Stockholm University

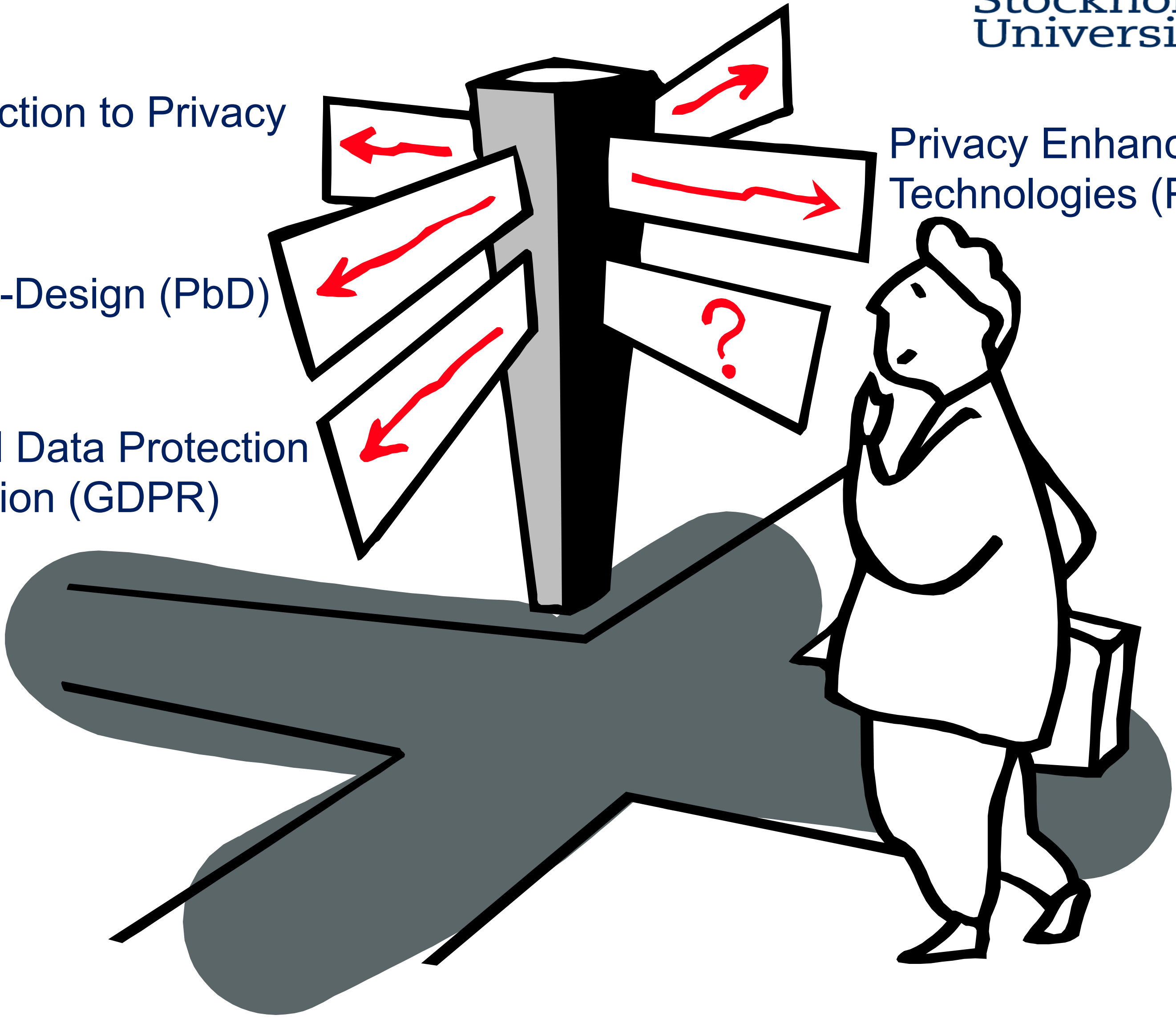
Privacy

Haris Mouratidis

Learning outcomes

- On completion of this session you should be able to
  - Demonstrate a basic understanding of information privacy;
  - Become familiar with Privacy-Enhancing Technologies (PETs);
  - Become familiar with Privacy-By-Design and its principles;
  - Understand the main principles and ideas behind GDPR.

Layout



An Introduction to Privacy

Privacy-by-Design (PbD)

General Data Protection Regulation (GDPR)

Privacy Enhancing Technologies (PETs)

Privacy

- Privacy
  - fundamental human right
    - UN Declaration of Human Rights,
    - the International Covenant on Civil and Political Rights
    - International and regional treaties

- *"the right to be left alone"* (Warren & Brandeis, 1890).

- *"The claim of individuals, groups and institutions to determine for themselves, when, how and to what extent information about them is communicated to others"* (Westin, 1967).

Privacy as a concept

- In general the concept of privacy incorporates three aspects:
  - *Territorial Privacy*, by protecting the close physical area surrounding a person;
  - *Privacy of the person*, by protecting a person against undue interference;
  - *Informational privacy*, by controlling whether and how personal data can be gathered, stored, processed or selectively disseminated.

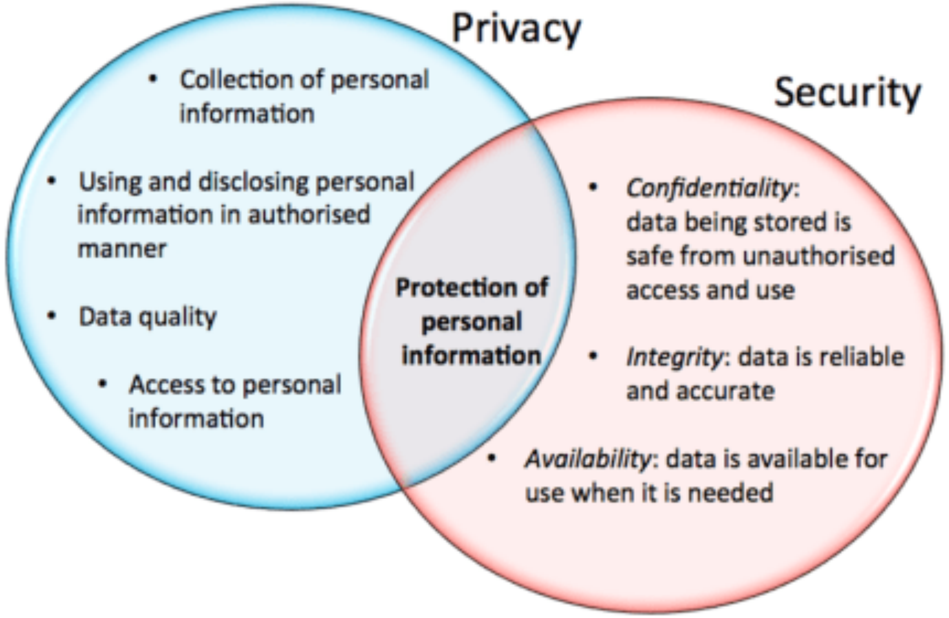
Why it is important?

- More and more sensitive and personal information on the public domain;

- Data Breaches
  - Swedish Transport Agency
  - Swedish HealthCare Guide service

Security vs Privacy

- Security proposals balance tangible harms
- Privacy proposals protect intangible harms that an intrusion on privacy would cause
- Implementation level: Interconnection between them
  - Security mechanisms are used as the basis for privacy implementation
- Requirements level: often there is conflict
  - Security requirements might conflict privacy requirements


Privacy
Security
• Collection of personal information
• Using and disclosing personal information in authorised manner
• Data quality
• Access to personal information
Protection of personal information
• *Confidentiality*: data being stored is safe from unauthorised access and use
• *Integrity*: data is reliable and accurate
• *Availability*: data is available for use when it is needed

Privacy Properties [Security based]

- Authentication:
  - provision of assurance that a claimed characteristic of an entity is correct
- Authorisation:
  - User’s private data should only be accessed by authorised users
- Identification:
  - the ability to identify uniquely a user of a system or an application that is running in the system.
- Data protection:
  - The protection of personal data in order to guarantee privacy

Privacy Properties

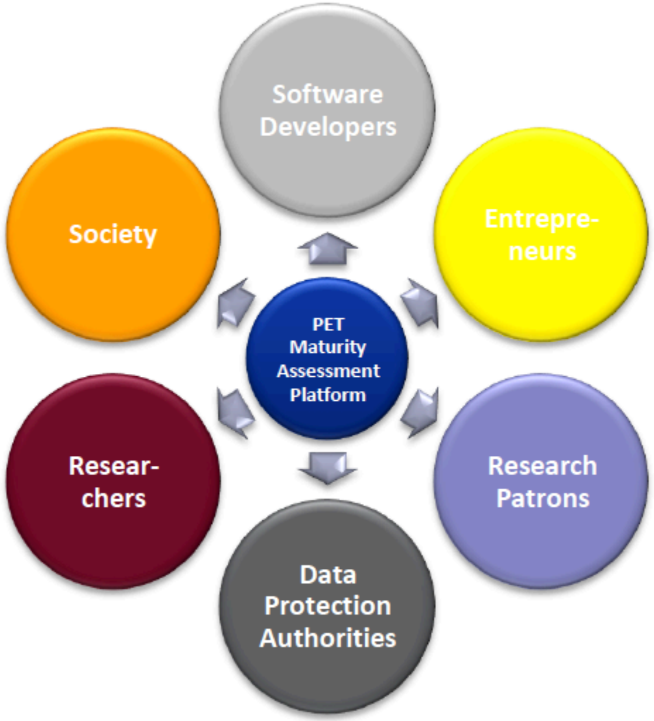
- Anonymity:
  - A characteristic of information that does not permit a personally identifiable information subject to be identified directly or indirectly.
- Pseudonymity:
  - The utilisation of an alias instead of personally identifiable information
- Unlinkability:
  - The use of a resource or a service by a user without a third party being able to link the user with the service
- Unobservability:
  - The inability of a third party to observe if a user is using a service
- Undetectability:
  - The inability for a third party to distinguish who is the user (among a set of potential users) using a service

Brief History of Privacy Engineering Research

- What is Privacy Engineering?
- 1970s: Privacy Technologies: anonymous electronic communication, transactions, and payment.
- 1990s, the idea of shaping technology according to privacy principles was discussed among Privacy and Data Protection Commissioners.
  - Main principles: data minimisation and identity protection by anonymisation or pseudonymisation.
- 2000s privacy is looked at a more multi-disciplinary level and from earlier in the process, not just from a technological perspective but also from a design perspective
  - Privacy-by-design/privacy-by-default

Privacy Enhancing Technologies

- PETs are a coherent system of ICT measures that protect privacy by eliminating or reducing personal data or by preventing unnecessary and/or undesired processing of personal data, all without losing the functionality of the information system.



[From ENISA]

PETs categorisation

- ENISA identifies four major categories of technology: secure messaging, virtual private networks, anonymizing networks and anti-tracking tools for online browsing

- Other researchers have categorised PETs “according to their technical contributions” (e.g., anonymous communication, and privacy preserving data mining)

Some examples of PETs

- Encryption
- Anonymisation
  - Differential Privacy
  - K-anonymity
  - Onion routing
- Pseudonymisation
  - Masking / obfuscation
  - Tokenisation

Privacy By Design

- Canada’s Information and Privacy Commissioner advocated the need for a privacy-by-design (PbD) approach

“privacy to be embedded into design as a preventive and proactive measure”.

  - At system design stage and at time of processing
  - Defined seven principles

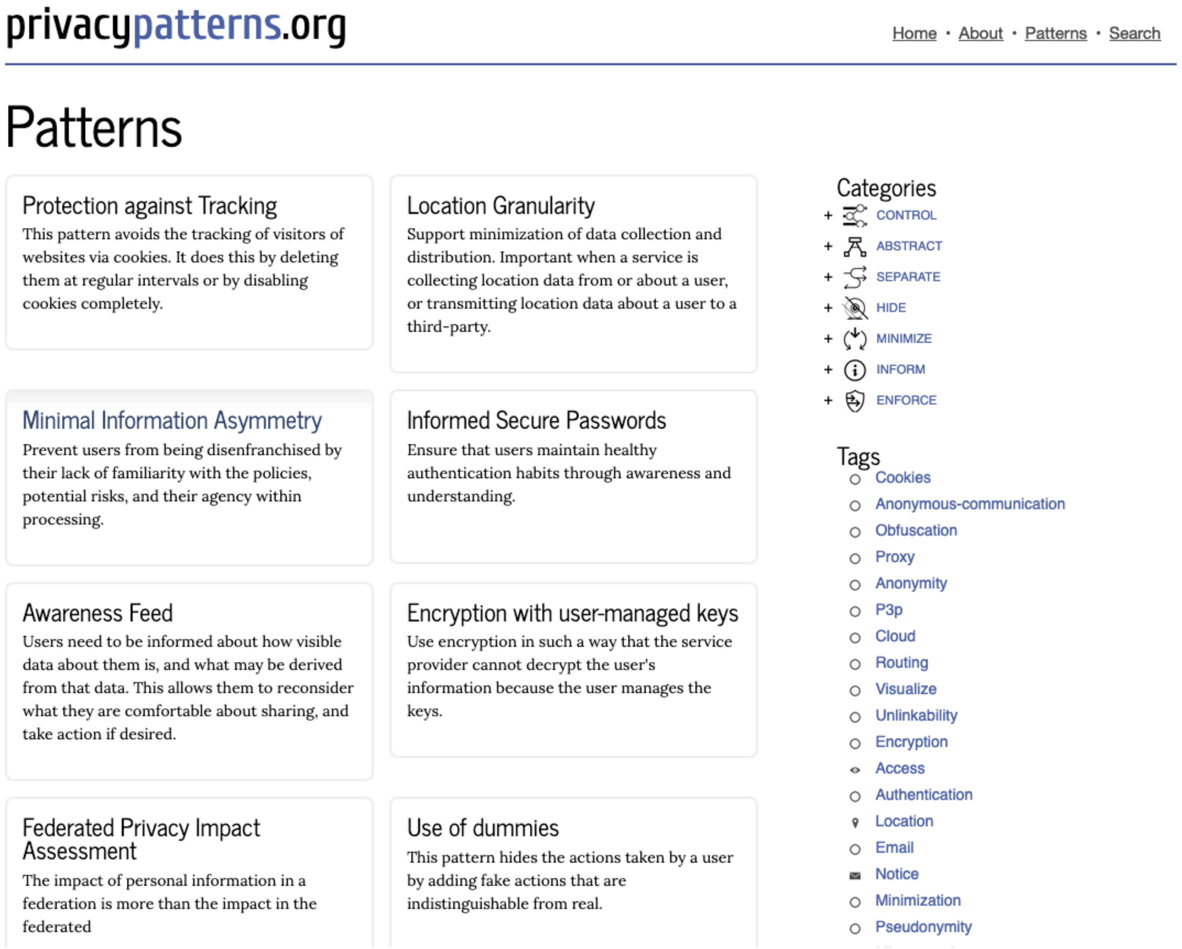
- Privacy as a design criterion received much attention in recent years

PbD Principles

- Proactive, not reactive; preventative, not remedial
- Privacy as the Default Setting
- Privacy Embedded into Design
- Full Functionality: Positive – Sum, not Zero – Sum
- End-to-End Security – Full Lifecycle Protection
- Visibility and Transparency
- Respect of User Privacy – Keep it User-Centric

Data Privacy Patterns

- Privacy patterns are design solutions to common privacy problems — a way to translate "privacy-by-design" into practical advice for software engineering.



privacypatterns.org

Home · About · Patterns · Search

Patterns

Protection against Tracking
This pattern avoids the tracking of visitors of websites via cookies. It does this by deleting them at regular intervals or by disabling cookies completely.

Location Granularity
Support minimization of data collection and distribution. Important when a service is collecting location data from or about a user, or transmitting location data about a user to a third-party.

Minimal Information Asymmetry
Prevent users from being disenfranchised by their lack of familiarity with the policies, potential risks, and their agency within processing.

Informed Secure Passwords
Ensure that users maintain healthy authentication habits through awareness and understanding.

Awareness Feed
Users need to be informed about how visible data about them is, and what may be derived from that data. This allows them to reconsider what they are comfortable about sharing, and take action if desired.

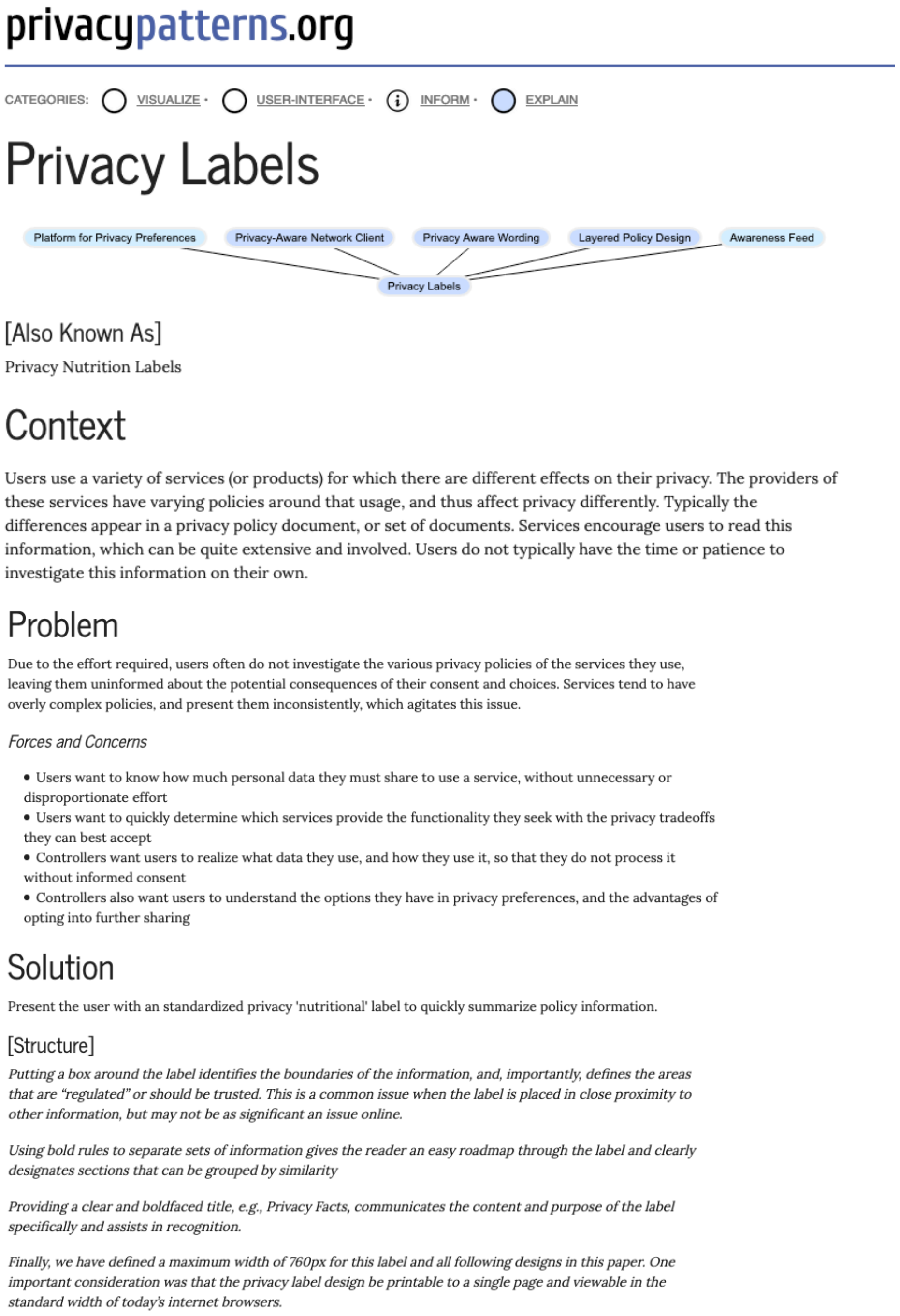
Encryption with user-managed keys
Use encryption in such a way that the service provider cannot decrypt the user's information because the user manages the keys.

Federated Privacy Impact Assessment
The impact of personal information in a federation is more than the impact in the federated

Use of dummies
This pattern hides the actions taken by a user by adding fake actions that are indistinguishable from real.

Categories
- + CONTROL
- + ABSTRACT
- + SEPARATE
- + HIDE
- + MINIMIZE
- + INFORM
- + ENFORCE

Tags
- ○ Cookies
- ○ Anonymous-communication
- ○ Obfuscation
- ○ Proxy
- ○ Anonymity
- ○ P3p
- ○ Cloud
- ○ Routing
- ○ Visualize
- ○ Unlinkability
- ○ Encryption
- ○ Access
- ○ Authentication
- ○ Location
- ○ Email
- ○ Notice
- ○ Minimization
- ○ Pseudonymity



privacypatterns.org

CATEGORIES: ○ VISUALIZE · ○ USER-INTERFACE · ⓘ INFORM · ○ EXPLAIN

Privacy Labels

[Also Known As]

Privacy Nutrition Labels

Context

Users use a variety of services (or products) for which there are different effects on their privacy. The providers of these services have varying policies around that usage, and thus affect privacy differently. Typically the differences appear in a privacy policy document, or set of documents. Services encourage users to read this information, which can be quite extensive and involved. Users do not typically have the time or patience to investigate this information on their own.

Problem

Due to the effort required, users often do not investigate the various privacy policies of the services they use, leaving them uninformed about the potential consequences of their consent and choices. Services tend to have overly complex policies, and present them inconsistently, which agitates this issue.

Forces and Concerns

- Users want to know how much personal data they must share to use a service, without unnecessary or disproportionate effort
- Users want to quickly determine which services provide the functionality they seek with the privacy tradeoffs they can best accept
- Controllers want users to realize what data they use, and how they use it, so that they do not process it without informed consent
- Controllers also want users to understand the options they have in privacy preferences, and the advantages of opting into further sharing

Solution

Present the user with an standardized privacy 'nutritional' label to quickly summarize policy information.

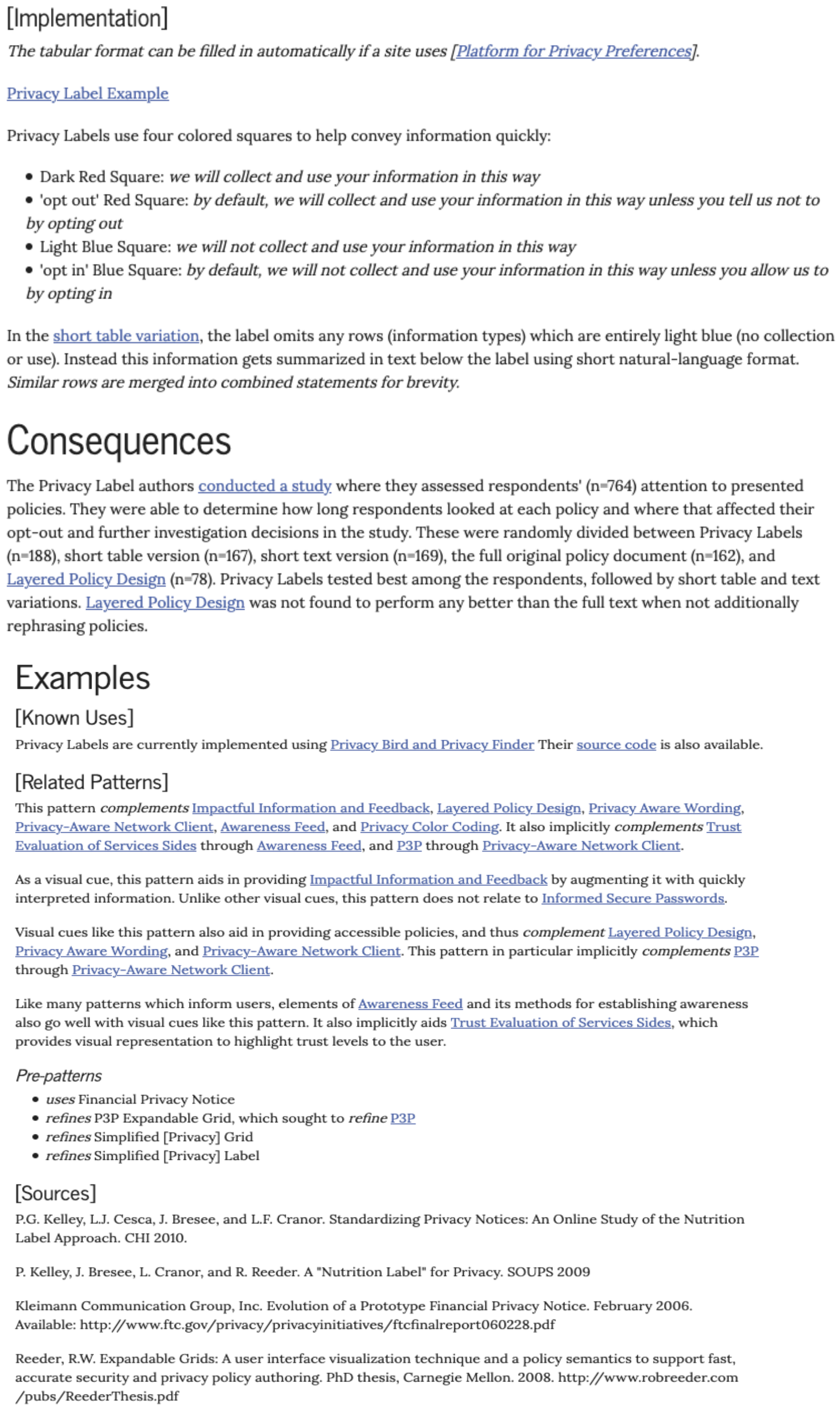
[Structure]

Putting a box around the label identifies the boundaries of the information, and, importantly, defines the areas that are "regulated" or should be trusted. This is a common issue when the label is placed in close proximity to other information, but may not be as significant an issue online.

Using bold rules to separate sets of information gives the reader an easy roadmap through the label and clearly designates sections that can be grouped by similarity

Providing a clear and boldfaced title, e.g., Privacy Facts, communicates the content and purpose of the label specifically and assists in recognition.

Finally, we have defined a maximum width of 760px for this label and all following designs in this paper. One important consideration was that the privacy label design be printable to a single page and viewable in the standard width of today's internet browsers.



[Implementation]

The tabular format can be filled in automatically if a site uses [Platform for Privacy Preferences].

Privacy Label Example

Privacy Labels use four colored squares to help convey information quickly:

- Dark Red Square: *we will collect and use your information in this way*
- 'opt out' Red Square: *by default, we will collect and use your information in this way unless you tell us not to by opting out*
- Light Blue Square: *we will not collect and use your information in this way*
- 'opt in' Blue Square: *by default, we will not collect and use your information in this way unless you allow us to by opting in*

In the short table variation, the label omits any rows (information types) which are entirely light blue (no collection or use). Instead this information gets summarized in text below the label using short natural-language format. *Similar rows are merged into combined statements for brevity.*

Consequences

The Privacy Label authors conducted a study where they assessed respondents' (n=764) attention to presented policies. They were able to determine how long respondents looked at each policy and where that affected their opt-out and further investigation decisions in the study. These were randomly divided between Privacy Labels (n=188), short table version (n=167), short text version (n=169), the full original policy document (n=162), and Layered Policy Design (n=78). Privacy Labels tested best among the respondents, followed by short table and text variations. Layered Policy Design was not found to perform any better than the full text when not additionally rephrasing policies.

Examples

[Known Uses]

Privacy Labels are currently implemented using Privacy Bird and Privacy Finder Their source code is also available.

[Related Patterns]

This pattern *complements* Impactful Information and Feedback, Layered Policy Design, Privacy Aware Wording, Privacy-Aware Network Client, Awareness Feed, and Privacy Color Coding. It also implicitly *complements* Trust Evaluation of Services Sides through Awareness Feed, and P3P through Privacy-Aware Network Client.

As a visual cue, this pattern aids in providing Impactful Information and Feedback by augmenting it with quickly interpreted information. Unlike other visual cues, this pattern does not relate to Informed Secure Passwords.

Visual cues like this pattern also aid in providing accessible policies, and thus *complement* Layered Policy Design, Privacy Aware Wording, and Privacy-Aware Network Client. This pattern in particular implicitly *complements* P3P through Privacy-Aware Network Client.

Like many patterns which inform users, elements of Awareness Feed and its methods for establishing awareness also go well with visual cues like this pattern. It also implicitly aids Trust Evaluation of Services Sides, which provides visual representation to highlight trust levels to the user.

Pre-patterns

- *uses* Financial Privacy Notice
- *refines* P3P Expandable Grid, which sought to *refine* P3P
- *refines* Simplified [Privacy] Grid
- *refines* Simplified [Privacy] Label

[Sources]

P.G. Kelley, L.J. Cesca, J. Bresee, and L.F. Cranor. Standardizing Privacy Notices: An Online Study of the Nutrition Label Approach. CHI 2010.

P. Kelley, J. Bresee, L. Cranor, and R. Reeder. A "Nutrition Label" for Privacy. SOUPS 2009

Kleimann Communication Group, Inc. Evolution of a Prototype Financial Privacy Notice. February 2006. Available: http://www.ftc.gov/privacy/privacyinitiatives/ftcfinalreport060228.pdf

Reeder, R.W. Expandable Grids: A user interface visualization technique and a policy semantics to support fast, accurate security and privacy policy authoring. PhD thesis, Carnegie Mellon. 2008. http://www.robreeder.com/pubs/ReederThesis.pdf

Privacy Legal dimension

- The General Data Protection Regulation (GDPR)
- Most fundamental change to data protection law in almost 20 years?
- The key changes from the current law are to strengthen rights of individuals and place more obligations on organisations in looking after personal data.
- Applies to:
  - Organisations operating within EU;
  - Non-EU organisations offering good/services within EU or processing EU data.

Financial Implications

- Up to €20M, or up to 4 % of the total worldwide annual turnover of the preceding financial year, whichever is higher
- The Swedish Data Protection Authority:
  - imposed a fine of 75M Kr. / €7M on Google for failure to comply with the GDPR.
  - issued a penalty of 200,000 Kr. to a school which used biometric facial recognition to record student attendance for violating GDPR.
- France: €50M Google
- Germany: €35,3M H&M
- Italy: €27,8M TIM

(Some) GDPR Terminology

- The **data controller** is the person or organisation who determines the how and what of data processing.
- The **data subject** is the person about whom personal data is being processed.
- A **data processor** is the person or organisation who takes an action with the personal data you control – this might be a 3rd party acting on the data controller behalf.
- **Processing** is anything done with/to personal data, including storing it.
- The **Data Protection Officer (DPO)** is a specific role which is a legal requirement for many organisations.

Personal vs Sensitive data

Personal data is defined as:

Any information, which directly or indirectly, could identify a person.

Sensitive personal data is defined as:

Any information relating to an individual's racial or ethnic origin, political opinions, *religious beliefs*, trade union membership, physical or mental health or condition, genetic data, biometric data and others.

GDPR Principles

- **Lawfulness, fairness and transparency** – as with Data Protection
- **Purpose limitation** – only collect for specific purposes and then don't use it for other purposes
- **Data minimisation** – only collect the data you need for the purpose you are using it
- **Accuracy** – as now, keep it up to date!
- **Storage limitation** – don't keep it for longer than you need to fulfil the purpose
- **Integrity and confidentiality** – keep it safe and secure e.g. encrypted if on a laptop or mobile phone.
- **Accountability** – you must be able to prove you have complied with the above.

Data Protection by Design and Default

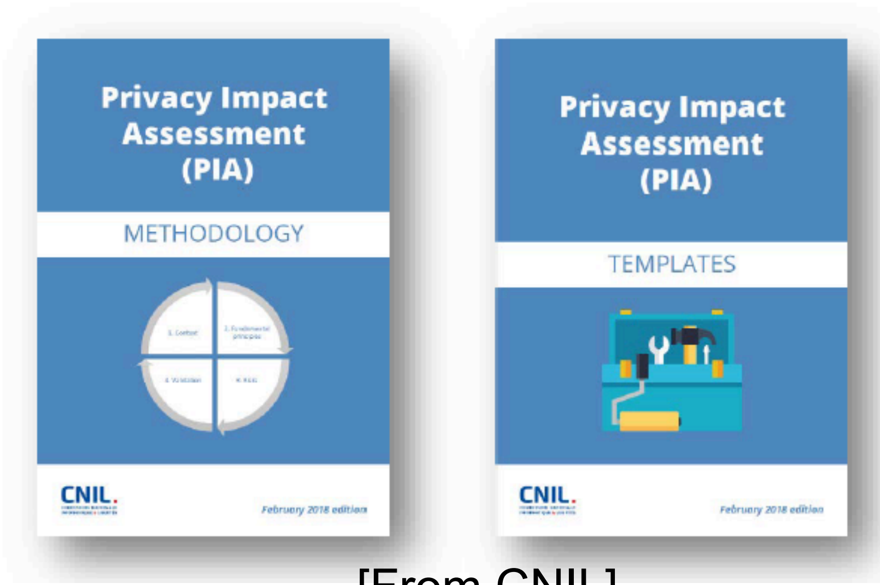
- Put in place appropriate technical and organisational measures to implement the data protection principles effectively; and

- Integrate safeguards into your processing so that you meet the GDPR’s requirements and protect individual rights

- Makes Privacy-by-design and Privacy-by-default legal requirement

Data Protection Impact Assessment

- A DPIA is a type of risk assessment.
- Your DPIA must:
  - describe the nature, scope, context and purposes of the processing;
  - assess necessity, proportionality and compliance measures;
  - identify and assess risks to individuals; and
  - identify any additional measures to mitigate those risks.



Privacy Impact Assessment (PIA)

METHODOLOGY

CNIL.

February 2018 edition

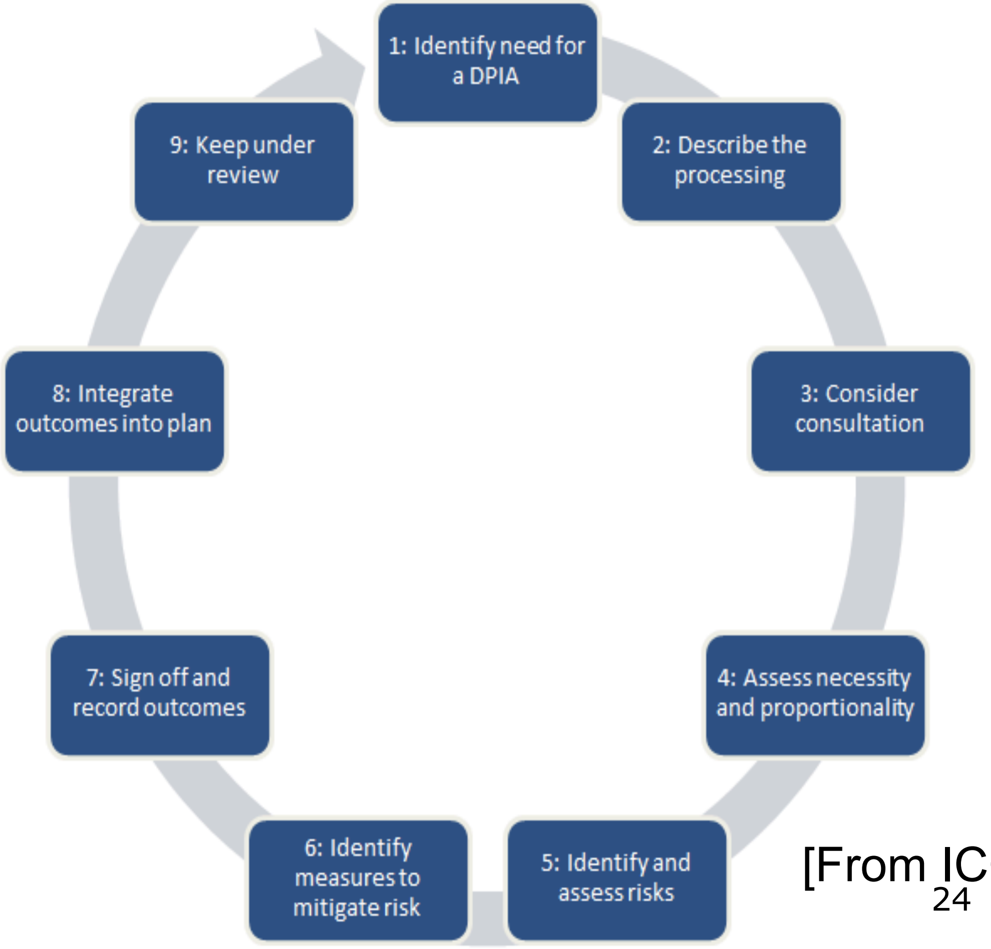
Privacy Impact Assessment (PIA)

TEMPLATES

CNIL.

February 2018 edition

[From CNIL]



1: Identify need for a DPIA

2: Describe the processing

3: Consider consultation

4: Assess necessity and proportionality

5: Identify and assess risks

6: Identify measures to mitigate risk

7: Sign off and record outcomes

8: Integrate outcomes into plan

9: Keep under review

[From ICO]

Lawful Processing

- Consent:
  - you can process personal data where the data subject has given consent to the processing for one or more specified purposes;
- Contract with individual:
  - you can process personal data, without consent, where required under a contract with the data subject E.g. employment contract, contract for sale of goods or services;
- Legal Obligations:
  - the processing is necessary for you to comply with the law (not including contractual obligations);

Lawful Processing II

- Vital interests:
  - – you can process personal data, without consent, if it is necessary to protect someone's life;
- Public Task:
  - – you can process personal data, without consent, to carry out official functions or a task in the public interest –and where you have legal basis for the processing under local law;
- Legitimate interest:
  - – you can process personal data without consent if you have a genuine and legitimate reason to do so.

Conclusions

- Information Privacy is important
- It is related to Security but it is not Security
- New Regulation requires compliance or risk of heavy fines