Protect

Digital Representation of Privacy Terms: a Case Study in Machine-Readable Policies for Healthcare and Genomics

Beatriz Esteves, Ontology Engineering Group – Universidad Politécnica de Madrid

besteves4.github.io/about-me

beatriz.gesteves@upm.es

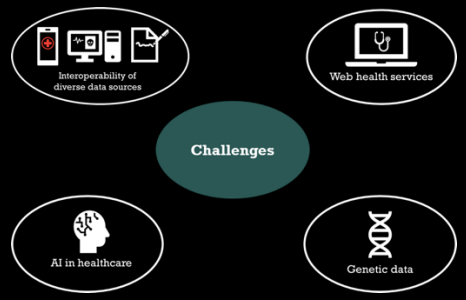
This project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement Nº813497

OUTLINE

- **Introduction**
  - Research motivation
  - State of the art

- **Debate**
  - Objectives
  - Hypotheses
  - Research questions

- **Main Goal**
  - Case study
  - Research methodology
  - Future steps



RESEARCH MOTIVATION

Promote data interoperability and portability to support individuals and companies navigating data protection regulations

Personal data

Laws

GDPR – General Data Protection Regulation

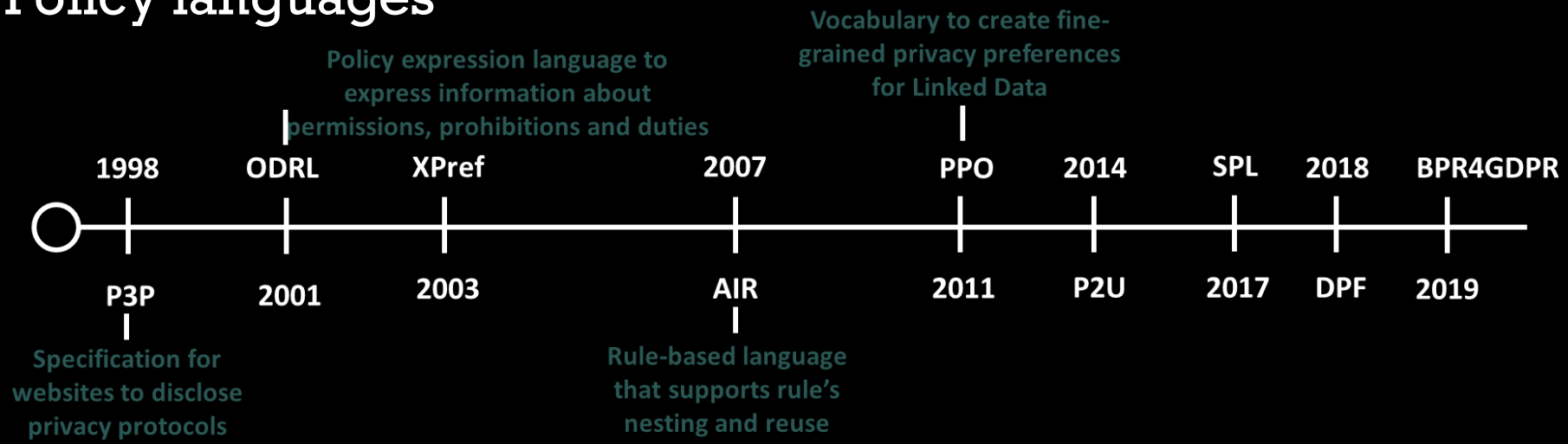
CCPA – California Consumer Privacy Act

LGPD - Brazilian General Data Protection Law

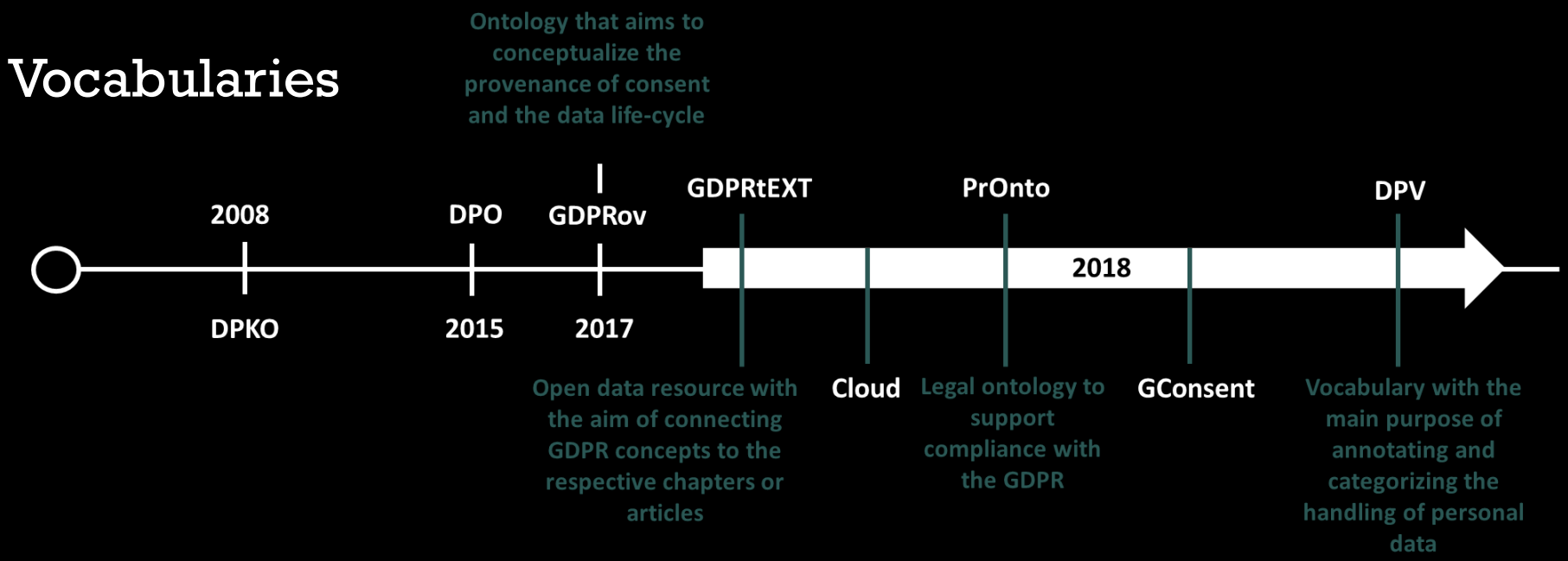
Controllers

STATE OF THE ART

Policy languages



Vocabularies



Semantic Web

W3C Semantic Web

RDF

Solid

OBJECTIVES

Design tools to assist with the digital representation of privacy terms, while supporting data subjects in the governance of their personal information and assisting data controllers to better handle compliance with the GDPR

OBJECTIVES

O1. Identify vocabularies on the **privacy** and **data protection** domain and extend these vocabularies where missing to deal with the **GDPR**.

O2. Facilitate the transition between **natural language terms** to its **digital representation**

O3. Build tools to assist data subjects in the establishment of their **privacy preferences** and data controllers in the formulation of GDPR-compliant **privacy policies**

O4. Design a **privacy negotiation** mechanism to assist in the establishment of privacy terms that satisfy both data subjects and controllers for the **disclosure of personal data**

O5. Design a service to assist with the representation of **information** connected with **GDPR rights** and **obligations**

HYPOTHESES

Use of **decentralized Web** technologies and **open standards** and **specifications** to achieve the aforementioned objectives

HYPOTHESES

H1. The use and extension of **vocabularies** and **policy languages** is suitable for the representation **of** **privacy preferences** and **policies** for the processing of personal data.

H2. Data interoperability on the data protection domain can be improved with the use of W3C standards and specifications.

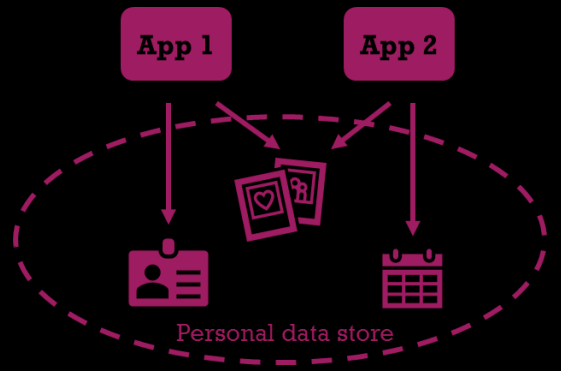
H3. These standards can also be used to establish more fine-grained **access control conditions**, by extending the WAC specification to manage the access to decentralized file systems.

H4. Semantic Web vocabularies can be used as a basis for establishing a **privacy negotiation process** for the **disclosure of personal data** to occur.

H5. The developed tools can be used to implement a service to facilitate the representation of **machine-readable information** related to **GDPR rights and obligations**.

RESEARCH QUESTIONS

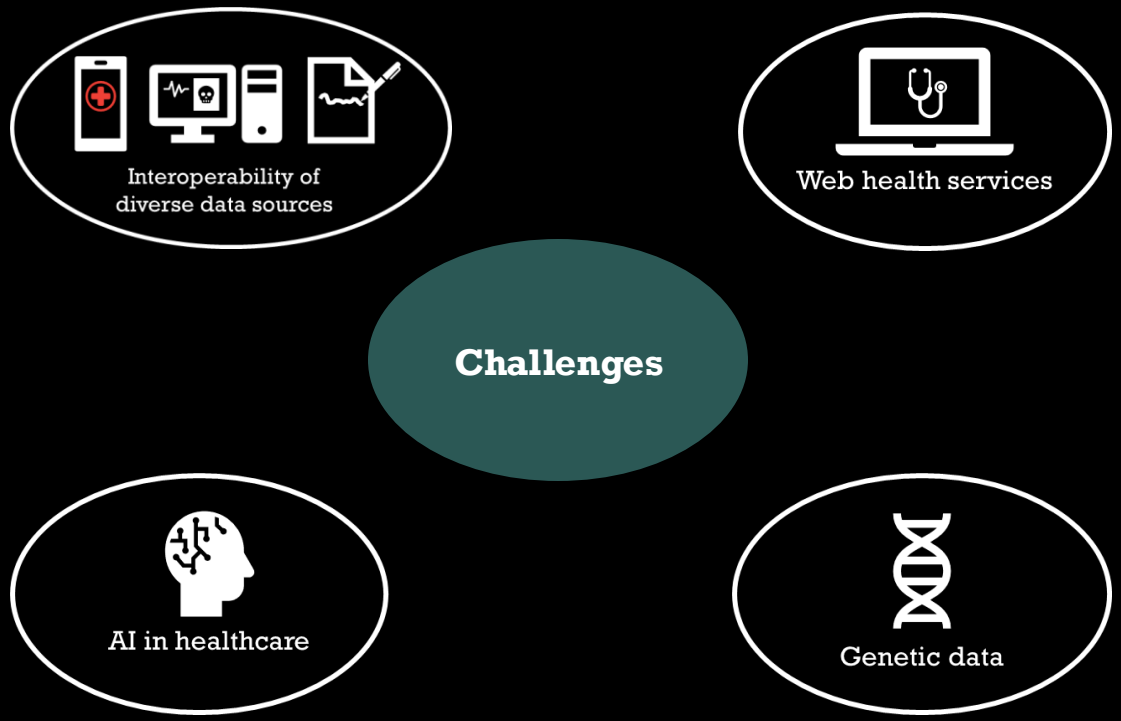
RQ1. Are **decentralized Web** technologies, **open standards and specifications** able to represent information related to consent forms and privacy terms in the domain of personal data protection, in order to **facilitate the communication** between GDPR stakeholders?



RQ2. Can Semantic Web vocabularies be used for the **governance** **of access** to personal data stored in decentralized file systems? And can they be extended to deal with specific use cases that involve **sensitive personal data**?

RQ3. Is it possible, using these technologies, to facilitate the **negotiation of privacy terms** and the **exercising of rights and obligations** in the light of the GDPR?

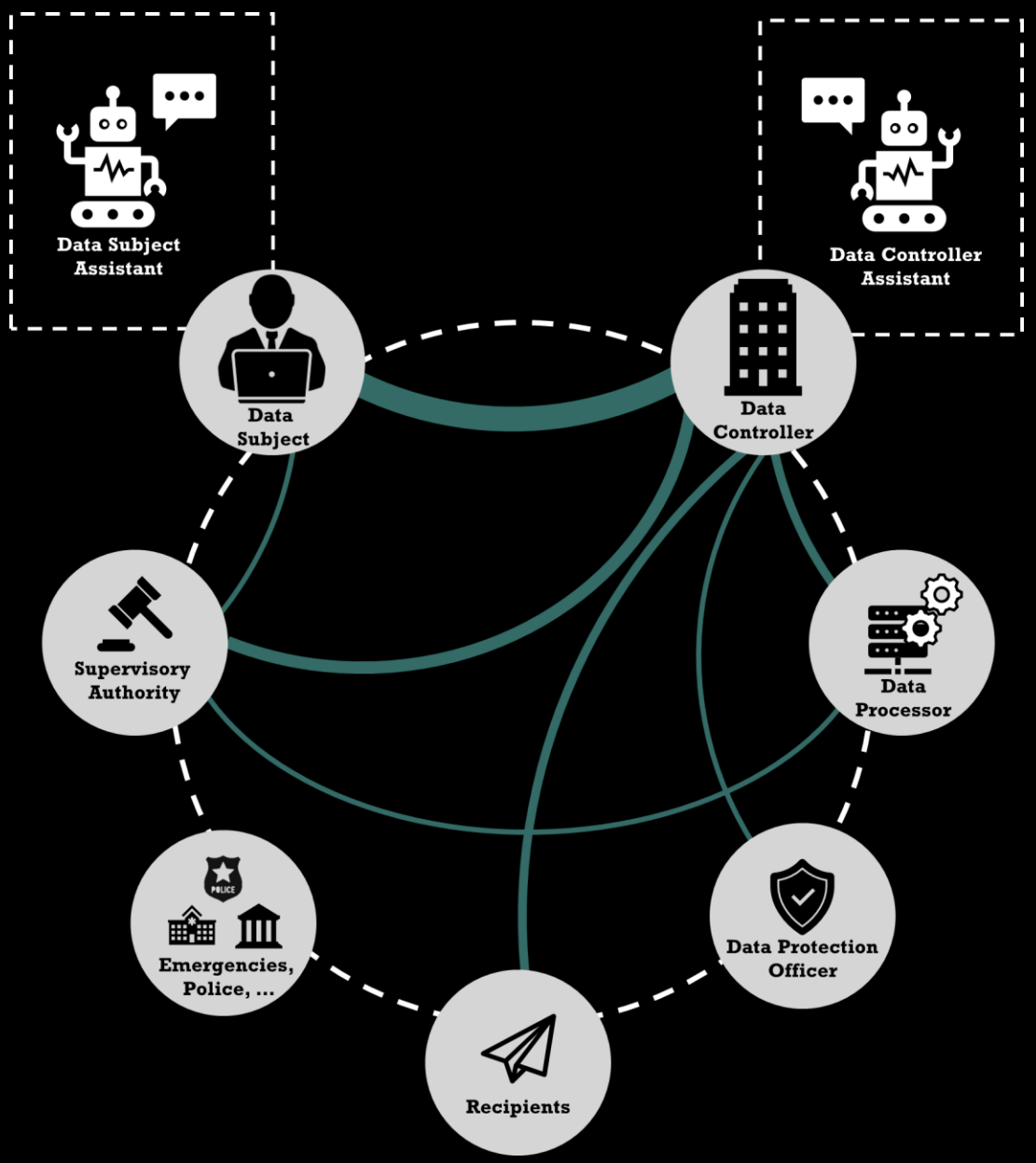
CASE STUDY: Machine-Readable Policies for Healthcare and Genomics



- Govern the access to personal data generated by **smart bio-sensors**.
- Share data generated in one **fitness app** with another.
- Grant access to **genetic data** for **academic research** purposes.
- Represent terms of a **DPIA** for a **large-scale study** to cure Alzheimer.

Semantic Web standards and specifications will be adopted, investigated and extended to govern the access to health-related and genetic data.

RESEARCH METHODOLOGY AND FUTURE STEPS



i. Analysis of existing open-access specifications
ii. Study of the architectures of research projects connected to the personal data protection domain
iii. Analysis of the concepts to be modeled to support the communication between GDPR stakeholders.

i. Establish the system's architecture and define the specific services to be implemented
ii. Extend ODRL and DPV for the purpose of defining privacy preferences and policies
iii. Understand how to use ODRL and DPV to extend WAC, to allow for more fine-grained access control to Solid Pods
iv. Extend ODRL and GDPRtEXT for the representation of GDPR rights and obligations
v. Extend DPV for the specific case of processing health-related and genetic data.

Protect

QUESTIONS

Beatriz Esteves, Ontology Engineering Group – Universidad Politécnica de Madrid

besteves4.github.io/about-me

beatriz.gesteves@upm.es

This project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement Nº813497