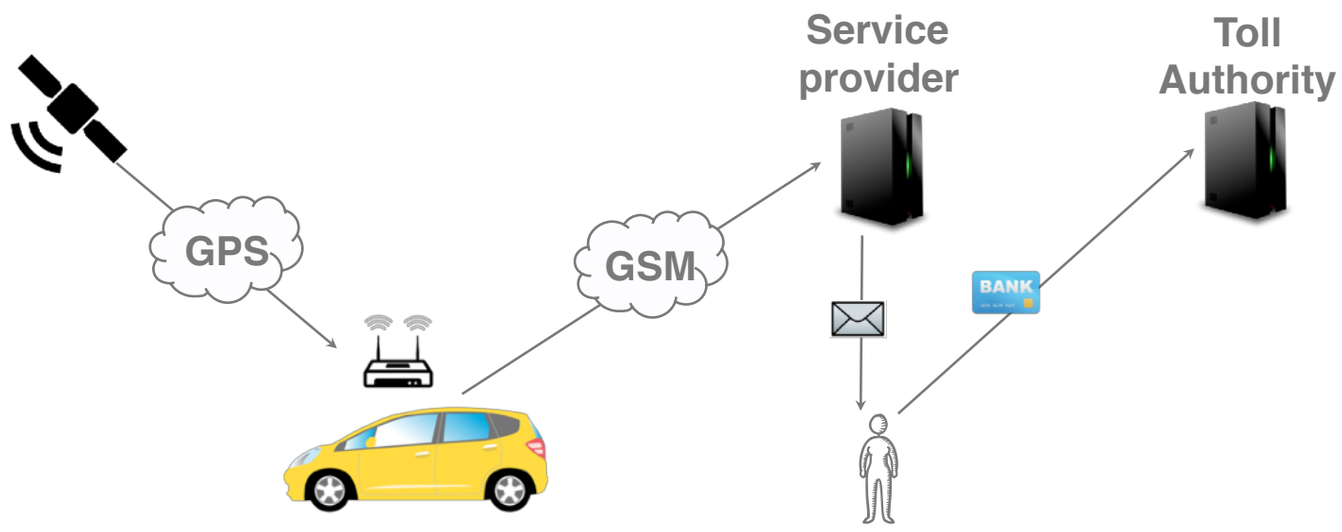
Charge depending on driving

2) Security, privacy & service integrity requirements

Users location should be private

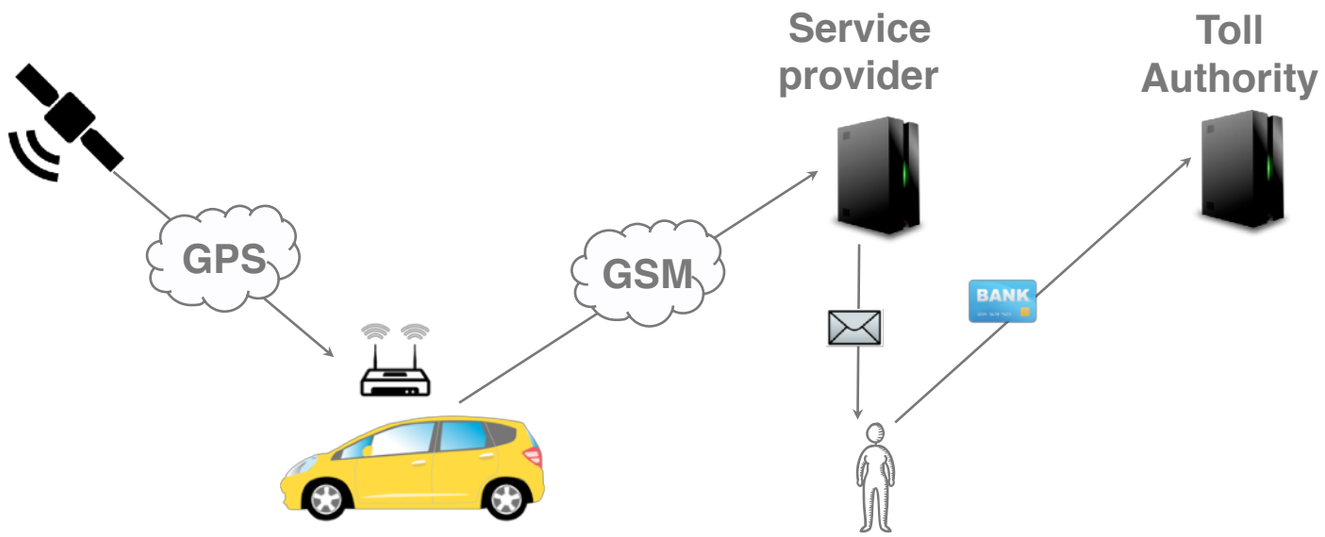
No cheating clients

3) Initial reference system



Commission Decision of 6 October 2009 on the definition of the European Electronic Toll Service and its technical elements
http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32009D0750

Case study: Electronic Toll Pricing

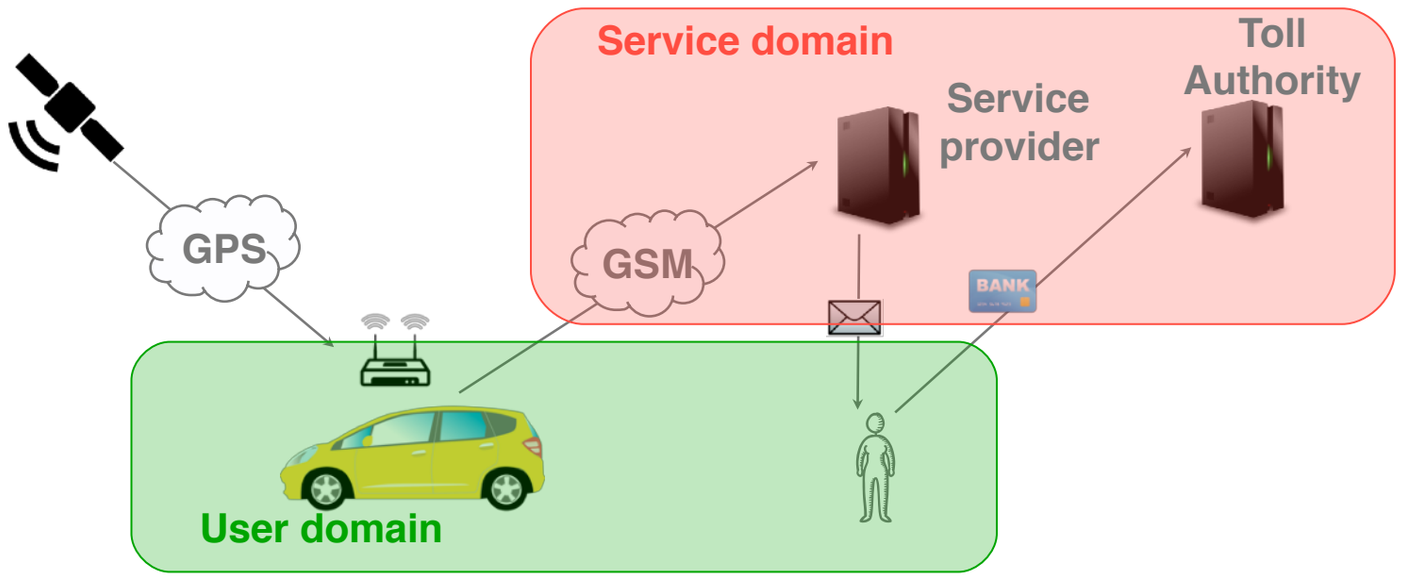


Activity 1: Classify Entities in domains

User domain: components under the control of the user, eg, user devices

Service domain: components outside the control of the user, eg, backend system at provider

Case study: Electronic Toll Pricing

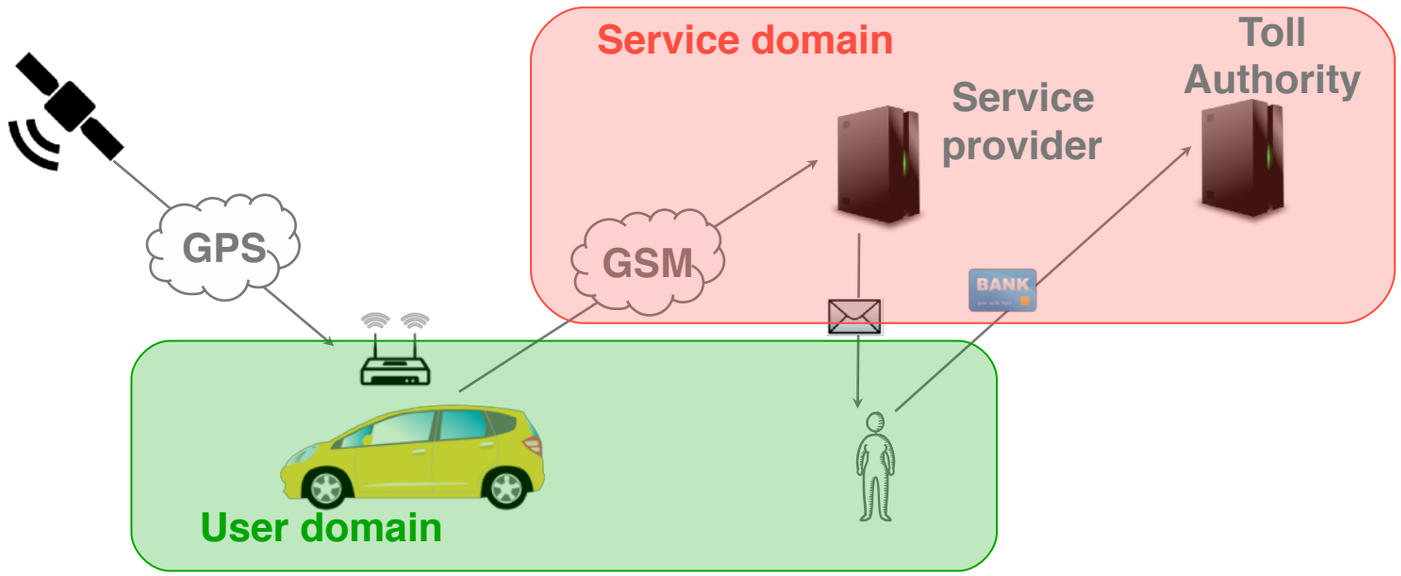


Activity 1: Classify Entities in domains

User domain: components under the control of the user, eg, user devices

Service domain: components outside the control of the user, eg, backend system at provider

Case study: Electronic Toll Pricing



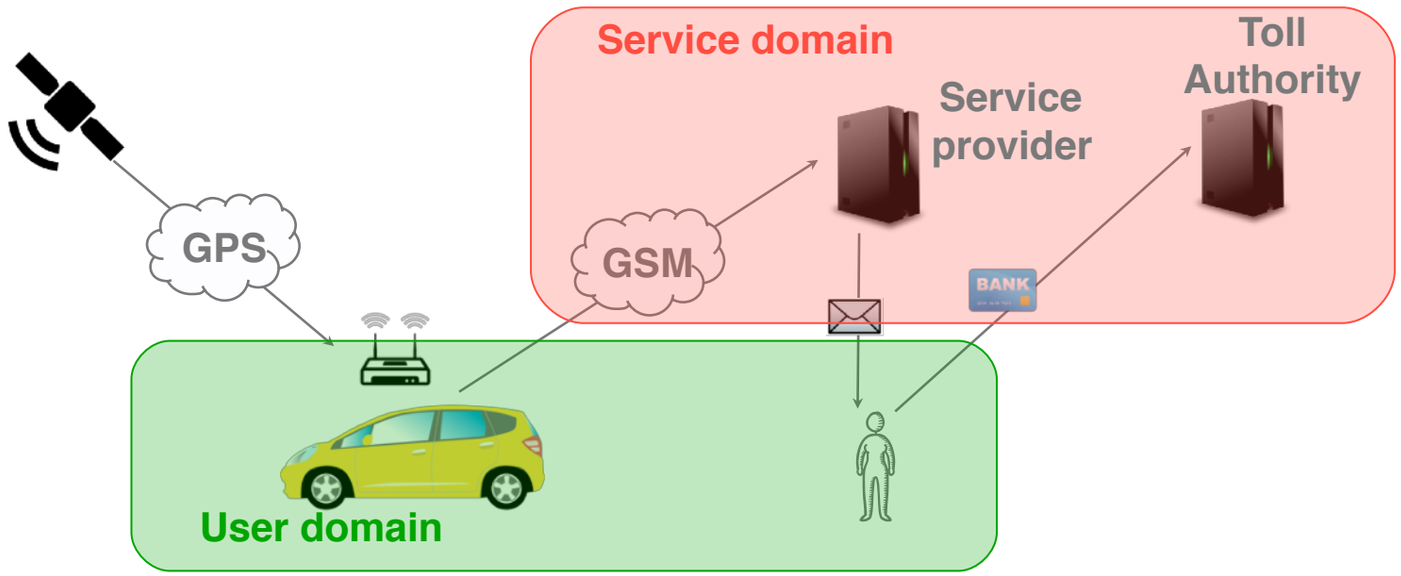
Activity 1: Classify Entities in domains

User domain: components under the control of the user, eg, user devices
Service domain: components outside the control of the user, eg, backend system at provider

Activity 2: Identify necessary data for providing the service

Location data – compute bill
Billing data – charge user
Personal data – send bill
Payment data – perform payment

Case study: Electronic Toll Pricing



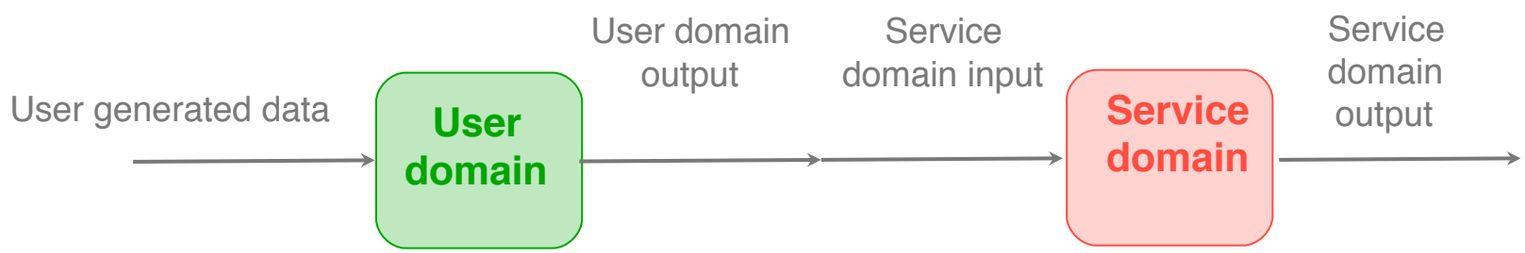
Activity 1: Classify Entities in domains

User domain: components under the control of the user, eg, user devices
Service domain: components outside the control of the user, eg, backend system at provider

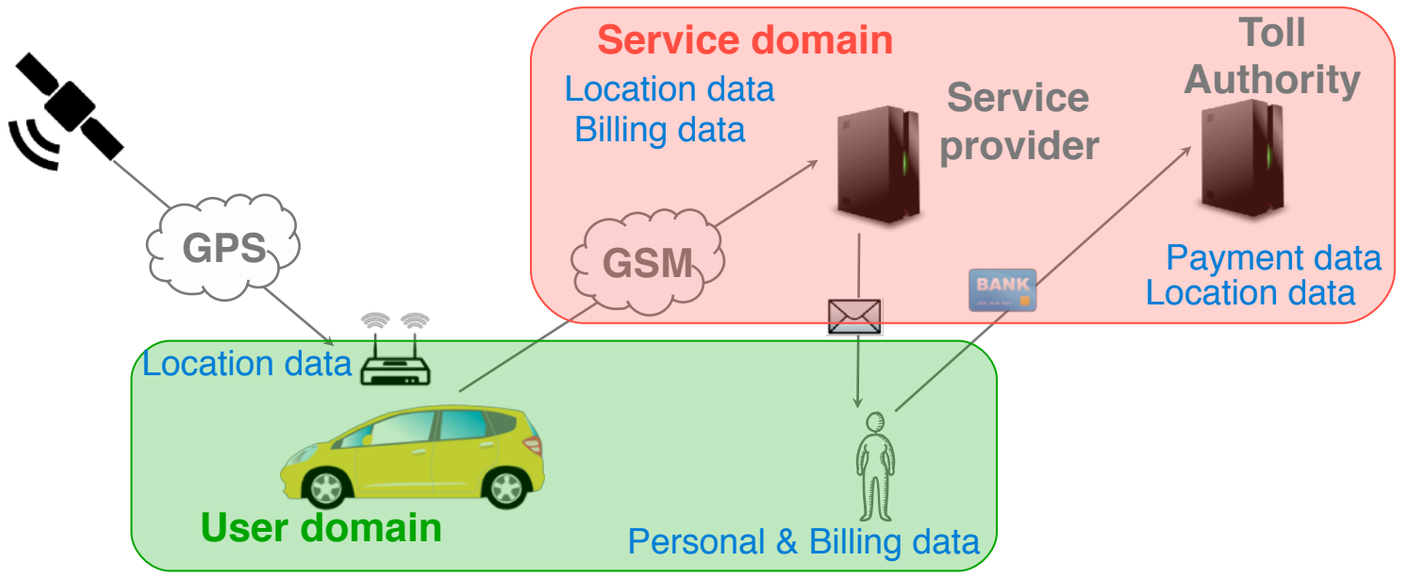
Activity 2: Identify necessary data for providing the service

Location data – compute bill
Billing data – charge user
Personal data – send bill
Payment data – perform payment

Activity 3: Distribute data in architecture



Case study: Electronic Toll Pricing



privacy as practice

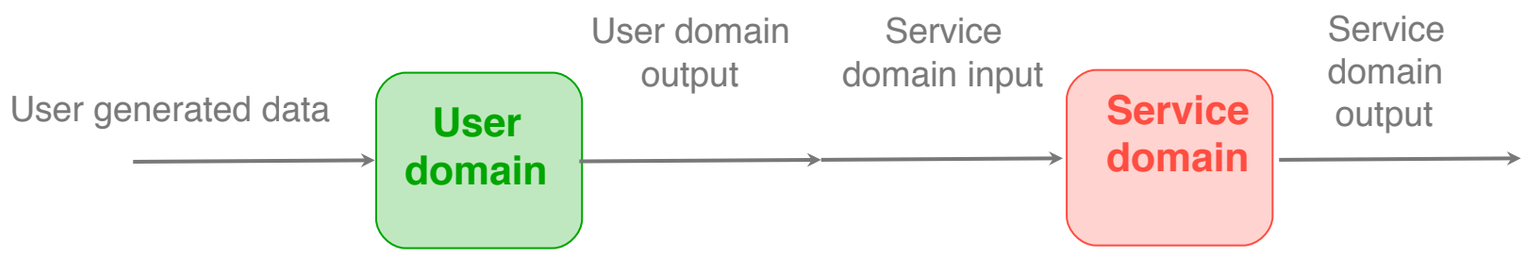
Activity 1: Classify Entities in domains

User domain: components under the control of the user, eg, user devices
Service domain: components outside the control of the user, eg, backend system at provider

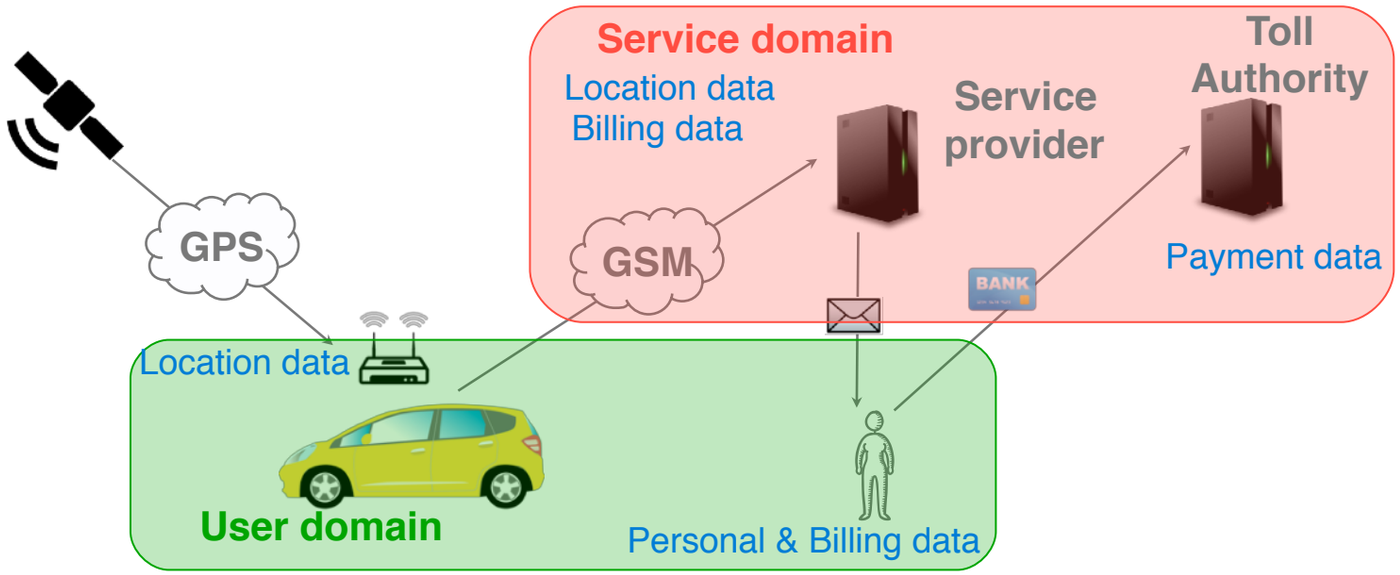
Activity 2: Identify necessary data for providing the service

Location data – compute bill
Billing data – charge user
Personal data – send bill
Payment data – perform payment

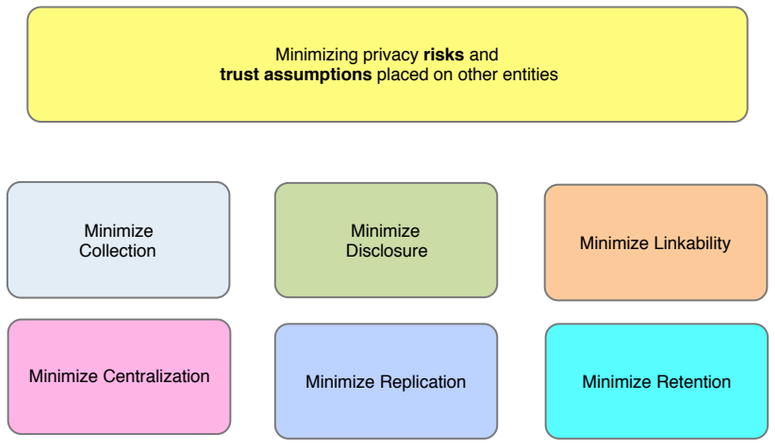
Activity 3: Distribute data in architecture



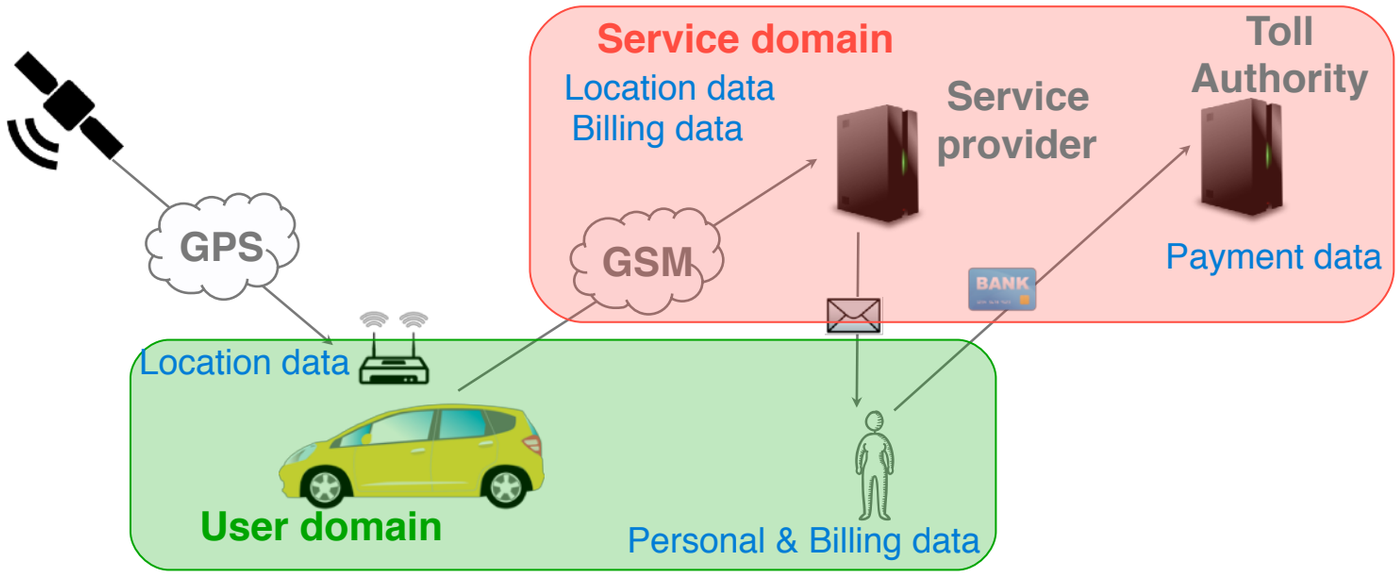
Case study: Electronic Toll Pricing



Activity 4: Select technological solutions following →

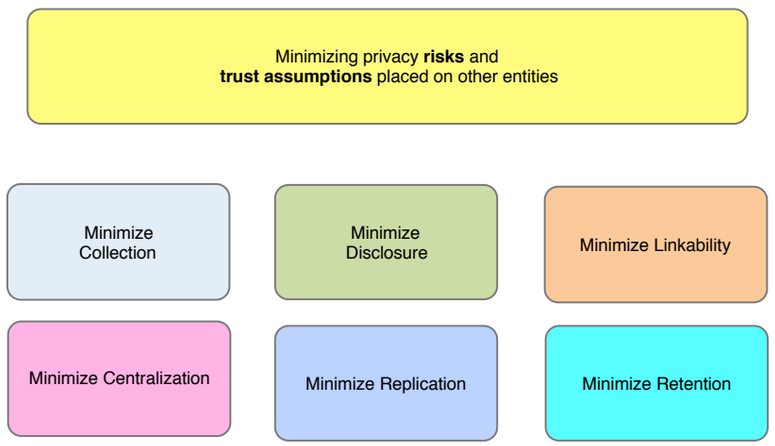


Case study: Electronic Toll Pricing

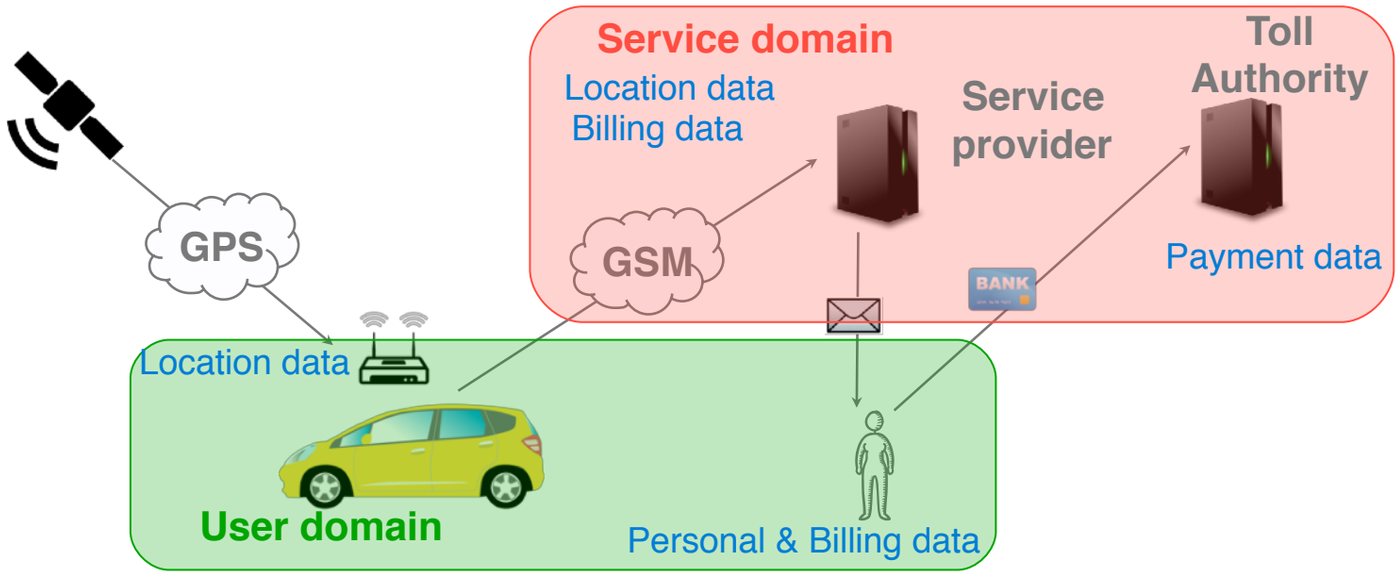


Activity 4: Select technological solutions following →

not sending the data (local computations)
encrypting the data
advanced privacy-preserving protocols
obfuscate the data
anonymize the data



Case study: Electronic Toll Pricing

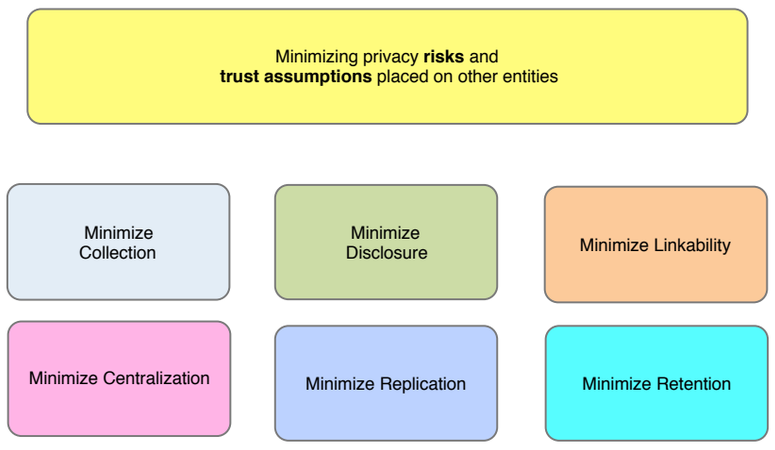


Trust Service to keep privacy of location data

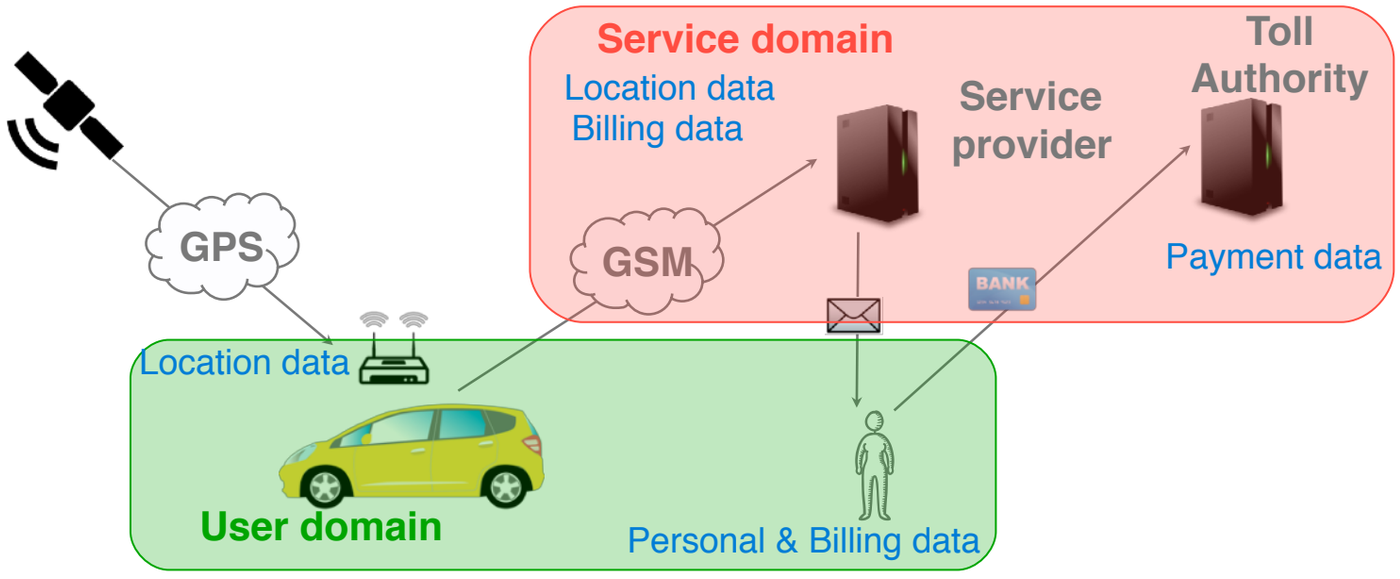
Risk of privacy breach

Activity 4: Select technological solutions following →

not sending the data (local computations)
encrypting the data
advanced privacy-preserving protocols
obfuscate the data
anonymize the data



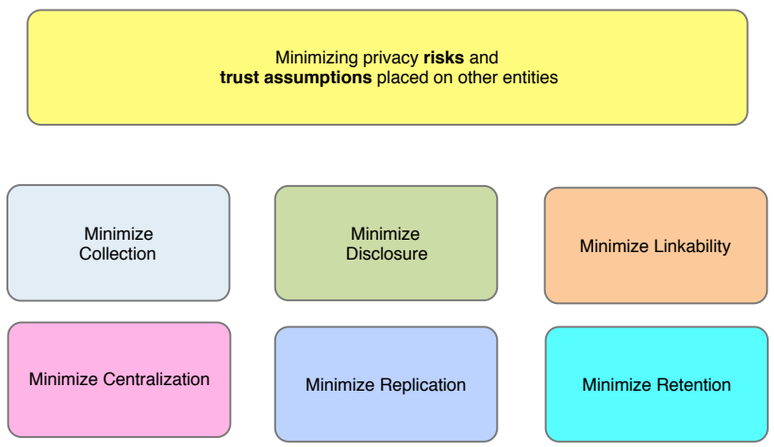
Case study: Electronic Toll Pricing



Location is not needed, only the amount to bill!

Activity 4: Select technological solutions following →

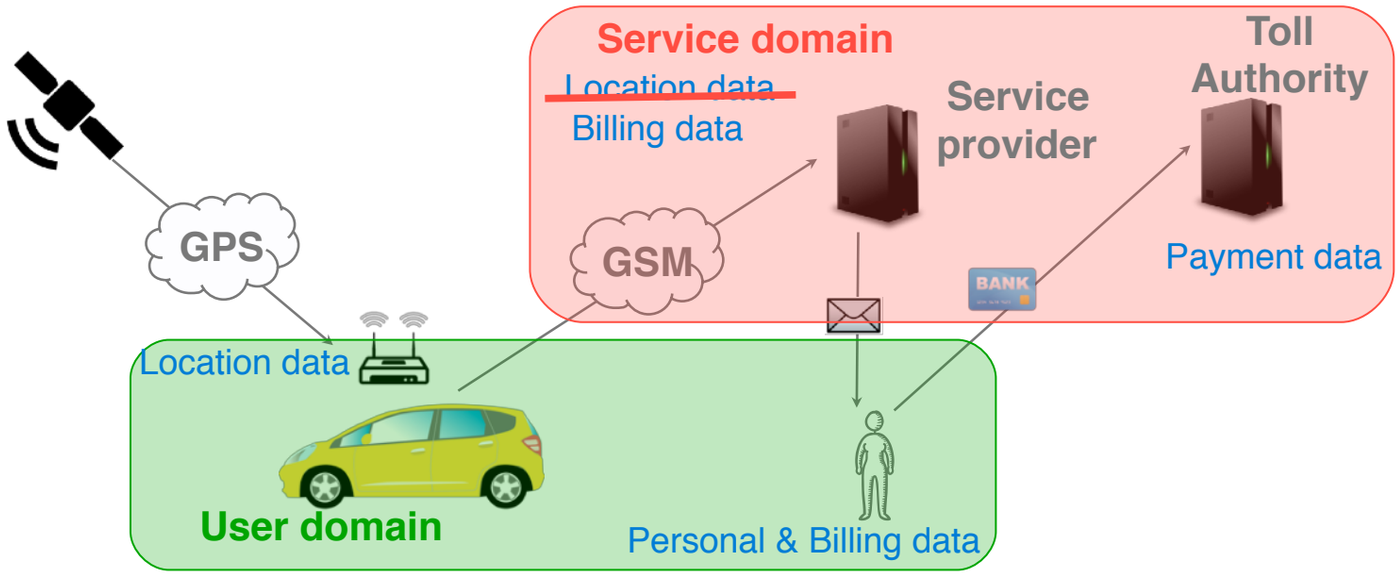
not sending the data (local computations)
encrypting the data
advanced privacy-preserving protocols
obfuscate the data
anonymize the data



J. Balasch, A. Rial, C. Troncoso, B. Preneel, I. Verbauwhede, C. Geuens. PrETP “Privacy-Preserving Electronic Toll Pricing” USENIX Security Symposium 2010

C. Troncoso, G. Danezis, E. Kosta, J. Balasch, B. Preneel. "PriPAYD. Privacy-Friendly Pay-As-You-Drive Insurance" IEEE TDSC 2011

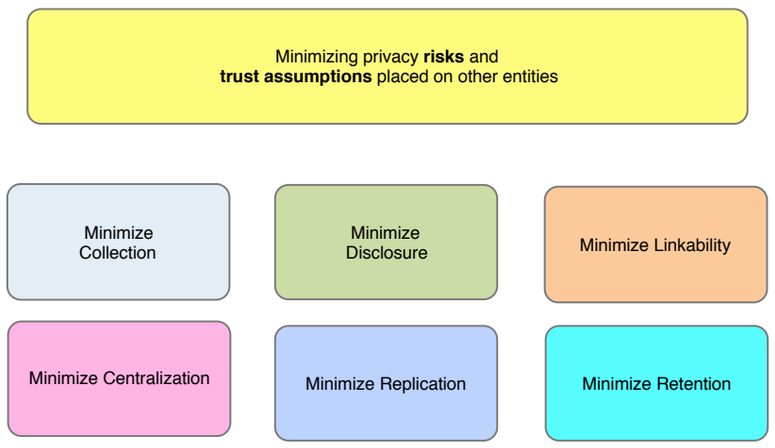
Case study: Electronic Toll Pricing



Location is not needed, only the amount to bill!

Activity 4: Select technological solutions following →

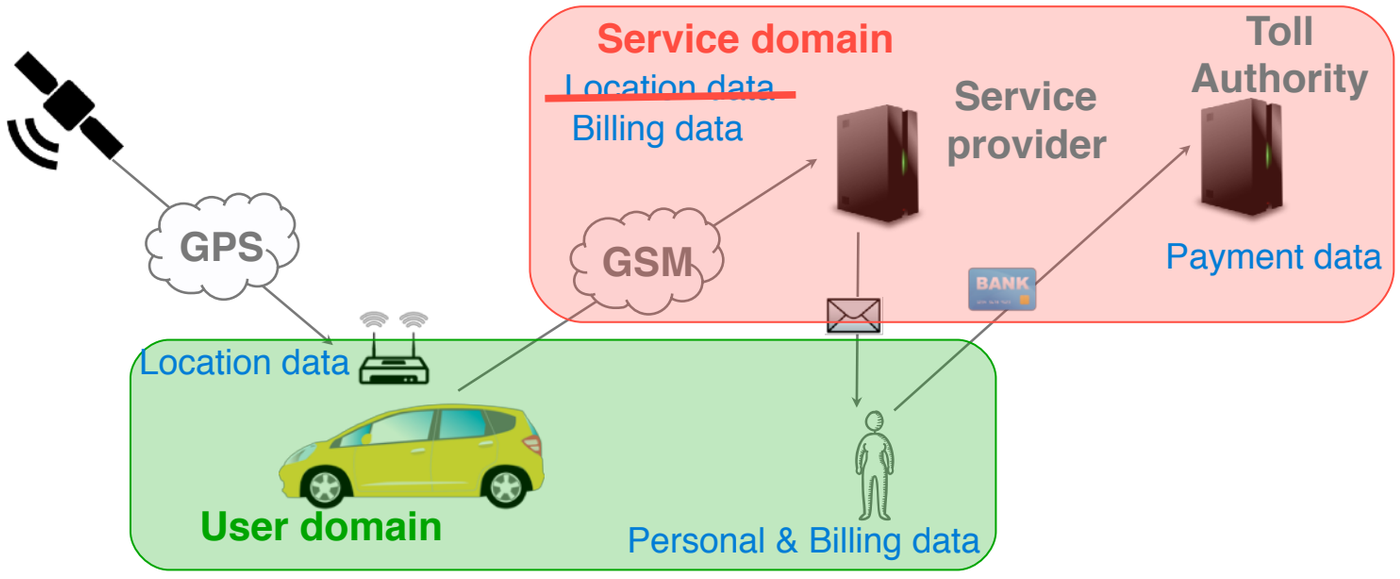
- not sending the data (local computations)
- encrypting the data
- advanced privacy-preserving protocols
- obfuscate the data
- anonymize the data



J. Balasch, A. Rial, C. Troncoso, B. Preneel, I. Verbauwhede, C. Geuens. PrETP “Privacy-Preserving Electronic Toll Pricing” USENIX Security Symposium 2010

C. Troncoso, G. Danezis, E. Kosta, J. Balasch, B. Preneel. "PriPAYD. Privacy-Friendly Pay-As-You-Drive Insurance" IEEE TDSC 2011

Case study: Electronic Toll Pricing

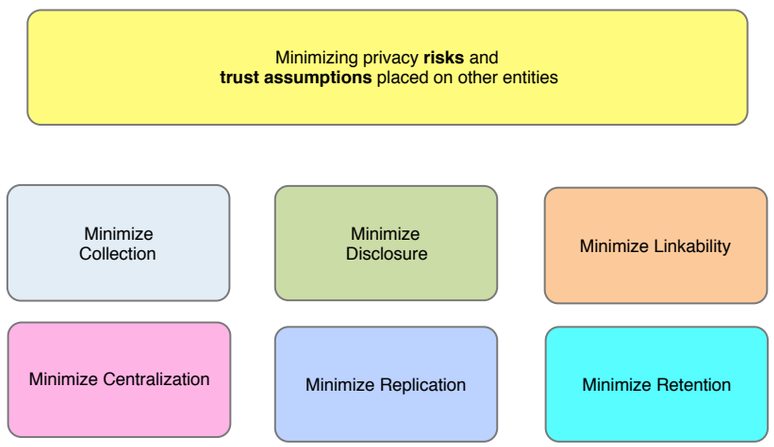


Location is not needed, only the amount to bill!

Service integrity?

Activity 4: Select technological solutions following →

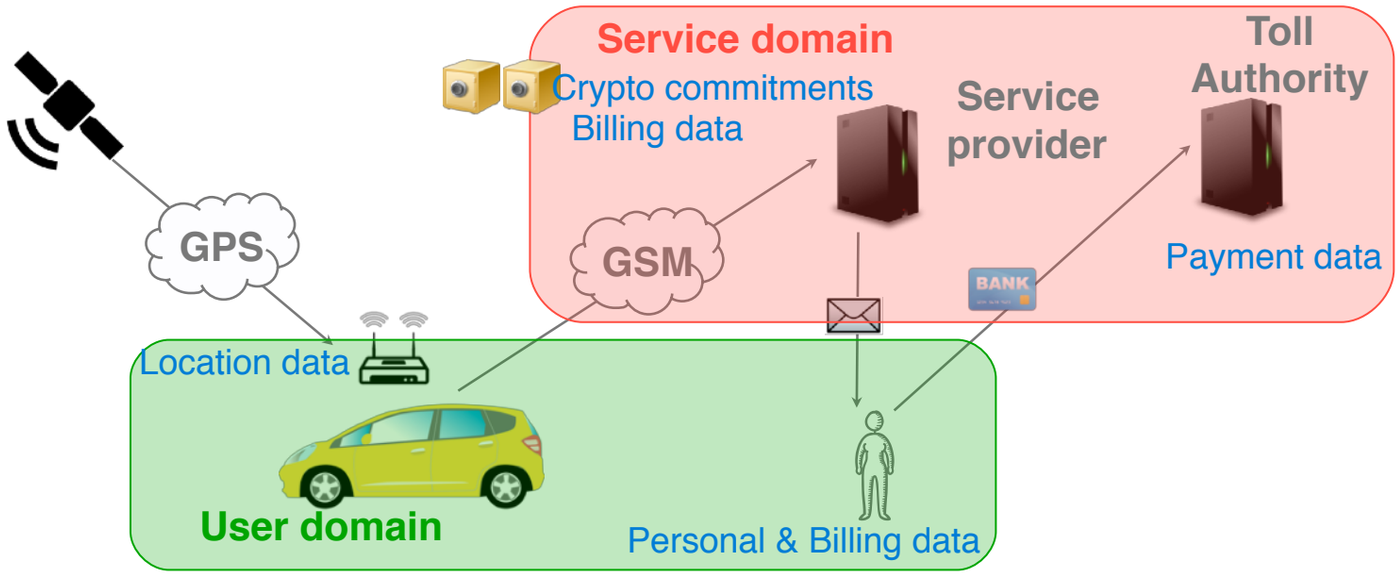
- not sending the data (local computations)
- encrypting the data
- advanced privacy-preserving protocols
- obfuscate the data
- anonymize the data



J. Balasch, A. Rial, C. Troncoso, B. Preneel, I. Verbauwhede, C. Geuens. PrETP “Privacy-Preserving Electronic Toll Pricing” USENIX Security Symposium 2010

C. Troncoso, G. Danezis, E. Kosta, J. Balasch, B. Preneel. "PriPAYD. Privacy-Friendly Pay-As-You-Drive Insurance" IEEE TDSC 2011

Case study: Electronic Toll Pricing

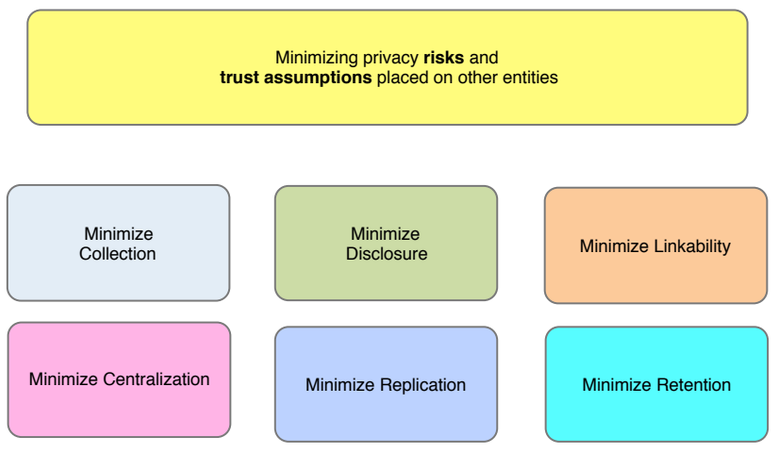


Location is not needed, only the amount to bill!

Service integrity?

Activity 4: Select technological solutions following →

- not sending the data (local computations)
- encrypting the data
- advanced privacy-preserving protocols
- obfuscate the data
- anonymize the data

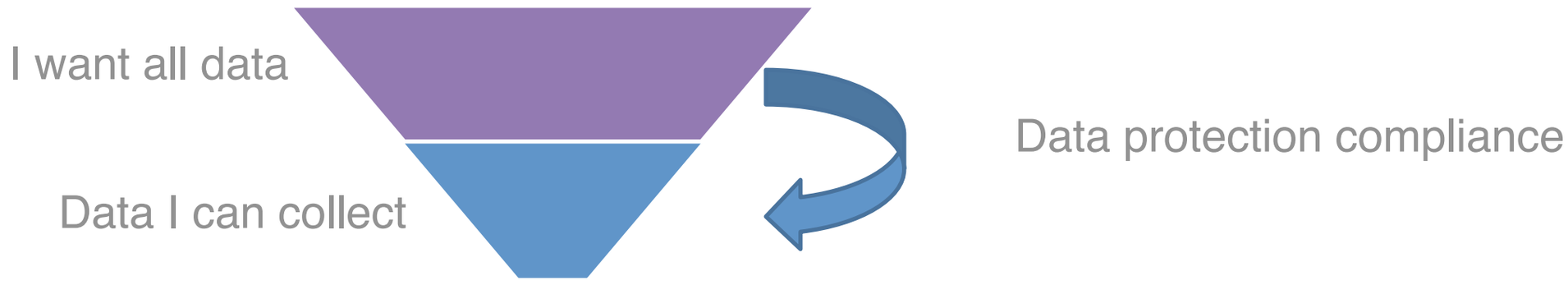


J. Balasch, A. Rial, C. Troncoso, B. Preneel, I. Verbauwhede, C. Geuens. PrETP “Privacy-Preserving Electronic Toll Pricing” USENIX Security Symposium 2010

C. Troncoso, G. Danezis, E. Kosta, J. Balasch, B. Preneel. "PriPAYD. Privacy-Friendly Pay-As-You-Drive Insurance" IEEE TDSC 2011

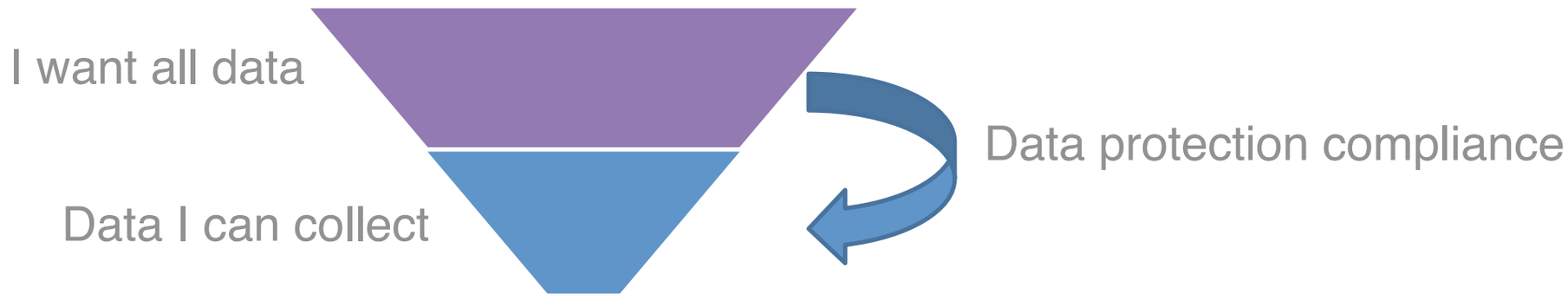
Change mental models for designing systems

The Usual approach

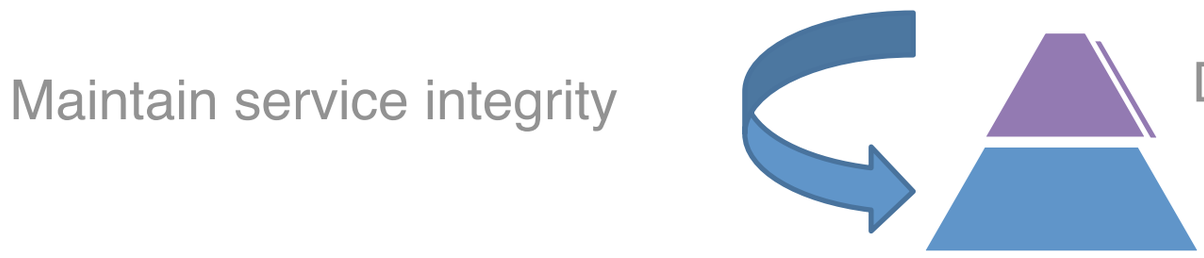


Change mental models for designing systems

The Usual approach

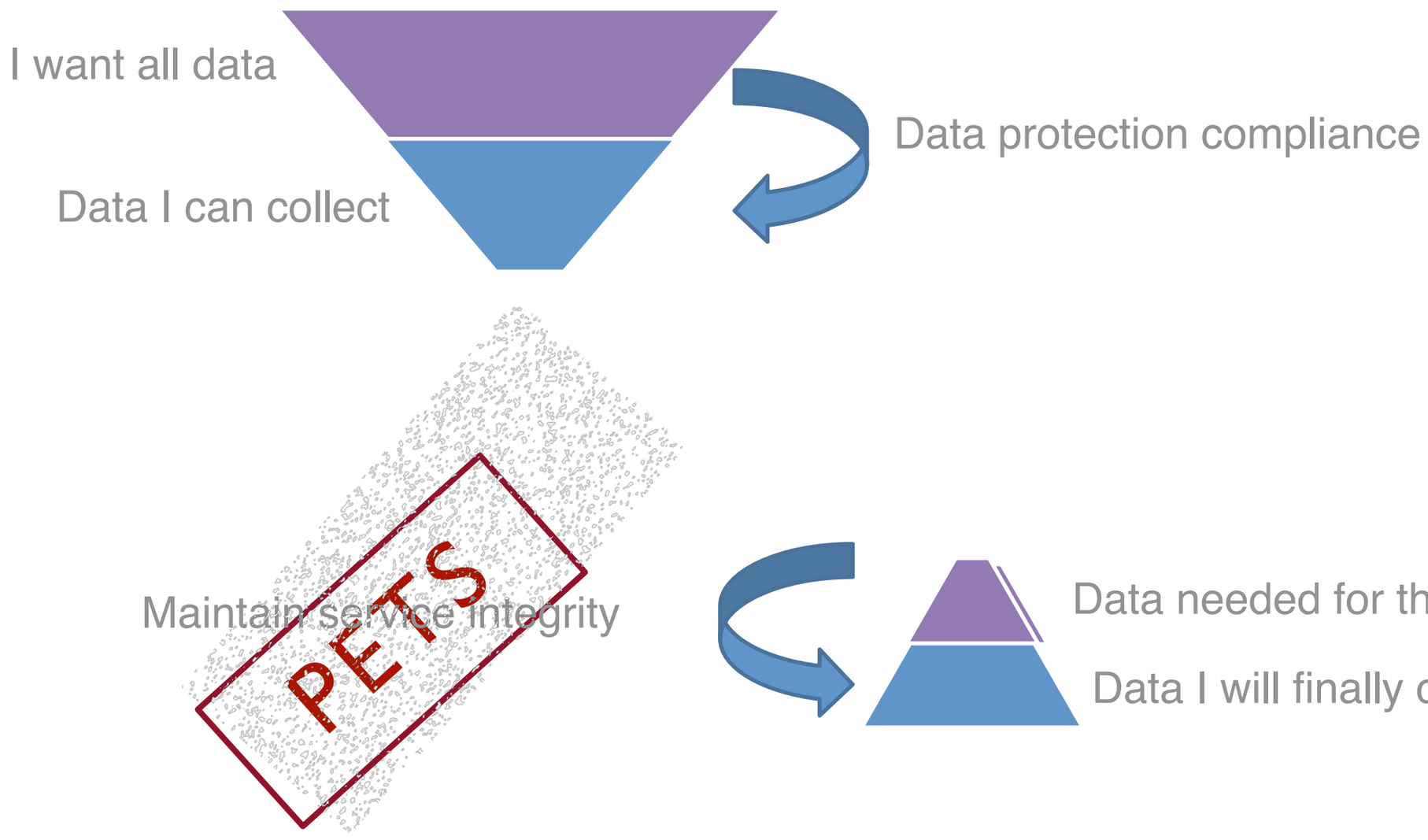


The PbD approach



Change mental models for designing systems

The Usual approach



	Engineering PbyD	Privacy Design Strategies	LINDDUN	PriPare
Lawfulness	x			
Data Subject Rights		x		
Sensitive Data	x	x	x	x
Transparency		x		x
Purpose limitation	x	x		x
Storage limitation	x	x		x
Accuracy	x	x		x
Security	x	x	x	x
Accountability	x	x	x	x

tools:
(automated) means that support privacy engineers during part of a privacy engineering process.

Tor Experimentation Tools

Fatemeh Shirazi
TU Darmstadt/KU Leuven
Darmstadt, Germany
fshirazi@cdc.informatik.tu-darmstadt.de

Matthias Goehring
TU Darmstadt
Darmstadt, Germany
de.m.goehring@ieee.org

Claudia Diaz
KU Leuven/iMinds
Leuven, Belgium
claudia.diaz@esat.kuleuven.be

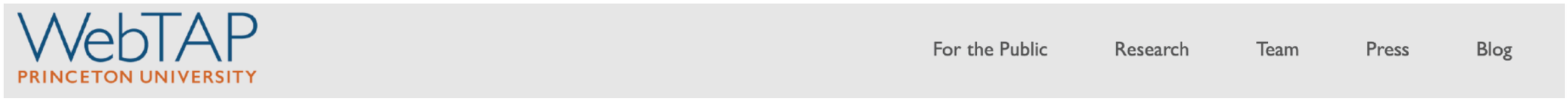
Comparison

TECHNISCHE UNIVERSITÄT DARMSTADT

Metric	Shadow	TorPS	ExperimenTor
1. Size / number of relays	downscaling, simulation with 500+ relays possible	no downscaling	limited by available resources
2. Routing approach	not using additional weighting in node	ignoring paths being dropped due to	

Web Transparency and Accountability Project

https://webtap.princeton.edu



Princeton Web Transparency & Accountability Project

Measure Threats

We monitor websites and services to find out what user data companies collect, how they collect it, and what they do with it. With our measurement platform, we study privacy, security, and ethics of consumer data usage.

Create Change

Our external oversight exposes the privacy practices of companies and forces them to make improvements. In addition, the data and studies that we produce assist regulators and privacy tool developers in their efforts.

Inform the Public

We translate our research into practical information for public consumption. We aim to improve the accuracy of media reports about online privacy and to provide useful advice for consumers on this website.

Web Transparency and Accountability Project
https://webtap.princeton.edu

Online tracking: A 1-million-site measurement and analysis is the largest and most detailed measurement of online tracking to date. We measure stateful (cookie-based) and stateless (fingerprinting-based) tracking, the effect of browser privacy tools, and "cookie syncing".

This measurement is made possible by our web measurement tool OpenWPM, a mature platform that enables fully automated web crawls using a full-fledged and instrumented browser.

Read the paper »

Lumen Privacy Monitor

https://www.haystack.mobi

NSF DATA TRANSPARENCY LAB

Keep control of your data

Lumen identifies apps leaking your privacy-sensitive data over the network so that you stay in control of your network fingerprint.

Find Online Trackers

Lumen reports the third-party organizations collecting your personal information.

HTTPS/TLS Support

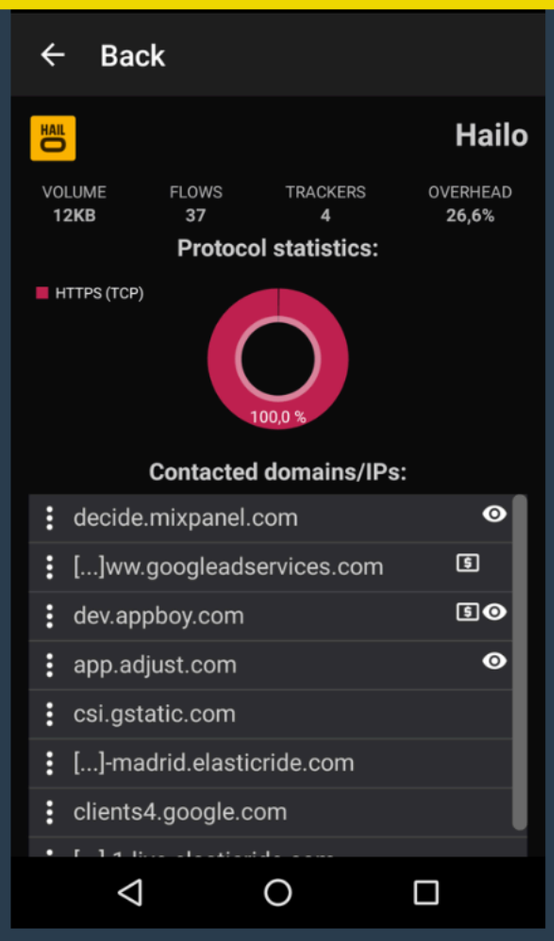
Lumen supports TLS interception so you can identify apps leaking privacy-sensitive information over encrypted traffic in real-time.

Be part of a research study!

Lumen comes from a research team at ICSI--UC Berkeley. By installing Lumen, you actively contribute to ongoing research efforts aiming to improve the operational transparency of mobile technologies.

Lumen Privacy Monitor

https://www.haystack.mobi

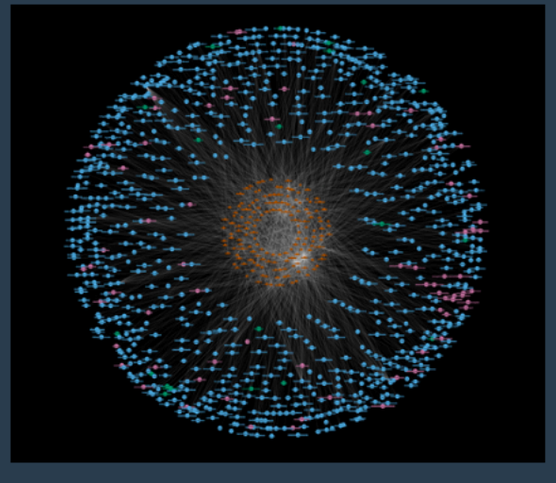


Detailed Reports

Apps may sometimes leak information to not only their own servers but also to online advertising networks or other online tracking services that monetize your metadata. Lumen aims to help you to understand many dynamics that may remain unknown for you! Lumen analyzes your mobile traffic and generates reports about the traffic patterns and the private data collected by each application and online service.

Illuminating App Behavior

Nearly 70% of Android apps leak personal data to third-party services such as analytics services and ad networks. The data provided by Lumen users is used to promote app and service transparency. For instance, you can play with our interactive ICSI panopticon tool to better understand the whole mobile ecosystem and how apps use third-party online trackers. **You can also contribute to our research efforts by installing and running our Lumen app**!

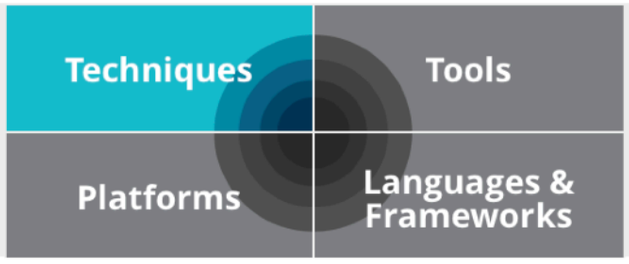


Differential Privacy

ThoughtWorks®

Clients Services Products Insights About us Careers

TECHNOLOGY RADAR



Search About the Radar Build your Radar Subscribe

The information in our interactive Radar is currently only available in English. To get information in your native language, please download the PDF here.

Techniques

Differential privacy

MAR 2017

ASSESS

It has long been known that "anonymized" bulk data sets can reveal information about individuals, especially when multiple data sets are cross-referenced together. With increasing concern over personal privacy, some companies—including Apple and Google—are turning to **differential privacy** techniques in order to improve individual privacy while retaining the ability to perform useful analytics on large numbers of users. Differential privacy is a cryptographic technique that attempts to maximize the accuracy of statistical queries from a database while minimizing the chances of identifying its records. These results can be achieved by introducing a low amount of "noise" to the data, but it's important to note that this is an ongoing research area. Apple has announced plans to incorporate differential privacy into its products—and we wholeheartedly applaud its commitment to customers' privacy—but the usual Apple secrecy has left some

NOT ON THE CURRENT EDITION

This blip is not on the current edition of the radar. If it was on one of the last few editions it is likely that it is still relevant. If the blip is older it might no longer be relevant and our assessment might be different today. Unfortunately, we simply don't have the bandwidth to continuously review blips from previous editions of the radar.
Understand more »

Data Subject Access Rights and Data Portability

right to receive: data subjects have a right to receive their data from a controller in a structured, commonly used, interoperable, and machine readable format.

right to transmit: data subjects have a right to move data between data controllers without hindrance, or where technically feasible have data moved directly between data controllers.

measure to provide users the ability to escape lock in (very much about competition!)

concerning which data?

not anonymized, but pseudonymized data. Not third party data, but sometimes, e.g., VoIP call records with third part numbers ok, CCTV images of others not.

“provided by the data subject’ includes
data entered by the data subject + observed data
but not inferred data??!

WE MADE IT!!! Thank you!

For further references and questions:
seda _AT_ esat.kuleuven.be