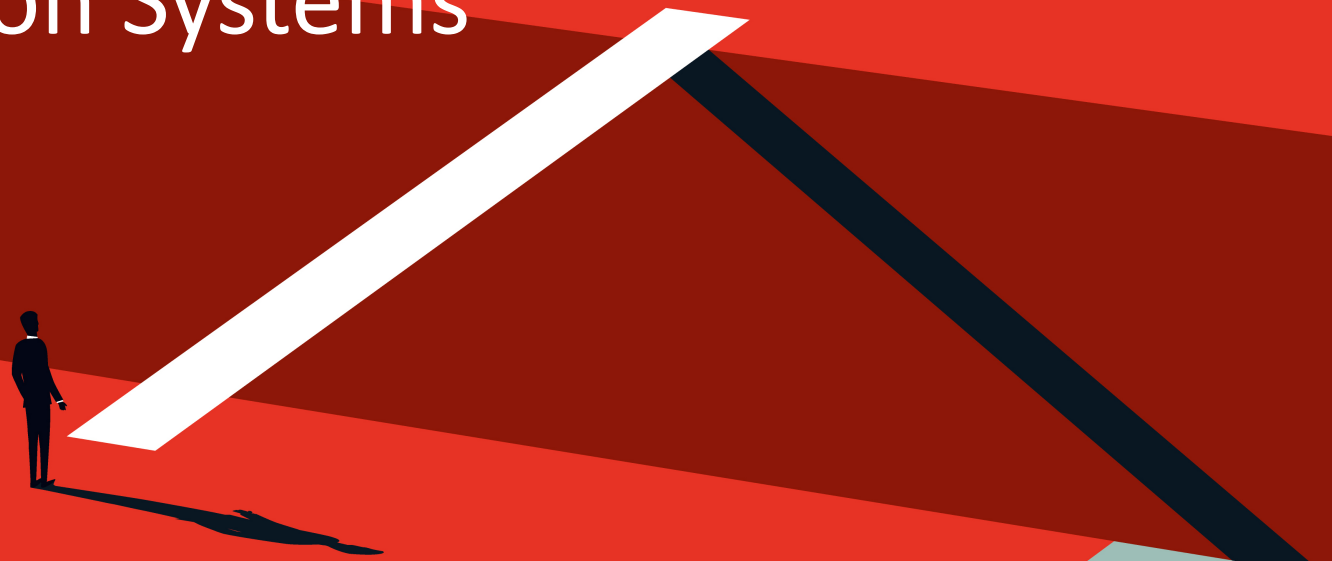
On the Design and Realization of Secure and Privacy-Protecting Information Systems



Keynote speech
at IARIA ICDS 2020
by Mortaza S. Bargh

Valencia, Spain
21-25 November 2020
Applied Research

Wetenschappelijk Onderzoek- en Documentatiecentrum
Ministerie van Justitie en Veiligheid

IARIA

ROTTERDAM UNIVERSITY
UNIVERSITY OF APPLIED SCIENCES

CREATING

About Mortaza S. (Shoae) Bargh

- Current positions
  - Research center of Dutch ministry of Justice & Security (WODC): Scientific researcher
  - Rotterdam University of Applied Sciences (RUAS): Research professor on Privacy & Cybersecurity

- Topics of interest
  - Privacy preserving by design (for socio-technological systems)
  - Responsible data analytics / sensor fusion / machine learning
  - Cyber security (secure distributed systems, security by design)

- Emails
  - WODC: m.shoae.bargh@qoedc.nl
  - RUAS: m.shoae.bargh@hr.nl

Bio Mortaza S. (Shoae) Bargh

Mortaza S. Bargh received his Ph.D. and M.Sc. degrees in Information Theory from Eindhoven University of Technology. Since 2013 he is a scientific researcher at the Research and Documentation Centre (WODC) of the Dutch Ministry of Justice and Security. Since 2017 he has been appointed as a part-time professor on Privacy and Cybersecurity Engineering at Rotterdam University of Applied Sciences (RUAS). His current research interests include privacy/security by design engineering, privacy preserving data mining and publishing, responsible (privacy preserving, fair, explainable, etc.) machine learning, access and usage control, collaborative security, usable security, and risk management.

Outline

- Introduction: Safeguarding the digital world
- **Privacy protection**
- Cybersecurity
- Privacy and security by design methodologies
- An example
- Reflection: Takeaways

SAFEGUARDING THE DIGITAL WORLD

CREATING

ROTTERDAM UNIVERSITY
UNIVERSITY OF APPLIED SCIENCES

The digital world

- Importance
  - Serving many purposes (healthcare, transportation, finance, e-government, …)
  - Providing many (new) business opportunities (Google, WhatsApp, Instagram, Amazon, bol.com and Spotify)
- Our dependence on its well-functioning
- The well-functioning: A challenge
  - Its complexity
  - Its potential for misuse (intentionally or unintentionally)

Information Systems

- Definition
  - Software, hardware, data, people, procedures & networks
  - Enabling the use of information resources in a setting (e.g., organization)

- Examples
  - An Internet of Thing (IoT) system
  - Sharing data based on some measures
    - Procedural measures: Request, decide
    - Technological measures: Data transformation, data delivery via email

- Being socio-technological
  - Multi-disciplines involved
  - ICT, people, organizations, society

Our focus

- Privacy risks
  - Due to proliferation of personal data via information systems
  - Attacks: Personal data disclosure attacks
- Cybersecurity risks
  - Due to vulnerability of information systems
  - Attacks: Hacking and denial of service attacks

Market share (cybersecurity & privacy)

- Compound Annual Growth Rate (CAGR)

- About 10.2 – 12 %
  - From 2018 till 2023 / 2024 / 2025
  - Expanding to about 250 – 300 billion (USD)

- Actual spending may be far more
  - As companies may understate their cybersecurity budgets to protect their reputations

Dutch market share

- A study commissioned by Dutch Ministry of Economic Affairs

- Size of the Dutch cybersecurity sector
  - About 10% of the whole ICT turnover in 2014
  - The sector's added value of 3.8 to 4.1 billion euros
  - About 0.6% of the Dutch GDP in 2014

- Growth of cybersecurity sector
  - About 14.5% faster than the ICT sector itself

Job market (cybersecurity & privacy)

- At the start of 2018
  - About half million job vacancies in the US
  - Bureau of Labor Statistics of the US department of Labor
- Job growth rate
  - 28% projected in 2016–2026
  - Much faster than the average for all other occupations
- Not just a job, but a job sector of the future
  - Many jobs, like: Data scientists, data security analysts, secure software developers, forensic analysts, penetration testers and chief security officers
- Like healthcare sector

Driving forces

- Emergence of disruptive ICT
  - Internet of Things (IoT)
  - Bring Your Own Device (BYOD)
- Rising needs for specific solutions
  - For cloud computing
- Strategic plans of businesses
  - Not to become a victim of these risks
  - To gain a cutting-edge business value out of being trustworthy
- Others
  - Increase in the frequency and sophistication of cyber threats (malware, ransomware and phishing messages)
  - Rising threat of global cyberterrorism
  - New regulations and laws coming into effect (like GDPR)

Approaching the field of privacy-protection & cybersecurity

- Remember being a job sector: Having a wide scope

- Individual (scientific) disciplines
  - Cryptography (e.g., to protect data integrity)
  - Criminology (e.g., to study motives of cybercriminals)

- System operation lifecycle/process
  - Risk management for an information system
  - Business continuity

- System development lifecycle/process
  - How to design and realize an information system

Our scope:
System development process

- How to design/realize privacy-protecting and secure information systems in practice?
  - A need for, among others, design methodologies
- Challenging: Requiring making trade-offs on many fronts
  - Data utility versus data privacy
  - Data subjects being in control versus ease-of-use
- Real innovation: To find balance among contending values
- Note: The other aspects also important, not our main focus

Scope of the talk

- Discuss a few existing challenges
  - Gaps between the current and desired situations
  - Gaps between relevant, but rather isolated, areas
- Bridging the gaps through
  - Practice-oriented and/or applied-research
  - Embodiment of the research results in education
- This talk: Portraying the research chair’s scope and activities

PRIVACY PROTECTION

Privacy

- A normative concept
- Deeply rooted in various disciplines
  - Philosophy
  - Law
  - Ethics
  - Politics
  - Sociology
- Aristotle
  - An early principled discussion of privacy
  - A distinction between public and private spheres of life

Definitions of privacy

1. The right to be let alone
2. Limited access to the self
3. Secrecy (of concealed information)
4. Control over personal information
5. Protection of personhood (personality integrity)
6. Intimacy (control over developing personal relationships like love, caring, loving, ...)

Source: Solove, **Understanding Privacy**, Harvard University Press, 2008

Def. 3: Secrecy

- Privacy is violated by the public disclosure of previously concealed information
  - No privacy issue, if a fact is previously known

- Too narrow
  - Failing to recognize group-privacy/selective-disclosure
    - Keeping things private from some people
  - Secret info is not always private (e.g., military plans)
  - Private info is not always secret (e.g., one's debts)
  - One expects privacy in public (e.g., the books we read)

Issues of privacy definitions

- Definitions
  - The right to be let alone
  - Limited access to the self
  - Secrecy
  - Control over personal information
  - Protection of personhood
  - Intimacy

- Issues
  - Solove: Being too narrow (over restrictive) or too broad (over inclusive)
  - Changing with (technological) developments

Evolution of privacy definition

- Aristotle (384-322 BC)
  - Making distinction between public and private spheres of life

Evolution of privacy definition

- Warren and Brandeis (1890)
  - The right to be let alone: To live one's life as one chooses, free from assault intrusion or invasion
  - Instantaneous photography via Kodak's new snap cameras (George Eastman, 1881-1889)
  - The widespread circulation of newspapers

Evolution of privacy definition

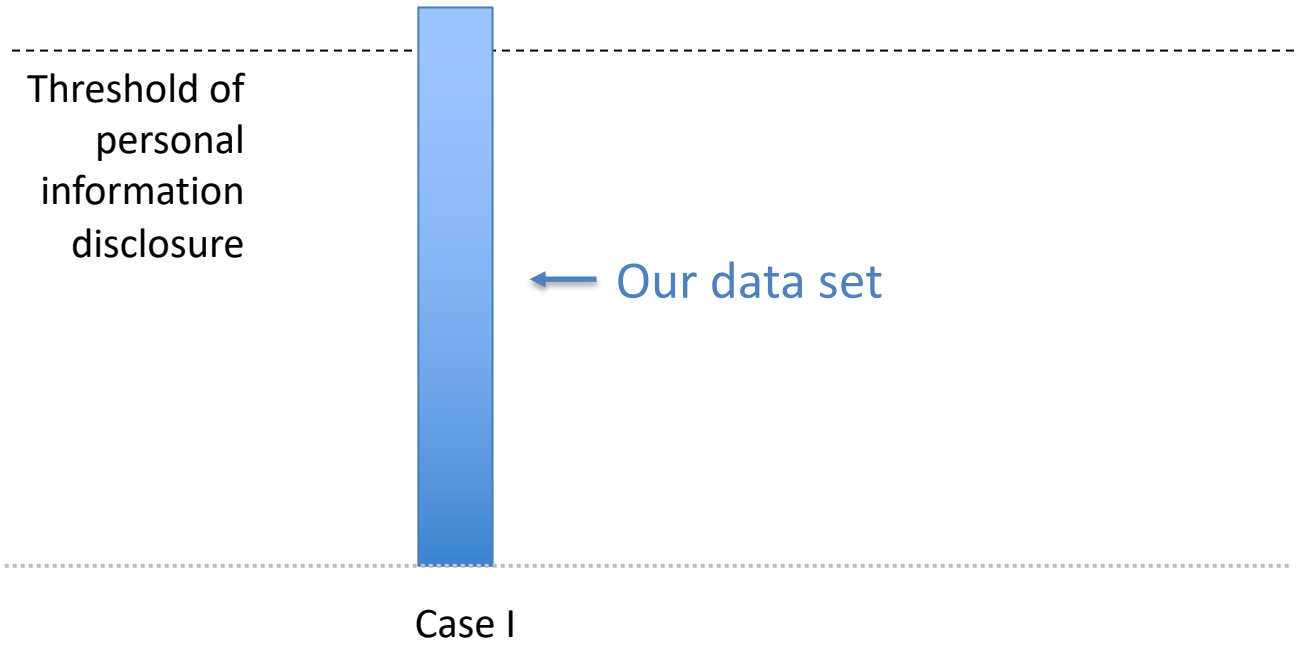
- Westin's definition (1968)
  - Control over personal information
  - Privacy is the claim of individuals, groups, or institutions to determine for themselves when, how, and to what extent information about them is communicated to others

Evolution of privacy definition

- From normative definition to formal definition
- Normative notion of privacy
  - Underlying many privacy regulations (e.g., GDPR)
- Example
  - Our dataset contains personal data if it can reveal personal information when it is combined with other datasets

Evolution of privacy definition

- From normative definition to formal definition
- Normative notion of privacy



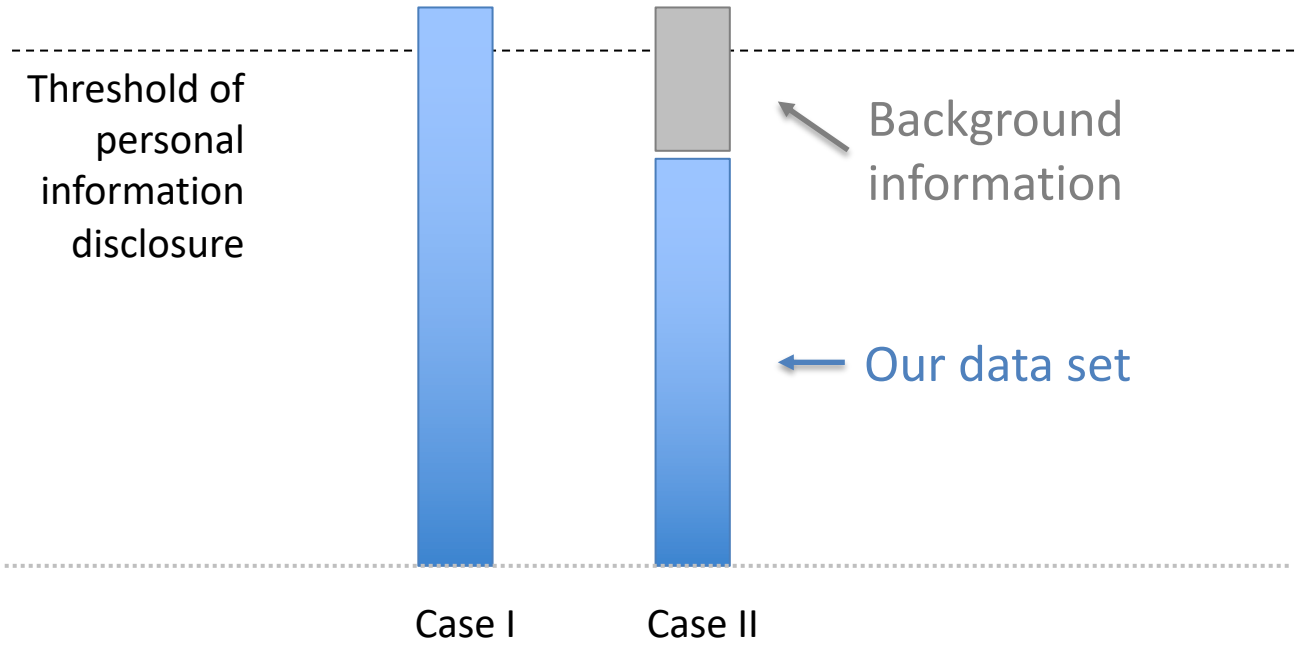
An example for case I

- Our data set: An unprotected patient table

name	**job**	**sex**	**age**	**disease**	**height (cm)**
Bob	engineer	male	35	hepatitis	184
Fred	engineer	male	38	HIV	180
Doug	lawyer	male	38	Flu	210
Alice	writer	female	30	Flu	172
Cathy	writer	female	33	HIV	170
Emily	dancer	female	31	HIV	169
Gladys	dancer	female	31	HIV	171

Evolution of privacy definition

- From normative definition to formal definition
- Normative notion of privacy



An example for case II

- A protected patient Table

name	job	sex	age	disease	Height (cm)
Bob	engineer	male	35	hepatitis	184
Fred	engineer	male	38	HIV	180
Doug	lawyer	male	38	Flu	210
Alice	writer	female	30	Flu	172
Cathy	writer	female	33	HIV	170
Emily	dancer	female	31	HIV	169
Gladys	dancer	female	31	HIV	171

An example for case II

- A protected patient Table

name	job	sex	age	disease	Height (cm)
	engineer	male	35	hepatitis	184
	engineer	male	38	HIV	180
	lawyer	male	38	Flu	210
	writer	female	30	Flu	172
	writer	female	33	HIV	170
	dancer	female	31	HIV	169
	dancer	female	31	HIV	171

An example for case II

- A protected patient Table

name	job	sex	age	disease	Height (cm)
	profess.	male	35-39	hepatitis	184
	profess.	male	35-39	HIV	180
	profess.	male	35-39	Flu	210
	writer	female	30	Flu	172
	writer	female	33	HIV	170
	dancer	female	31	HIV	169
	dancer	female	31	HIV	171

An example for case II

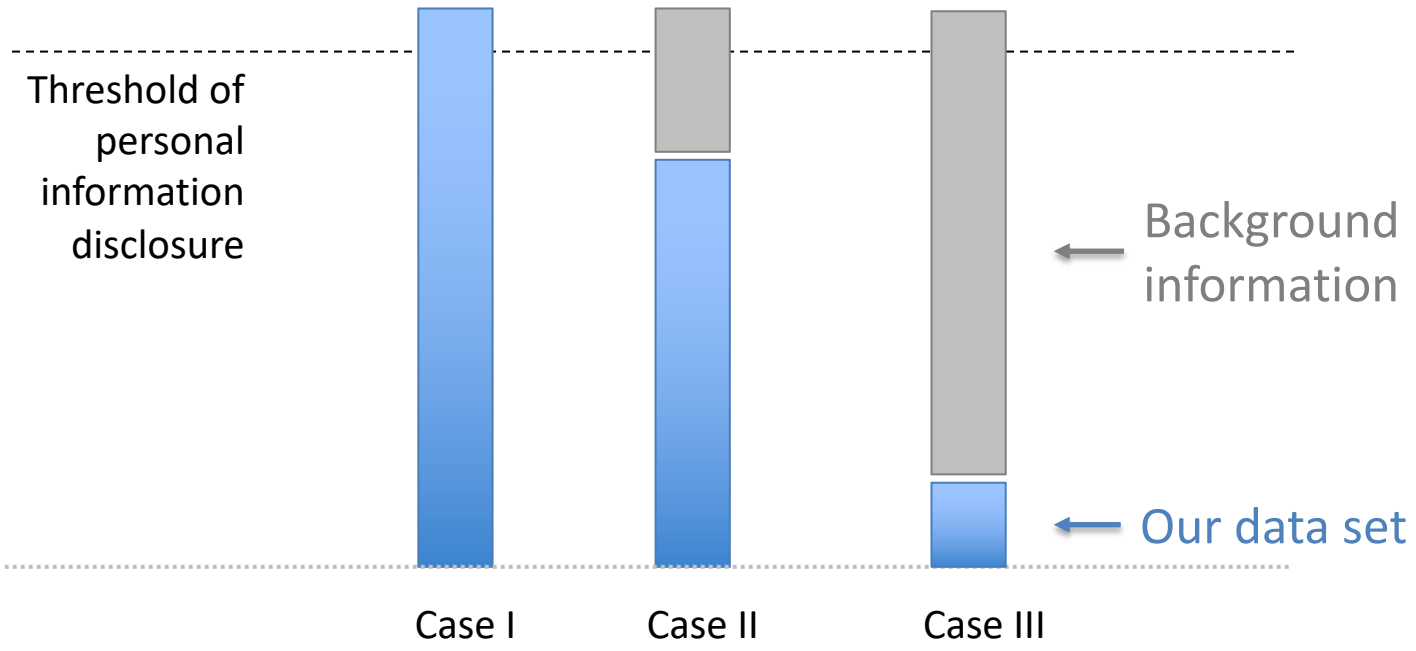
- A protected patient Table

name	job	sex	age	disease	Height (cm)
	profess.	male	35-39	hepatitis	184
	profess.	male	35-39	HIV	180
	profess.	male	35-39	Flu	210
	artist	female	30-34	Flu	172
	artist	female	30-34	HIV	170
	artist	female	30-34	HIV	169
	artist	female	30-34	HIV	171

Background information = This table belongs to people in this room

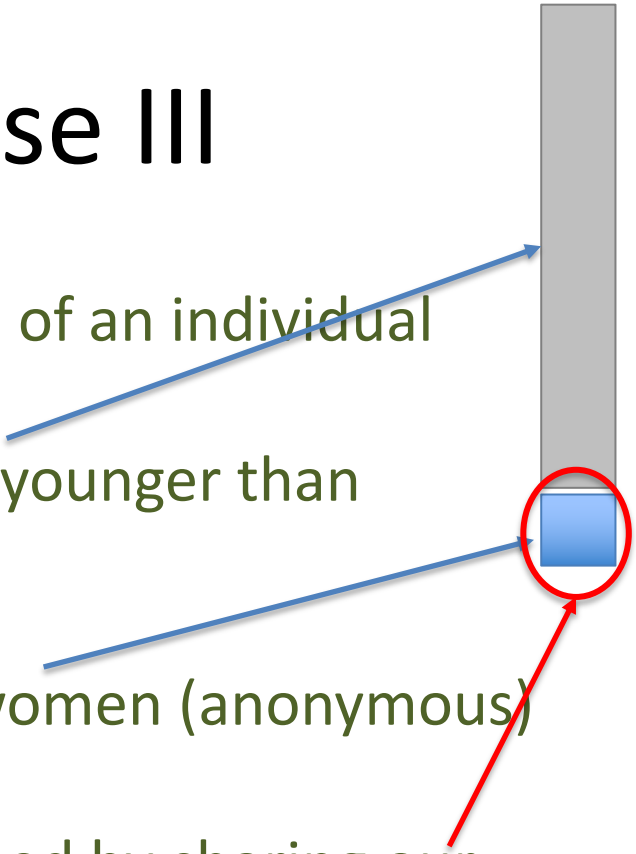
Evolution of privacy definition

- From normative definition to formal definition
- Normative notion of privacy



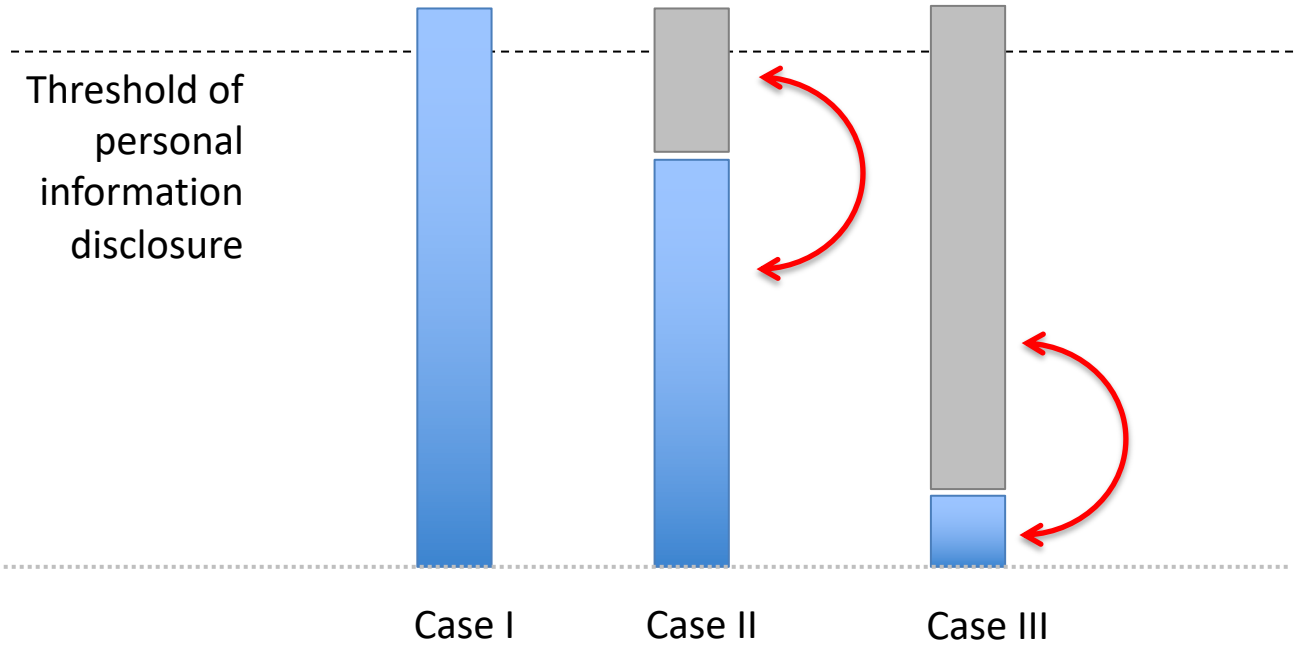
An example for case III

- **Sensitive personal information**: The age of an individual
- **Background knowledge**: Alice is 5 years younger than average American women
- **Our data set**: The ages of all American women (anonymous)
- **Question:** Is Alice's privacy is compromised by sharing our data set?
- What if Alice is not American (i.e., Alice isn't in the data set)



Evolution of privacy definition

- From normative definition to formal definition
- Normative notion of privacy



Evolution of privacy definition

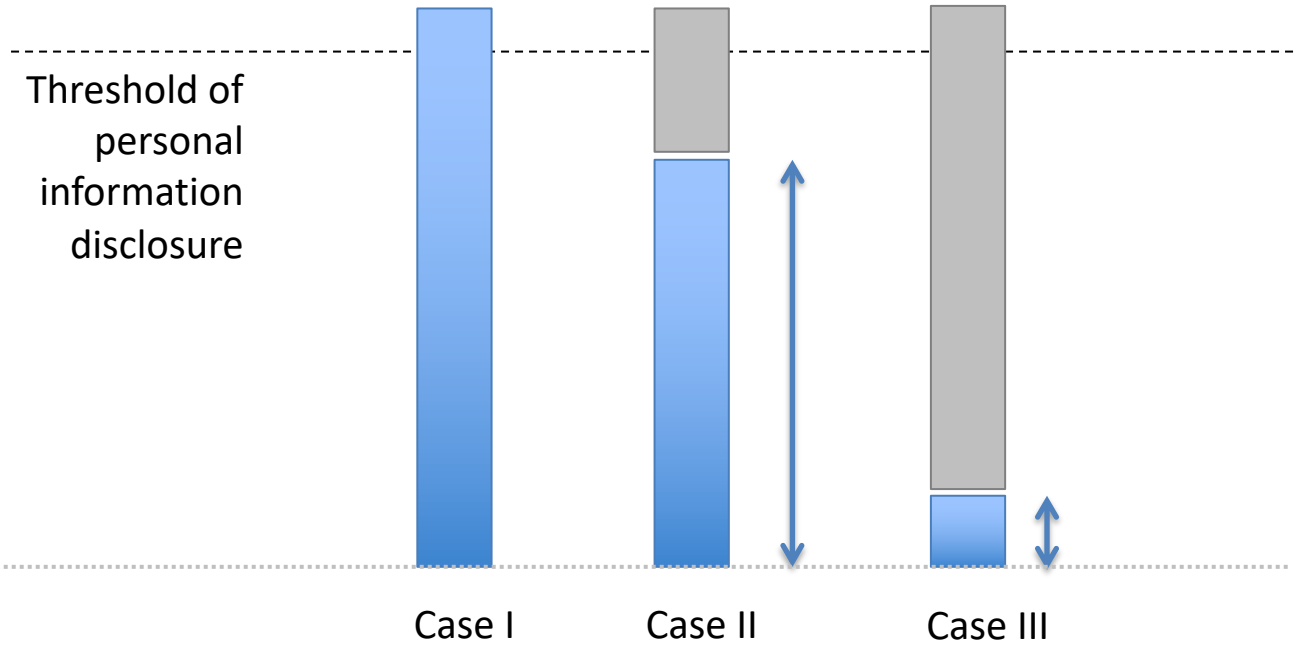
- From normative definition to formal definition

- Formal notion of privacy:
  - Dwork et al. (2006) differential privacy
    - The presence or absence of the data of an individual in a dataset must not have an observable impact on the output of a computation over the data set
  - Already in use by Google, Apple, Uber, and the U.S. Census Bureau

Nessim et al. (2018, 2019)

Evolution of privacy definition

- From normative definition to formal definition
- Normative notion of privacy



Issues of privacy definitions

- Definitions
  - The right to be let alone
  - Limited access to the self
  - Secrecy
  - Control over personal information
  - Protection of personhood
  - Intimacy

- Issues
  - Solove: Being too narrow (over restrictive) or too broad (over inclusive)
  - Changing with (technological) developments

Solove's conclusions

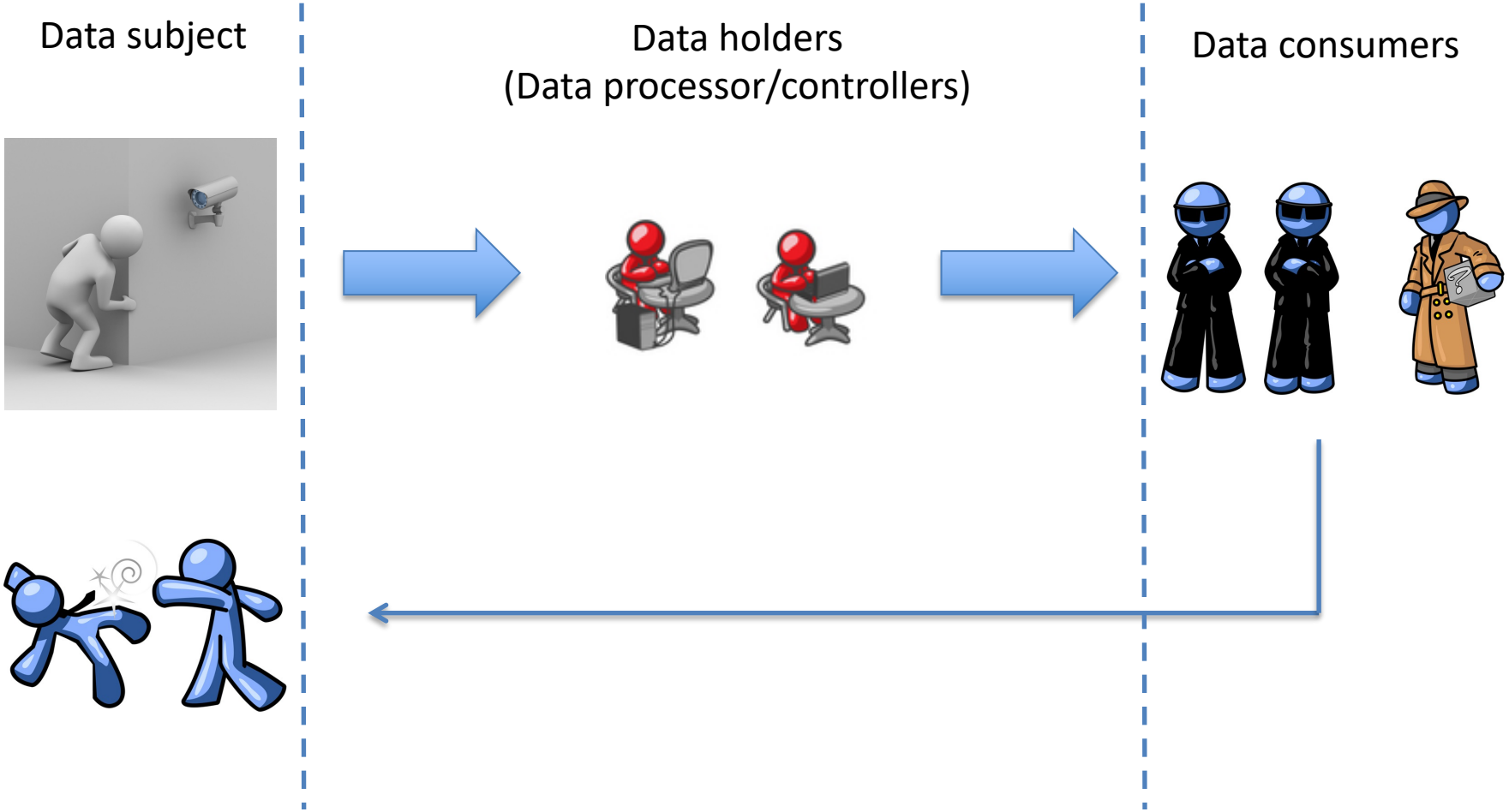
- Privacy cannot be conceptualized in a definition
  - Focusing on necessary and sufficient conditions (inclusion and exclusion game)

Solove's model

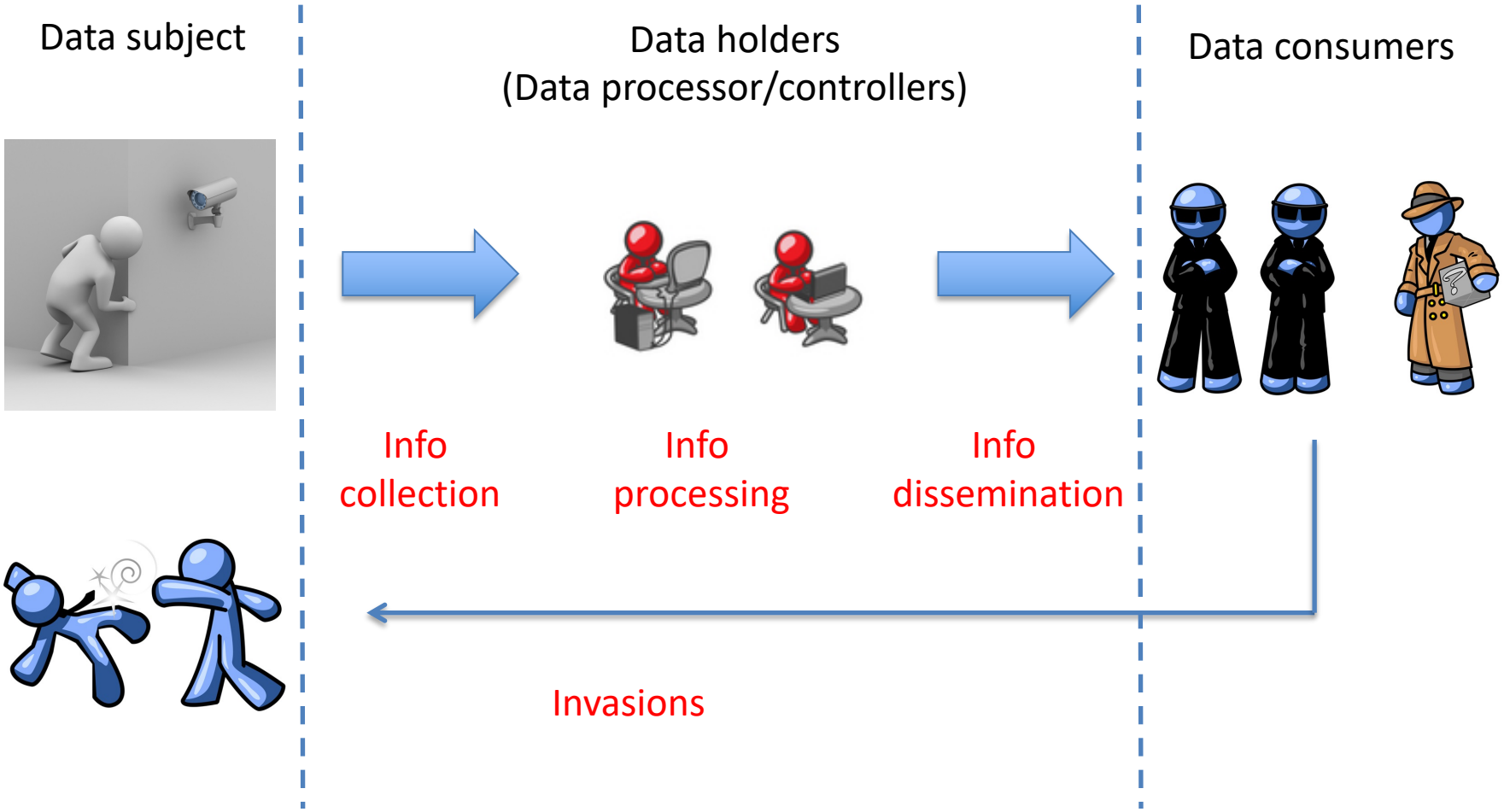
- A bottom up approach
- Focusing on
  - Harmful activities for privacy (privacy problems)
  - Rather than on what privacy is

- Our opinion
  - It can be relevant for realizing privacy by design

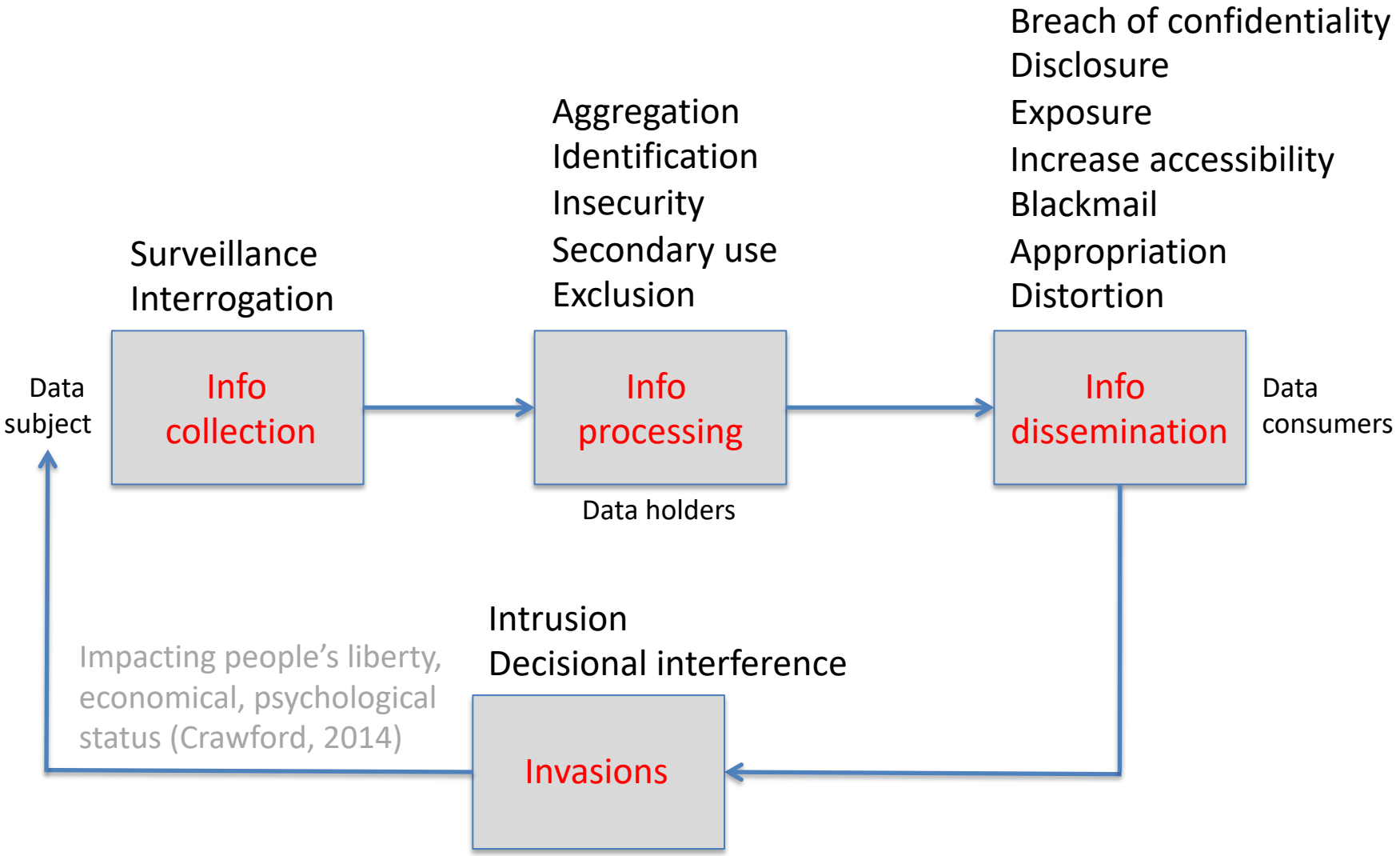
Solove's model (slightly modified)



Solove's model (slightly modified)



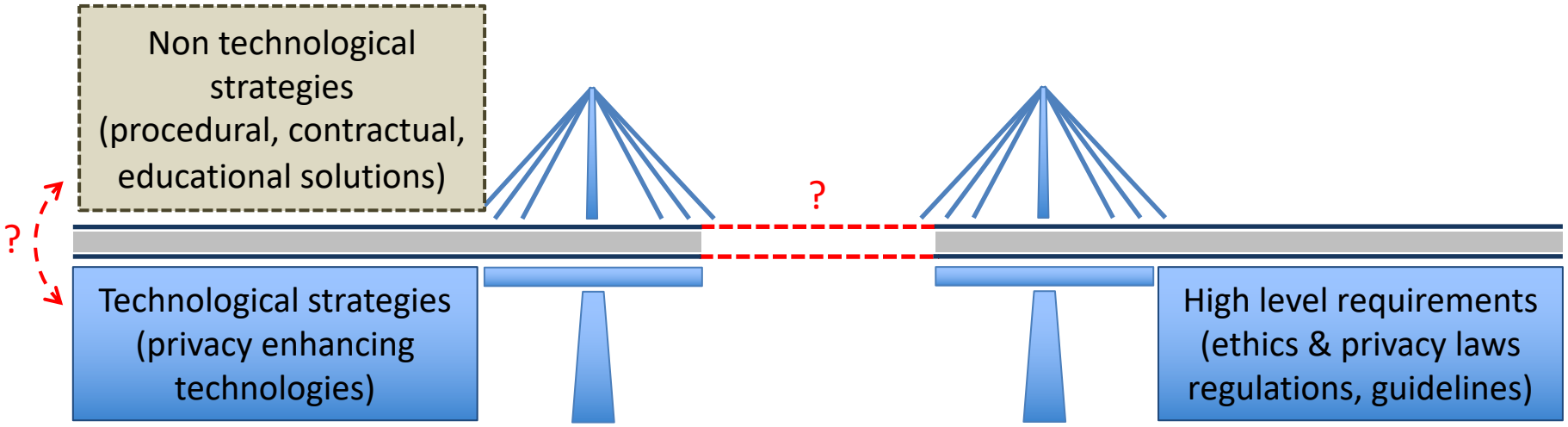
Solove's model for privacy risks



Eliciting privacy requirements

- Risk oriented
  - Identify the assets, the threats, & risks (e.g., probability × impact)

- Goal oriented
  - Privacy principles as goals that the system must fulfill
  - Example: Ensure accountability
    - Demonstrating compliance with data protection principles
  - Each high-level goal → guidelines → a set of operational requirements


Non technological strategies (procedural, contractual, educational solutions)
?
?
Technological strategies (privacy enhancing technologies)
High level requirements (ethics & privacy laws regulations, guidelines)
Solution space
Problem space

CYBERSECURITY

Cybersecurity

- “Protection of information and its critical elements, including the ICT systems (software and hardware) that use, store, and transmit that information”

  US Committee on National Security Systems (CNSS)

- Aiming at protecting a number of the, so-called, critical characteristics of information assets, whether in storage, processing, or transmission

Critical characteristics of info assets

- Confidentiality
  - To protect information from disclosure or exposure to unauthorized individuals or systems
  - For example, passwords are confidential information
- Integrity
  - To protect information so that it is complete and uncorrupted
  - For example, bank account information should not be modified
- Availability
  - To enable authorized entities to access to information without interference or obstruction, and with the required data quality
  - For example: Denial of Service (DoS) attacks prevent people accessing their bank accounts

CIA triangular

Interplay between privacy protection cybersecurity

- Dependency of privacy protection on cybersecurity
  - Two pivotal privacy principles in legal domain: data integrity and confidentiality principles (CIA)

- Dependency of cybersecurity on privacy
  - Less known
  - Relevant
    - For distributed defense against cyber attacks
    - For distributed systems (like IoT)

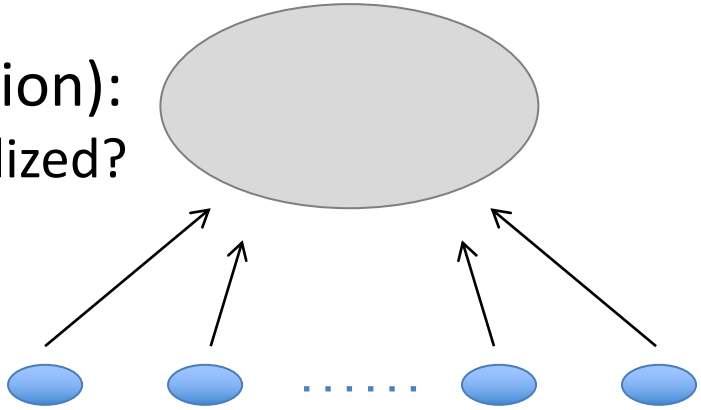
Information sharing

- A pillar of collaborative cybersecurity

- Especially in distributed settings
  - The Internet itself
  - The Internet of Things (IoT) systems
  - Distributed intrusion detection systems
  - Identity management systems
  - Etc.

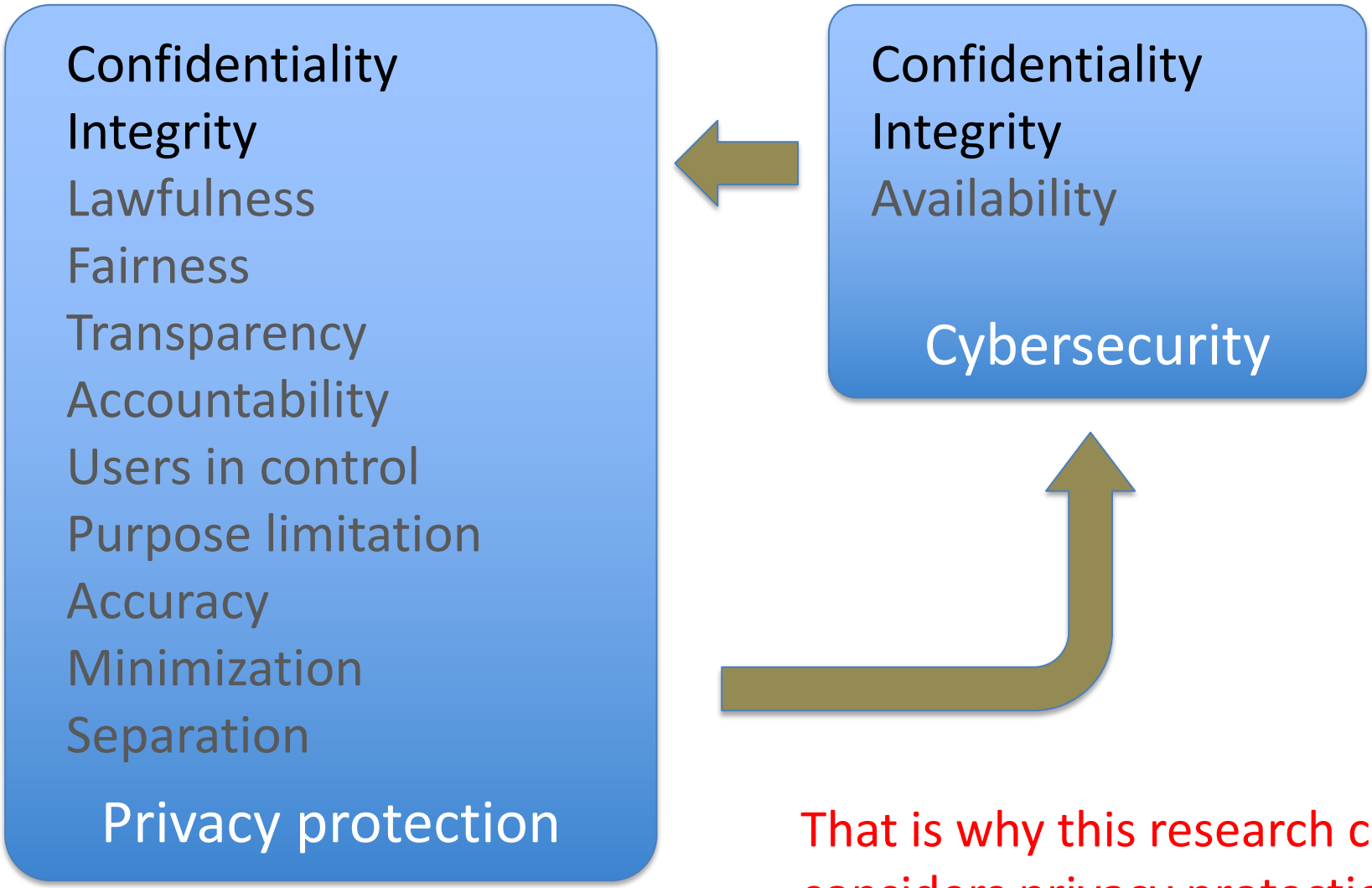
Centralized information sharing

- Data processing (e.g., privacy protection):
  - How much local and how much centralized?



heavily-centralized data processing

- Privacy issues
- Of victims being under attack (from local organizations)
- Of suspects being seen as the attacker
  - Like the IP-addresses of potential attackers
  - If done inappropriately, may lead to imposing sanctions against alleged, but not proven, cyber attackers



That is why this research chair considers privacy protection and cybersecurity together

TOWARDS A PRIVACY AND SECURITY BY DESIGN METHODOLOGY

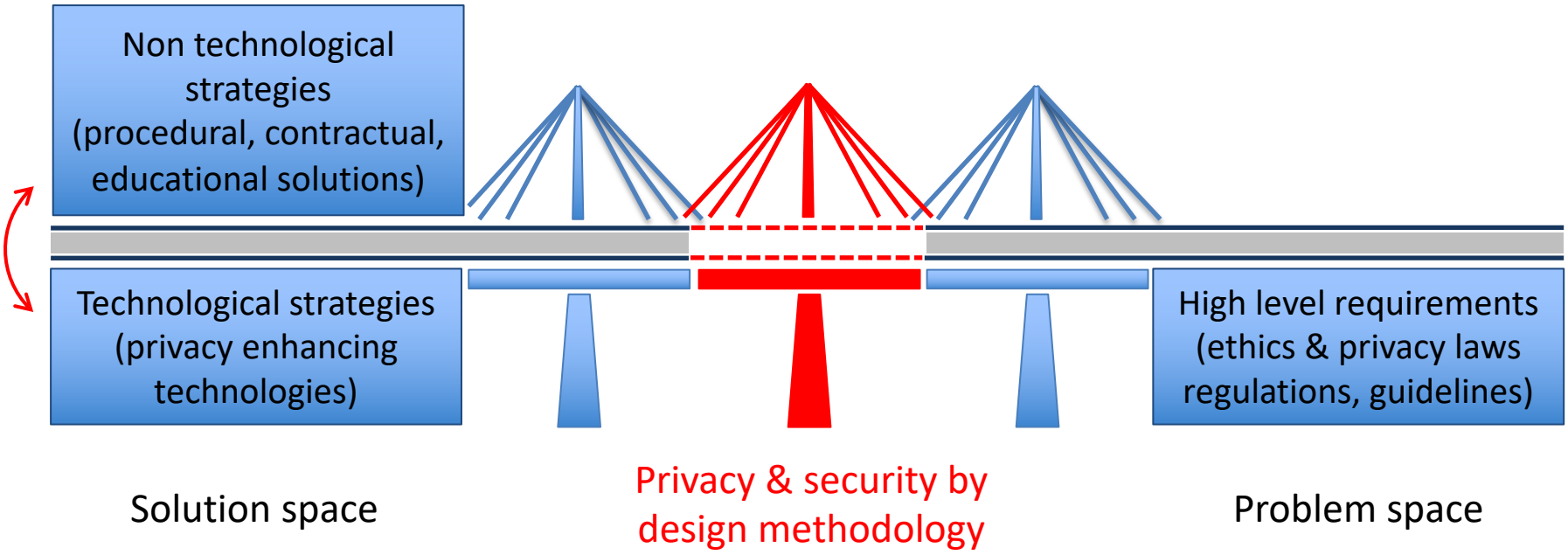
CREATING

ROTTERDAM UNIVERSITY

UNIVERSITY OF APPLIED SCIENCES

Bridging the gap

- Our goal
  - Privacy-protecting and secure info system design
- Privacy & security by design



Security & privacy by design

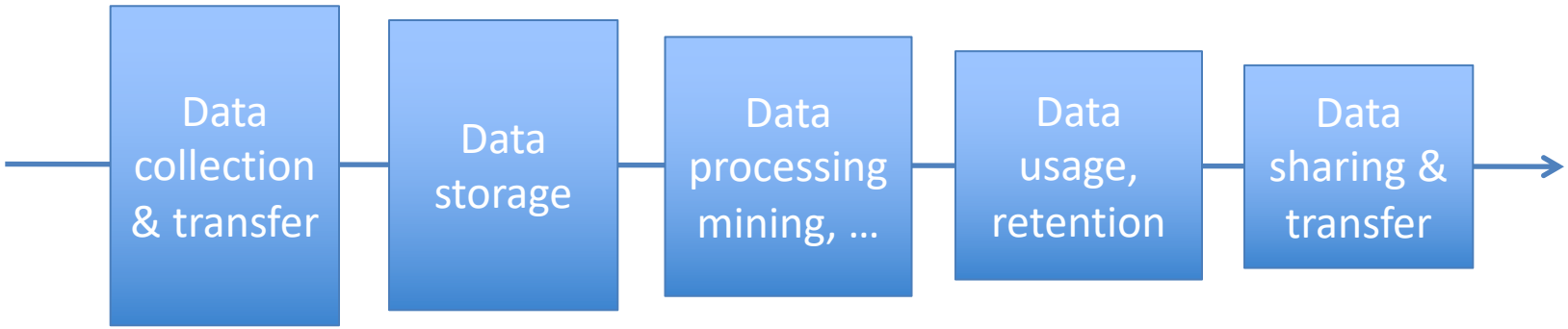
- Future research

- Vision on possible approaches
  - Engineering
  - Design thinking
  - Mix of engineering & design thinking

Spiekermann et al. 2009

Engineering privacy

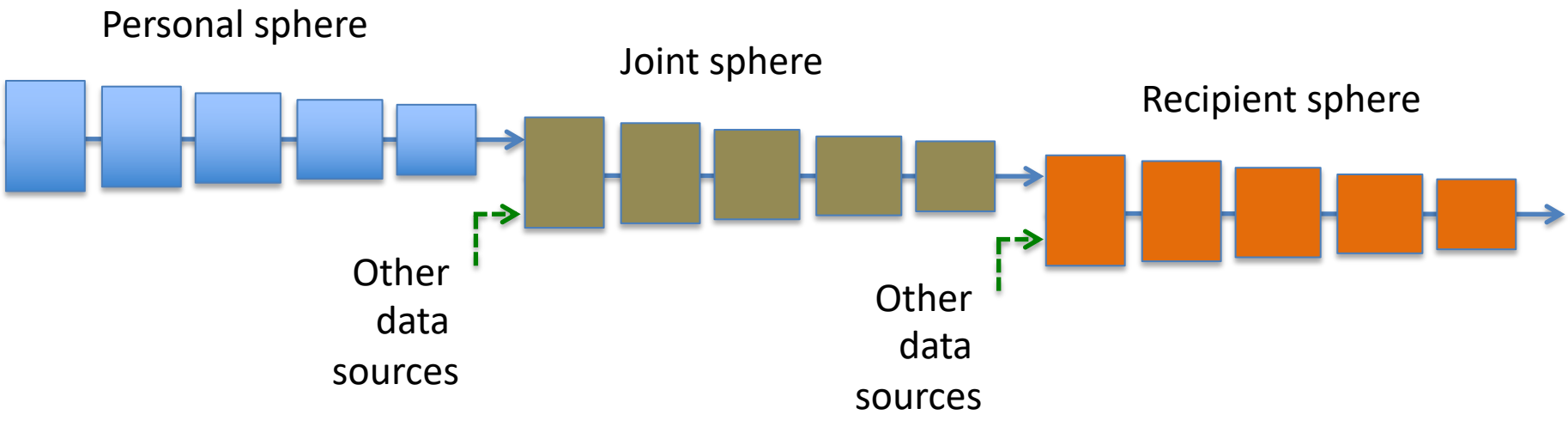
- Information systems' tasks
  - Data transfer
  - Data storage
  - Data processing
- How they are performed
- What type of data is involved
- Who uses the data



Engineering privacy

- Information systems' tasks
  - Data transfer
  - Data storage
  - Data processing
- How they are performed
- What type of data is involved
- Who uses the data
- In which sphere
  - User sphere (under control of the data subject)
  - Joint sphere (under the joint control of the data subject and service providers)
  - Recipient sphere (not under control of the data subject)

Data lifecycle (data journey)



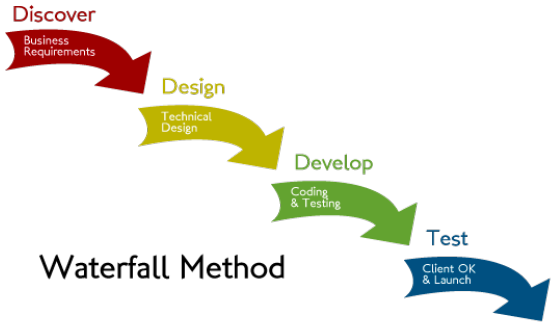
What to do do next?

Eliciting privacy requirements

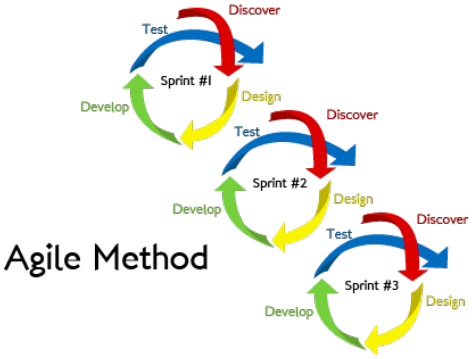
- Risk oriented
  - Identify the assets, the threats, & risks (e.g., probability × impact)
- Goal oriented
  - Like accountability
  - High-level goal → guidelines → requirements

Engineering Approaches

Waterfall model



Scrum model

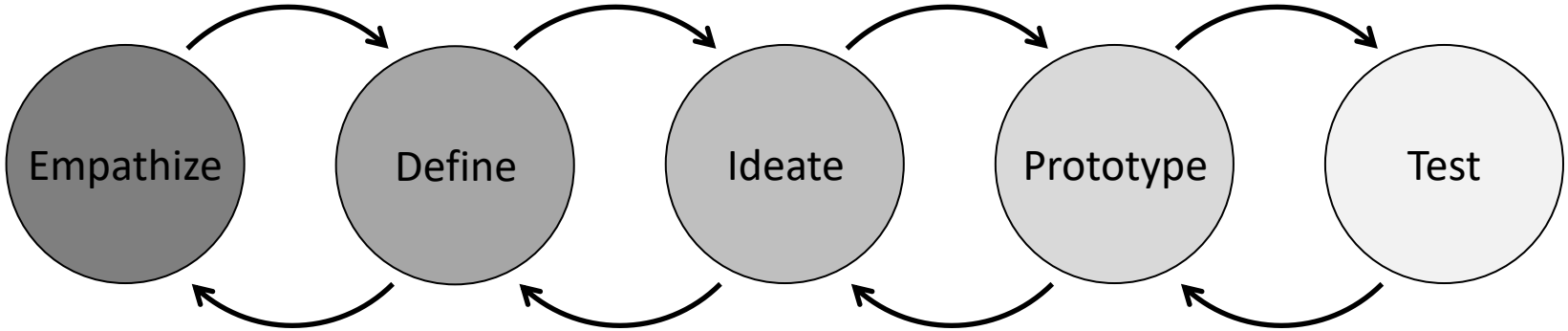


Design thinking approach

- Initially used for product and service design
- Also applied to other areas with interacting
  - People
  - Organizations
  - Technologies
- Shown useful where user needs and concerns are insufficiently communicated and formulated (being hidden in tacit knowledge)

Process of design thinking

- Empathize: Discover and understand the real concerns, problems, and experiences of stakeholders
- Define: Find out the deeper roots of the needs of stakeholders (esp. those of directly involved end-users)
- Ideate: Explore and generate solutions for the identified needs
- Prototype: Make prototypes tangible for (a subset of) the ideated viable solutions
- Test: Experiment and evaluate the prototypes with the end-users and learn from them



Characteristics of design thinking

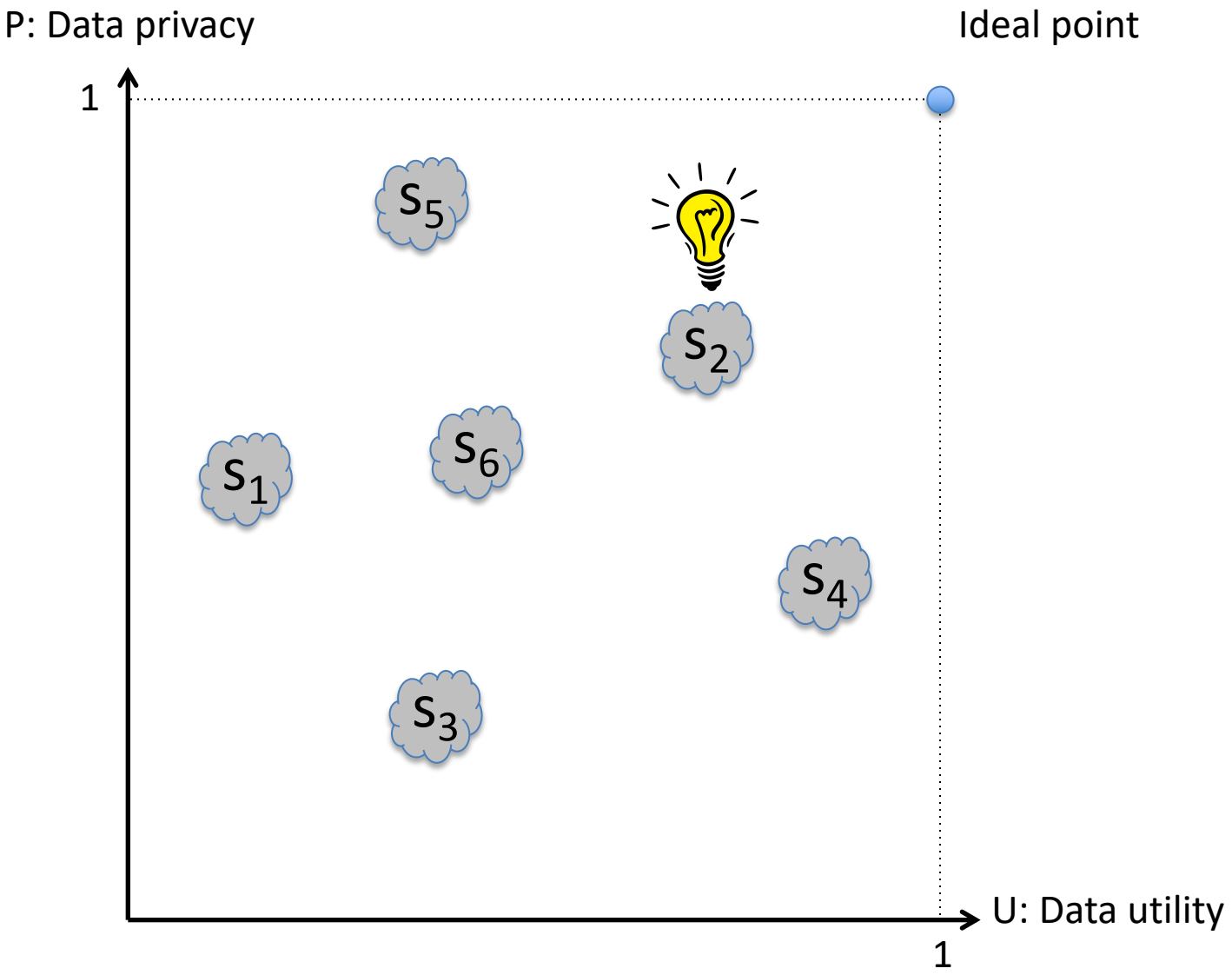
- Design process is highly collaborative and multidisciplinary
- Involving end-users to prevent disappointments
  - That the artifacts do not cater the real needs of users
- Fail fast approach to push the design process towards producing viable products
- Creating human-centric solutions
  - Being innovative
  - Based on real end-user needs
  - Being holistic in considering the contextual circumstances
  - Capable of having social impacts and changing mindsets

Design thinking in development of information systems

- Proposed for
  - Creating innovative mobile apps
  - Designing complex embedded/IoT systems
  - Devising social information systems to have positive social changes

- Proposed for improving privacy protection & security
  - Putting risk awareness into practical and collaborative action within organizations
  - Delivering more user-focused security

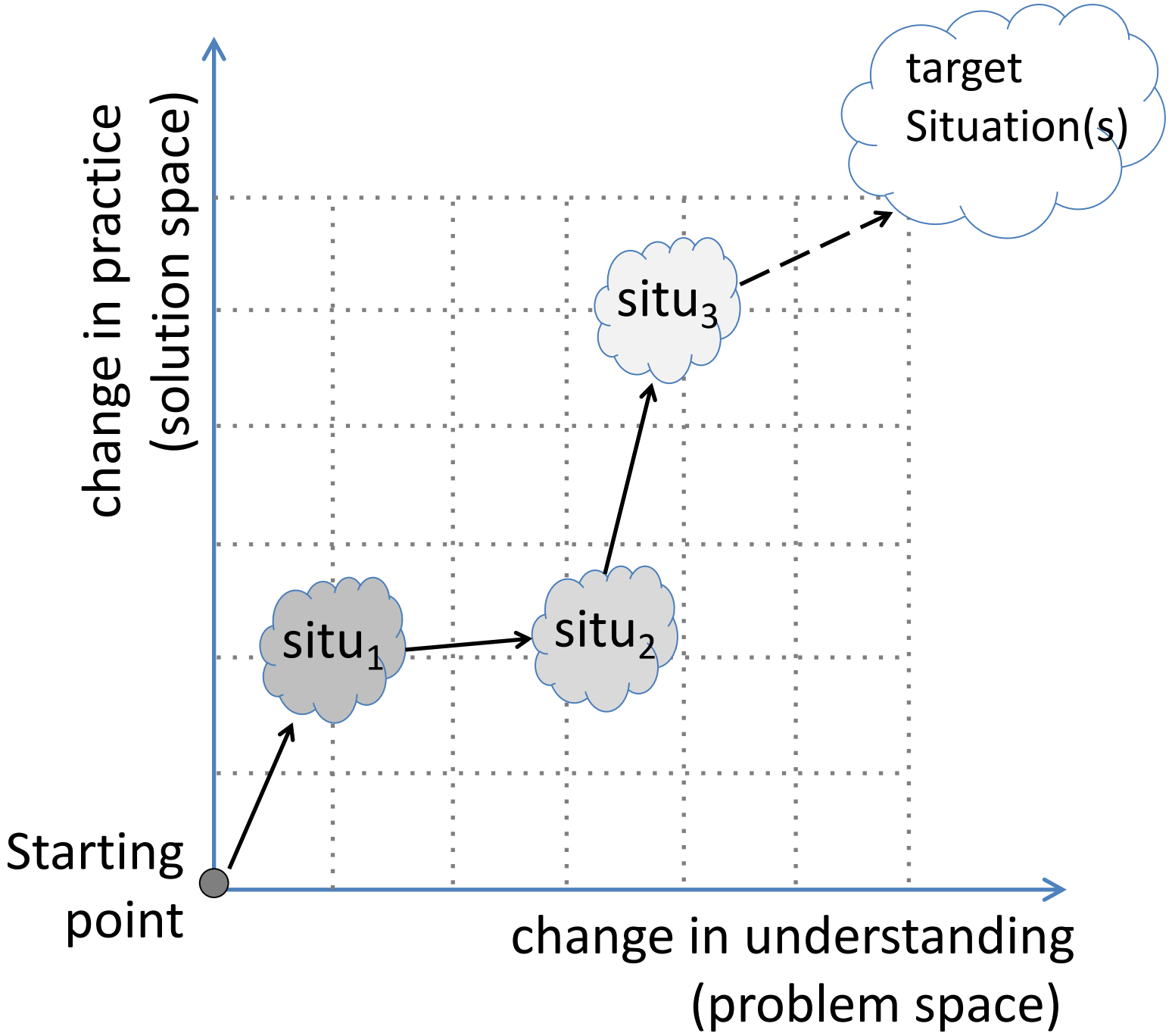
(a) Making multi-dimensional design trade-offs



(b) Making trade-offs among actionable decisions

- Those data protection measures that are going to be operationalized in a complex and possibly unpredictable social context

- Example: Open data scenarios
  - Transparency versus privacy
  - Small steps in right direction


target
Situation(s)
situ3
situ2
situ1
change in practice
(solution space)
Starting
point
change in understanding
(problem space)

AN EXAMPLE

Uitdagingen van de digitale transformatie

Problem class		Problem goals and values	
		Consensus among stakeholders	Dissensus among stakeholders
Special knowledge needed to address the problem	Certainty about facts and cause-effect	(1) Tamed or structured problems (debating on the technicalities)	(3) Weakly structured problems (debating goals and values)
	Uncertainty about facts and cause-effect	(2) Weakly structured problems (debating cause–effects and optimizing fact collection)	(4) Wicked or unstructured problems (endless debate)

Wicked Problems

1. *There is no definitive formulation of a wicked problem.*
2. *Wicked problems have no stopping rule.*
3. *Solutions to wicked problems cannot be true-or-false, only good-or-bad.*
4. *There is no immediate and no ultimate test of a solution to a wicked problem.*
5. *Every solution to a wicked problem is a "one-shot operation"; because there is no opportunity to learn by trial-and-error, every attempt counts significantly*
6. *Wicked problems do not have an enumerable set of potential solutions*
7. *Every wicked problem is essentially unique.*
8. *Every wicked problem can be considered to be a symptom of another problem.*
9. *The existence of a discrepancy representing a wicked problem can be explained in numerous ways [depending on the Weltanschauung of the designer]. The choice of explanation determines the nature of the problem's resolution.*
10. *The wicked problem solver]has no right to be wrong – they are fully responsible for their actions.*

Objectives of Corona app (CoronaMelder)

- To automatize (part of) the source and contact research (the BCO), currently conducted by the Dutch Health services (GGD)
- To alert the contacts of an infected person about a possible contamination
- To quickly test potentially infected persons, regardless of symptoms
- To notify faster those who have contacted a positively tested person
- To make more targeted use of available test capacity

Disclaimer

- No intention to evaluate (the effectiveness of) the app
  - Neither its development process
- Looking into the development process
  - From distance, not in a scientific way (like via an extensive case study)
- Used for illustrative purposes
  - As an example to show the similarity of the adopted approach to ours

CoronaMelder: A socio-technical system

- Complex
  - Many diverse stakeholders
  - Technology (i.e., the app) + the surrounding ecosystem
- High impact on society & individuals
  - May concern highly sensitive personal information (health, location, social behaviours)
- Contention between values
  - Economy vs privacy (or even rule of law / democracy)

Stakeholders

- Officially: Contributors of requirement elicitation and evaluation
  - Central government organizations (like ministry of VWS, RIVM, and GGD GHOR Nederland)
  - Civil rights public organizations (like the Netherlands Institute for Human Rights and the DPA/AP)
  - National security organizations, civil rights activists (like individuals associated with Bits of Freedom)
  - Independent legal experts (like privacy lawyers) and system developers (like software engineers, system architects and cyber security experts)
  - An advisory board to supervise the app development (15 board members, being epidemiology, virology, technology, privacy and security experts)

- How about
  - Citizens?
  - Local governments?

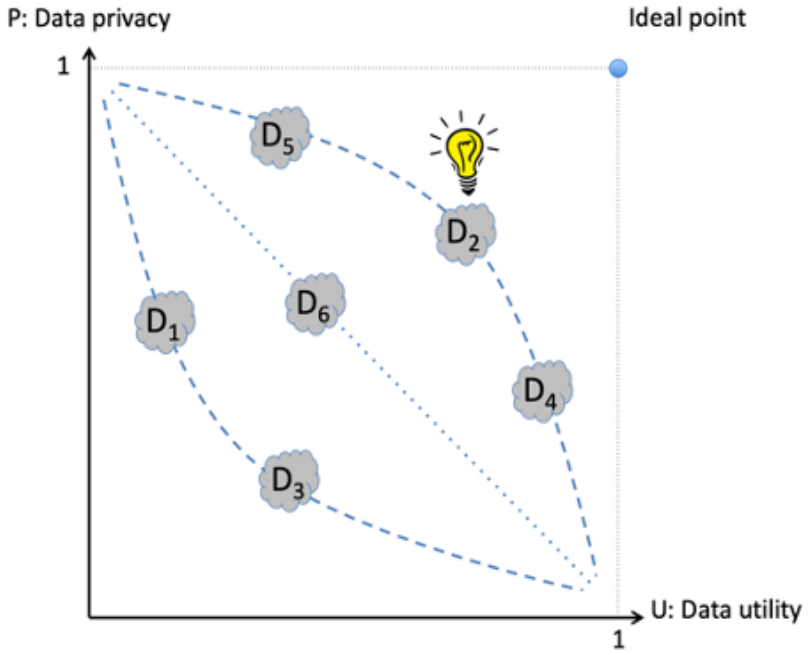
Source: https://www.frankwatching.com/archive/2020/10/08/coronamelder-making-of/

Multiple design options & choosing one

- Multiple design concepts

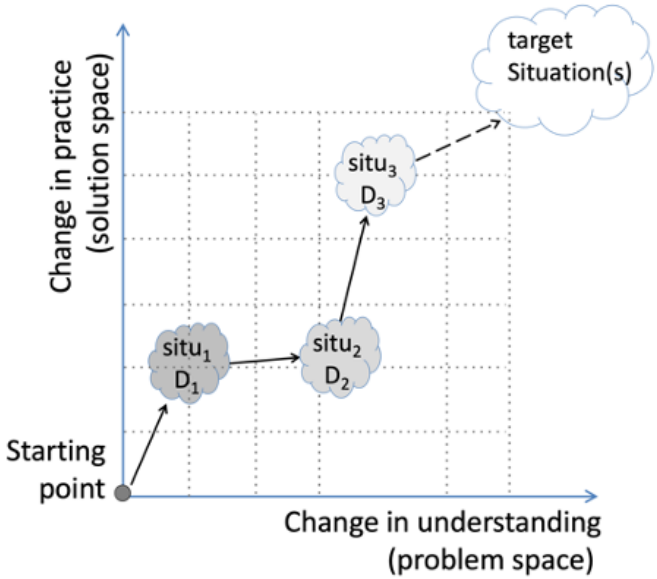
- Appathon events
  - On 18 and 19 April
  - Organized by the min VWS
  - Accessible to the public

- Designerly approach to
  - Come up with design options
  - Understand these design options
  - Choose a promising design



Incremental problem-solution specification

- Uncertainty about the chosen option
- Going through some pilots to
  - Test its capabilities
  - Identify its limitations
  - Discover its side effects
- In action



Being a wicked problem?

- The sketched cases so far are not perceived as wicked
  - Because almost all parties agree on the bigger problem (the pandemic management) and use of technology to address that
  - However, being concerned about which technology to adopt and how as well as about the unforeseeable side effects of these technologies

- Whan can it be perceived as a wicked one?
  - Rule of law vs public health
  - Conducting contact research vs other measures

REFLECTION

Takeaways

- Privacy protection and cybersecurity field
  - Covering a wide spectrum of expertise areas
  - Facing a large shortage of human capital, while its market share growing
- Privacy cannot be conceptualized in a definition
  - Should aim at identifying privacy risks
- Interplay between privacy and cybersecurity
  - Two intertwined concepts nowadays

Takeaways

- Realizing privacy/security by design principles
  - Linking the solution space and problem space
  - Solution space
    - Technological measures
    - Non-technological measures (e.g., procedural, educational and contractual)
- A need for a systematic design methodology
  - Design-thinking
  - Conventional engineering

UNIVERSITY OF
APPLIED SCIENCES

exceed expectations