Haojian Jin

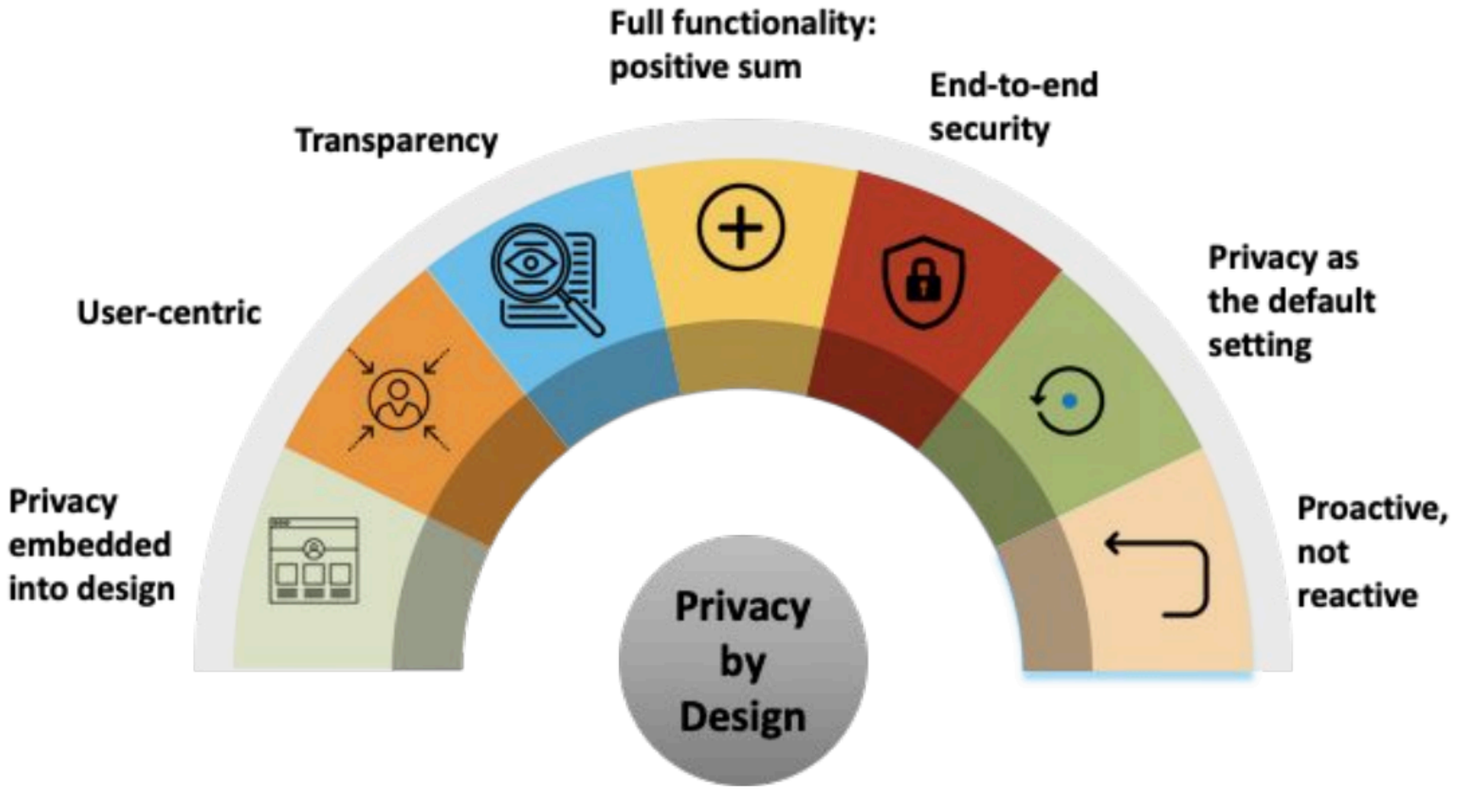
Where are we in the class?

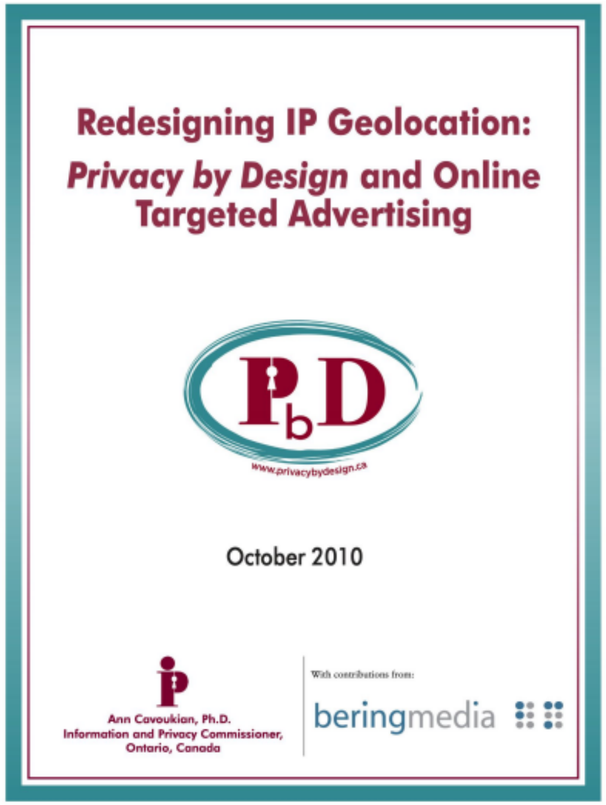
- System
  - Location privacy; Permissions for Privacy; Policies for Privacy
- User
  - Privacy Norms/Contextual Integrity
  - Individual Privacy (CogSci); User Agency
- Developer | Auditors
  - Privacy designs

Agenda

- Privacy by Design
- Design Privacy
- Privacy Threat Modeling
- Execute Designs

Privacy By Design (since 1995)





What is Privacy by Design

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 Life Cycle Protection.
6. Visibility and Transparency—Keep it Open.
7. Respect for User Privacy—Keep it Individual and User-Centric.

Challenges of Privacy By Design

- Vague
- Not practical
- Lack incentives
- Expensive

Privacy by design

10 languages

Article Talk | Read Edit View history Tools

From Wikipedia, the free encyclopedia

Privacy by design is an approach to systems engineering initially developed by Ann Cavoukian and formalized in a joint report on privacy-enhancing technologies by a joint team of the Information and Privacy Commissioner of Ontario (Canada), the Dutch Data Protection Authority, and the Netherlands Organisation for Applied Scientific Research in 1995.[1][2] The privacy by design framework was published in 2009[3] and adopted by the International Assembly of Privacy Commissioners and Data Protection Authorities in 2010.[4] Privacy by design calls for privacy to be taken into account throughout the whole engineering process. The concept is an example of value sensitive design, i.e., taking human values into account in a well-defined manner throughout the process.[5][6]

Cavoukian's approach to privacy has been criticized as being vague,[7] challenging to enforce its adoption,[8] difficult to apply to certain disciplines,[9][10] challenging to scale up to networked infrastructures,[10] as well as prioritizing corporate interests over consumers' interests[7] and placing insufficient emphasis on minimizing data collection.[9] Recent developments in computer science and data engineering, such as support for encoding privacy in data[11] and the availability and quality of Privacy-Enhancing Technologies (PET's) partly offset those critiques and help to make the principles feasible in real-world settings.

The European GDPR regulation incorporates privacy by design.[12]

GDPR Art 25: Which factors?

Appropriate Measures and Principles 	State of the art 	Nature, scope, context and purpose 
Effective (Demonstrate with metrics) 	Costs 	Risk for the individual 

JULY 19, 2021

Privacy is an afterthought in the software lifecycle. That needs to change.

The key to combining privacy and innovation is baking it into the SDLC. Analogous to application security's (AppSec) upstream shift into the development cycle, privacy belongs at the outset of development, not as an afterthought. Here's why.

5 replies

SR

It is interesting how more things keep being moved to "the outset of development". Don Norman said that understanding the users should come before any project definition whatsoever. But if everything is important and should be moved to the beginning, don't we just end up back at Big Design Up Front? It sounds like a generic call to become smarter and more knowledgeable.
Don Norman argues that this goal has become impossible due to how knowledge has expanded in our lifetimes. We have outgrown the brain, and the ability to design automated assistance for it. We brought this on ourselves.

Jul 20, 2021 at 07:35 Reply

https://stackoverflow.blog/2021/07/19/privacy-is-an-afterthought-in-the-software-lifecycle-that-needs-to-change/

Takeaway Msgs

- Privacy Design is a spectrum.
- Two aspects on privacy designs.
  - Consequential and procedural.
  - Due Diligence.
- The life cycle of privacy design.
  - Articulating the system design
  - Modeling the potential threats
  - Navigating the design space
  - Implementing

Haojian Takeaway Msgs

- Privacy Design is a spectrum.

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*

Privacy by policy vs. architecture

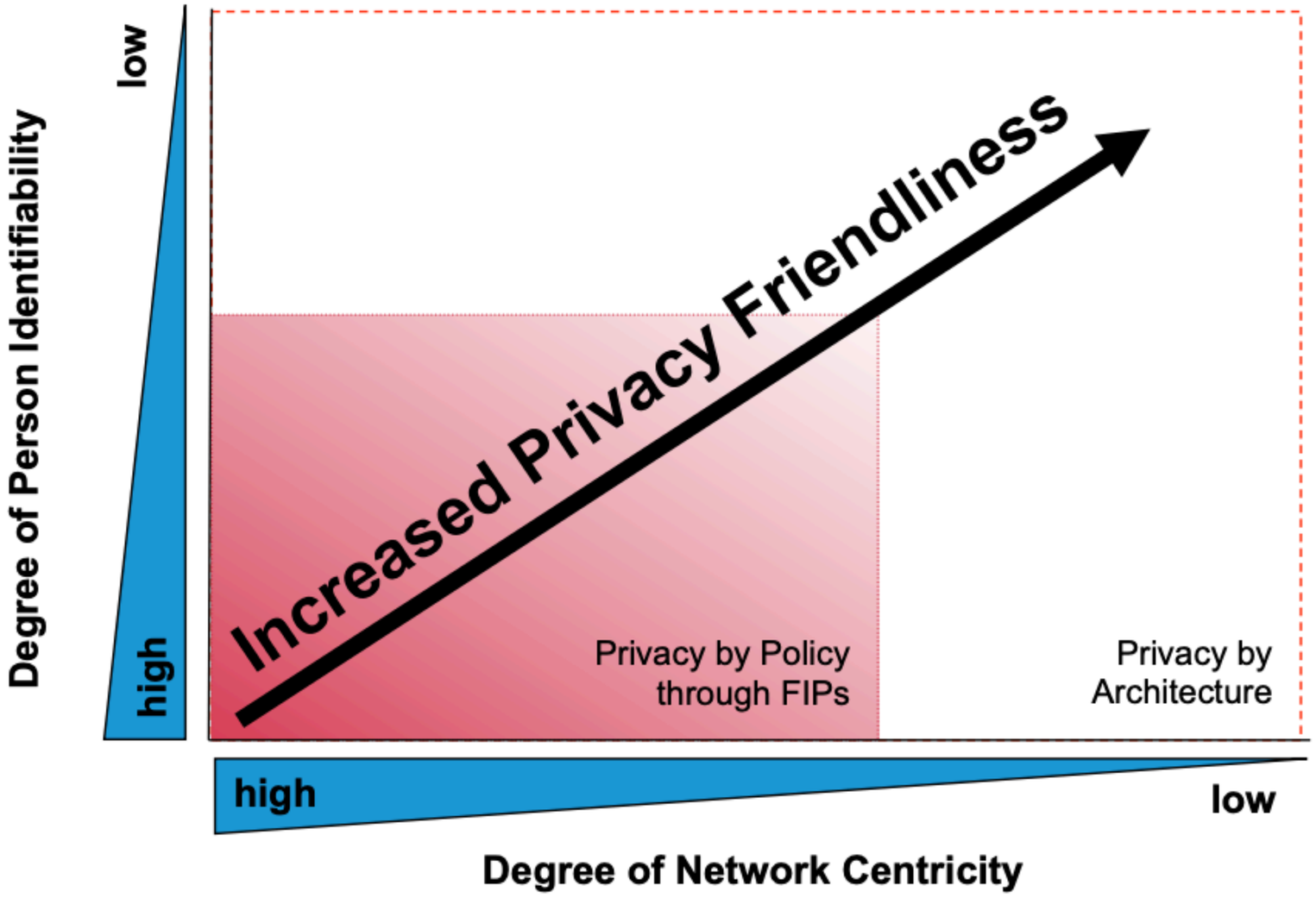
Privacy by Policy

- Through laws and policies
- Requires enforcement, tech 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 tech fails
- May be viewed as too expensive or restrictive

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



Haojian Takeaway Msgs

- Privacy Design is a spectrum.
- Two aspects on privacy designs.
  - Consequential and procedural.
  - Due Diligence.

What is privacy by design?

"Privacy by Design (PbD) represents a significant shift from traditional approaches to protecting privacy, which focus on setting out minimum standards for information management practices and providing remedies for privacy breaches, after-the-fact.

Advocating privacy as a core requirement of systems, right from the outset, it is a proactive approach to privacy protection which seeks to avoid data breaches and their attendant harm. "

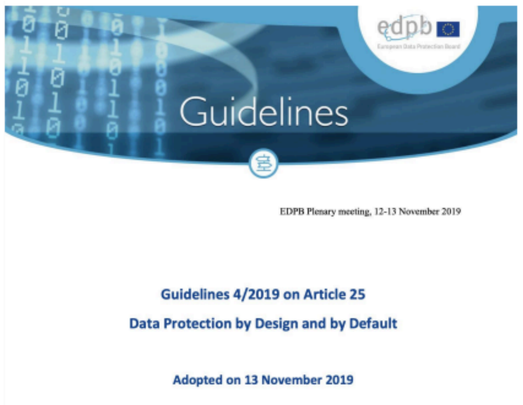
Haojian Takeaway Msgs

- Privacy Design is a spectrum.
- Two aspects on privacy designs.
  - Consequential and procedural.
  - Due Diligence.
- The life cycle of privacy design.
  - Articulating the system design
  - Modeling the potential threats
  - Navigating the design space
  - Implementing

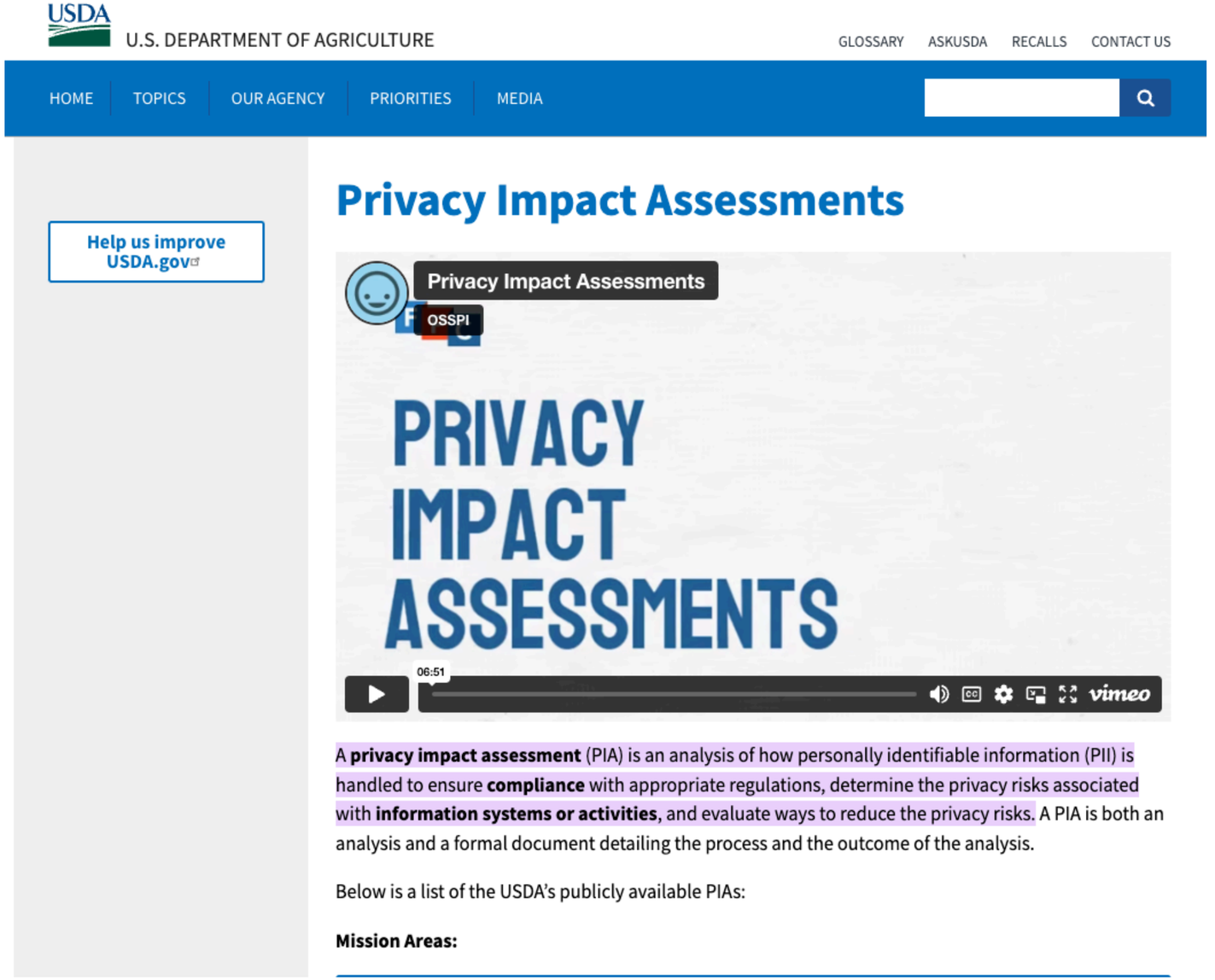
FIPs are not self-executing. Rather, privacy by design requires the translation of FIPs into engineering and usability principles and practices.

When to implement the controls?

- Decision making on the tech stack, vendor
- Abstract, engineering design document, prototypes
- Implementation is effective
- Periodic vendor reviews
- Data breaches
- Data deletion



Privacy Impact Assessment



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

PIA is a process

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

Privacy Analysis Worksheet

SECTION A

Summary Information

1. Name of project or system:
   <Please enter the project or system name here.>

2. Description of project or system and its purpose:
   <Please provide a general description of the project or system, and its purpose using a non-technical description, if statutory, provide citation.>

3. Contact Name, Title, Telephone Number and Organization:
   <Please provide information here.>

Specific Questions

1. Does this project or system collect, maintain, retrieve or share personal information that can be used to directly or indirectly identify an individual?

   ☐ NO. A PIA is not required for this project. Skip to Signature Page.
   ☐ YES. A PIA is required for this project.
   <Please provide a specific description of the information that might be collected or maintained.>

2. Does this project or system retrieve information using a personal identifier?

   ☐ NO. A Privacy Act SORN is not required for this project. Skip to Signature Page.
   ☐ YES. A Privacy Act SORN is required for this project.
   <Please provide a description of the data fields that might be used to retrieve the information.>

   Is there an existing Privacy Act System of Records Notice (SORN)?
   ☐ NO. <Contact privacyhelp@sec.gov for assistance.>
   ☐ YES. The existing SORN may need to be modified to reflect changes.
   <Please provide the system notice number.>

Why carry out a PIA?

To manage risks

- Negative media attention
- Reputation damage
- Legal violations
- Fines, penalties
- Privacy harms
- Opportunity costs

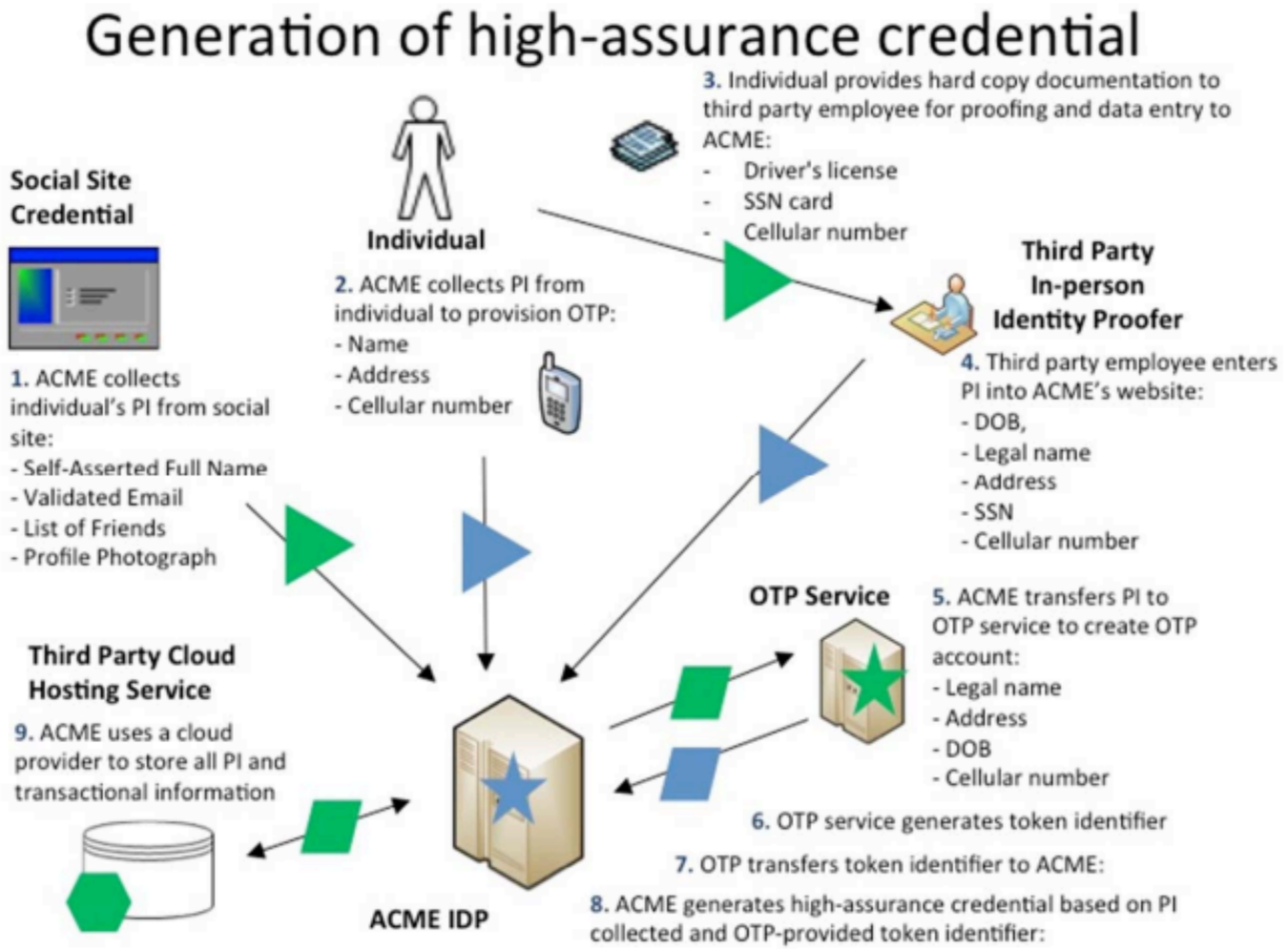
To derive benefits

- Increase trust
- Avoid future liability
- Early warning system
- Facilitate privacy by design early in design process
- Enforce or encourage accountability

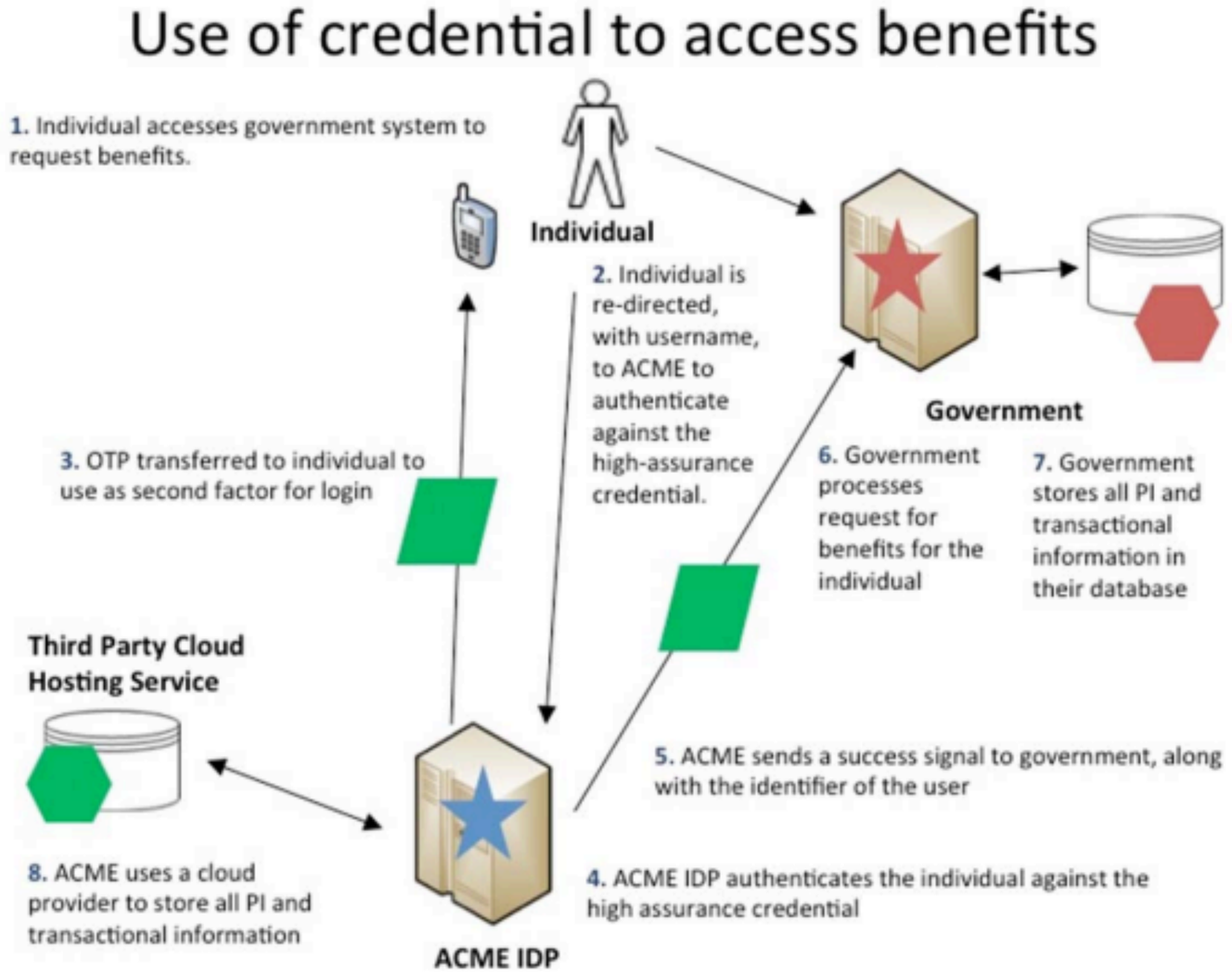
Who has to carry out PIAs?

- 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

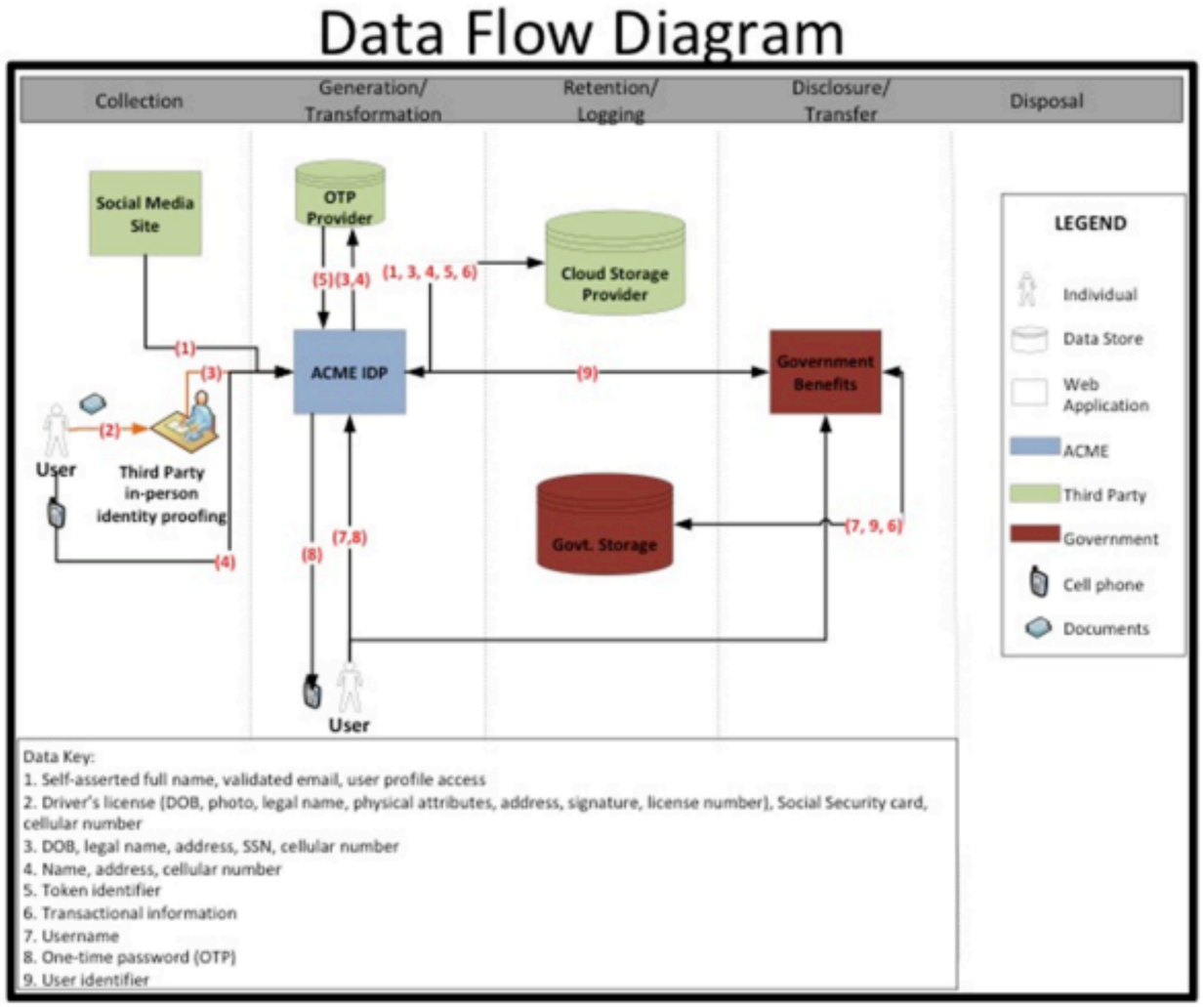
PRAM - Task 1 - Map data processing within the system



PRAM - Task 1 - Map data processing within the system



PRAM - Task 1 - Map data processing within the system



PRAM - Task 2 - Catalog general contextual factors

Data Action	Personal Information	Specific Context	Summary Issues
Collection from the Social Media Site	-Self-Asserted Full Name -Validated Email -List of Friends -Profile Photograph	-One-time action (per user) between social credential and ACME IDP, but establishes an ongoing relationship between user's social media presence and ACME IDP -Social credential linking is visible to user -Linking of social credential simplifies access to government benefits system -User profile may contain information the user considers sensitive -User profile may contain information from other users not participating in the system -User profile includes information unrelated to the purpose and operations of the system -Access to PI is consented by user -Nature of the API: full profile access is granted (by default: name, validated email, profile photograph, and list of friends)	-Full social credential profile access (including picture and list of friends) is not necessary for fulfilling operational purpose. -Will users understand the eventual high-assurance credential is controlled by ACME and not by their social credential provider? -How will perception of the social media organization's privacy practices impact users' willingness to consent to this data action? -Will the user understand ACME will have ongoing access to information stored in their social profile? -Will users' social media privacy settings allow this data action?

PRAM - Task 3 - Likelihood

Data Actions	Summary Issues	Problematic Data Actions	Potential Problems for Individuals	Likelihood
Collection from the social media site	Full social credential profile access (including picture and list of friends) is not necessary for fulfilling operational purpose.	-Appropriation -Induced disclosure -Surveillance -Unanticipated revelation	Stigmatization: Information is revealed about the individual that they would prefer not to disclose.	7
			Power Imbalance: People must provide extensive information, giving the acquirer an unfair advantage.	2
	Will users understand the eventual high-assurance credential is controlled by ACME and not by their social credential provider?	-The summary issue will be associated with another data action.		N/A
	How will perception of the social media organization's privacy practices impact users' willingness to consent to this data action?	-Induced disclosure -Surveillance	Loss of Trust: Individuals lose trust in ACME due to a breach in expectations about the handling of personal information.	6

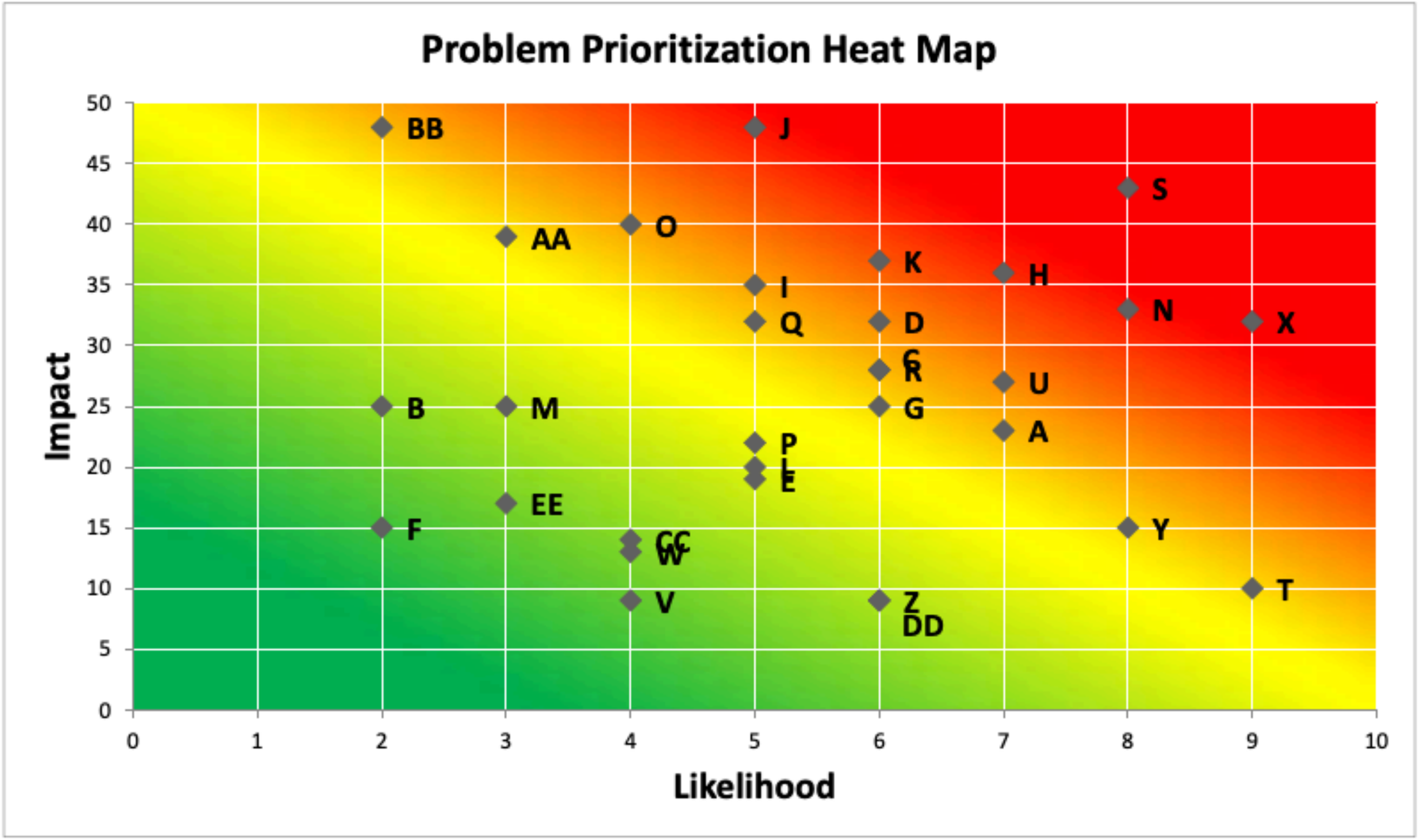
PRAM - Task 4- Likelihood X Impact = Risk

Data Actions	Summary Issues	Problematic Data Actions	Potential Problems for Individuals	Business Impact Factors					Total Business Impact
				Noncompliance Costs	Direct Business Costs	Reputational Costs	Internal Culture Costs	Other	
Collection from the social media site	Full social credential profile access (including picture and list of friends) is not necessary for fulfilling operational purpose.	-Appropriation -Induced disclosure -Surveillance -Unanticipated revelation	Stigmatization	7	6	6	4		23
			Power Imbalance	7	6	8	4		25
	How will perception of the social media organization's privacy practices impact users' willingness to consent to this data action?	-Induced disclosure -Surveillance	Loss of Trust	7	6	8	7		28

PRAM - Task 5 - Problem prioritization

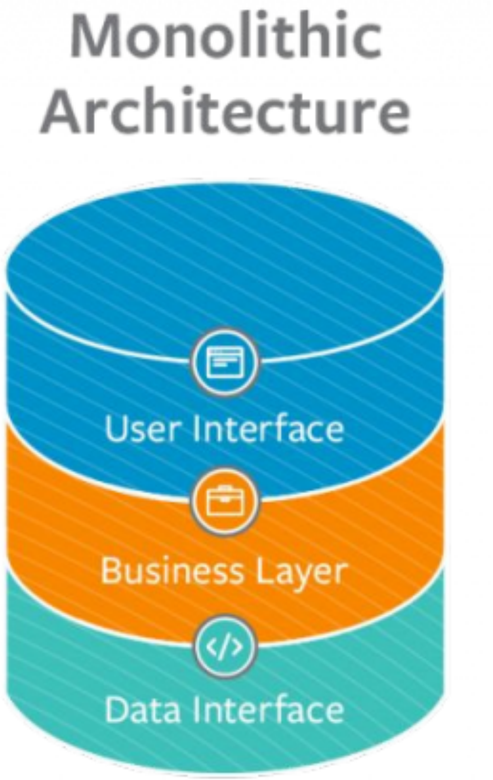
Data Actions	Potential Problems	Point Label	Likelihood	Business Impact
Collection from the social media site	Stigmatization	A	7	23
	Power Imbalance	B	2	25
	Loss of Trust	C	6	28
DA2	Economic Loss	D	6	32
	Loss of Autonomy	E	5	19
	Exclusion	F	2	15
DA3	Loss of Trust	G	6	25
	Stigmatization	H	7	36
	Loss of Liberty	I	5	35
DA4	Loss of Trust	J	5	48
DA5	Economic Loss	K	6	37
	Loss of Autonomy	L	5	20
	Power Imbalance	M	3	25
	Exclusion	N	8	33
	Stigmatization	O	4	40
DA6	Loss of Trust	P	5	22
	Loss of autonomy	Q	5	32
	Exclusion	R	6	28
DA7	Loss of Autonomy	S	8	43
	Stigmatization	T	9	10
	Power Imbalance	U	7	27
	Exclusion	V	4	9
DA8	Loss of autonomy	W	4	13
	Stigmatization	X	9	32
	Power Imbalance	Y	8	15
	Exclusion	Z	6	9
DA9	Loss of Trust	AA	3	39
	Loss of Liberty	BB	2	48
DA10	Loss of Trust	CC	4	14
	Power Imbalance	DD	6	9
	Stigmatization	EE	3	17

PRAM - Task 5 - Problem prioritization



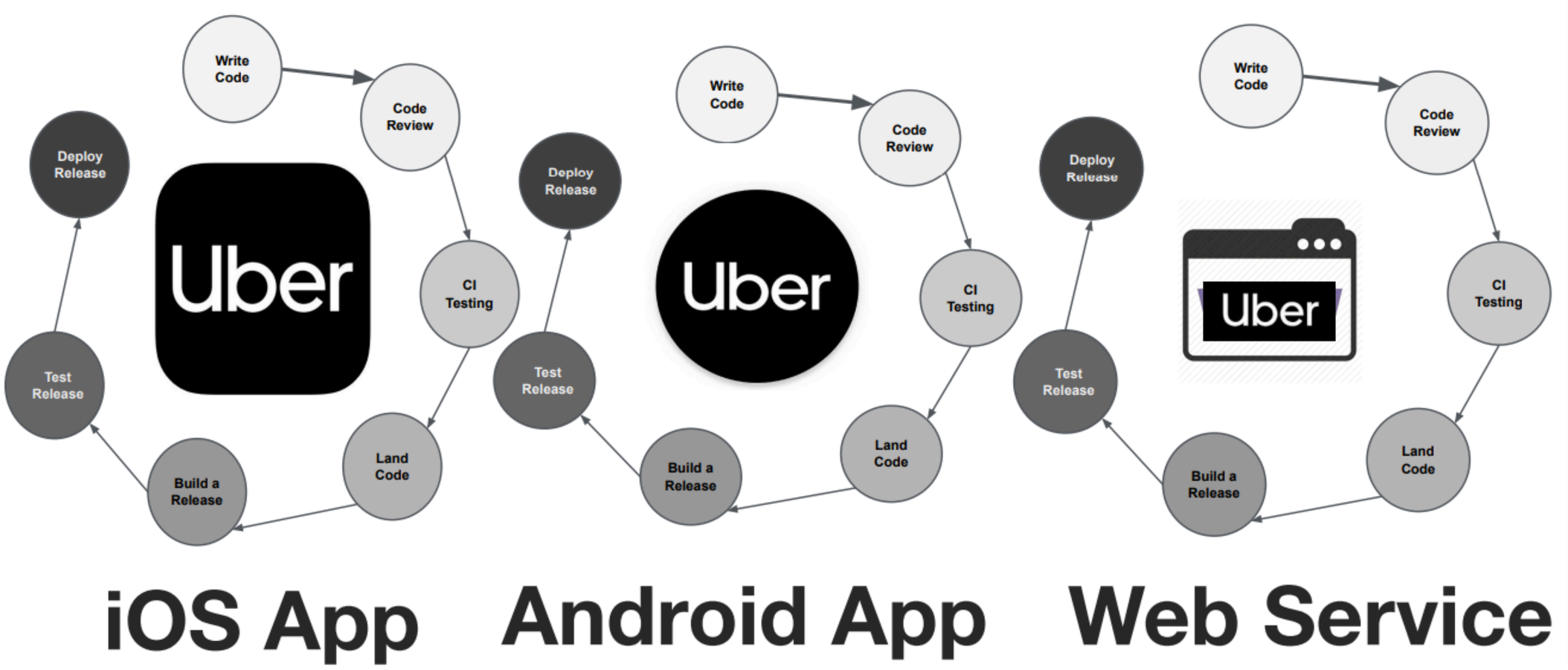
Practices in the industry

Existing Guidelines

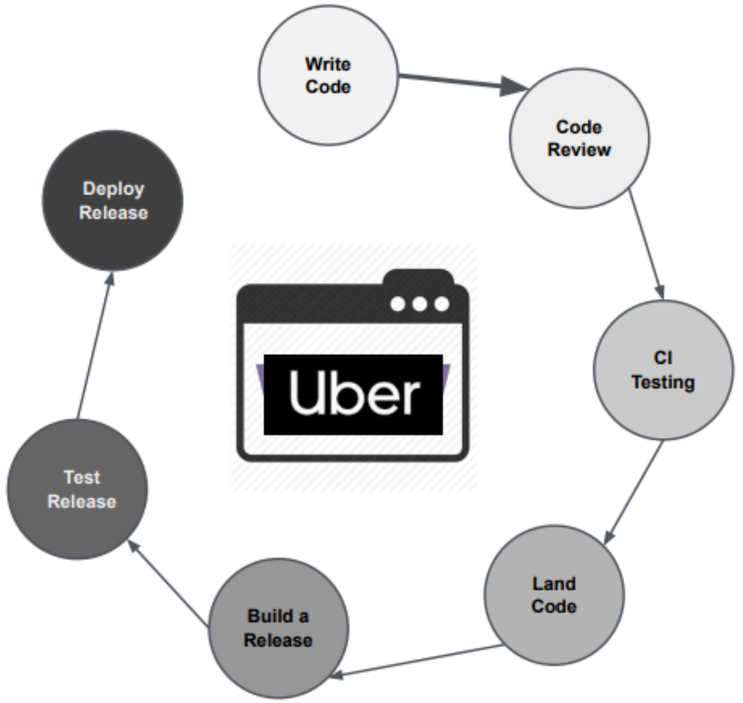


Transparency	Accuracy	Data Minimization	Privacy Rights
Lawfulness	Purpose Limitation	Storage Limitation	Privacy by Default

Microservices



Many more services



another service

another service

another service

another service

Uber iOS App

another service

another service

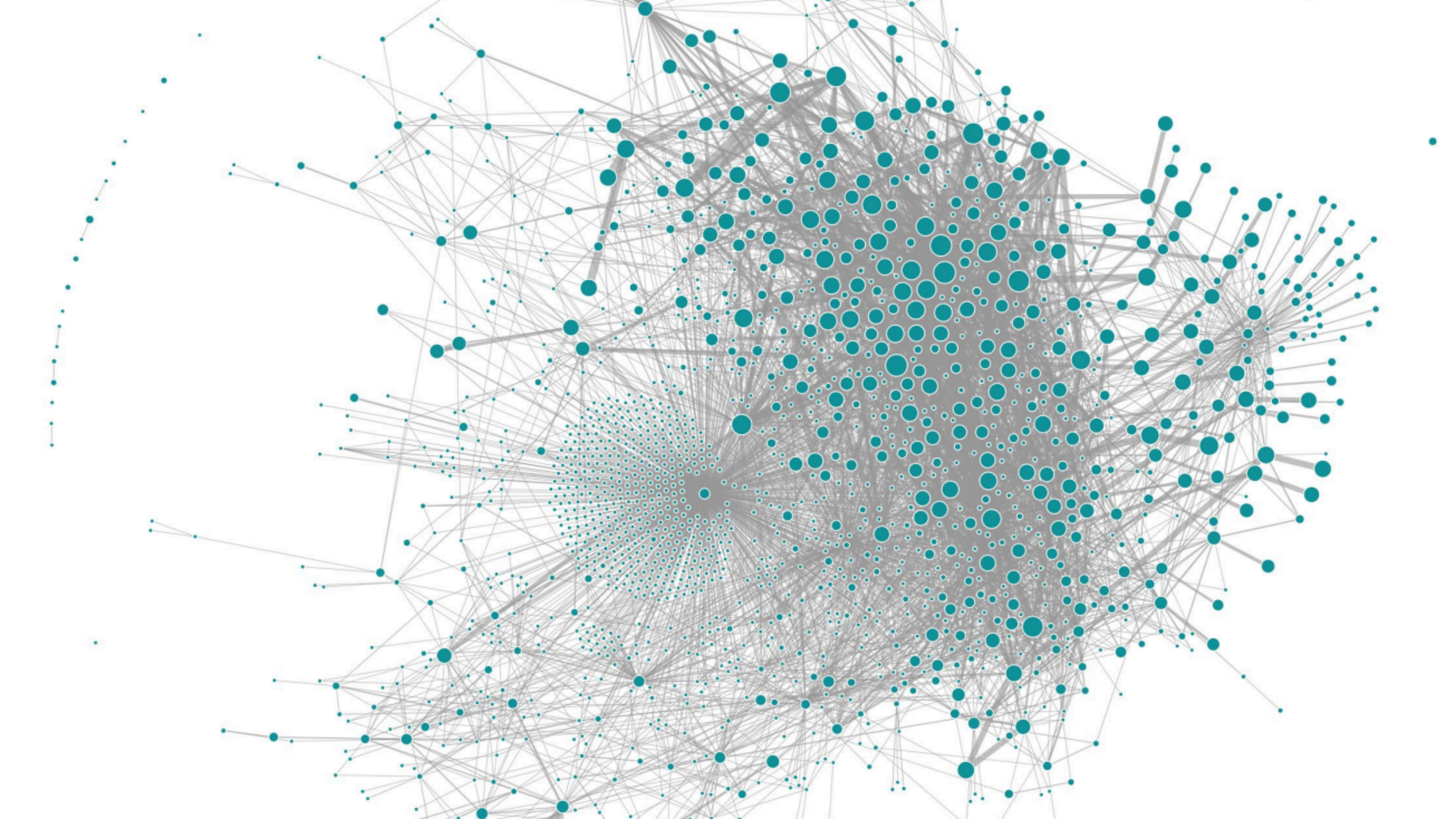
another service

Uber Android App

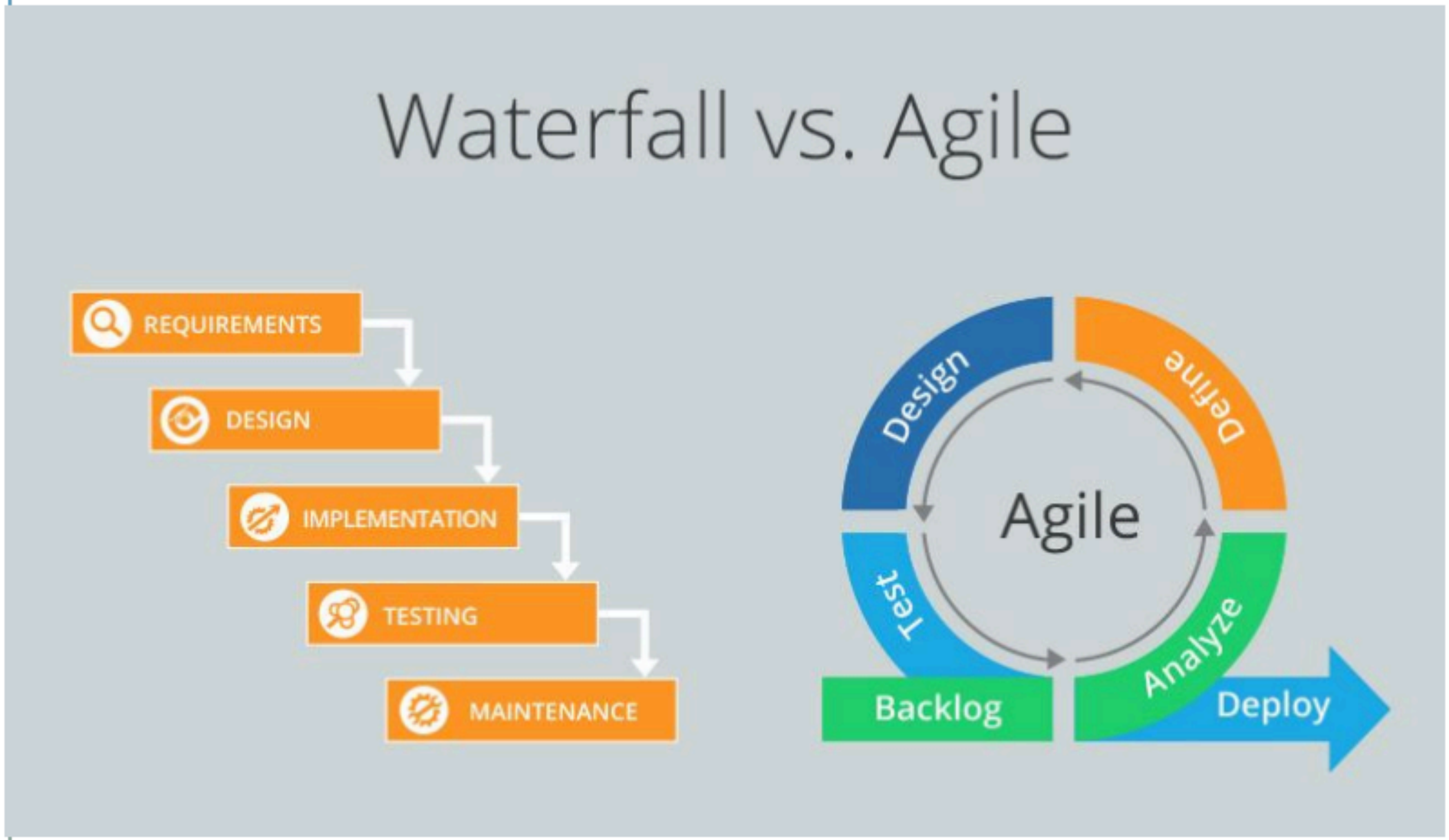
another service

another service

uber-web service

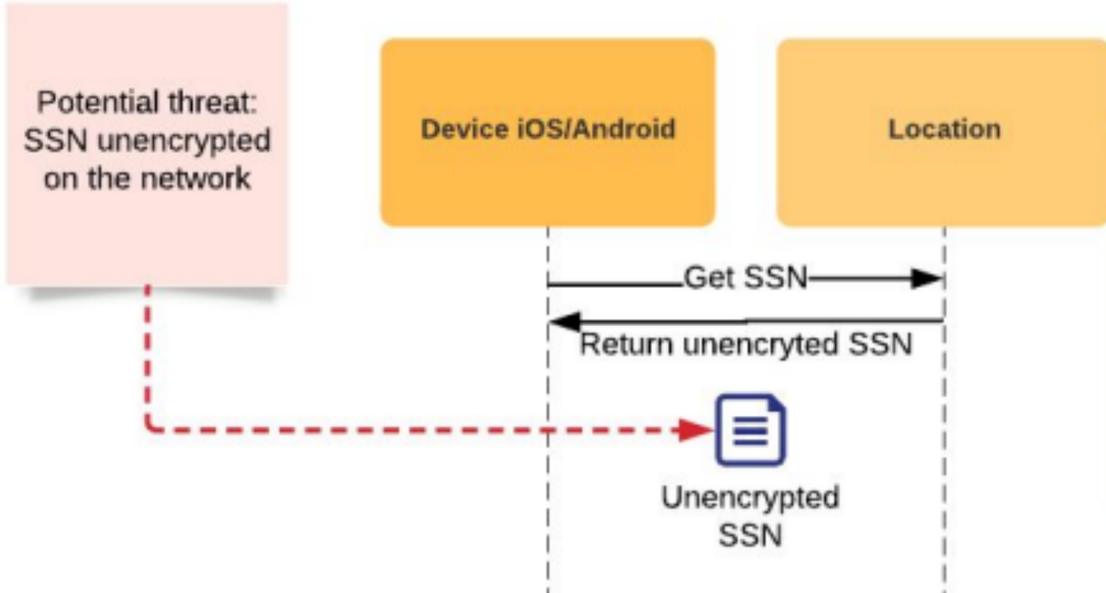


Agile Development



Challenge 1: System characterization

- Distributed data
- Mix of structured + unstructured data
- Off-the-shelf tools do not scale
- No (stable) architectural documentation



Data Classification	Example Category	Example Data Sets
Tier 1: Highly Restricted	Government Identifier & location	Driver's License
Tier 2: Restricted	Vehicle Data	License Plate Number Proof of Insurance
Tier 3: Confidential	Non-Identifying Vehicle Data	Make and Model Color
Tier 4: Public	Public Information	Product Brochures

Stakeholders

Data inventory

- Unified data category tags
- Automatic tagging and verification
- Maturity Levels: Tagging at DB level, tagging at column level, identify ALL data of an individual
- Use Data Inventory results to improve processes

LegalEase

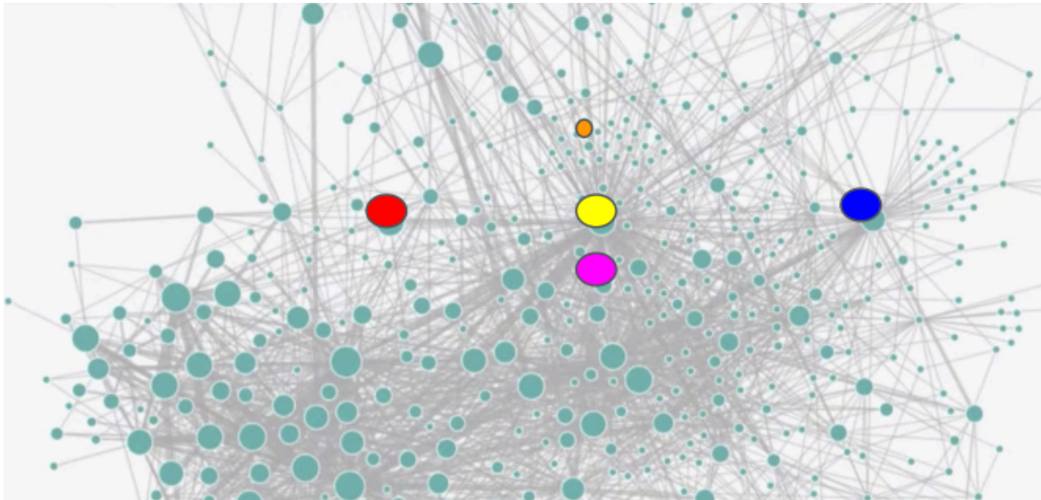
Bootstrapping Privacy Compliance in Big Data Systems

Shayak Sen, Saikat Guha, Anupam Dutta, Sriram Rajamani, Janice Tsai, Jeannette Wing
Proceedings of the 35th IEEE Symposium on Security & Privacy (Oakland) | May 2014
Published by IEEE

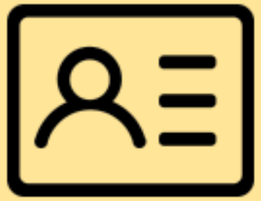
With the rapid increase in cloud services collecting and using user data to offer personalized experiences, ensuring that these services comply with their privacy policies has become a business imperative for building user trust. However, most compliance efforts in industry today rely on manual review processes and audits designed to safeguard user data, and therefore are resource intensive and lack coverage. In this paper, we present our experience building and operating a system to automate privacy policy compliance checking in Bing. Central to the design of the system are (a) L EGALEASE —a language that allows specification of privacy policies that impose restrictions on how user data is handled; and (b) G ROK —a data inventory for Map-Reduce-like big data systems that tracks how user data flows among programs. G ROK maps code-level schema elements to datatypes in LEGALEASE , in essence, annotating existing programs with information flow types with minimal human input. Compliance checking is thus reduced to information flow analysis of big data systems. The system, bootstrapped by a small team, checks compliance daily of millions of lines of ever-changing source code written by several thousand developers.

Challenge 2: Threats and Mitigation

- Privacy threat of the service vs the whole chain
- Where to place the control?
- Privacy can be slow vs Agile fast
- Legacy systems and privacy debt
- Resource and costs



Modular Reviews: Technical Privacy Consulting

Which personal data? 	Purpose and Access 	Retention 
Source And Recipients 	Storage Location 	Onboarding to PrivEng Services 

Modular Reviews: Technical Privacy Consulting

- Outcome:
  - Technical privacy requirements for this specific project
  - Mitigation prioritization
  - Input to Privacy Legal
- After the review:
  - Further analysis of platforms based on discovered knowledge
  - Embed privacy into platforms
  - Update Data Classification and Handling Standard

IRB Review


IRB REVIEW PROCESS
1
Complete IRB Request for Review
2
Submit Request, Documents, and CITI to irb@ualr.edu
3
Receive IRB Reviewer Feedback in 10 Business Days
4
Possible Decisions:
Approved
NHPR
Further Review
Full Board
IRB Decision is Required Prior to Contacting Participants or Collecting Data
CLICK TO EXPAND

IRB Review


IRB Levels of Review
Full Board
• Greater than minimal risk*
(e.g. drug or domestic violence study)
• Reviewed by convened IRB Committee
Expedited
• Minimal risk* (e.g. blood draw/alcoholism)
• Reviewed by IRB chair or designee
Exempt
• Not greater than minimal risk*
(e.g. blood pressure/educational tests)
• Reviewed by IRB chair, designee, or staff
Coded Data/ Specimens
• Analyzing items like discarded tissue
• Reviewed by IRB staff
Not Human Subjects Research
• Does not meet federal definition or is coded data
• Reviewed by IRB staff
*Minimal risk means that the probability and magnitude of harm or discomfort anticipated in the research are not greater than those ordinarily encountered in daily life or during routine physical or psychological examinations or tests.

Homeland Security

Topics | How Do I? | Get Involved | News | About DHS

Home > About DHS > Organization > Office of the Secretary > Privacy Office > **Chief Privacy Officer's Authorities and Responsibilities**

Share / Email

Privacy Office

Authorities and Responsibilities

Contacts

Authorities and Responsibilities of the Chief Privacy Officer

The activities of the Privacy Office serve to build privacy into departmental programs. The following is a framework of privacy laws through which the Privacy Office accomplishes its activities and mission:

- Privacy Act of 1974, as amended (5 U.S.C. § 552a): Embodies a code of fair information principles that governs the collection, maintenance, use, and dissemination of personally identifiable information by federal agencies;
- E-government Act of 2002 (Public Law 107-347): Mandates Privacy Impact Assessments (PIAs) for all Federal agencies when there are new collections of, or new technologies applied to, personally identifiable information;

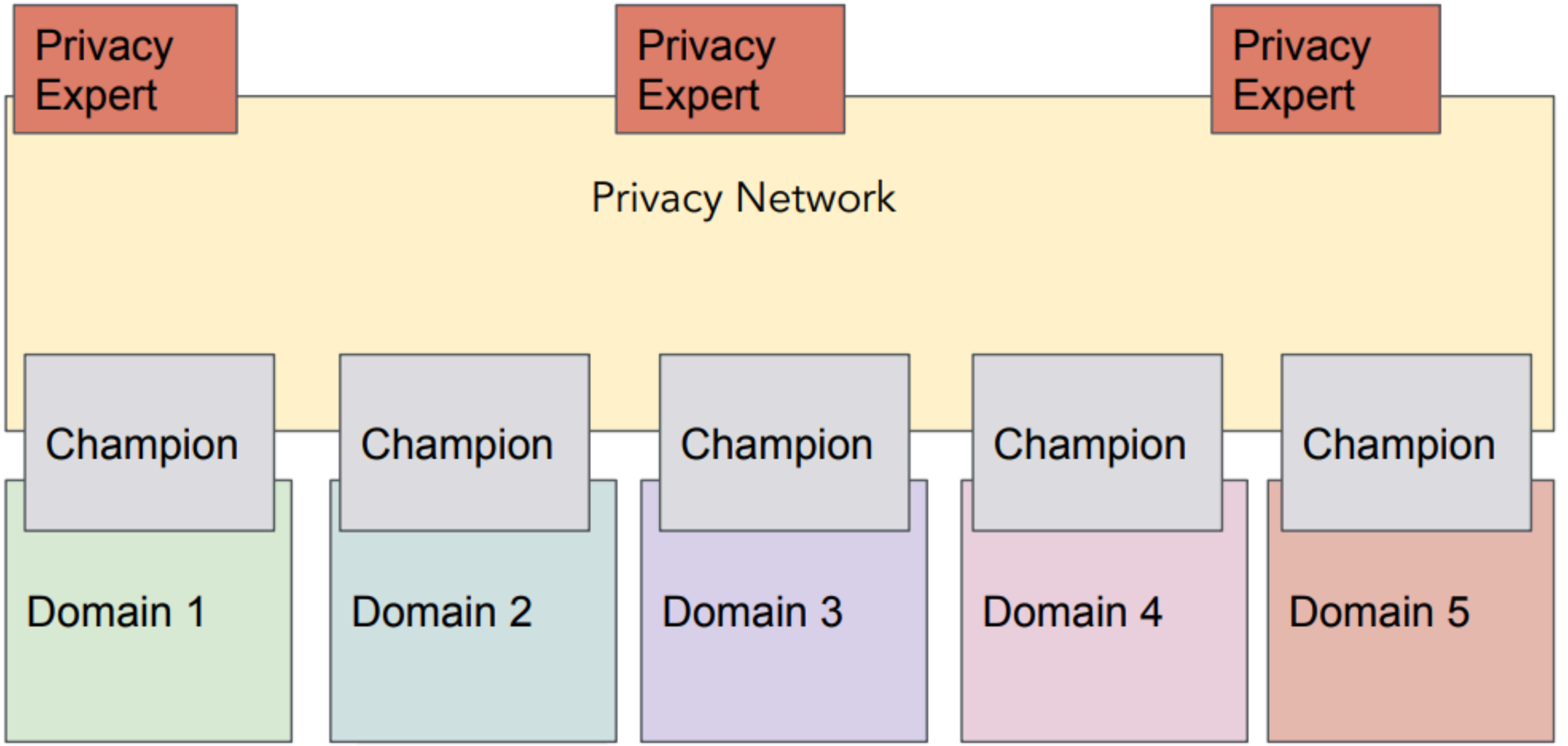
Home > Security

SECURITY ADVISER

By Roger A. Grimes | Follow

Why your company needs a chief privacy officer

Uber Approach: Education and Privacy Champions



Chief Privacy Officers

- Companies are increasingly appointing CPOs to have a central point of contact for privacy concerns
- Role of CPO varies in each company
  - Draft privacy policy
  - Respond to customer concerns
  - Educate employees about company privacy policy
  - Review new products and services for compliance with privacy policy
  - Develop new initiatives to keep company out front on privacy issue
  - Monitor pending privacy legislation

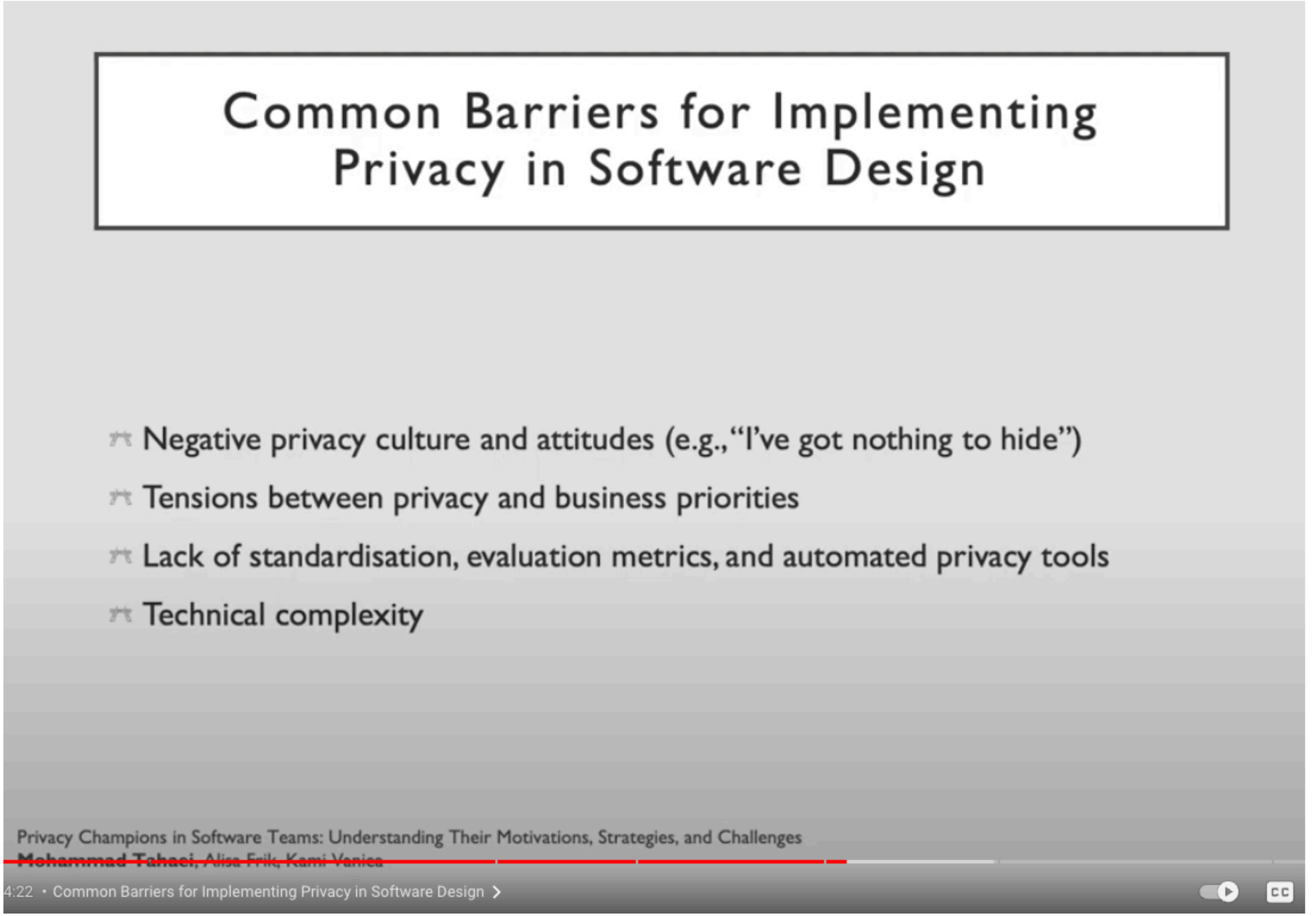
Motivations: Champions

- "Champions" advocate for a cause (e.g., an innovation or idea), encourage others to engage, and aid with overcoming barriers that a new idea could face
- Literature in software engineering shows champions' value in promoting both security and new software technologies
- And what about privacy?

Privacy Champions in Software Teams: Understanding Their Motivations, Strategies, and Challenges
Mohammad Tahaei, Alisa Frik, Kami Vaniea

Privacy Champions

- Formally or informally promote best practices for users' privacy, educate others, persuade, and advocate for privacy adoption throughout the software development process
- Have an official or unofficial role within their team acting as the "voice" of users' privacy for the product or team, for example by giving privacy-related advice that can influence decisions and privacy practices


Common Barriers for Implementing Privacy in Software Design
Negative privacy culture and attitudes (e.g., "I've got nothing to hide")
Tensions between privacy and business priorities
Lack of standardisation, evaluation metrics, and automated privacy tools
Technical complexity
Privacy Champions in Software Teams: Understanding Their Motivations, Strategies, and Challenges
Mohammad Tahaei, Alisa Frik, Kami Vaniea
4:22 • Common Barriers for Implementing Privacy in Software Design
CC

Strategies for Promoting Privacy

Effective

- Regular privacy-focused meetings and informal discussions
- Management support, facilitation of communication among stakeholders (e.g., between legal and product teams)
- Appropriate privacy documentation and guidelines
- Incorporating privacy considerations into design reviews

Not effective

- Punishing developers for not implementing privacy features
- Company-wide awareness programs or on-boarding privacy training for new hires

Privacy Champions in Software Teams: Understanding Their Motivations, Strategies, and Challenges
Mohammad Tahaei, Alisa Frik, Kami Vaniea

How to Support and Attract Privacy Champions

- Embed privacy values in the organisational culture
- Include privacy topics in university degree and online learning curricula
- Offer hands-on privacy-oriented projects (e.g., privacy hackathons)
- Acknowledge privacy-oriented efforts (by both colleagues and management)
- Provide resources (compensate extra time, or let spend 10-20% of their time on privacy work)

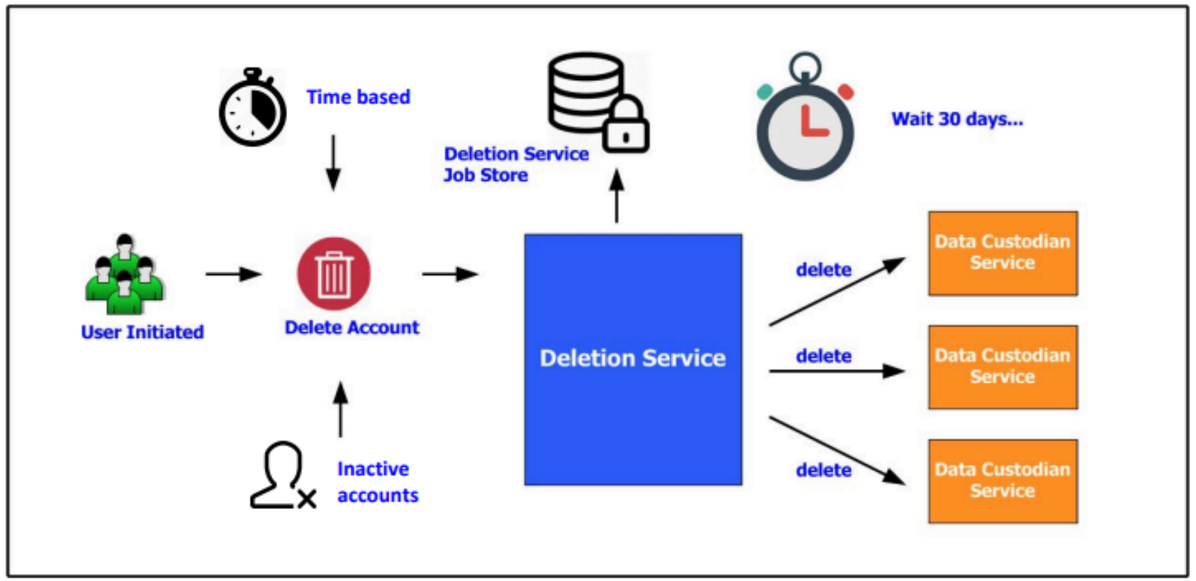
Privacy Champions in Software Teams: Understanding Their Motivations, Strategies, and Challenges
Mohammad Tahaei, Alisa Frik, Kami Vaniea

Challenge 3: Doing Privacy at Scale: Deletion

- Multiple use cases: user initiated account deletion, inactive account deletion, time-based deletion
- Variety of data stores
- Scalable, reliable, adaptable, demonstrable

Uber's Approach to Data Deletion

- Support scale of data, data stores, and microservices
- Privacy Impact Assessment and Technical Privacy Reviews
- Vetting process combines legal and technical privacy
- Automate onboarding process for new services



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

Takeaway Msgs

- Privacy Design is a spectrum.
- Two aspects on privacy designs.
  - Consequential and procedural.
  - Due Diligence.
- The life cycle of privacy design.
  - Articulating the system design
  - Modeling the potential threats
  - Navigating the design space
  - Implementing

Credits

1. PRIVACY AT SPEED: PRIVACY BY DESIGN FOR AGILE DEVELOPMENT AT UBER, https://www.usenix.org/conference/enigma2020/presentation/bozdag
2. Privacy Policy, Law, and Technology, Lorrie Cranor, https://cups.cs.cmu.edu/courses/pplt-fa15/