A conceptual approach to AI and data protection

Cecilia Magnusson Sjöberg
Professor of Law & Informatics
Stockholm University
Or getting the grip on GDPR

Conceptual approach

- Legal system management
  - Supplementing the sequential checklist approach
  - Broad scope
    - From beginners to advanced

- Towards AI
  - Big data
  - Machine learning
    - Models & algorithms
Conceptual context

- **Transparency** is a condition for **privacy** in the context of **personal data processing** based on **AI** methods
  - Article 5.1 (a) GDPR
    - Data protection principles

- Processing governed by principles of **openness** might still not provide transparency due to lack of **access rights** and/how they are **implemented**
  - Article 25 GDPR
    - Data protection by design and by default
Clusters of concepts

**Formal**
Legal definitions
Ex. controller, personal data, consent

**Functional**
Processing
Ex. anonymisation, pseudonymisation, reidentification

**Steering**
Algorithms
Ex. self learning, training data, models
Clusters of concepts cont.

**Topical**
Information security
Ex. confidentiality, integrity and availability

**Historical**
Legacy data/applications
Ex. unstructured data, misuse
A Nordic law perspective

Data minimisation vs long term storage of Big data in national e-archives

- Article 5.1 (c) GDPR
  - Processing for archiving purposes in the public interest
- Swedish National Archives legislation
  - Cultural heritage etc.
Transparency by way of “routine measures”

- Article 5.1 (a) GDPR
- Swedish Freedom of the Press Act
  - Public access to official documents
    - compilations included
Technical interpreters

Digital perspective
- Article 15.1 (h) GDPR, Article 22 GDPR
  - Right of access by the data subject
  - Meaningful information about logic

Natural language perspective
- 13 § the Swedish Public Administration Act (2017:900)
DataLEASH Project

- The acronym stands for Learning and Sharing under Privacy constraints (DataLEASH) opening up for letting law play a proactive role instead of the traditional reactive one when matters already have gone wrong. In this context the GDPR (General Data Protection Regulation) requires attention, in particular Article 5 (principles relating to processing of personal data) and Article 25 (data protection by design and by default).

- Read more about the project: [https://www.kth.se/en/2.89968/forskningsprojekt/learning-and-sharing-under-privacy-constraints-dataleash-1.950497](https://www.kth.se/en/2.89968/forskningsprojekt/learning-and-sharing-under-privacy-constraints-dataleash-1.950497)

- DataLEASH will develop and test methods for more open data. In practice, it will consist of risk analysis for privacy protection (key indicators, methodology, legal requirements) and privacy configured learning systems (mechanisms, considerations of integrity and user value). It will support public organizations who are required to have open data and therefore need managing methods that are fast, reliable and uncomplicated. Based on these methods they are able to make well-informed decisions on how and if data should be shared (limiting access and various security settings) and choose what form of data conversion that should be aligned with a certain level of privacy and use.