

An ODP for Describing Personal Data in Privacy Policies

Harshvardhan J. Pandit, Declan O'Sullivan, Dave Lewis

ADAPT Centre, Trinity College Dublin, Dublin, Ireland

{ harshvardhan.pandit | declan.osullivan | dave.lewis } @ adaptcentre.ie

@coolharsh55 <https://openscience.adaptcentre.ie/>

Problem Domain

- Privacy Policy is a legal document that specifies information regarding the use of Personal Data amongst other things.
- This information needs to be represented as machine-readable metadata to enable machine-based approaches over it.
- Approaches that work with Privacy Policies have their own internal representation of this metadata specific to their aims.
- It is difficult to garner consistency and standardisation on privacy policies as a document itself.

Information We Collect

There are **three** general categories of **information** we collect.

- data collected from user**

1.1 Information You Give to Us.

1.1.1 Information that is **necessary** for provision of services

We ask for and collect the following personal information **legitimate interest** when you use our service. This information is necessary for the adequate performance of the contract between you and us and to allow us to comply with our legal obligations. Without it, we may not be able to provide you with **data category** services.

- Account Information** When you **sign up for an account**, we require certain information such as your **first name**, **last name**, **email address**, and **date of birth**.
- Profile and Listing Information** To use certain features, we may ask you to provide additional information, which may include your **id address**, **phone number**, and a **profile picture**.
- Identity Verification Information** To help create and maintain a trusted environment, we may collect identity verification information (such as **images of your government issued ID**, **passport**, **national ID card**, or **driving license**, as permitted by applicable laws) or other **authentication information**.
- Payment Information** To use certain features of the such as **booking**, we may require you to provide certain **financial information** (like your **bank account** or **credit card information**) in order to **enable us to process the payments**.

- process**
- consent**

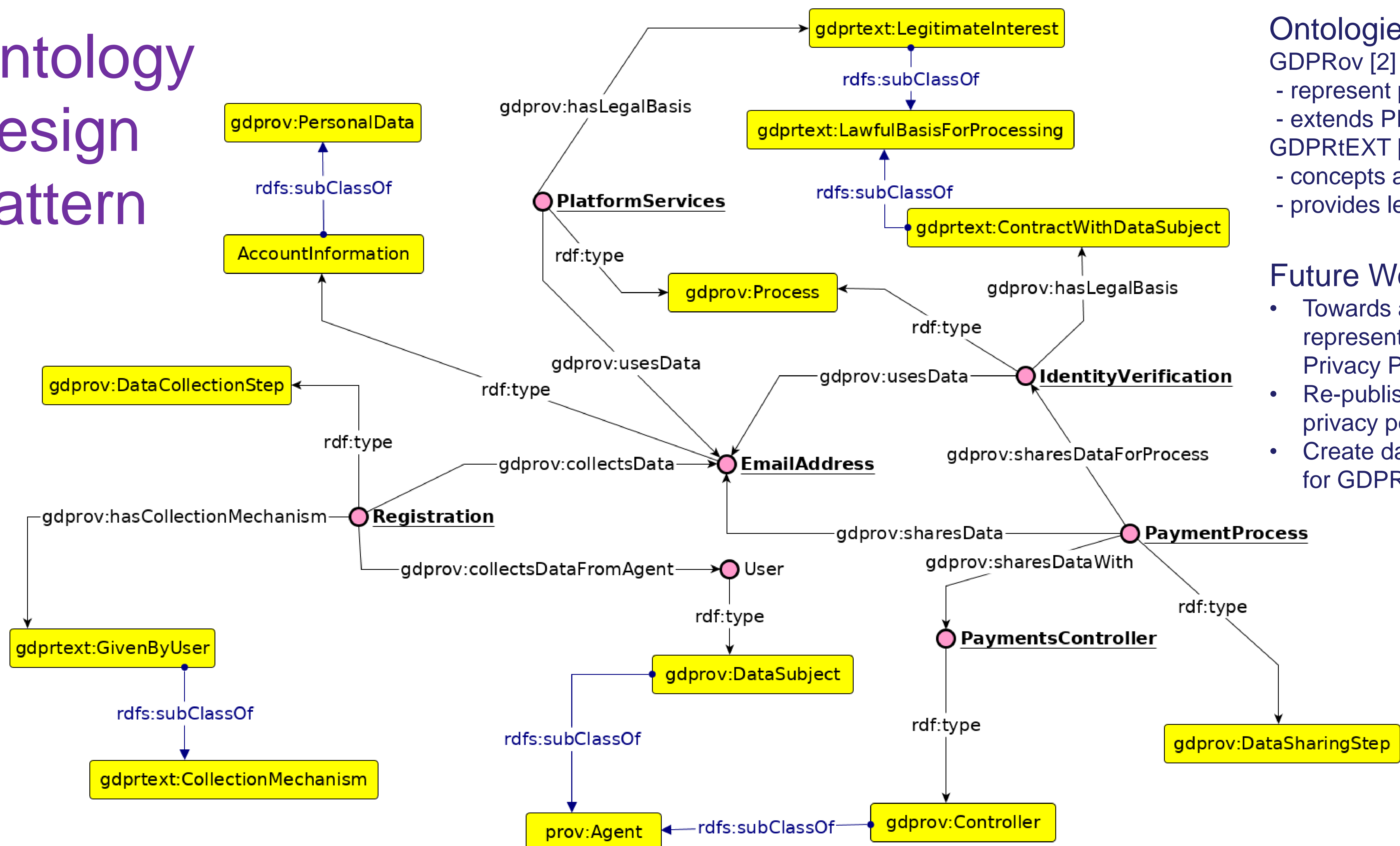
1.1.2 Information you **choose** to give us

Annotated Airbnb Ireland Privacy Policy [1]

Solution Sphere

1. Identify relevant information that is common to all privacy policies.
2. Incorporate legal requirements from General Data Protection Regulation (GDPR).
3. Use existing vocabularies to represent concepts and relations.
4. Define an ODP that can be incorporated into ontologies or used as a mapping between approaches.
5. Create a corpus of privacy policy metadata using ODP → future work!

Ontology Design Pattern



Ontologies Re-used

GDPRov [2]

- represent provenance
 - extends PROV-O [3] ; P-Plan [4]
- GDPRtEXT [5]
- concepts and terms from GDPR
 - provides legal basis

Future Work

- Towards an ontology for representing the entirety of Privacy Policy as a document.
- Re-publish existing corpus of privacy policies using ODP.
- Create dataset from new policies for GDPR (post 25 May 2018)

[1] Privacy Policy <http://openscience.adaptcentre.ie/privacy-policy/personalise/demo/policy.html> Re-used under CC-by-4.0

[2] GDPRov – GDPR Provenance Vocabulary <https://w3id.org/GDPRov>

[3] PROV-O: The PROV Ontology. W3C Recommendation 30 April 2013. <http://www.w3.org/TR/prov-o/>

[4] The P-PLAN Ontology. Release 12 March 2014. <http://purl.org/net/p-plan>

[5] GDPRtEXT – GDPR Text Extensions. <https://w3id.org/GDPRtEXT>