**LECTURE 6 🡪 PRIVACY AND PRIVACY ENHANCING TECHNIQUES**

Goal of this lecture is to

* Introduce me to privacy
* Overview of taxonomy, privacy by design, standards, regulations, techniques …
* Understand how privacy taxonomy helps with identifying possible privacy requirements for new AI solutions
* Overview of graph modelling (not related to privacy)

What is privacy?

Control of personal info? Possession of personal info? A combination?

There is not a clear definition. A possible one, though, could be

*The right to be let alone*

*The quality or state of being apart from company or observation*

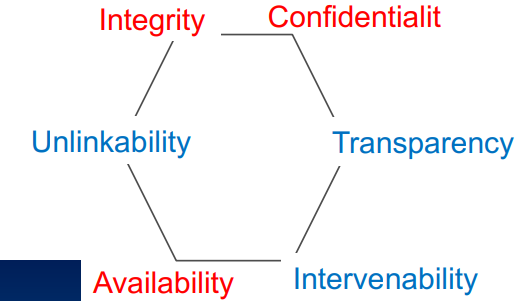
*Freedom from unauthorized intrusion, private, or matter*

*“Privacy is the claim of individuals, groups and institutions to determine for themselves, when, how and to what extent information about them is communicated to others”*

Privacy has many dimensions:

1. Data protection-related privacy
   1. Info privacy: control how and whether your data is collected, stored and processed
   2. Privacy of communications: protection of privacy and security of communications
2. Beyond data protection privacy
   1. Spatial privacy: a person’s virtual space against intrusion
   2. Territorial privacy: protection of physical area surrounding individuals, such as public and workplace
   3. Bodily privacy: protection of your physical body against procedures such as drug testing

6 protection goals



Unlikability dictates that data should not be linked from one domain to others

Transparency refers to provide access to information about data (metadata)

Standards and regulations

There are many, such as ISO/IEC 29100

There is one thing called **GDPR** (GENERAL DATA PROTECTION REGULATION)

The objectives are:

* **Harmonization** of data protection across UE
* **Modernization** of data protection rules
* **Empowering** citizens to have more control
* Improved levels of **compliance**

It is a regulation applicable in EU law on data protection and privacy for all residents of EU member states

Also deals with personal data being transported through communications channels leaving EU

So, what is personal data???

* Any information that relates to an alive person

What is sensitive personal data???

* Political views
* Ethnic or religious
* Racial, philosophic
* Health, sexual

It is forbidden to process sensitive data unless consent is obtained from data subjects or owners

And what is processing data? Collection, recording, storing, organization, structing, alteration and dissemination

Okay, and what rights do people have (in EU)?

1. Right to be informed
2. Right to access
3. Right of rectification, removing information
4. Restrict processing
5. Using one’s data
6. Object the use of personal data by other entities

Taxonomy of privacy

The taxonomy of privacy enables you to identify potential harms arising from infringements of privacy

There are 4 groups of activities:

* Information collection
* Information processing
* Information dissemination
* Invasion

Privacy By Design PbD

Let’s say that ‘privacy’ is a social construct hard to interpret, but **data protection** is more tangible

Privacy by Design is a white paper for regulators, decision makers and policy makers. The motivation behind it is:

* The need to manage information responsibility has grown
* Challenges of informational privacy
* Privacy has evolved
* Privacy must be approached as being an integral part of systems
* Privacy must be embedded into protocols and processes that touch our lives

STEPS:

1. Preventive instead of remedial, proactive and not reactive (mas vale prevenir que curar)

This means anticipation (using maybe the taxonomy of privacy to identify possible threats)

1. Privacy as the default setting

This means communicating the purpose of PbD before, limit the data collection and usage

1. Privacy embedded into design
2. Full functionality; rejecting weighing some interests to others. All interests and objectives must be clearly documented, functions articulated and so on.
3. End to end security; privacy must be protected across the entire domain (assure confidentiality, integrity and availability of data)
4. Visibility and transparency; compensation methods, information available, documentation
5. Respect for user privacy 🡪 keep it user centric

Privacy preserving techniques

Sooo many data are being gathered by corporate and public information systems (including sensible data). Protection is necessary, and technical solutions are needed. Some examples are:

* Regulations, standards and protocols
* Tools and mechanisms (encryption, anonymization)
* Data security and integrity
* Identity management

The technical solutions include:

* Statistical disclosure control. For the protection of sensitive data from being disclosed while using sensitive data in data-driven research. Ensures confidentiality of data released.
* Privacy-preserving machine learning/data mining. Data is modified before being published to the public for data mining
* Privacy enhancing techniques (PETs). Focus on hiding techniques (anonymity, etc). There are many types, like standard protocols
* Evaluations

A privacy impact assessment (PIA) is an analysis of how personally identifiable information (PII) is handled to ensure compliance with appropriate regulations, determine the privacy risks associated with information systems or activities, and evaluate ways to reduce the privacy risks.

Graph modelling (not related)

How can we do graph modelling? Generating codes from graphs

* Create a graph model with labels, properties, relationships, etc.
* Generate cypher codes
* Run cypher codes on Neo4J

To Answer

How does privacy taxonomy help you in designing an AI solution?

How does privacy by design help you in designing an AI solution?