Privacy Research Paradigms
Privacy Engineering
and the Agile Turn

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

Interdisciplinary Privacy
Summer School
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?
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?
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 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 as confidentiality

Secure messaging

Anonymous communications
<table>
<thead>
<tr>
<th></th>
<th>Encrypted in transit?</th>
<th>Encrypted so the provider can’t read it?</th>
<th>Can you verify contacts’ identities?</th>
<th>Are past comms secure if your keys are stolen?</th>
<th>Is the code open to independent review?</th>
<th>Is security design properly documented?</th>
<th>Has there been any recent code audit?</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Off-The-Record Messaging for Mac (Adium)</strong></td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>❌</td>
</tr>
<tr>
<td><strong>Off-The-Record Messaging for Windows (Pidgin)</strong></td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td><strong>PGP for Mac (GPGTools)</strong></td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>❌</td>
<td>✔️</td>
<td>✔️</td>
<td>❌</td>
</tr>
<tr>
<td><strong>PGP for Windows Gpg4win</strong></td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
<td>❌</td>
<td>✔️</td>
<td>✔️</td>
<td>❌</td>
</tr>
</tbody>
</table>
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 as control

privacy policy languages

purpose based access control
<table>
<thead>
<tr>
<th>Information we collect</th>
<th>Ways we use your information</th>
<th>Information sharing</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>to provide service and maintain site</td>
<td>marketing</td>
</tr>
<tr>
<td>contact information</td>
<td>opt in</td>
<td></td>
</tr>
<tr>
<td>cookies</td>
<td></td>
<td></td>
</tr>
<tr>
<td>demographic information</td>
<td></td>
<td></td>
</tr>
<tr>
<td>financial information</td>
<td></td>
<td></td>
</tr>
<tr>
<td>health information</td>
<td></td>
<td></td>
</tr>
<tr>
<td>preferences</td>
<td></td>
<td></td>
</tr>
<tr>
<td>purchasing information</td>
<td></td>
<td></td>
</tr>
<tr>
<td>social security number &amp; gov't ID</td>
<td></td>
<td></td>
</tr>
<tr>
<td>your activity on this site</td>
<td></td>
<td></td>
</tr>
<tr>
<td>your location</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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
Are you sure you want to make your photo public?

No  Yes

slide: Lorrie Cranor
heat in the moment

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

heat in the moment

Your post will be published in 3 seconds. Post Now | Edit It | Cancel

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.
privacy theory

methods

techniques

tools
privacy theory

CONTEXTUAL INTEGRITY
privacy theory

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

Modal phrase “will” indicates an assumed permission

Transfer keyword

Datum

Purpose

Target

We will provide your information to third party companies to perform services on our behalf, including payment processing, data analysis, e-mail delivery, hosting services, customer service and to assist us in our marketing efforts.
(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

<table>
<thead>
<tr>
<th>Metric</th>
<th>Shadow</th>
<th>TorPS</th>
<th>ExperimenTor</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Size / number of relays</td>
<td>downscaling, simulation with 500+ relays possible</td>
<td>no downscaling</td>
<td>limited by available resources</td>
</tr>
<tr>
<td>2. Path length</td>
<td>not using additional weighting in node</td>
<td>ignoring paths being dropped due to</td>
<td></td>
</tr>
</tbody>
</table>
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
the turn to agile

shrink wrap

waterfall model

PC

services

agile programming

cloud
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
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)
fullstory in top 1 million sites

waterfall model

requirements analysis and specification

architectural design

implementation and integration

verification

operation and maintenance
agile manifesto

- individuals and interactions
- working software
- customer collaboration
- responding to change

- process and tools
- comprehensive documentation
- contract negotiation
- 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
- short iterations
- simplicity
- server - thin client model

- user centric development
- data centric development
- rapid feature development
- reuse and modularity
feature inflation

rapid feature development

where do features come from?
- product manager
- boss/VC said so
- designers said so
- competitor did it

where do features go?
- behavioral analytics
data centric development

- data products
- user/behavioral analytics
- data centric development
- predictive modeling 4 pricing
- user churn

anecdotes

metrics
new information panel

website
perspective 3: behavior and data centricity

- 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?
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?
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

<table>
<thead>
<tr>
<th>Information we collect</th>
<th>Ways we use your information</th>
<th>Information sharing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Contact information</td>
<td>To provide service and</td>
<td></td>
</tr>
<tr>
<td></td>
<td>maintain site</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Marketing</td>
<td>Opt in</td>
</tr>
<tr>
<td></td>
<td>Telemarketing</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Profiling</td>
<td></td>
</tr>
<tr>
<td>Cookies</td>
<td>Opt in</td>
<td></td>
</tr>
<tr>
<td>Demographic information</td>
<td>Opt in</td>
<td></td>
</tr>
<tr>
<td>Financial information</td>
<td>Opt out</td>
<td></td>
</tr>
<tr>
<td>Health information</td>
<td>Opt out</td>
<td></td>
</tr>
<tr>
<td>Preferences</td>
<td>Opt in</td>
<td></td>
</tr>
<tr>
<td>Purchasing information</td>
<td>Opt in</td>
<td></td>
</tr>
<tr>
<td>Social security number &amp; gov't ID</td>
<td>Opt out</td>
<td></td>
</tr>
<tr>
<td>Your activity on this site</td>
<td>Opt in</td>
<td></td>
</tr>
<tr>
<td>Your location</td>
<td>Opt out</td>
<td></td>
</tr>
</tbody>
</table>

**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?
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.

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 Statcounter 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.
• 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 Leekker, There is no software, just services, Meson Press, 2015.
capture

everyday activities

1. analysis

2. articulation

3. imposition

4. instrumentation

system

5. elaboration

optimization

grammars of action
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!