Privacy engineering, privacy by design, and privacy governance

Lorrie Faith Cranor

November 17, 2015

8-533 / 8-733 / 19-608 / 95-818:
Privacy Policy, Law, and Technology

Carnegie Mellon University
CyLab

institute for SOFTWARE RESEARCH

Engineering & Public Policy

Today's agenda

- Quiz
- Questions about midterm
- Homework 7 discussion
- Beam case study
- Privacy engineering
- Privacy by design
- Privacy governance

By the end of class you will be able to:

- Understand how to apply various approaches to privacy engineering and privacy by design to design problems

Beam

- https://www.suitabletech.com/


beam
privacy
ENGINEERING


beam
smartpresence.com
Café
TAXIM
Gyro
Kebab
ISTANBUL STREET FOOD
Hot Food
Lavazza Cafe
FREE WIRELESS
CAUTION
PUSH TO OPEN

Beam discussion

- https://www.youtube.com/watch?v=-uUb4TrPyxs
- What privacy issues does this technology raise in the home environment? How might these issues be addressed?

Privacy by policy vs. architecture

- What techniques are used in each approach?
- What are the advantages and disadvantages of each approach?

How rights are protected

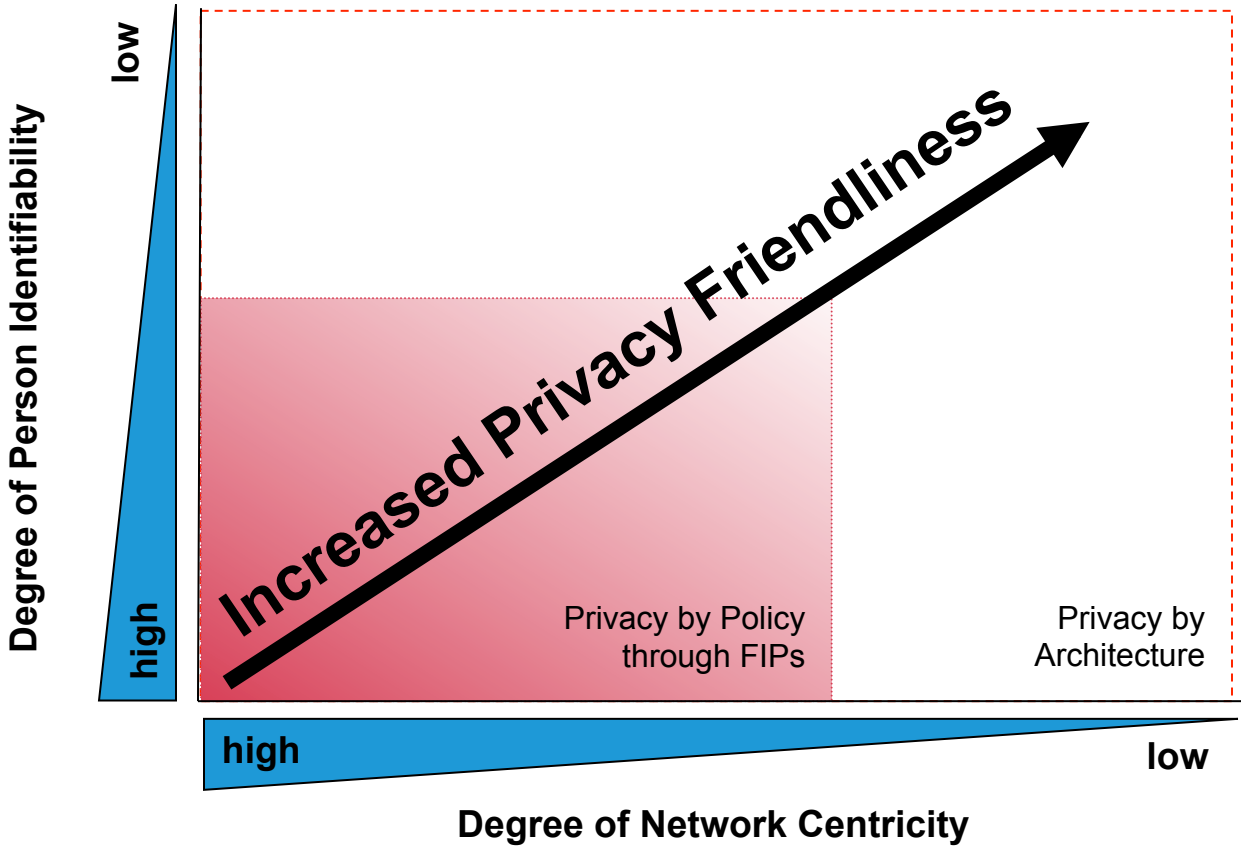
Privacy by Policy

- Through laws and policies
- Requires enforcement, technology can facilitate compliance
- Violations possible due to bad actors, mistakes, government mandates

Privacy by Architecture

- Through technology
- Reduces need to rely on trust & external enforcement
- Violations possible if technology fails or availability of new data or technology defeats protections
- May be viewed as too expensive or restrictive

What system features tend to lead to more or less privacy?



Privacy by policy techniques

- Notice
- Choice
- Security safeguards
- Access
- Accountability
  - Audits
  - Privacy policy management technology
    - Enforcement engine

Privacy by architecture techniques

- Best
  - No collection of contact information
  - No collection of long-term person characteristics
  - k-anonymity with large value of k
- Good
  - 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 personal characteristics w/ low granularity
  - Technically enforced deletion of profile details at regular intervals

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*

De-identification and re-identification

- Simplistic de-identification: remove obvious identifiers
- Better de-identification: also k-anonymize and/or use statistical confidentiality techniques
- Re-identification can occur through linking entries within the same database or to entries in external databases

Examples

- When RFID tags are sewn into every garment, how might we use this to identify and track people?
- What if the tags are partially killed so only the product information is broadcast, not a unique ID?
- How can a cellular provider identify an anonymous pre-paid cell phone user?

Privacy by Design Principles (PbD)

1. Proactive not Reactive; Preventative not Remedial
2. Privacy as the Default Setting
3. Privacy Embedded into Design
4. Full Functionality—Positive-Sum, not Zero-Sum
5. End-to-End Security—Full Lifecycle Protection
6. Visibility and Transparency—Keep it Open
7. Respect for User Privacy—Keep it User-Centric

Ann Cavoukian
https://www.privacybydesign.ca/content/uploads/2009/08/7foundationalprinciples.pdf

Data governance

- People, process, and technology for managing data within an organization
- Data-centric threat modeling and risk assessment
- Protect data throughout information lifecycle
  - Including data destruction at end of lifecycle
- Assign responsibility

Privacy Impact Assessment

A methodology for

- assessing the impacts on privacy of a project, policy, program, service, product, or other initiative which involves the processing of personal information and,
- in consultation with stakeholders, for taking remedial actions as necessary in order to avoid or minimize negative impacts

D. Wright and P. De Hert, eds. *Privacy Impact Assessment*. Springer 2012.

PIA is a process

- Should begin at early stages of a project
- Should continue to end of project and beyond

Why carry out a PIA?

- To manage risks
  - Negative media attention
  - Reputation damage
  - Legal violations
  - Fines, penalties
  - Privacy harms
  - Opportunity costs
- To derive benefits
  - Increase trust
  - Avoid future liability
  - Early warning system
  - Facilitate privacy by design early in design process
  - Enforce or encourage accountability

Who has to carry out PIAs?

- US administrative agencies, when developing or procuring IT systems that include PII
  - Required by E-Government Act of 2002
- Government agencies in many other countries
- Sometimes done by private sector
  - Case studies from Vodaphone, Nokia, and Siemens in PIA book

Carnegie Mellon University
CyLab

institute for SOFTWARE RESEARCH

Engineering & Public Policy