GDPR Legal Principles and Privacy By Design Strategies

Josep Domingo-Ferrer

Universitat Rovira i Virgili, Tarragona
CYBERCAT-Center for Cybersecurity Research of Catalonia

josep.domingo@urv.cat

Bamberg, March 2019

1 Introduction

2 Privacy principles from the law

3 Privacy design strategies

General bibliography

- G. Danezis, J. Domingo-Ferrer, M. Hansen, J.-H. Hoepman, D. Le Métayer, R. Tirtea and S. Schiffner, *Privacy and Data Protection by Design From Policy to Engineering*, European Union Agency for Network and Information Security-ENISA, 2015.
- A. Hundepool, J. Domingo-Ferrer, L. Franconi, S. Giessing, E. Schulte Nordholt, K. Spicer and P.-P. de Wolf, *Statistical Disclosure Control*, Wiley, 2012.

Basic privacy concepts

- Privacy is a fundamental human right:
  - Art. 12, Universal Declaration of Human Rights
  - Arts. 7 and 8, Charter of Fundamental Rights of the EU
- However, in the digital world,
  - Data collectors determine what and how data are processed;
  - Data subjects whose data is at stake need privacy protection.

Privacy-enhancing tecnologies

- Technologies must not only be designed to collect and process more data more efficiently but also to protect privacy (privacy by design, PbD).
- Privacy-enhancing technologies (PETs) seek to minimize the processing of personal data.
- Yet privacy cannot be guaranteed by just technology
  $\Longrightarrow$ A legal framework is needed.

Personal data protection principles in the EU law

Personal data, or more precisely personally identifiable information (PII) mean any information related to an identified or identifiable natural person.
Principles applicable to PII (Art. 29 DP Working party, European General Data Protection Regulation-GDPR[1]):

- **Lawfulness** (consent obtained or processing needed for: a contract or legal obligation or the subject's vital interests or a public interest or legitimate processor's interests compatible with the subject's rights)
- **Consent** (simple, specific, informed and explicit)
- **Purpose limitation** (legitimate and specified before collection)

[1] https://gdpr-info.eu

Personal data protection principles in the EU law (II)

- **Necessity and data minimization** (collect only what is needed and keep only as long as needed)
- **Transparency and openness** (subjects need to get info about collection and processing in a way they understand)
- **Individual rights** (to access, rectify, erase/be forgotten)
- **Information security** (collected data protected against unauthorized access and processing, manipulation, loss, destruction, etc.)
- **Accountability** (ability to demonstrate compliance with principles)
- **Data protection by design and by default** (privacy built-in from the start rather than added later)

Personal big data conflict with principles

- Big data result from collecting and linking data from several sources, often in a continuous way
- Unless personal data are anonymized, potential conflicts with the above principles:
  - **Purpose limitation**. Big data often used secondarily for purposes not even known at collection time.
  - **Consent**. If purpose is not clear, consent cannot be obtained.
  - **Lawfulness**. Without purpose limitation and consent, lawfulness is dubious.
  - **Necessity and data minimization**. Big data result precisely from accumulating data for potential use.
  - **Individual rights**. Individuals do not even know which data are stored on them.
  - **Accountability**. Compliance does not hold and hence cannot be demonstrated.

Eight privacy design strategies

1. **MINIMIZE**
2. **HIDE**
3. **SEPARATE**
4. **AGGREGATE**
5. **INFORM**
6. **CONTROL**
7. **ENFORCE**
8. **DEMONSTRATE**

Strategy #1: MINIMIZE

- **The amount of personal data that is processed should be restricted to the minimum possible (relates to minimization).**
- By avoiding collection of unnecessary data, the possible privacy impact of a system is limited.
- Decide whether:
  - The processing of personal data is proportional to the purpose;
  - No other, less invasive means exist to achieve the same purpose.
- Design patterns: Select before you collect; anonymization, pseudonyms.

Strategy #2: HIDE

- **Any personal data and their interrelationships should be hidden from plain view (relates to information security).**
- Doing so ensures personal data cannot be easily abused.
- The HIDE strategy seeks unlinkability and unobservability (not easy given the many automatic identifiers: IP addresses, RFID tags, wi-fi SSID, etc.).
- Design patterns: encryption of data (stored or in transit), mix networks (to hide traffic patterns), attribute-based credentials for unlinkability, anonymization, pseudonyms.

Strategy #3: SEPARATE

- **Personal data should be processed in a distributed fashion, in separate compartments whenever possible.**
- By separating the processing or storage of several sources of personal data that belong to the same person, complete profiles of that person cannot be made.
- Distributed rather than centralized processing.
- Data from separate sources to be stored in separate databases.
- Data to be processed and stored locally as much as possible.
- Database tables to be split when possible and rows in those table to be made difficult to link (by removing identifiers and using table-specific pseudonyms).
- E.g. decentralized social networks like Diaspora more privacy-friendly than Facebook or Google+

Strategy #4: AGGREGATE

- **Personal data should be processed at the highest level of aggregation and with the least possible detail in which it is (still) useful.**
- Aggregating information over groups of attributes or individuals reduces the detail of the resulting information $\Longrightarrow$ data become less sensitive.
- If data are general enough that they can fit more than one person, they are not attributable to any single person $\Longrightarrow$ privacy is protected.
- Design patterns: k-anonymity (via microaggregation or generalization), aggregation over time (e.g. in smart metering), location coarsening (in location-based services), differential privacy and other privacy models/anonymization techniques.

Strategy #5: INFORM

- **Whenever data subjects use a system, they should be informed about which information is processed, for what purpose and by which means, and also how is that information protected (system security).**
- Relates to transparency and openness.
- Clear documentation must be provided.
- In case information is shared with third parties, subjects should be informed about it.
- Subjects should also be informed about their data access rights and how to exercise them.
- Design patterns: Platform for Privacy Preferences (P3P), transparency-enhancing techniques.

Strategy #6: CONTROL

- **Data subjects should be given agency over the processing of their personal data (relates to individual rights).**
- INFORM and CONTROL are intertwined:
  - Informing makes little sense if the subject has no control.
  - Control is impossible without the subject being informed.
- Access rights include the subject viewing, updating or even deleting her data.
- CONTROL interfaces should be easy to use.
- Design patterns: User-centric identity management, end-to-end encryption support control, intervenability techniques.

Strategy #7: ENFORCE

- **A privacy policy compatible with legal requirements should be in place and should be enforced (for accountability).**
- The policy ensures that the system respects privacy.
- The privacy level should meet all legal requirements.
- Technical protection mechanisms must exist that prevent policy violations.
- Governance structures to enforce the policy must be established.
- Design patterns: Access control, privacy rights management (a form of digital rights management involving licenses to personal data).

Strategy #8: DEMONSTRATE

- **The data controller must be able to demonstrate compliance with the privacy policy and applicable legal requirements (for accountability).**
- This strategy goes beyond ENFORCE: not only must privacy be enforced but it must be done demonstrably.
- Design patterns: Privacy management systems, logging, auditing.