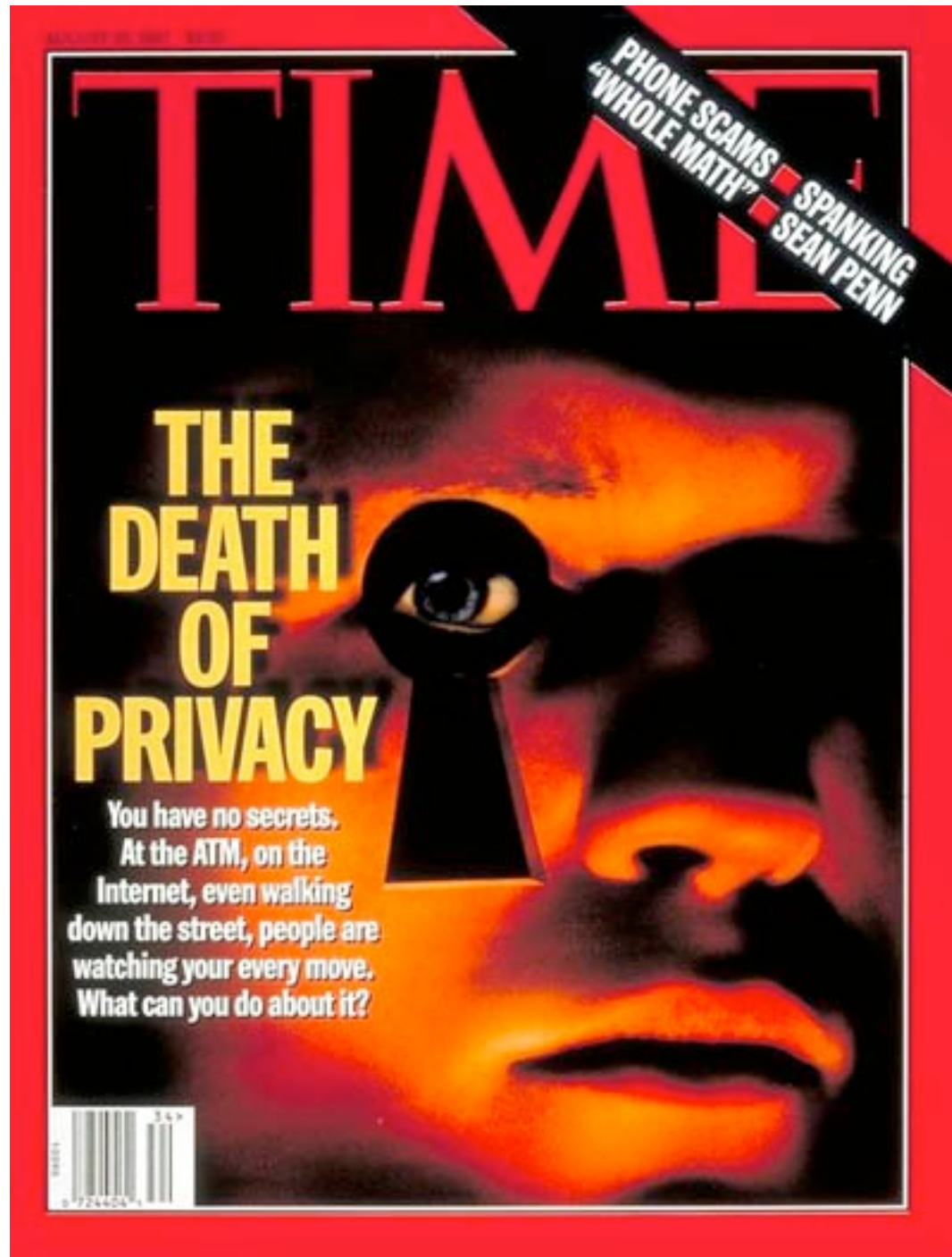


Privacy Research Paradigms Privacy Engineering and the Agile Turn

Seda Gürses
fgurses @ princeton.edu
CITP, Princeton University
COSIC, University of Leuven

13. July 2016
Interdisciplinary Privacy
Summer School



PET SEMATARY



©2009 Last Legion Games, LLC. All Rights Reserved.

getting privacy engineering right?

getting privacy engineering right?

**privacy
research**



**software
engineering
practice**

**privacy
research**

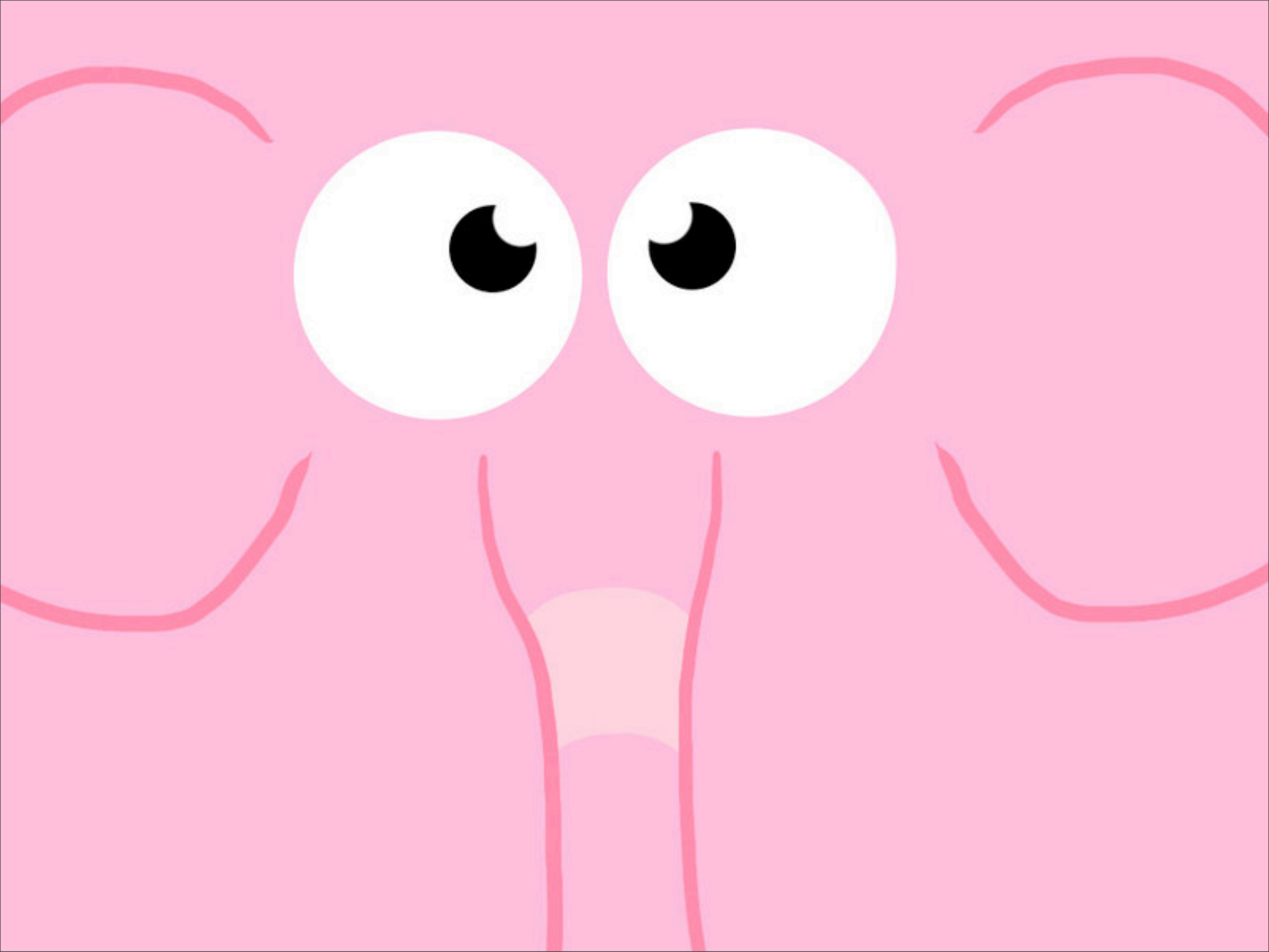


**software
engineering
practice**

**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**



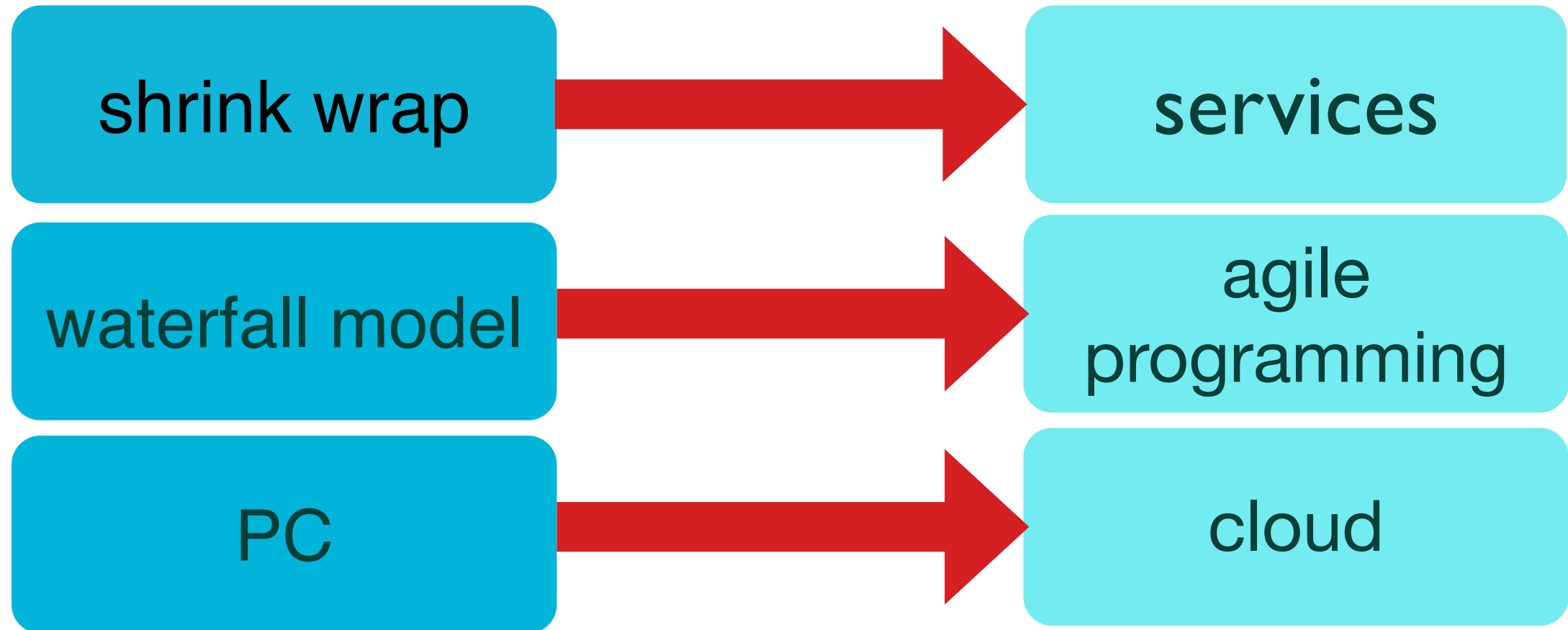
Wurstküche

How the Sausage Gets Made



matters?

the turn to agile



**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?**

**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?**

**what is the
impact of**

**the turn to
agile in
software
engineering
practice**

**on
computer
science
research in
privacy?**

SOK
lit review
42 interviews
events/papers

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

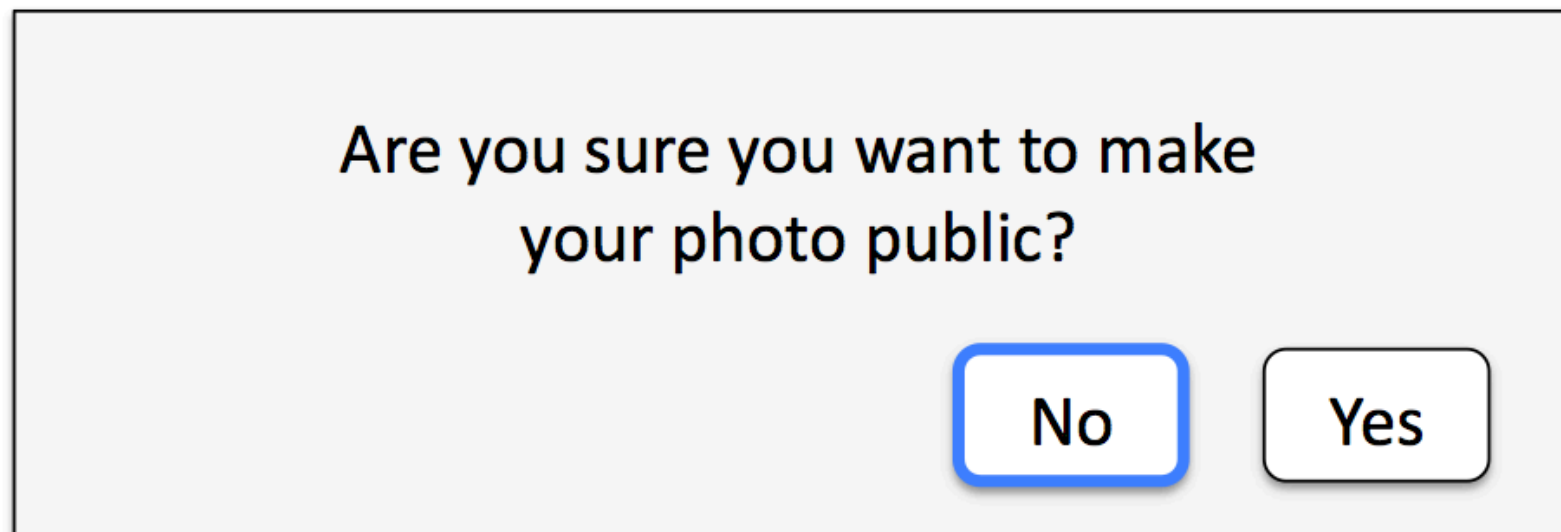
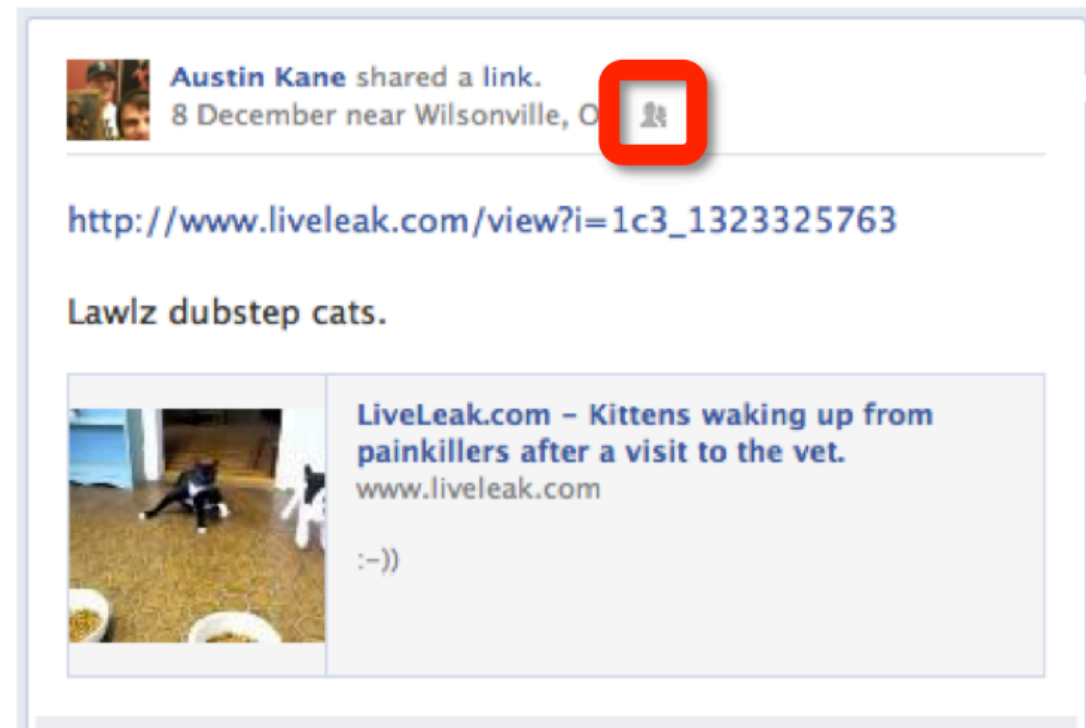
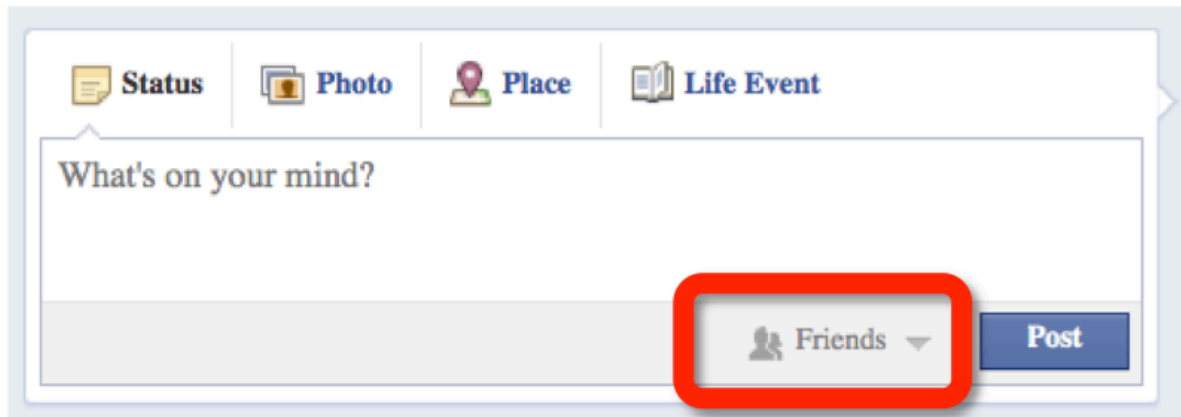
transparency of social impact

PRIVACY RESEARCH PARADIGMS

privacy as
practice

feedback &
awareness design




privacy nudges



slide: Lorrie Cranor

 Update Status  Add Photo / Video  Ask Question




heat in the moment|

   Friends ▼ [Post](#)

You will have 10 seconds to cancel after you post the update

 Update Status  Add Photo / Video  Ask Question

heat in the moment

   Friends ▼ [Post](#)

Your post will be published in **3 seconds**. [Post Now](#) | [Edit It](#) | [Cancel](#)

slide: Lorrie Cranor



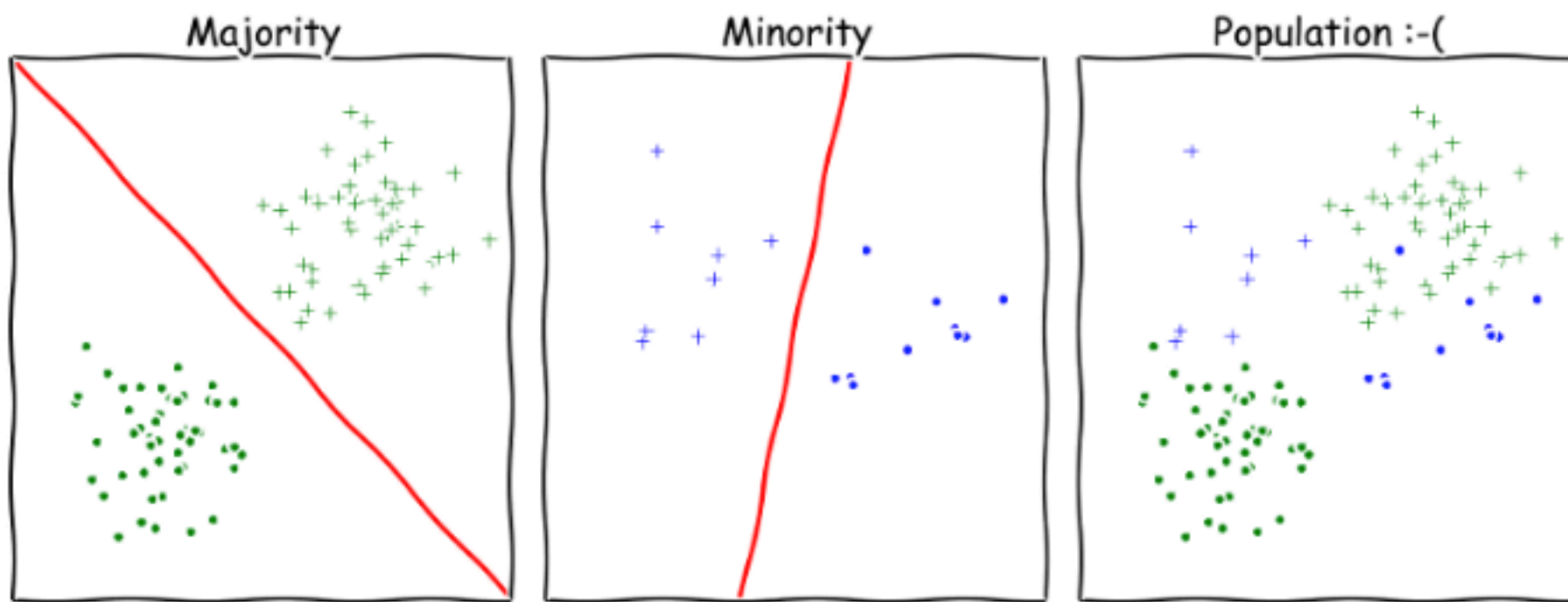
Moritz Hardt [Follow](#)

Researcher. Machine learning, optimization, privacy and social questions in computation.

Sep 26, 2014 · 8 min read

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.

privacy theory

methods

techniques

tools

privacy theory

CONTEXTUAL
INTEGRITY

privacy theory

privacy

non-absolute

contextual

relational

opacity of the individual

data protection
FIPPs

procedural safeguards

accountability

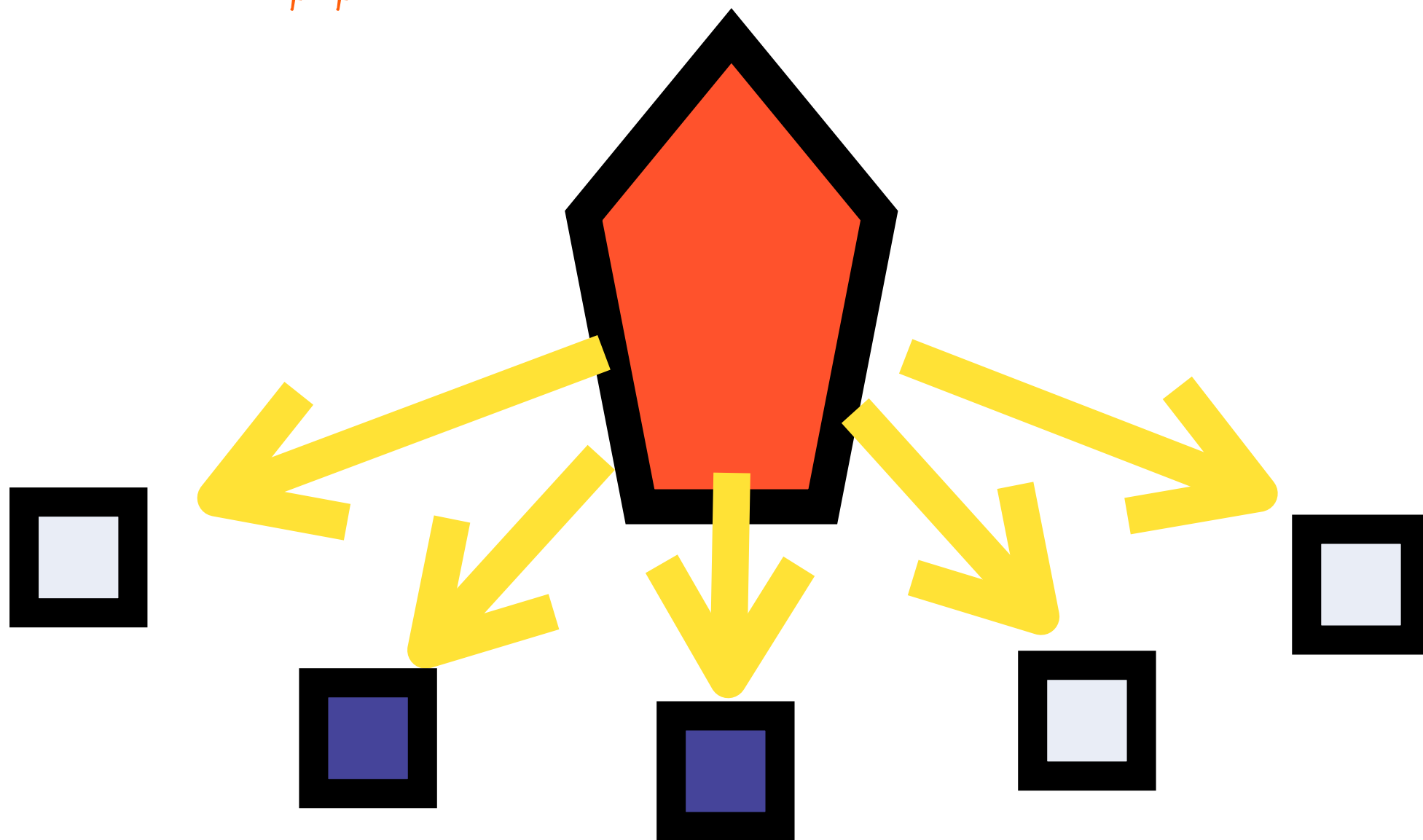
transparency

personal data

data minimization

privacy theory

Surveillance



privacy theory

methods

techniques

tools

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*

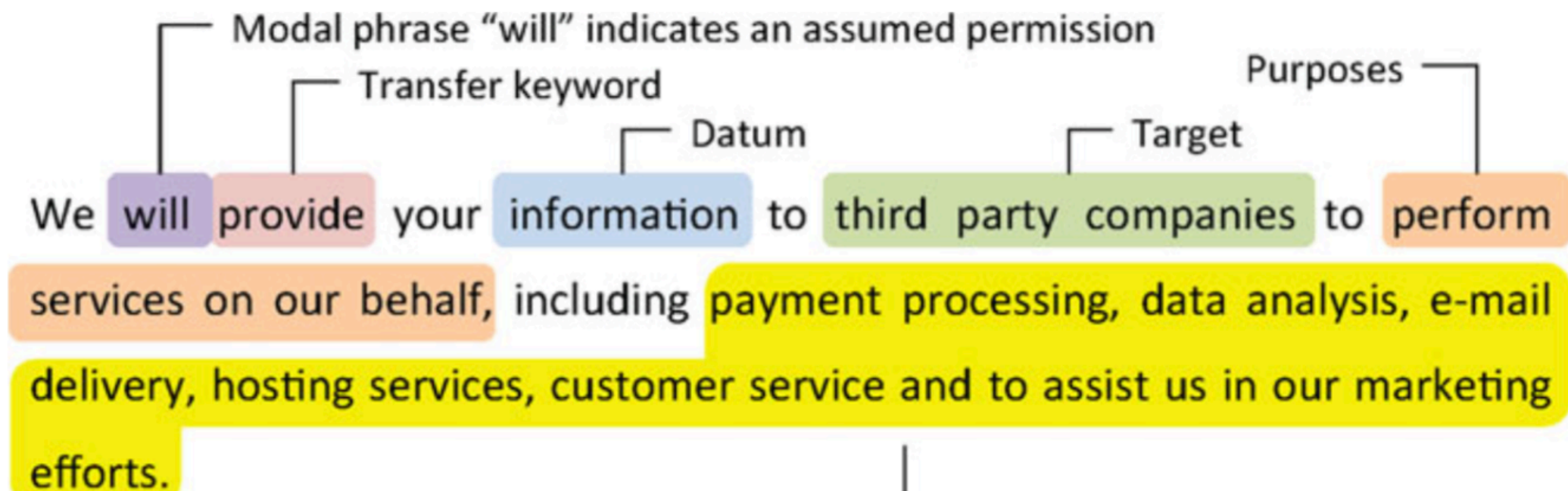
Privacy stages	identifiability	Approach to privacy protection	Linkability of data to personal identifiers	System Characteristics
0	identified	privacy by policy (notice and choice)	linked	<ul style="list-style-type: none"> • unique identifiers across databases • contact information stored with profile information
1	pseudonymous		linkable with reasonable & automatable effort	<ul style="list-style-type: none"> • 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	<ul style="list-style-type: none"> • 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	<ul style="list-style-type: none"> • 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

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

Tor/PreTP

privacy
enhancement of
system or function

privacy policy languages

research into
privacy violations

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?**

methodology

- **exploratory study (work in progress)**
 - develop and shape an agenda for further study
- **interviews and chats**
 - devs, devops, product managers, a/b testers, AI/data product developers, data engineers, privacy officers
- **industry white papers**
- **legal and policy literature**

shrink wrap software



A Venn diagram with three overlapping circles. The top-left circle is pink and labeled 'cloud IaaS/PaaS'. The top-right circle is light blue and labeled 'SOA'. The bottom circle is yellow and labeled 'agile methods'. The intersection of the pink and blue circles is purple and labeled 'SaaS'. The intersection of the pink and yellow circles is orange. The intersection of the blue and yellow circles is green. The intersection of all three circles is a darker brownish-purple.

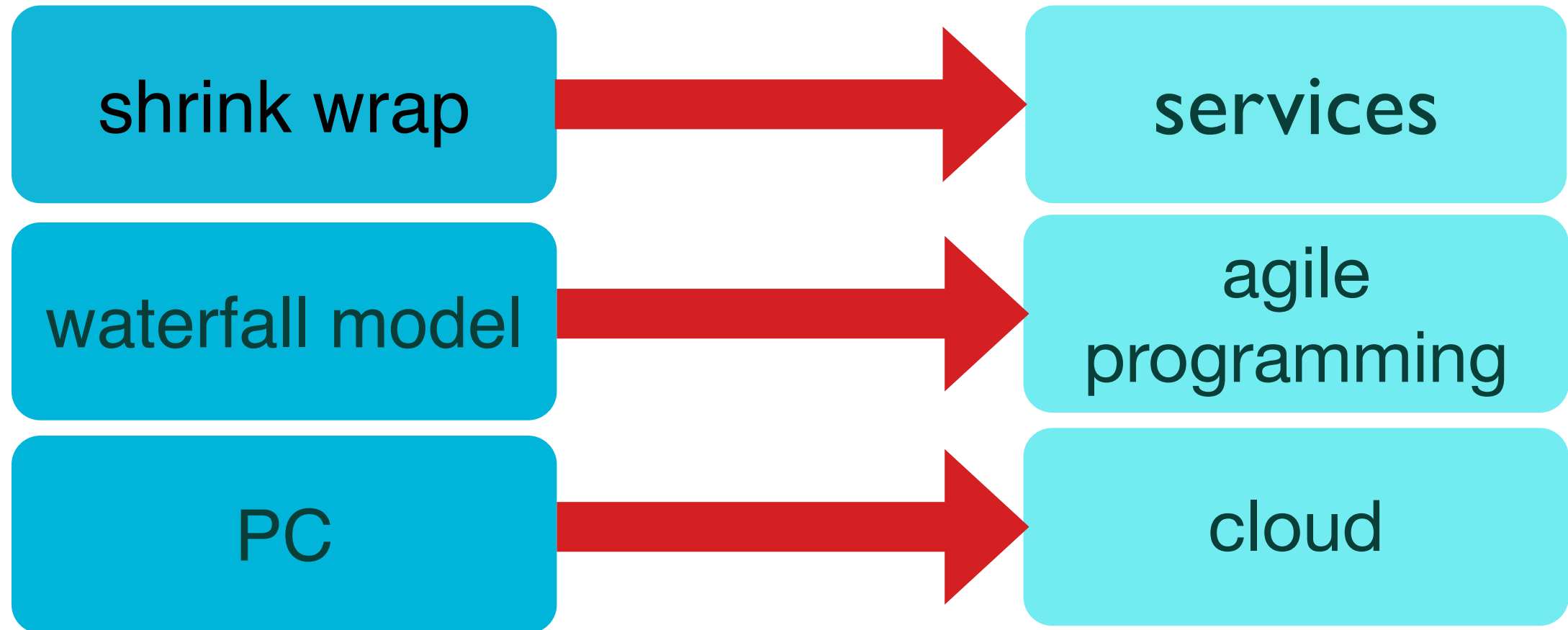
cloud
IaaS/PaaS

SOA

SaaS

agile methods

the turn to agile



shrink wrap



services



1) All teams will henceforth expose their data and functionality through service interfaces.

2) Teams must communicate with each other through these interfaces.

3) There will be no other form of interprocess communication allowed: no direct linking, no direct reads of another team's data store, no shared-memory model, no back-doors whatsoever. The only communication allowed is via service interface calls over the network.

4) It doesn't matter what technology they use. HTTP, Corba, Pubsub, custom protocols – doesn't matter. Bezos doesn't care.

5) All service interfaces, without exception, must be designed from the ground up to be externalizable. That is to say, the team must plan and design to be able to expose the interface to developers in the outside world. No exceptions.

**6) Anyone who doesn't do this will be fired.
~2001/2002**

shrink wrap

binary runs solely on client side

requires matching soft & hardware

updates & maintenance cumbersome

user has control (oh no!)

pay in advance

Microsoft Word

enterprise

apps

services

server (thin) client model

data "secured" by service

updates and maintenance server side

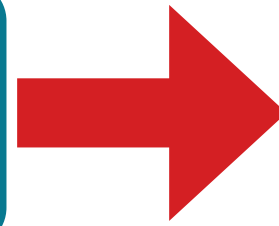
collaborative

pay as you use/trial

office 365

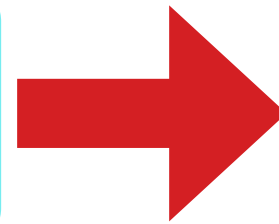
implications of the shift to services

server - thin client model



transaction throughout use

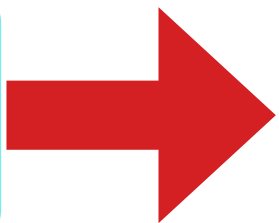
bundled services



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

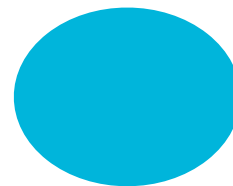
payment

maps

social

embedded media

picture album creation service



See what your users see.

FullStory lets your company easily record, replay, search, and analyze each user's actual experience with your website. Think of it as your team's super-searchable DVR for all customer interactions.

Start your free 14-day trial today!

you@widgetco.com [LET'S GO!](#)

[▶ Watch the video \(1:13\)](#)

The screenshot shows the FullStory dashboard interface. On the left is a sidebar with navigation options: FullStory, Account, YOUR SEGMENTS, and TEAM SEGMENTS. Under 'YOUR SEGMENTS', there is a segment named 'Everyone' with the subtext 'Search & create segments'. The main content area is titled 'All time' and features a search bar labeled 'Search Everyone'. Below the search bar is a table with three columns: PERSON, RECENTLY, and CONTEXT. The table lists two users: Jonny Appleseed and Lauren Anne.

PERSON	RECENTLY	CONTEXT
Jonny Appleseed 17 SESSIONS • SINCE 11/6/14	Online 3 EVENTS • 0:01 • WWW.FULLSTORY.COM	Royal Oak OS X • CHROME
Lauren Anne 50 SESSIONS • SINCE JAN 12	Online 2 EVENTS • 0:02 • /ADMIN	Atlanta OS X • CHROME

fullstory in top 1 million sites

<http://uservoice.com>

<http://remitly.com>

<http://moosejaw.com>

<http://sproutvideo.com>

<http://wahoofitness.com>

<http://clickminded.com>

<http://startapp.com>

<http://wayup.com>

<http://keen.io>

<http://fitocracy.com>

<http://tieks.com>

<http://samcart.com>

<http://meuspedidos.com.br>

<http://referralcandy.com>

<http://thebouqs.com>

<http://oyorooms.com>

<http://codeschool.com>

<http://mymove.com>

<http://urbanclap.com>

<http://owler.com>

<http://scripted.com>

<http://himalayastore.com>

<http://surfdome.com>

<http://namely.com>

<http://travelport.com>

<http://autopilothq.com>

<http://shethinx.com>

<http://credomobile.com>

<http://conte.it>

<http://castorama.pl>

<http://deputy.com>

<http://autoeurope.com>

<http://nexojournal.com.br>

waterfall model



agile
programming



waterfall
model

spiral
model

agile programming

Xtreme programming

waterfall model

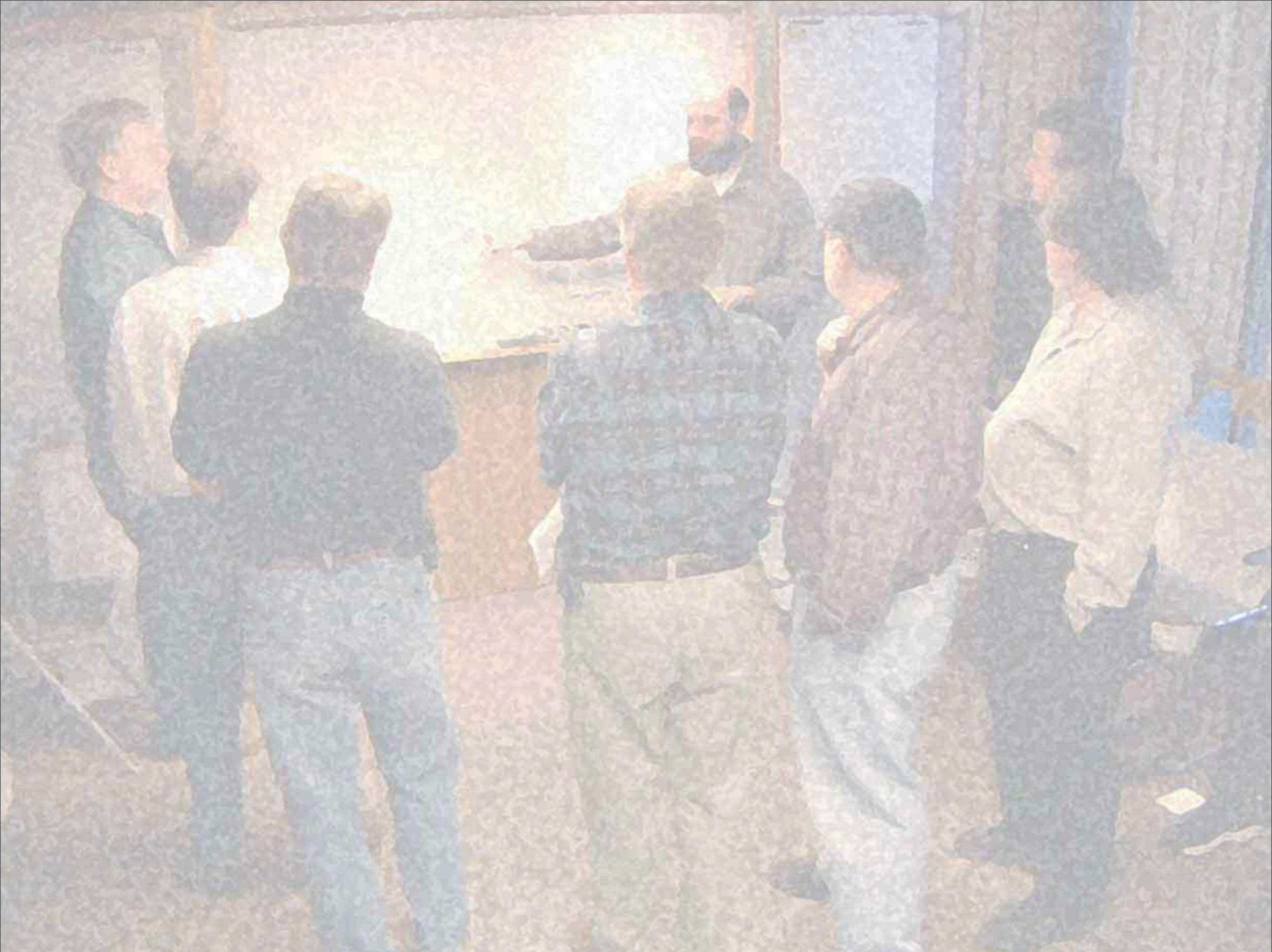
requirements analysis and
specification

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

eXtreme Programming

if short iterations are good, make them as short as possible

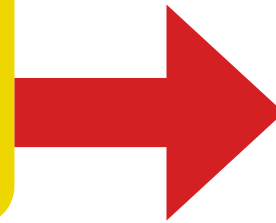
if simplicity is good, do the simplest thing that can work

if testing is good, test all the time

if code reviews are good, review code continuously

implications of the shift to agile dev

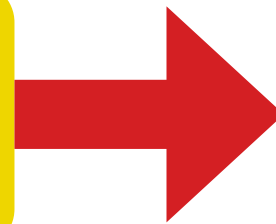
testing testing testing



user centric development

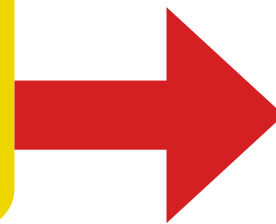
data centric development

short iterations



rapid feature development

simplicity



reuse and modularity

server - thin client model

feature inflation

product manager

behavioral analytics

rapid feature development

where do features come from?

where do features go?

boss/VC said so

designers said so

competitor did it

data centric development

anecdotes

data products

metrics

user/behavioral analytics

data centric development

predictive modeling 4 pricing

user churn

new information
panel

website

perspective 3: behavior and data centrality

- recursively keeping track:
 - capturing behavior of users
 - capturing behavior of service components
 - capturing behavior of your capture models
 - QA and continuous monitoring become one thing



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

Philip Agre: Two models of privacy

These systems capture knowledge of people's behavior, and they reconfigure them through rapid development of features that are able to identify, sequence, reorder and transform human activities.

This also means that they open these human activities to evaluation in terms of economic efficiency.

Philip Agre.

**what is the
impact of**

**the turn to
agile in
software
engineering
practice**

**on
computer
science
research in
privacy?**

Moving Targets: Security and Rapid-Release in Firefox

Sandy Clark
saender@cis.upenn.edu
University of Pennsylvania

Michael Collis
mcollis@cis.upenn.edu
University of Pennsylvania *

Matt Blaze
mab@crypto.com
University of Pennsylvania

Jonathan M. Smith
jms@cis.upenn.edu
University of Pennsylvania

can't apply security
frameworks

no threat
modeling

no risk
assessment

code maturity?
lol

rapid feature development

++ vulnerability
density

++ immature
code

honeymoon

defies attackers
learning curve

impact of the agile turn?

privacy as
confidentiality

data minimization

properties with mathematical guarantees

avoid single point of failure



Reflections: The ecosystem is moving

[moxie0](#) on 10 May 2016

Software exists as part of an ecosystem, and **the ecosystem is moving**. The platform changes out from under it, the networks evolve, security threats and countermeasures are in constant shift, and the collective UX language rarely sits still. As more money, time, and focus has gone into the ecosystem, the faster the whole thing has begun to travel.

One of the controversial things we did with Signal early on was to build it as an unfederated service. Nothing about any of the protocols we've developed requires centralization; it's entirely possible to build a federated Signal Protocol based messenger, but I no longer believe that it is possible to build a *competitive* federated messenger at all.

impact of the agile turn?

privacy as
control

data protection/FIPPS compliance

transparency and accountability

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

impact of the agile turn?

privacy as
practice

privacy integral to collective info practices

improve user agency in negotiating privacy

transparency of social impact

**what is the
impact of**

**the turn to
agile in
software
engineering
practice**

**on
computer
science
research in
privacy?**



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



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.

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.

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](#)

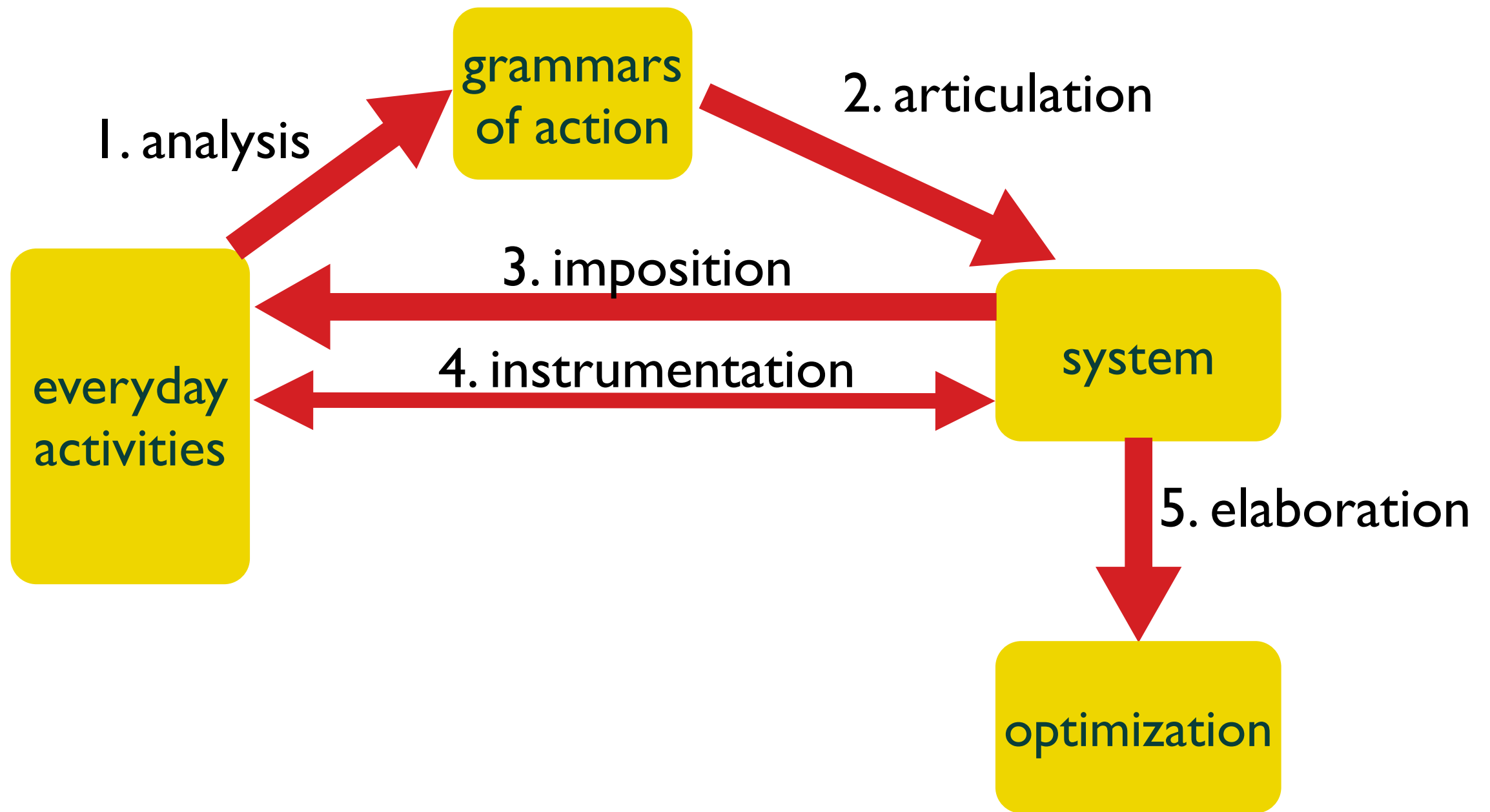
OUTLOOK

- 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!!!

references

- Please contact me for further references
- Philip E. Agre, Surveillance and capture: Two models of privacy, The Information Society, Vol. 10, Iss. 2, 1994
- Irina Kaldrack and Martina Leeker, There is no software, just services, Meson Press, 2015.

capture



computers can only compute what they capture

what would a total reorganization of all spheres of life in accord with the capture model look like?

Capture speaks to current landscape

But he wrote in time of shrink-wrap!