Protect

Fostering trust with transparency in the data economy era: an integrated ethical, legal, and knowledge engineering approach

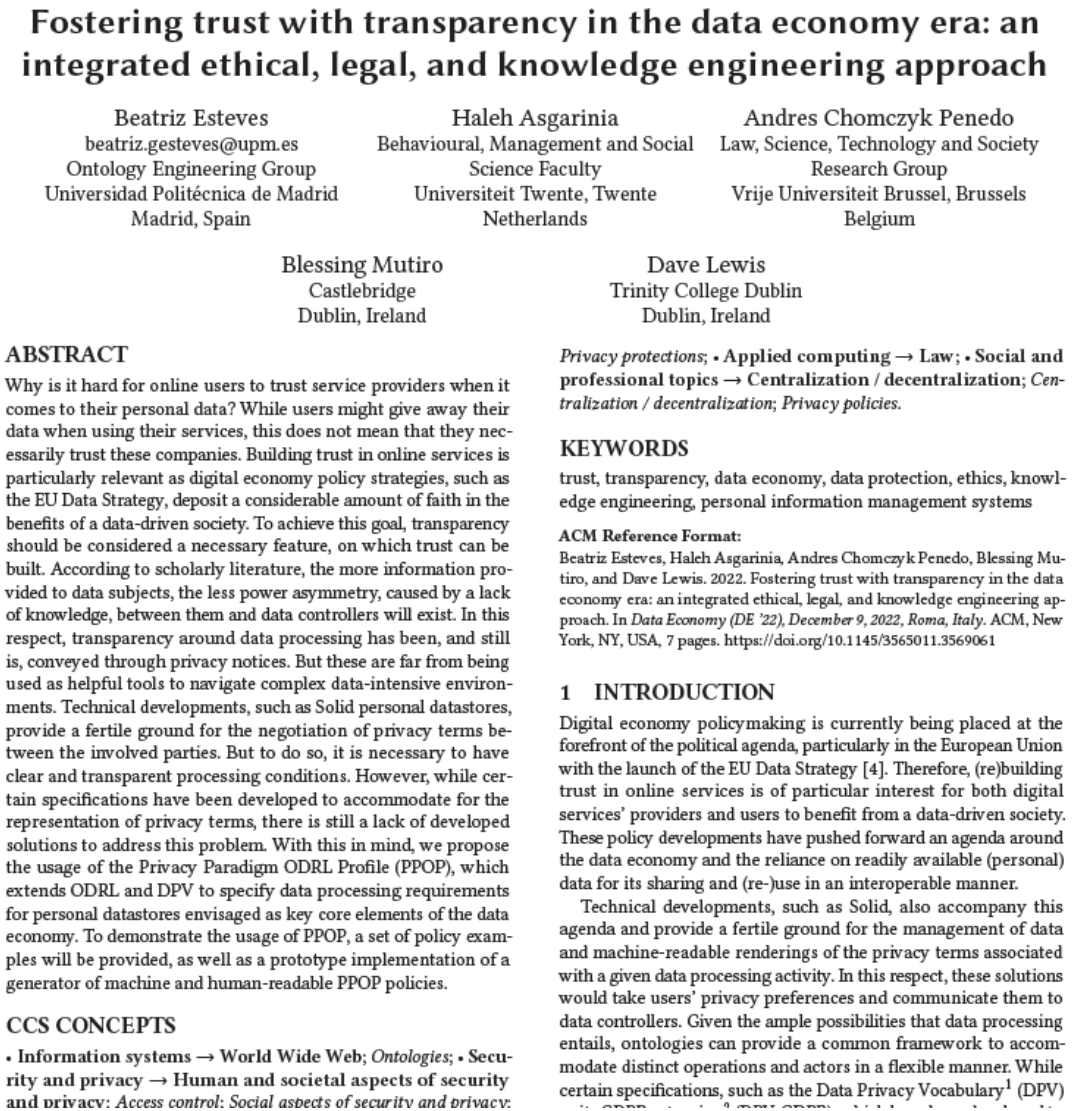
Beatriz Esteves (UPM), Haleh Asgarinia (UT), Andrés Chomczyk (VUB), Blessing Mutiro (Castlebridge), Dave Lewis (TCD)

beatriz.gesteves@upm.es | h.asgarinia@utwente.nl

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 813497.

Overview

- Introduction
- Related Work
  - Legal and Ethical Privacy Notices
  - Machine-readable Policies
- Motivation
- Privacy Paradigm ODRL Profile
- Examples
- Conclusions and Future Work



Fostering trust with transparency in the data economy era: an integrated ethical, legal, and knowledge engineering approach

Beatriz Esteves
beatriz.gesteves@upm.es
Ontology Engineering Group
Universidad Politécnica de Madrid
Madrid, Spain

Haleh Asgarinia
Behavioural, Management and Social Science Faculty
Universiteit Twente, Twente
Netherlands

Andres Chomczyk Penedo
Law, Science, Technology and Society Research Group
Vrije Universiteit Brussel, Brussels
Belgium

Blessing Mutiro
Castlebridge
Dublin, Ireland

Dave Lewis
Trinity College Dublin
Dublin, Ireland

ABSTRACT

Why is it hard for online users to trust service providers when it comes to their personal data? While users might give away their data when using their services, this does not mean that they necessarily trust these companies. Building trust in online services is particularly relevant as digital economy policy strategies, such as the EU Data Strategy, deposit a considerable amount of faith in the benefits of a data-driven society. To achieve this goal, transparency should be considered a necessary feature, on which trust can be built. According to scholarly literature, the more information provided to data subjects, the less power asymmetry, caused by a lack of knowledge, between them and data controllers will exist. In this respect, transparency around data processing has been, and still is, conveyed through privacy notices. But these are far from being used as helpful tools to navigate complex data-intensive environments. Technical developments, such as Solid personal datastores, provide a fertile ground for the negotiation of privacy terms between the involved parties. But to do so, it is necessary to have clear and transparent processing conditions. However, while certain specifications have been developed to accommodate for the representation of privacy terms, there is still a lack of developed solutions to address this problem. With this in mind, we propose the usage of the Privacy Paradigm ODRL Profile (PPOP), which extends ODRL and DPV to specify data processing requirements for personal datastores envisaged as key core elements of the data economy. To demonstrate the usage of PPOP, a set of policy examples will be provided, as well as a prototype implementation of a generator of machine and human-readable PPOP policies.

CCS CONCEPTS

• **Information systems** → **World Wide Web**; *Ontologies*; • **Security and privacy** → **Human and societal aspects of security and privacy**; *Access control*; *Social aspects of security and privacy*; *Privacy protections*; • **Applied computing** → **Law**; • **Social and professional topics** → **Centralization / decentralization**; *Centralization / decentralization*; *Privacy policies*.

KEYWORDS

trust, transparency, data economy, data protection, ethics, knowledge engineering, personal information management systems

ACM Reference Format:

Beatriz Esteves, Haleh Asgarinia, Andres Chomczyk Penedo, Blessing Mutiro, and Dave Lewis. 2022. Fostering trust with transparency in the data economy era: an integrated ethical, legal, and knowledge engineering approach. In *Data Economy (DE '22), December 9, 2022, Roma, Italy.* ACM, New York, NY, USA, 7 pages. https://doi.org/10.1145/3565011.3569061

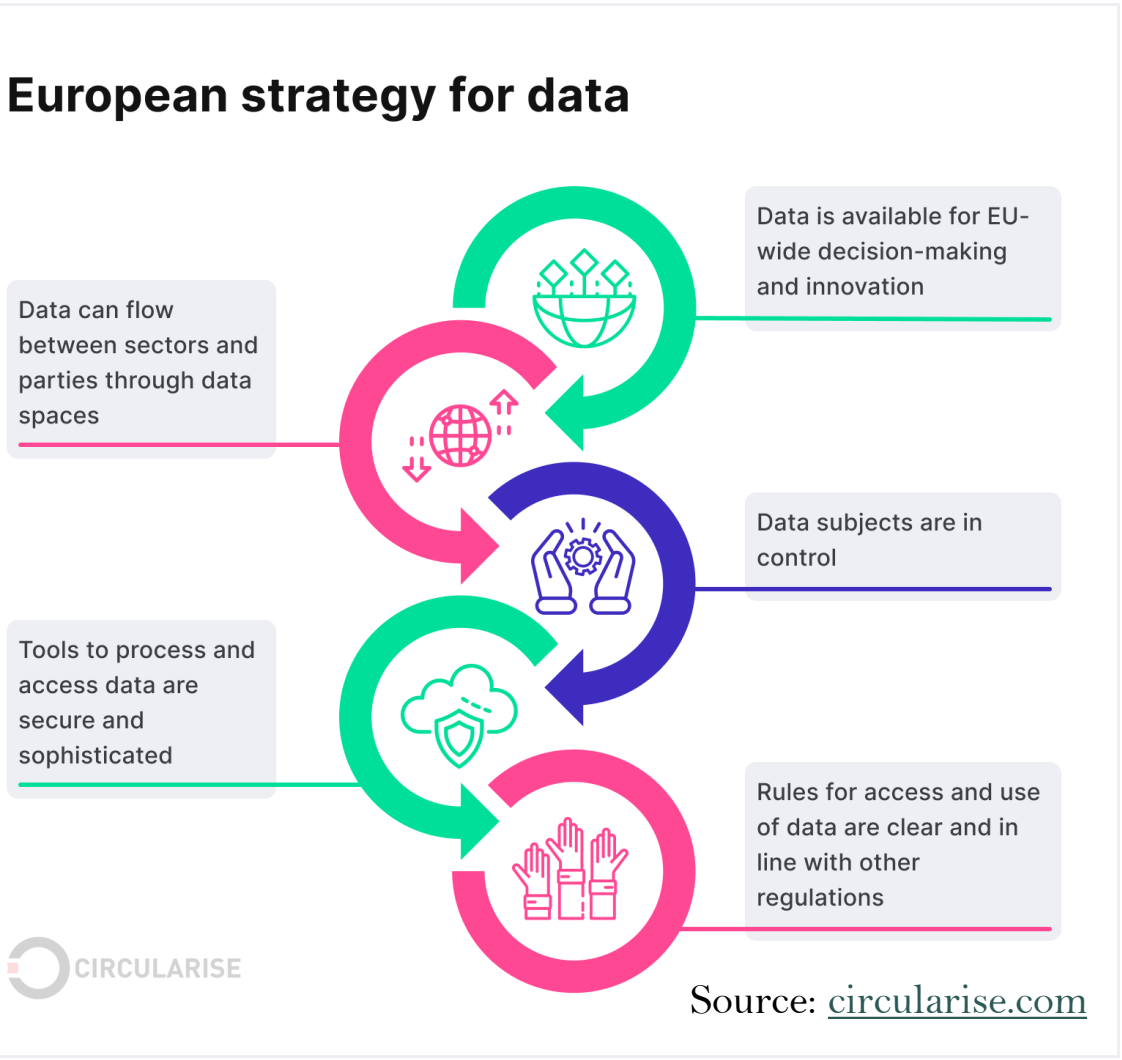
1 INTRODUCTION

Digital economy policymaking is currently being placed at the forefront of the political agenda, particularly in the European Union with the launch of the EU Data Strategy [4]. Therefore, (re)building trust in online services is of particular interest for both digital services' providers and users to benefit from a data-driven society. These policy developments have pushed forward an agenda around the data economy and the reliance on readily available (personal) data for its sharing and (re-)use in an interoperable manner.

Technical developments, such as Solid, also accompany this agenda and provide a fertile ground for the management of data and machine-readable renderings of the privacy terms associated with a given data processing activity. In this respect, these solutions would take users' privacy preferences and communicate them to data controllers. Given the ample possibilities that data processing entails, ontologies can provide a common framework to accommodate distinct operations and actors in a flexible manner. While certain specifications, such as the Data Privacy Vocabulary[1] (DPV)

https://doi.org/10.1145/3565011.3569061

Introduction

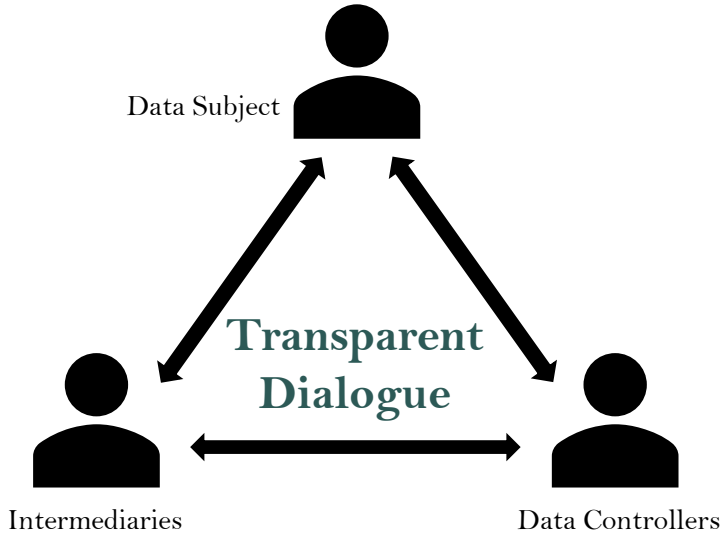


Source: circularise.com

New EU-level regulations

New technologies

- PIMS
- Policy languages
- Data protection vocabularies



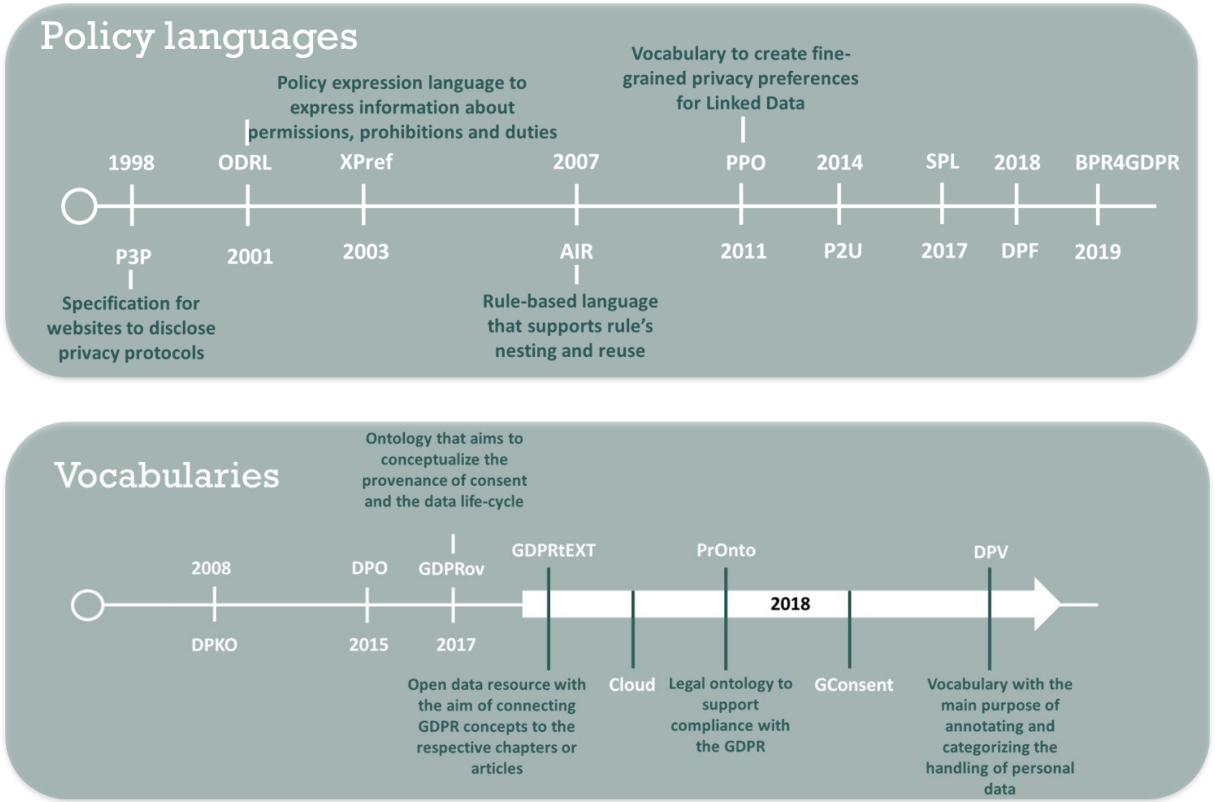
https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age_en

Related Work

Legal and Ethical Privacy Notices

Source: newamerica.org

Machine-readable Policies



Esteves, B. and Rodríguez-Doncel, V., "Analysis of Ontologies and Policy Languages to Represent Information Flows in GDPR" (2022). Semantic Web Journal : 1 – 35. https://content.iospress.com/articles/semantic-web/sw223009

Motivation

Requirements

R1. Classify transparency practices of intermediary services

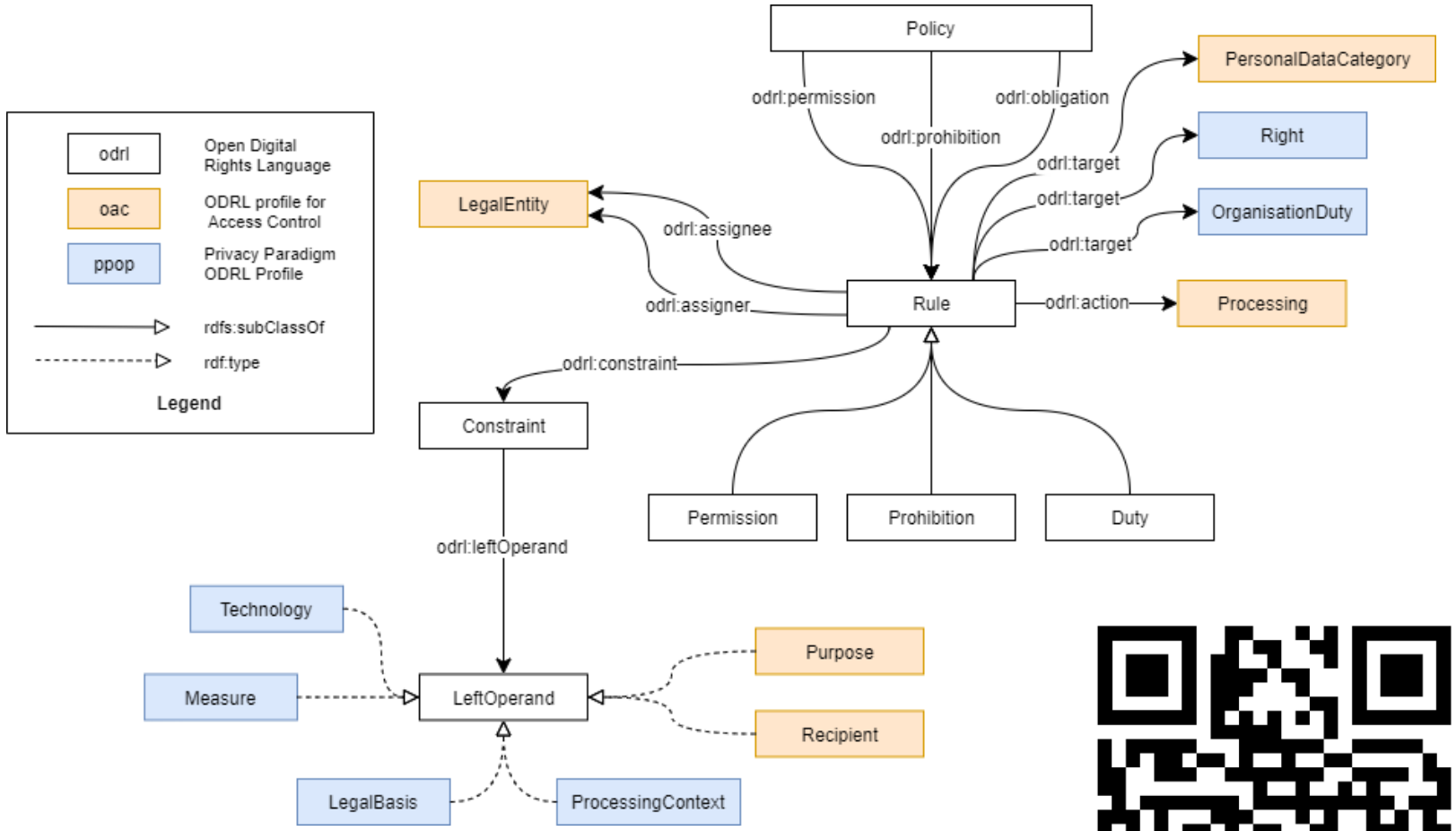
R2. Define access control policies for legal and ethical access to group and individual personal data stores

R3. Model safeguards for the trustworthiness of AI systems and respective rights and duties

Ethical Sources

Regulatory Sources

Privacy Paradigm ODRL Profile



https://w3id.org/ppop

- Developed following the LOT methodology
- ORSD: https://w3id.org/ppop#orsd
- Extends previous work with ODRL and DPV

Main concepts
Measure, TransparencyMeasure, SafeguardForTrustworthiness
Right, GroupRight, DataSubjectRight, OrganisationDuty, RightExemption
Group, DataSharingEntity, DataTrustProvider
Technology, PIMS

Examples

```
ex:family-pod a odrl:Policy ;
    odrl:profile ppop:, oac: ; dc:issued "2022-02-22" ;
    odrl:uid <https://pod-provider/familyA/policy1> ;
    odrl:permission [
        odrl:assigner ex:family-pool ;
        odrl:assignee [ a ppop:DataSharingServiceProvider ] ;
        odrl:target oac:MedicalHealth ;
        odrl:action [
            rdf:value oac:Use ;
            odrl:refinement [ odrl:and (ex:RD-purpose, ex:tech) ] ] ] .
ex:RD-purpose odrl:leftOperand oac:Purpose ;
    odrl:operator odrl:isA ;
    odrl:rightOperand dpv:ResearchAndDevelopment .
ex:tech odrl:leftOperand ppop:Technology ;
    odrl:operator odrl:isA ;
    odrl:rightOperand ex:PersonalDataStore .
ex:PersonalDataStore a ppop:PersonalDataStore ;
    dpv:hasStorage [ dpv:hasLocation <https://pod-provider/familyA/> ] .
ex:family-pool a ppop:Group ;
    ppop:hasVoluntaryMembership ex:Parent1, ex:Parent2 ;
    ppop:hasNonVoluntaryMembership ex:Child1, ex:Child2 .
ex:Parent1 a ppop:DataHolder, dpv:DataSubject ;
    ppop:isDataHolderFor ex:Child1, ex:Child2, ex:Parent1 .
ex:Parent2 a ppop:DataHolder, dpv:DataSubject ;
    ppop:isDataHolderFor ex:Child1, ex:Child2, ex:Parent2 .
ex:Child1 a dpv:Child .
ex:Child2 a dpv:Child .
```

Family sharing policy related to its health data.

A family, with two parents and two children, has created a policy to allow the use of their **medical health data** that is stored on their family Pod, for the **purpose of research and development** and to have a **data-sharing service provider** as a data intermediary.

Since the data in the Pod can belong to any of the four members of the family, the **assigner** of the policy is a **ppop:Group** which is composed by both the parents and the children, where the **parents are data holders** and **act as data holders for the children**.

Examples

```
ex:transparency-policy a odrl:Policy ;
    odrl:profile ppop:, oac: ;
    odrl:uid <https://application.com/policy1> ;
    dc:issued "2022-03-13" ;
    odrl:target oac:WorkHistory ;
    odrl:action oac:Use ;
    odrl:assignee ex:data-intermediary ;
    odrl:duty [
        odrl:action [
            rdf:value ppop:implement ;
            odrl:refinement [
                odrl:leftOperand ppop:Measure ;
                odrl:operator odrl:isA ;
                odrl:rightOperand ppop:SafeguardForExplainability ] ] ] .
ex:data-intermediary a ppop:DataIntermediary ;
    ppop:hasChargePrice "false" ;
    ppop:convertsData "false" .
```

Modeling transparency practices and safeguards for trustworthiness.

A **data intermediary** published a policy related to the use of **work history data** from their users. The intermediary's transparency practices are reflected in this policy: **no price is charged** for the service and the **data is not converted to other formats**. The intermediary also states its **duty to implement a ppop:SafeguardForExplainability measure** in its service to provide details on how the data is being used to its users.

Conclusions and Future Work

PPOP generator

UI to generate human-readable and machine-readable privacy policies as ODRL policies using PPOP

- Further development of prototype implementation of a generator of machine and human-readable PPOP policies
- Test the work in a real-world setting
- Mantain and improve the work as new regulations and ethical guidelines are published

Protect

Thanks for your attention and looking forward to your comments!

Beatriz Esteves (UPM), Haleh Asgarinia (UT), Andrés Chomczyk (VUB), Blessing Mutiro (Castlebridge), Dave Lewis (TCD)

beatriz.gesteves@upm.es | h.asgarinia@utwente.nl

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 813497.