Engaging Content
Engaging People

CITYSPIN

SPECIAL

FFG
Forschung wirkt.

Creating A Vocabulary for Data Privacy

The First-Year Report of Data Privacy Vocabularies and Controls Community Group (DPVCG)

Harshvardhan J. Pandit[1], Axel Polleres[2], Bert Bos, Rob Brennan, Bud Bruegger, Fajar J. Ekaputra, Javier D. Fernández, Roghaiyeh Gachpaz Hamed, Elmar Kiesling, Mark Lizar, Eva Schlehahn, Simon Steyskal, and Rigo Wenning

DPV vocabulary: https://w3.org/dpv

[1]Presenter: **Harshvardhan J. Pandit** ADAPT Centre, Trinity College Dublin, Ireland
email: pandith@tcd.ie | twitter: @coolharsh55
[2] Axel Polleres Vienna University of Economics and Business, Austria
email: axel.polleres@wu.ac.at

This work was supported by the European Union's Horizon 2020 research and innovation programme under grant 731601 (SPECIAL), by the Austrian Research Promotion Agency (FFG) under the projects "EXPEDiTE" and "CitySpin", by the ADAPT Centre for Digital Excellence funded by SFI Research Centres Programme (Grant 13/RC/2106), and co-funded by European Regional Development Fund.

Ireland's European Structural and Investment Funds Programmes 2014-2020
Co-funded by the Irish Government and the European Union

European Union
European Regional Development Fund

The ADAPT Centre is funded under the SFI Research Centres Programme (Grant 13/RC/2106) and is co-funded under the European Regional Development Fund.

Overview

1. Background
2. DPVCG
3. Methodology
4. Existing & Related Work
5. Data Privacy Vocabulary
   a) Base Ontology
   b) Purposes
   c) Processing
   d) Technical & Organisational Measures
   e) Personal Data Categories
   f) Consent (Legal Base)
6. Future Work

Authors:

- Harshvardhan J. Pandit (Trinity College Dublin)
- Axel Polleres (Vienna University of Economics and Business)
- Bert Bos (W3C/ERCIM)
- Rob Brennan (Dublin City University)
- Bud Bruegger (Unabhängige Landeszentrum für Datenschutz Schleswig-Holstein)
- Fajar J. Ekaputra (Vienna University of Technology)
- Javier D. Fernández (Vienna University of Economics and Business)
- Ramisa Gachpaz Hamed (Trinity College Dublin)
- Elmar Kiesling (Vienna University of Technology)
- Mark Lizar (OpenConsent/Kantara Initiative)
- Eva Schlehan (Unabhängige Landeszentrum für Datenschutz Schleswig-Holstein)
- Simon Steyskal (Siemens)
- Rigo Wenning (W3C/ERCIM)

We thank all members of the W3C DPVCG for their feedback and input to this work: a preliminary outline of the goals of CG has been presented in ISWC2018's SWSG workshop [5] where we also gathered valuable feedback by the participants; this work is the first complete presentation of the resulting, proposed vocabulary elaborated by the DPVCG since.

Overview

- Laws such as GDPR (EU, 2016) & CCPA (CA, 2018) regulate processing of personal data
- Obligations and compliance requirements provide motivation for adopting technical solutions
- Momentum towards interoperable privacy solutions (e.g. ISO 27701)

However, there are still some gaps to address:

1) Lack of standardised vocabularies for representing
   - personal data categories (what?)
   - purposes of processing (why? & how?)
2) No aligned terminology within privacy laws

Problem: How to represent information about handling (or processing) of personal data in a machine-readable format for compliance with data privacy laws such as GDPR and CCPA?

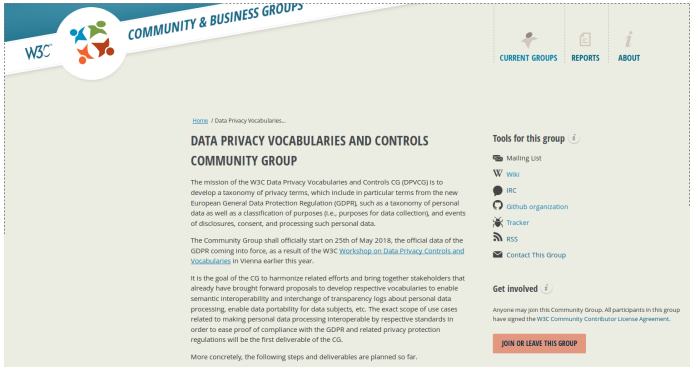
DPVCG

- To address the issue, a W3C community group (CG) was started "**Data Privacy Vocabularies & Controls CG**"

- Aim: To create vocabularies for representing personal data, purposes/processing, disclosure/consent, etc. in an interoperable format towards standardisation

- Started: 25th May 2018
- Members: 58 participants to date

- Deliverable of work done till date:
  Data Privacy Vocabulary
  http://w3.org/ns/dpv



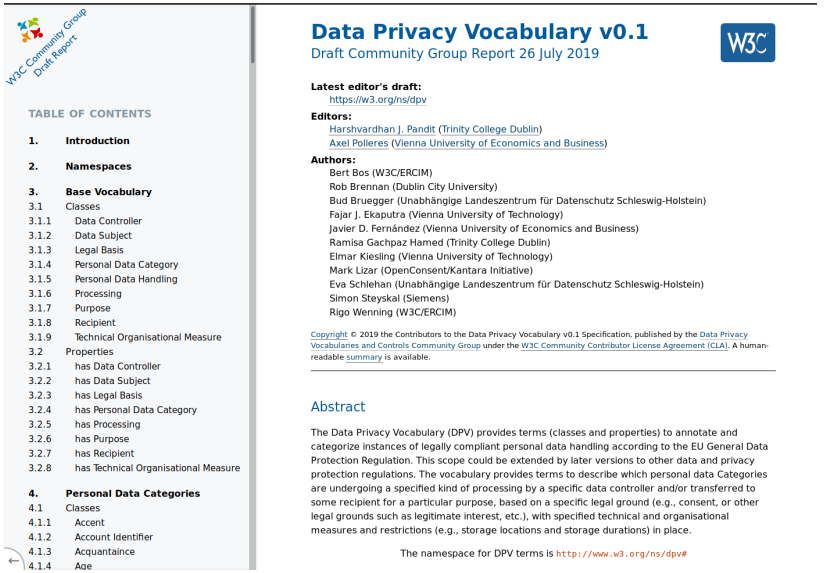
https://www.w3.org/community/dpvcg/

Public Mailing List
https://lists.w3.org/Archives/Public/public-dpvcg/

Data Privacy Vocabulary (DPV)

- Modular vocabulary for representing *personal data handling*
- Represented as an ontology using OWL2
- Published: 26th July 2019

- Models:
  1) Personal Data Categories
  2) Purposes
  3) Processing Categories
  4) Technical & Organisational Measures
  5) Legal Basis
  6) Consent
  7) Recipients, Data Controllers, Data Subjects



W3C Community Group Draft Report

TABLE OF CONTENTS

1. Introduction
2. Namespaces
3. Base Vocabulary
   - 3.1 Classes
     - 3.1.1 Data Controller
     - 3.1.2 Data Subject
     - 3.1.3 Legal Basis
     - 3.1.4 Personal Data Category
     - 3.1.5 Personal Data Handling
     - 3.1.6 Processing
     - 3.1.7 Purpose
     - 3.1.8 Recipient
     - 3.1.9 Technical Organisational Measure
   - 3.2 Properties
     - 3.2.1 has Data Controller
     - 3.2.2 has Data Subject
     - 3.2.3 has Legal Basis
     - 3.2.4 has Personal Data Category
     - 3.2.5 has Processing
     - 3.2.6 has Purpose
     - 3.2.7 has Recipient
     - 3.2.8 has Technical Organisational Measure
4. Personal Data Categories
   - 4.1 Classes
     - 4.1.1 Accent
     - 4.1.2 Account Identifier
     - 4.1.3 Acquantaince
     - 4.1.4 Age

Data Privacy Vocabulary v0.1

Draft Community Group Report 26 July 2019

Latest editor's draft:
https://w3.org/ns/dpv

Editors:
Harshvardhan J. Pandit (Trinity College Dublin)
Axel Polleres (Vienna University of Economics and Business)

Authors:
Bert Bos (W3C/ERCIM)
Rob Brennan (Dublin City University)
Bud Bruegger (Unabhängige Landeszentrum für Datenschutz Schleswig-Holstein)
Fajar J. Ekaputra (Vienna University of Technology)
Javier D. Fernández (Vienna University of Economics and Business)
Ramisa Gachpaz Hamed (Trinity College Dublin)
Elmar Kiesling (Vienna University of Technology)
Mark Lizar (OpenConsent/Kantara Initiative)
Eva Schlehan (Unabhängige Landeszentrum für Datenschutz Schleswig-Holstein)
Simon Steyskal (Siemens)
Rigo Wenning (W3C/ERCIM)

Abstract

The Data Privacy Vocabulary (DPV) provides terms (classes and properties) to annotate and categorize instances of legally compliant personal data handling according to the EU General Data Protection Regulation. This scope could be extended by later versions to other data and privacy protection regulations. The vocabulary provides terms to describe which personal data Categories are undergoing a specified kind of processing by a specific data controller and/or transferred to some recipient for a particular purpose, based on a specific legal ground (e.g., consent, or other legal grounds such as legitimate interest, etc.), with specified technical and organisational measures and restrictions (e.g., storage locations and storage durations) in place.

The namespace for DPV terms is http://www.w3.org/ns/dpv#

Link: http://w3.org/ns/dpv

GDPR Legal Basis Vocabulary
https://www.w3.org/ns/dpv-gdpr

Methodology

- (informally based) NeOn methodology for ontology development [28]
- Decision to create top-level hierarchy of concepts
- Analysis of existing and relevant work for reuse
- Use of OWL for expressing relations and logic
- Base Vocabulary adopted from SPECIAL Usage Policy [6]
- Terms were proposed/resolved using spreadsheet in Google Sheets
- Documentation generated using W3C ReSpec
- Repo on Github
- Discussions via F2F and public mailing list

Column Name	Description	Representation
Class/Property	If term is Class or Property	*rdfs:Class/rdfs:Property*
term	The IRI of the term	as IRI
description	Description or definition	*dct:description*
domain	Domain if it is a property	*rdfs:domain*
range	Range if it is a property	*rdfs:range*
super classes/properties	Parent classes or properties	*rdfs:isSubClassOf*
sub classes/properties	Child classes or properties	N/A
related terms	Terms relevant to this	*rdfs:seeAlso*
how related?	Nature of relation	use as is
comments	Comments used for discussion	N/A
source	The source of the term	*rdfs:isDefinedBy*
date	Date of creation	*dct:created*
status	Status e.g. accepted,proposed	*sw:term_status*
comments	Comments to be recorded	*rdfs:comment*
contributor	dc:creator	*dct:creator*
date-accepted	Date of acceptance	*dct:date*-accepted
resolution	Record e.g. minutes of meeting	as IRI

[6] Bonatti, P.A., Kirrane, S., Petrova, I.M., Sauro, L., Schlehahn, E.: The SPECIAL Usage Policy Language, V0.1. Tech. rep. (2018), https://www.specialprivacy.eu/vocabs
[28] Suárez-Figueroa, M.C., Gómez-Pérez, A., Fernández-López, M.: The NeOn Methodology for Ontology Engineering. In: Suárez-Figueroa, M.C., Gómez- Pérez, A., Motta, E., Gangemi, A. (eds.) Ontology Engineering in a Net- worked World, pp. 9–34. Springer Berlin Heidelberg, Berlin, Heidelberg (2012). https://doi.org/10.1007/978-3-642-24794-1_2, http://link.springer. Com/10.1007/978-3-642-24794-1_2

Existing & Related Work

1) Document using Wiki https://www.w3.org/community/dpvcg/wiki/Main_Page
2) Identify relevant concepts and relationships in vocabularies

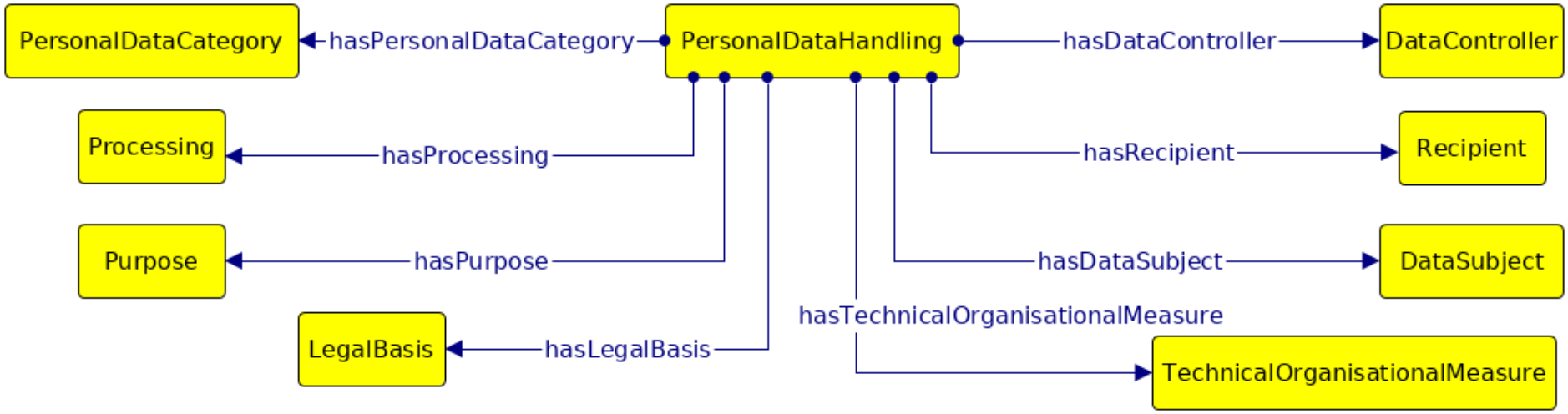
Standards
- PROV-O
- P-Plan
- ODRL
- vCard
- ActivityStreams
- COEL
- CoreVocabularies (ISA2)
- IEEE P7012
- Consent Receipt v1.1
- P3P

Vocabularies
- SPECIAL project
- MIREL project
- DAPRECO project
- GDPRtEXT
- GDPRov
- GConsent
- Privacy Preference Ontology (PPO)

Reasons for not reusing concepts:
- ✗ Importing all external semantics into DPV
- ✗ Choice of OWL2 was made at a later stage of development
- ✗ Not all concepts matched or fit in together

DPV – Base Ontology

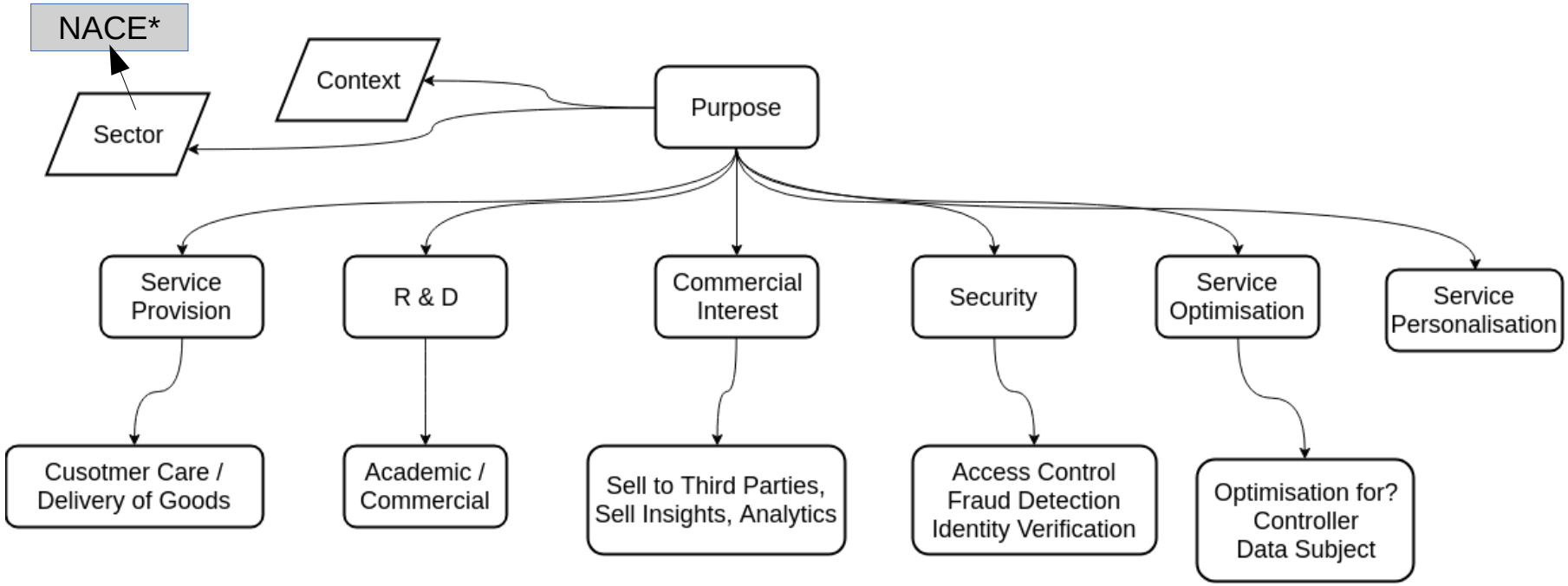
- Top-level classes defining a policy for legal personal data handling
- Represented by class ***PersonalDataHandling***
- Connects different modular components



DPV – Purposes


```
:SomePurpose a dpv:Purpose ;
      rdfs:label "Some Purpose" ;
      dpv:hasSector dpv-nace:M72 .
```

Top-level categorisation of purposes

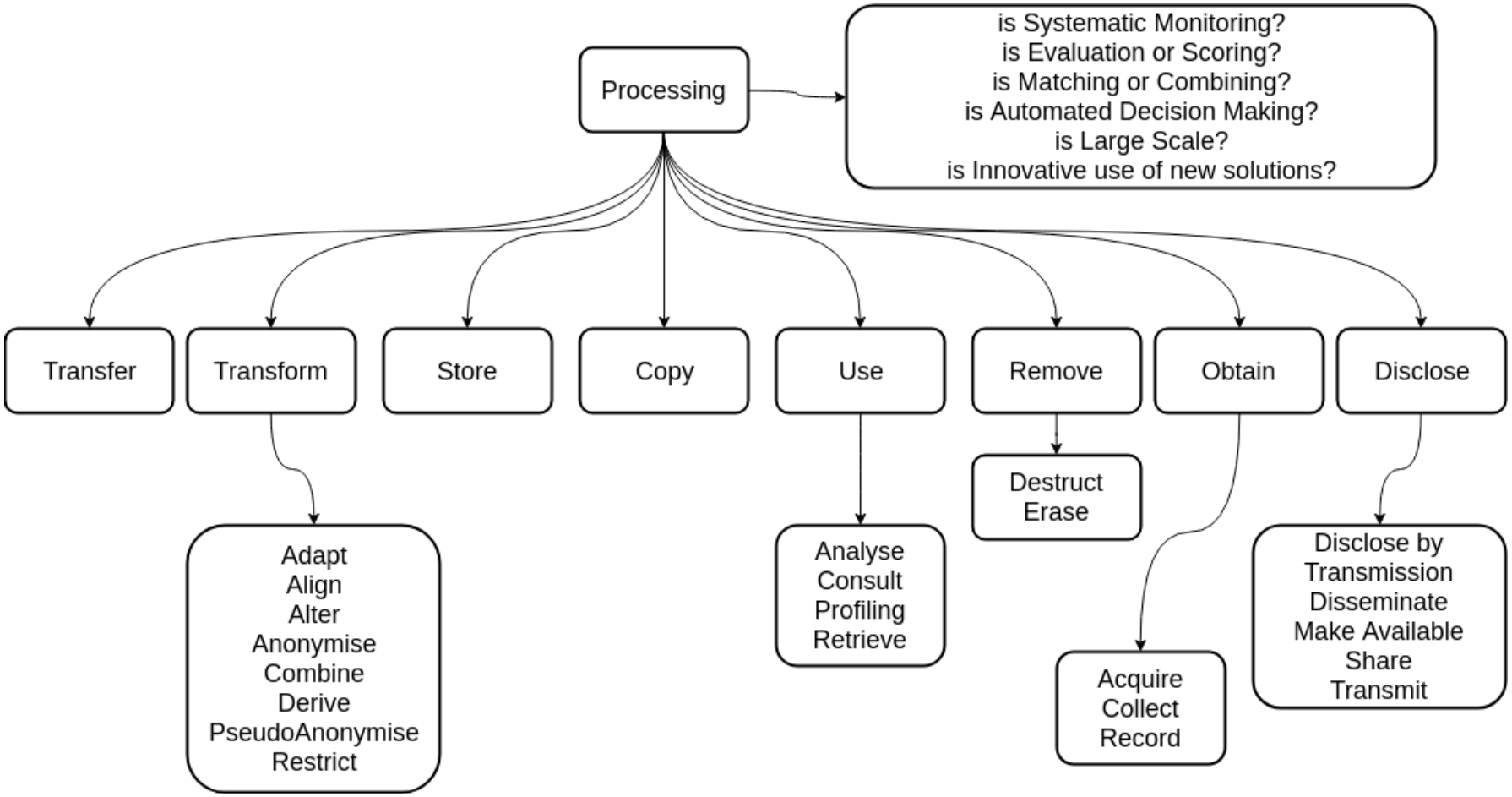



```
:NewPurpose
    rdfs:subClassOf dpv:DeliveryOfGoods, dpv:FraudPreventionAndDetection ;
    rdfs:label "New Purpose" ;
    rdfs:comment "Intended delivery of goods with fraud prevention" .
```

* RDFS representation of NACE https://github.com/dpvcg/dpv-nace

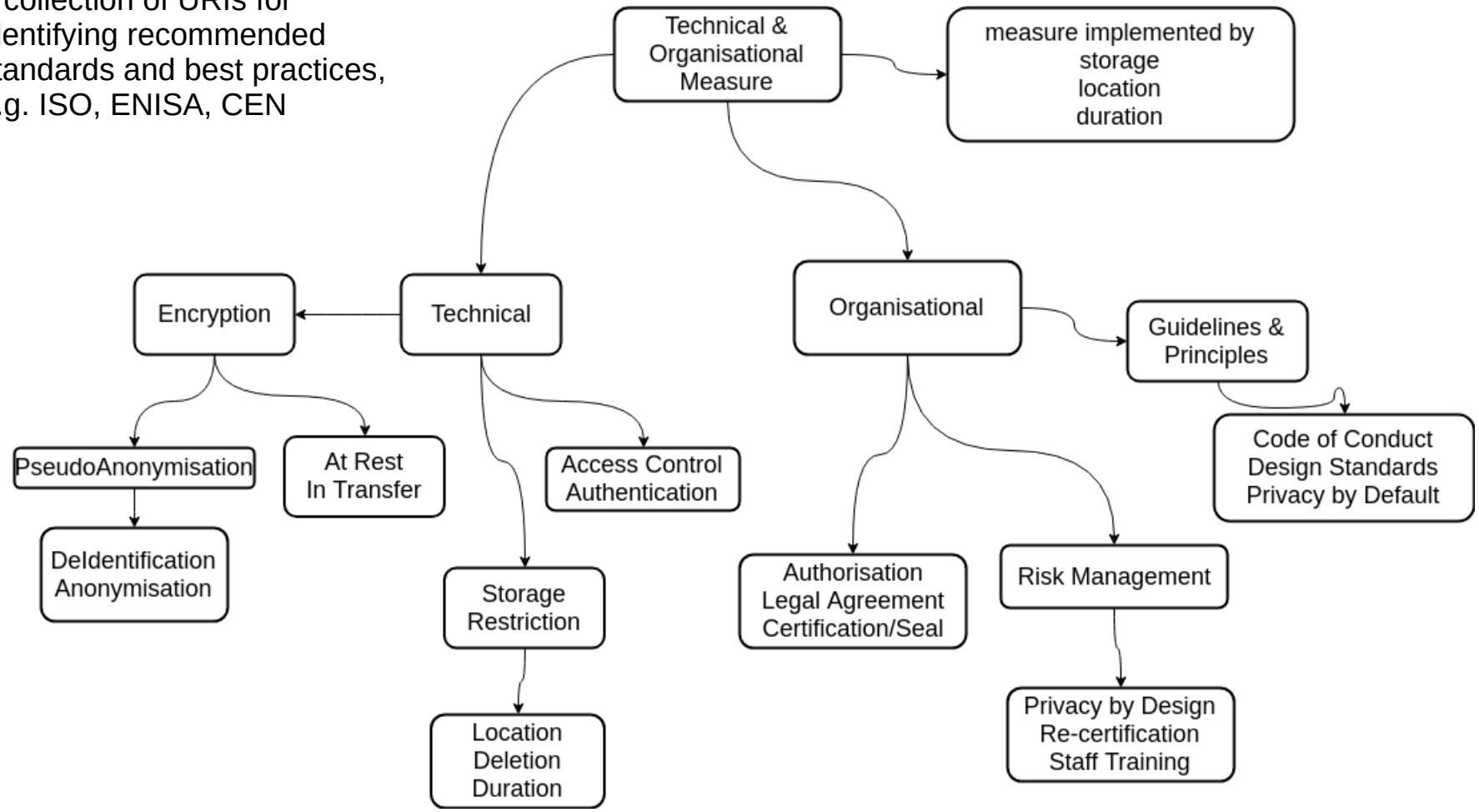
DPV – Processing

Hierarchy of Processing Categories



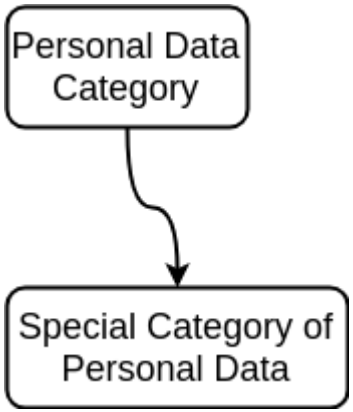
DPV – Technical & Organisational Measures

In the future, we plan to provide a collection of URIs for identifying recommended standards and best practices, e.g. ISO, ENISA, CEN



DPV – Personal Data

- Utilise taxonomy provided by Enterprivacy
- Extend for identified use-cases
- Add required classes e.g. Special Categories or Sensitive Categories

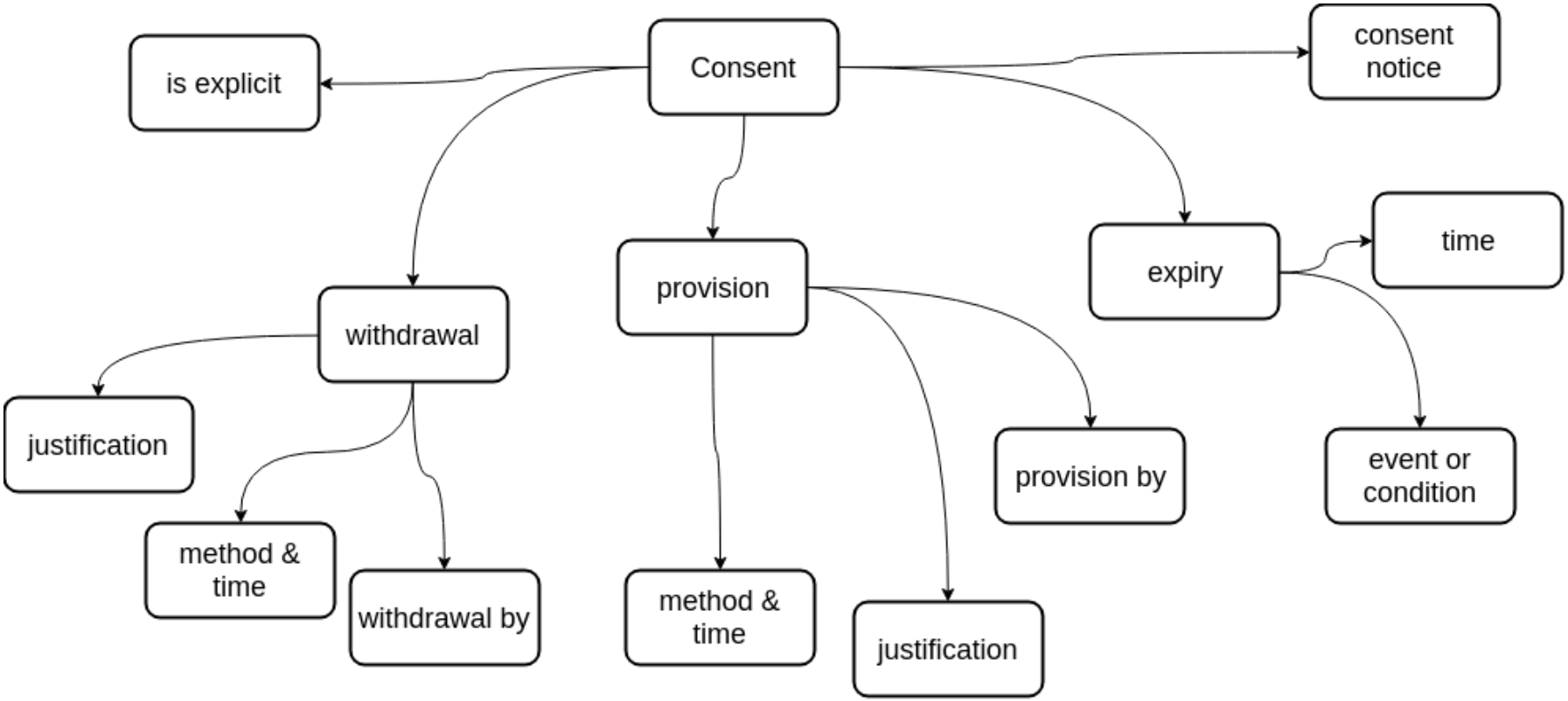



```
:PresentationsLiked
    rdfs:subClassOf dpv:Like ;
    rdfs:label “Presentations Liked” .

Like → Interest → Preference → Internal
```

DPV – Legal Bases – Consent (GDPR)

Attributes required to demonstrate given consent satisfies all legal requirements – based on Consent Receipt v1.1 standard



Future Work

- ‘Primer Document’: non-technical introduction to DPV
- Use-Cases and Examples documentation for adoption
- Open & ongoing feedback/suggestion
  - Github Issues / Pull-Requests
  - Public Mailing Lists
- Agreement towards v1 ‘stable’ vocabulary
- Provide JSON-LD fragments for adoption on Web
- Resolve open issues
  https://www.w3.org/community/dpvcg/track/issues/open

  Invitation to join and participate!

~ end of presentation ~

Creating A Vocabulary for Data Privacy

The First-Year Report of Data Privacy Vocabularies and Controls Community Group (DPVCG)

Harshvardhan J. Pandit, Axel Polleres, Bert Bos, Rob Brennan, Bud Bruegger, Fajar J. Ekaputra, Javier D. Fernández, Roghaiyeh Gachpaz Hamed, Elmar Kiesling, Mark Lizar, Eva Schlehahn, Simon Steyskal, and Rigo Wenning

Presenter: **Harshvardhan J. Pandit**

ADAPT Centre, School of Computer Science & Statistics, Trinity College Dublin, Ireland

email: pandith@tcd.ie | twitter: @coolharsh55

DPV vocabulary: https://w3.org/dpv

This work was supported by the European Union's Horizon 2020 research and innovation programme under grant 731601 (SPECIAL), by the Austrian Research Promotion Agency (FFG) under the projects "EXPEDiTE" and "CitySpin", by the ADAPT Centre for Digital Excellence funded by SFI Research Centres Programme (Grant 13/RC/2106), and co-funded by European Regional Development Fund.

The ADAPT Centre is funded under the SFI Research Centres Programme (Grant 13/RC/2106) and is co-funded under the European Regional Development Fund.