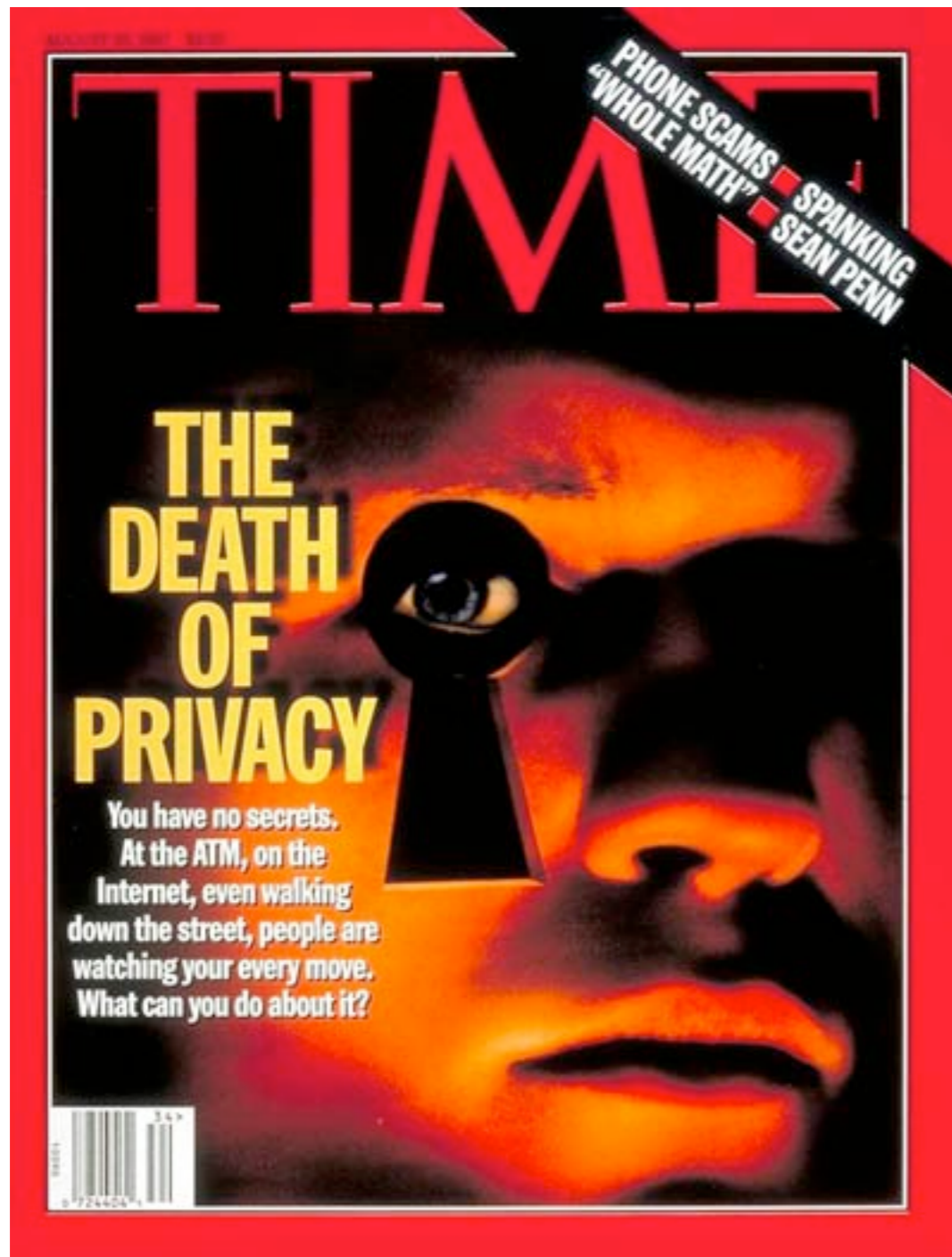# Privacy Research Paradigms
Privacy Engineering
and the Agile Turn

Seda Gürses
sguerses @ esat.kuleuven.be
COSIC, University of Leuven
CITP, Princeton University

19. September 2016
ULD Sommerakademie

PET SEMATARY

getting privacy engineering right?

getting privacy engineering right?

privacy research ←→ software engineering practice

can it be that the practices around the production of software are an important element of privacy research?
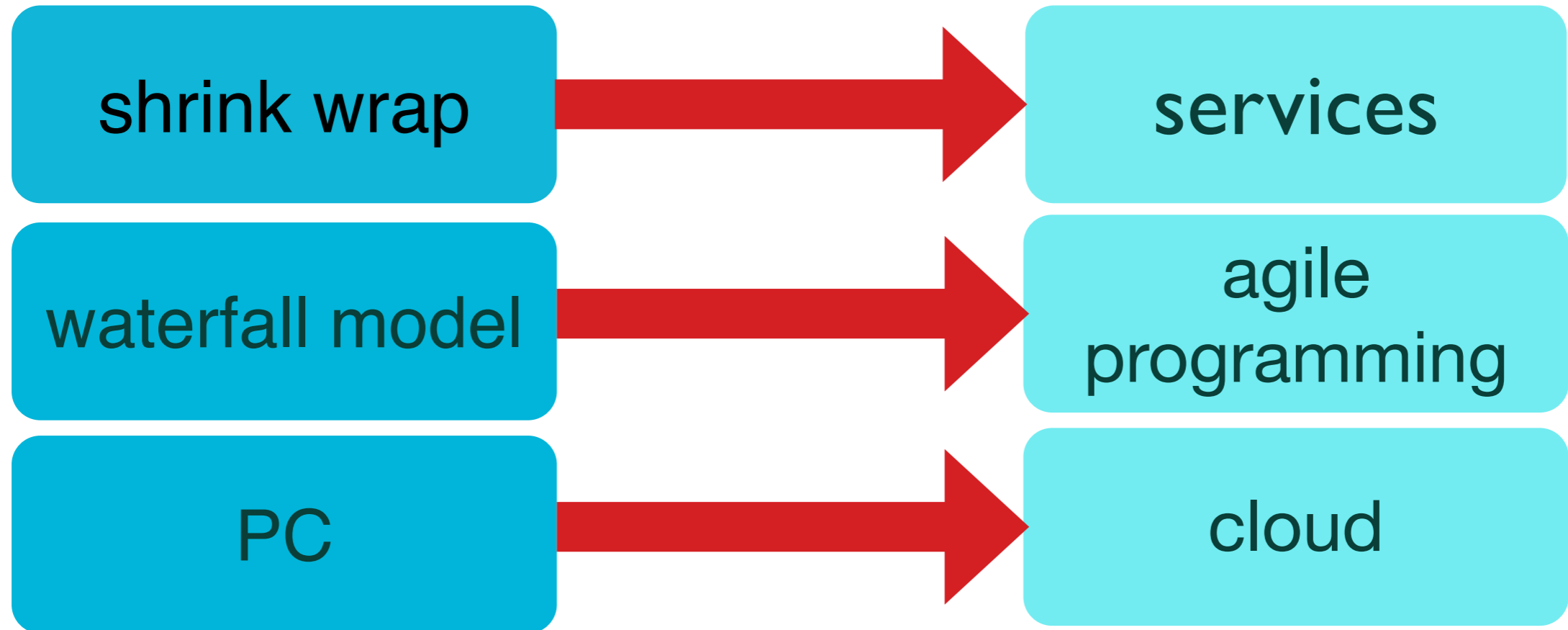
privacy research ← software engineering practice

Wurstküche
How the Sausage Gets Made

matters?

the turn to agile

shrink wrap ➡ services

waterfall model ➡ agile programming

PC ➡ cloud

what is the impact of

the turn to agile in software engineering practice

on computer science research in privacy?

what is the impact of

the turn to agile in software engineering practice

on

computer science research in privacy?

PRIVACY RESEARCH PARADIGMS

privacy as confidentiality

privacy as control

privacy as practice

PRIVACY RESEARCH PARADIGMS

privacy as confidentiality

"the right to be let alone"
Warren and Brandeis

data minimization

properties with mathematical guarantees

avoid single point of failure

open source - it takes a village to keep it secure

PRIVACY RESEARCH PARADIGMS

privacy as confidentiality

secure messaging

anonymous communications

All Tools	Encrypted in transit?	Encrypted so the provider can't read it?	Can you verify contacts' identities?	Are past comms secure if your keys are stolen?	Is the code open to independent review?	Is security design properly documented?	Has there been any recent code audit?
Off-The-Record Messaging for Mac (Adium)	✅	✅	✅	✅	✅	✅	🚫
Off-The-Record Messaging for Windows (Pidgin)	✅	✅	✅	✅	✅	✅	✅
PGP for Mac (GPGTools)	✅	✅	✅	🚫	✅	✅	🚫
PGP for Windows Gpg4win	✅	✅	✅	🚫	✅	✅	🚫

PRIVACY RESEARCH PARADIGMS

privacy as control

"right of the individual to decide what information about himself should be communicated to others and under what circumstances" Westin

data protection/FIPPS compliance

transparency and accountability

individual participation and control

PRIVACY RESEARCH PARADIGMS

privacy as control

privacy policy languages

purpose based access control

Bell Group

information we collect	ways we use your information				information sharing	
	to provide service and maintain site	marketing	telemarketing	profiling	other companies	public forums
contact information	■	opt in		■	opt out	
cookies	■			■		
demographic information	■	opt in		■	opt out	
financial information						
health information						
preferences	■					
purchasing information	■	opt in		■	opt out	
social security number & gov't ID						
your activity on this site	■	opt in		■	opt out	
your location						

Access to your information
This site gives you access to your contact data and some of its other data identified with you

How to resolve privacy-related disputes with this site
Please email our customer service department

bell.com
5000 Forbes Avenue
Pittsburgh, PA 15213 United States
Phone: 800-555-5555
help@bell.com

PRIVACY RESEARCH PARADIGMS

privacy as practice

"the freedom from unreasonable constraints on the construction of one's identity" Agre

improve user agency in negotiating privacy

privacy integral to collective info practices
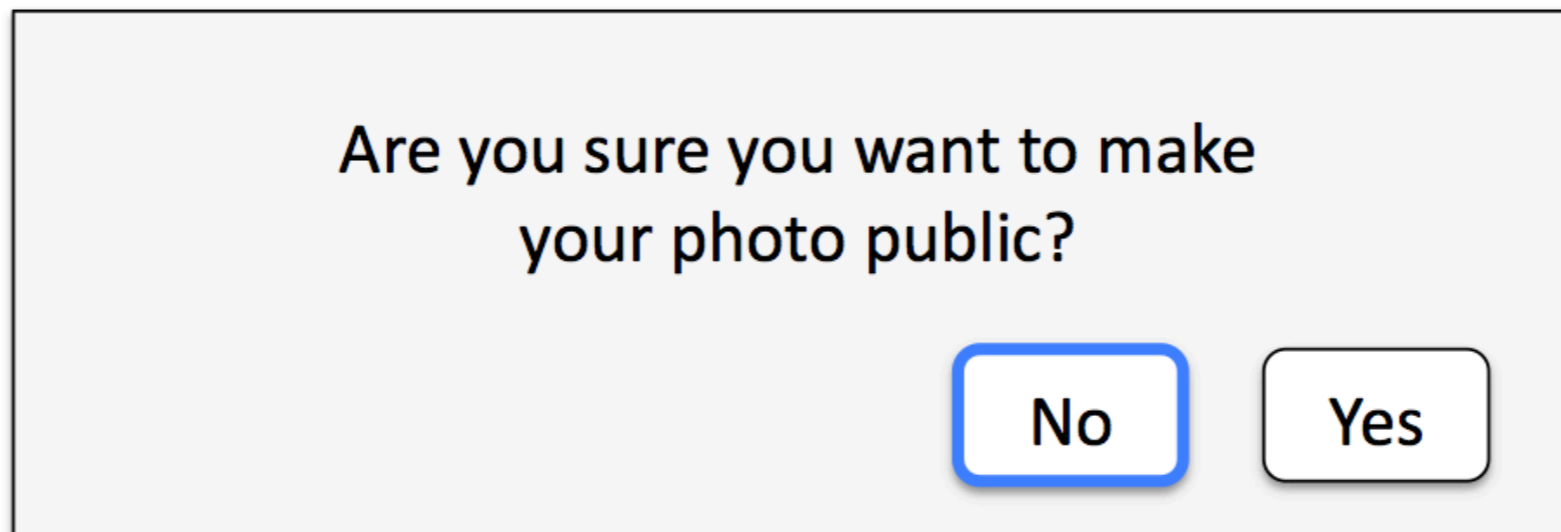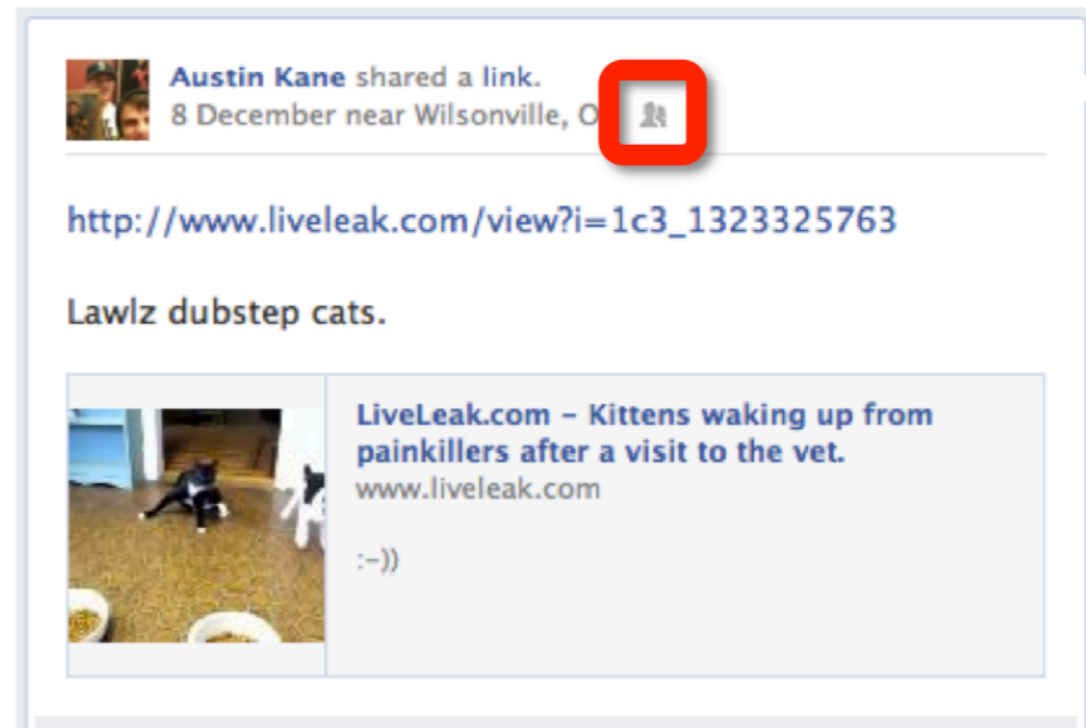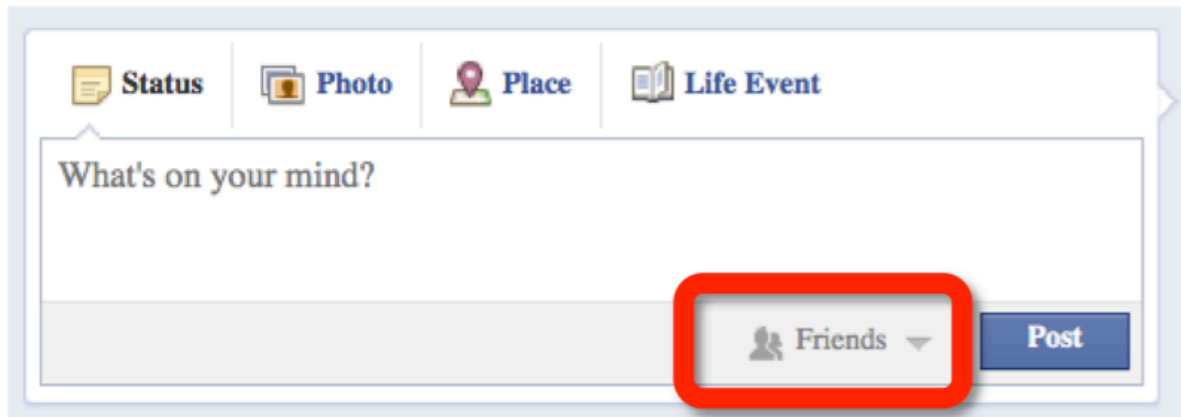
aid in privacy decision making

transparency of social impact

PRIVACY RESEARCH PARADIGMS

privacy as practice

feedback & awareness design

privacy nudges

Status Photo Place Life Event

What's on your mind?

Friends ▾   Post

Austin Kane shared a link.
8 December near Wilsonville, O

http://www.liveleak.com/view?i=1c3_1323325763

Lawlz dubstep cats.

LiveLeak.com – Kittens waking up from
painkillers after a visit to the vet.
www.liveleak.com

:-))

Are you sure you want to make
your photo public?

No   Yes

slide: Lorrie Cranor

How big data is unfair

Understanding unintended sources of unfairness in data driven decision making
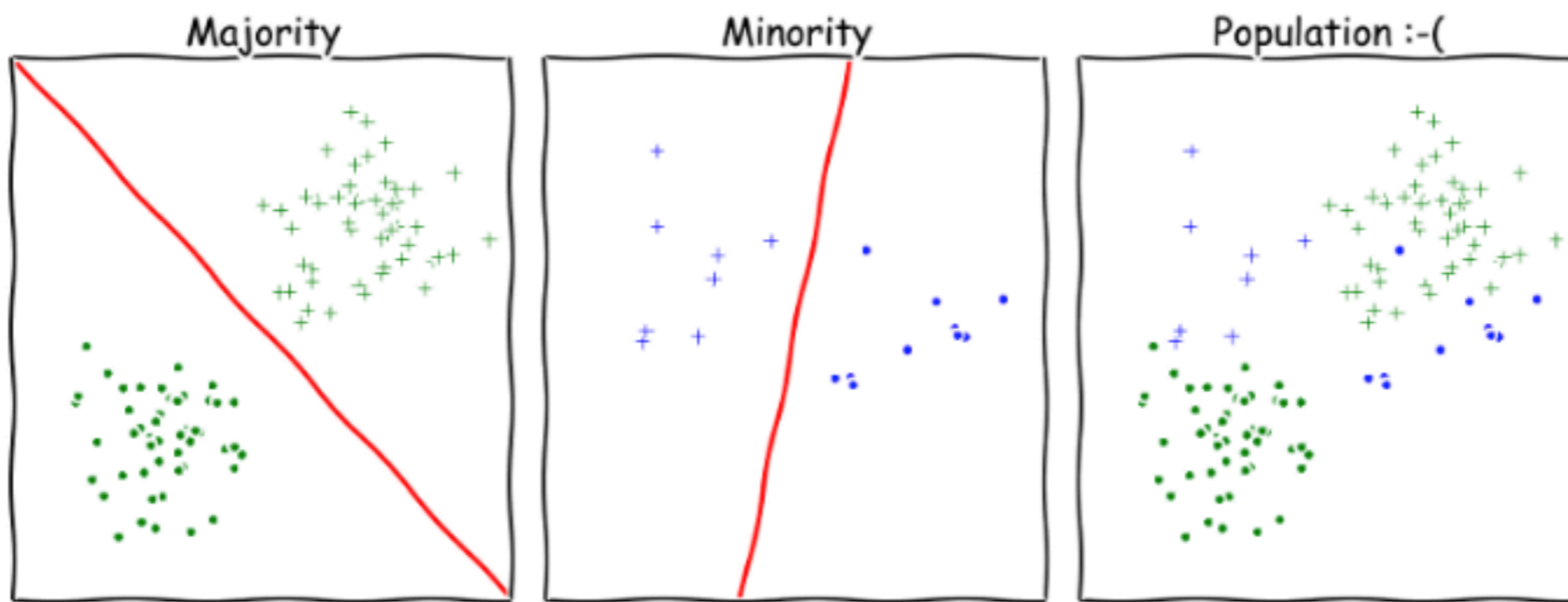
Even if two groups of the population admit simple classifiers, the whole population may not.

PRIVACY RESEARCH PARADIGMS

privacy as confidentiality

privacy as control

privacy as practice

diversity in problems & solutions

integration

systematization

generalization

practice

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.

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

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			linkable with reasonable & automatable effort	• no unique identifies across databases • common attributes across databases • contact information stored separately from profile or transaction information
2	pseudonymous	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

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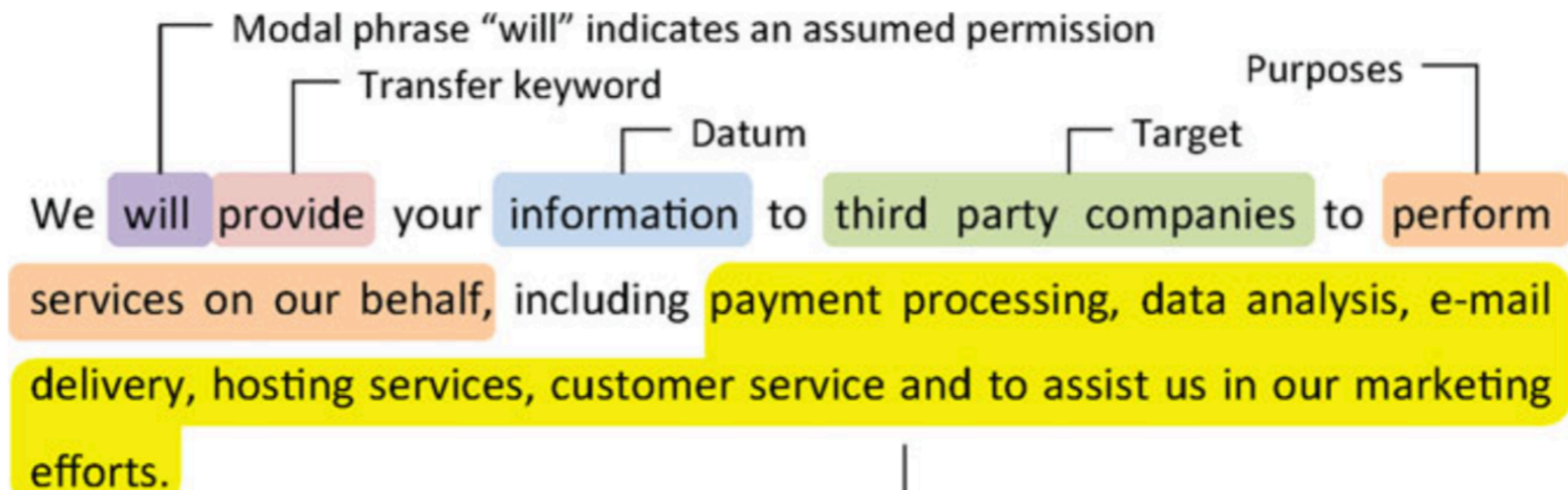
tools:
(automated) means that support privacy engineers during part of a privacy engineering process.

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

TECHNISCHE
UNIVERSITÄT
DARMSTADT

Comparison

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	

socio-technical systems

standalone privacy technology	privacy enhancement of system or function	research into privacy violations
Tor/PreTP	privacy policy languages	web census

future research needs

empirical studies:
how are privacy issues being addressed in engineering contexts?

machine learning and engineering:
methods, techniques and tools to address privacy, fairness and semantic power

frameworks and metrics:
for evaluating efficacy of privacy engineering methods, techniques and tools
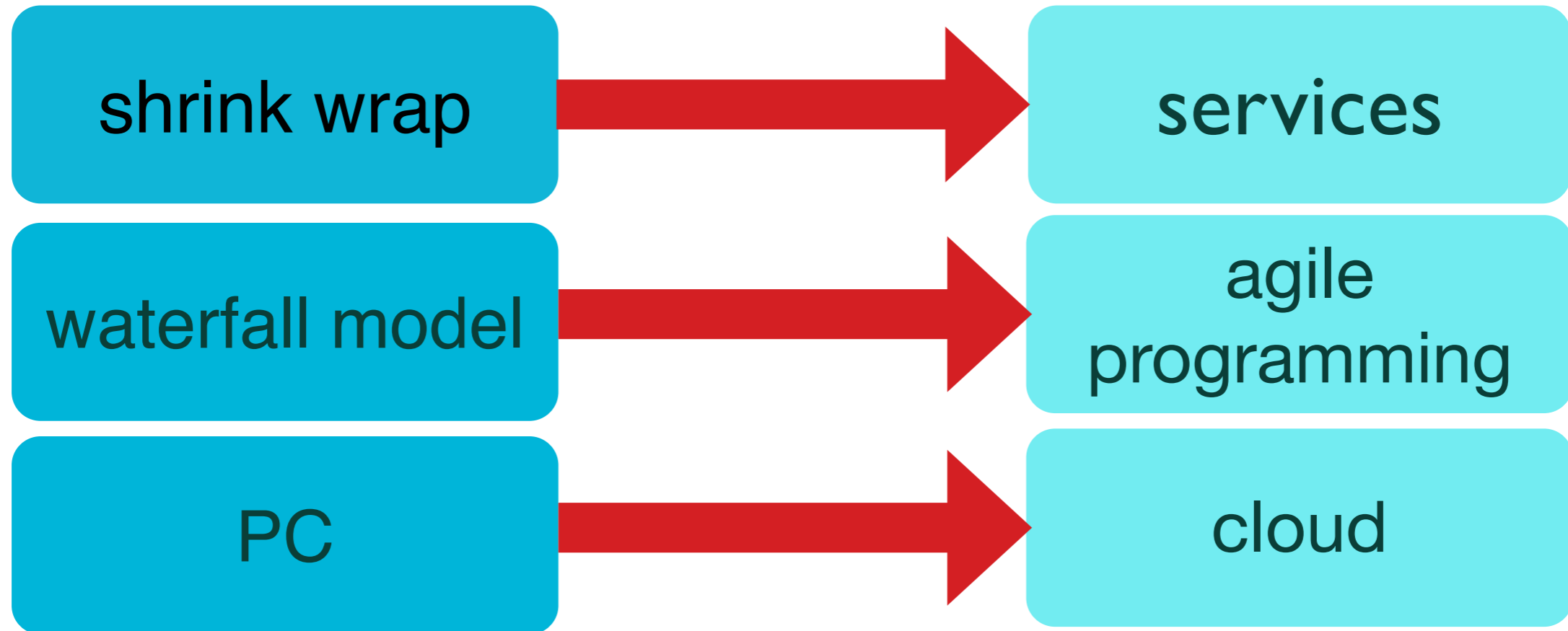
what is the impact of the turn to agile in software engineering practice on computer science research in privacy?

shrink wrap software

the turn to agile

shrink wrap	→	services
waterfall model	→	agile programming
PC	→	cloud

shrink wrap → services

shrink wrap	enterprise	apps	services
binary runs solely on client side			server (thin) client model
requires matching soft & hardware			data "secured" by service
updates & maintenance cumbersome			updates and maintenance server side
user has control (oh no!)			collaborative
pay in advance			pay as you use/trial
Microsoft Word			office 365

implications of the shift to services
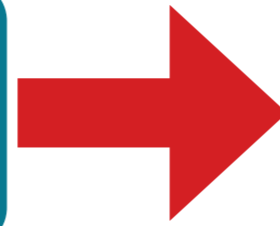
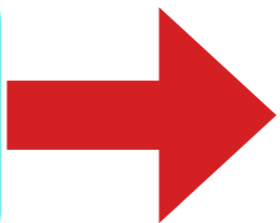|		
server – thin client model	→	transaction throughout use
bundled services	→	agile service integration
		pooling of data
licensing and pricing models	→	intensified tracking
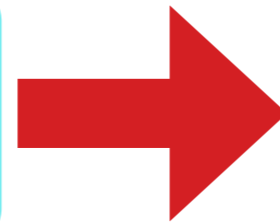
shrink wrap software production

version + purchase

use

time

service bundle

pay per use

use

team integration | SDK/PaaS | cybersecurity | performance

CRM | data brokers | analytics | AB Testing | UX capture
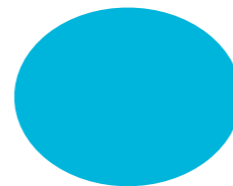
production tools

advertisement

authentication | payment | maps | social | embedded media

picture album creation service

fullstory in top 1 million sites

http://uservoice.com

http://sproutvideo.com

http://startapp.com

http://fitocracy.com

http://meuspedidos.com.br

http://oyorooms.com

http://urbanclap.com

http://himalayastore.com

http://travelport.com

http://credomobile.com

http://deputy.com

http://remitly.com

http://wahoofitness.com

http://wayup.com

http://tieks.com

http://referralcandy.com

http://codeschool.com

http://owler.com

http://surfdome.com

http://autopilothq.com

http://conte.it

http://autoeurope.com

http://moosejaw.com

http://clickminded.com

http://keen.io

http://samcart.com

http://thebouqs.com

http://mymove.com

http://scripted.com

http://namely.com

http://shethinx.com

http://castorama.pl

http://nexojornal.com.br

waterfall model → agile programming

waterfall model

requirements analysis and specification

architectural design

implementation and integration

verification

operation and maintenance

agile manifesto

individuals and interactions

process and tools

working software

comprehensive documentation

customer collaboration

contract negotiation

responding to change

following a plan

implications of the shift to agile dev

testing testing testing → user centric development
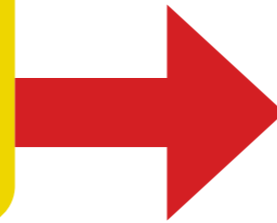
data centric development
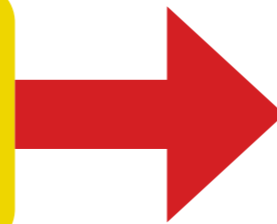
short iterations → rapid feature development
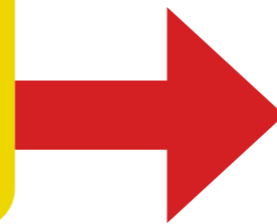
simplicity → reuse and modularity

server - thin client model

data centric development

anecdotes

data products

metrics

user/behavioral analytics

data centric development

predictive modeling 4 pricing

user churn

time

feature space

service bundle

pay per use

consent

use

how is all this fluffy management stuff relevant to privacy research?

INFORMATION PROVIDERS GUIDE
The EU Internet Handbook

European Commission > IPG > Basics > Legal requirements > 3rd party tools

RESOURCES: Our tools | Our services | Library | Standards | Quality assurance | Procedures | Training | Rules   SEARCH

FOLLOW THE STEPS: HOME   BASICS   PLAN   CONTENT   DESIGN   BUILD   GO LIVE   MAINTAIN

Basics

What is EUROPA?

Structure of EUROPA

Web addresses/URLs

Management

Legal requirements

- Legal notices and copyright

- Cookies

- Data Protection

- Sensitive information 🔒

- **Third party tools**

EUROPA digital transformation

Use of third-party tools and services

IPG rule   Third party services **are not allowed on EUROPA. Webmasters must use in-house solutions and not third party tools.**

View all IPG Rules

Third-party tools and services carry considerable continuity, accuracy and privacy risks and their use on EUROPA websites is therefore not allowed. Webmasters must use in-house solutions.

On this page

→ Description

→ Use on EUROPA websites

→ Risks
  → Privacy and data protection
  → Business continuity is not guaranteed
  → Legal uncertainty
  → Dependency on third party
  → Limited accuracy assurance, dubious data comparability
  → Internet security risks
  → Endorsement

Description

Many companies offer "free" tools, services, plug-ins or widgets that provide extra features and functionalities on websites. Use of these tools generally requires registration on the site and acceptance of the companies' terms of use. Examples include Google analytics or Statscounter to analyse site traffic; Bing maps for geographical information; AddThis to share or bookmark; YouTube for videos; Facebook social plug-ins an extension of Facebook in other site; Twitter plug-ins, etc.

OUTLOOK

- Privacy research will need to speak to existing SE approaches

  - domain specificity not enough: SE practices matter

- Future research: systemic empirical study of the agile turn

  - evaluate the paradigmatic principles that guide privacy research

  - study feature inflation and its impact on activities/privacy

  - behavioral analytics role in software engineering

  - the politics of new service metrics

- Investigate policy implications:

  - DP was developed during the time of mainframes!!!

Publications

- Privacy Engineering: Shaping an emerging field of Research and Practice, w. Jose del Alamo, https://bit.ly/27Te955

- Privacy after the Agile Turn, w. Joris van Hoboken https://osf.io/ufdvb/

PRIVACY RESEARCH PARADIGMS

privacy as confidentiality

negative

website fingerprinting (Juarez et al., CCS 2014)

positive

obfuscation (location: Shokri, query:Nissenbaum)

anonymouth (McDonald et al., PETs, 2012)

co-evolution

differentially private recommender systems (McSherry et al, SIGKDD, 2009)

privacy preserving deep learning (Shokri & Shmatikov, CCS, 2015)

integrating PETs into agents, (Such et al., Knowledge Engineering Review, 2013)

PRIVACY RESEARCH PARADIGMS

privacy as control

negative

bypassing access control (PowerSpy, Michalevsky et al., USENIX, 2015)

positive

automatically analyzing privacy policies (Zimmeck, USENIX, 2014)

mining privacy goals from policies (Bhatia et al., TOSEM, 2016)

co-evolution

discrimination discovery, characterization and prevention (FATML)

A multidisciplinary survey on discrimination analysis
(Romei and Ruggieri, Knowledge Engineering Review, 2013)

PRIVACY RESEARCH PARADIGMS

privacy as practice

negative

facebook emotional contagion study
(Kramer et al. Proc. of National Academy of Sciences, 2014)

positive

improve privacy decision making and management
(Knijnenburg and Kobsa, TiiS, 2013; Lin et al., USENIX, 2014)

privacy agents

co-evolution

transparency through quantitative input influence (Datta et al. IEEE S&P, 2016)

explanatory debugging to personalize interactive machine learning
(Kulesza et al., ICIUI, 2015)