The W3C Data Privacy Vocabularies and Controls Community Group (DPVCG) proudly presents…

Data Privacy Vocabulary (DPV)

Version 2.0

Bigger!

Better!

Boundless!

https://w3id.org/dpv

Harshvardhan J. Pandit
ADAPT Centre, Dublin City University, Dublin, Ireland me@harshp.com

Beatriz Esteves
IDLab, Ghent University – imec, Ghent, Belgium beatriz.esteves@ugent.be

Georg P. Krog
Signatu AS, Oslo, Norway georg@signatu.com

Paul Ryan
ADAPT Centre, Dublin City University, and Uniphar PLC, Dublin, Ireland paul.ryan76@mail.dcu.ie

Delaram Golpayegani
ADAPT Centre, Trinity College Dublin, Dublin, Ireland delaram.golpayegani@adaptcentre.ie

Julian Flake
University of Koblenz, Koblenz, Germany flake@uni-koblenz.de

AVAILABLE NOW!
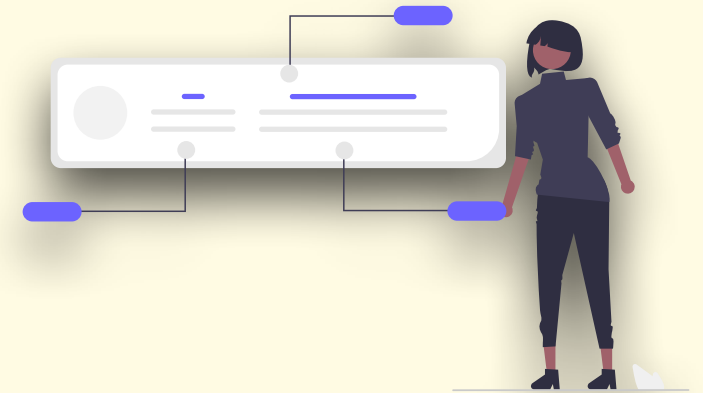
Presented at

ISWC
Baltimore, USA

2016

Why do we need 'Legal Metadata' ???

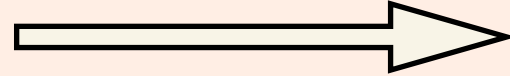**You need it to express processes and control them at a granular level**

You need it to model documents and information like contracts

You need it to enable user agents for *humans* to manage decisions

**And of course you need it to manage legal compliance**

2018

S PECIAL

→ **The W3C Data Privacy Vocabularies and Controls Community Group (DPVCG)**

The SPECIAL project launched the W3C Data Privacy Vocabularies and Controls Community Group (DVPCG) to develop a community that can create and maintain such legal metadata, based on EU GDPR

The DPVCG quickly attracted ontologists, computer scientists, lawyers, authorities, NGOs, academics, industry stakeholders, students, researchers, and basically formed a melting pot of expertise.
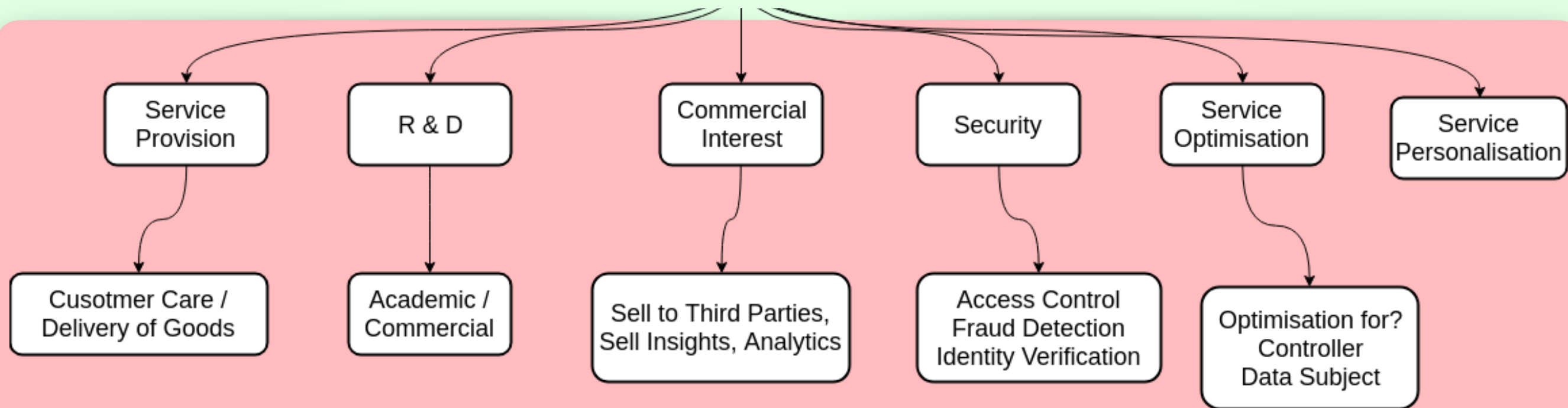
'Data Privacy Vocabulary' (DPV)
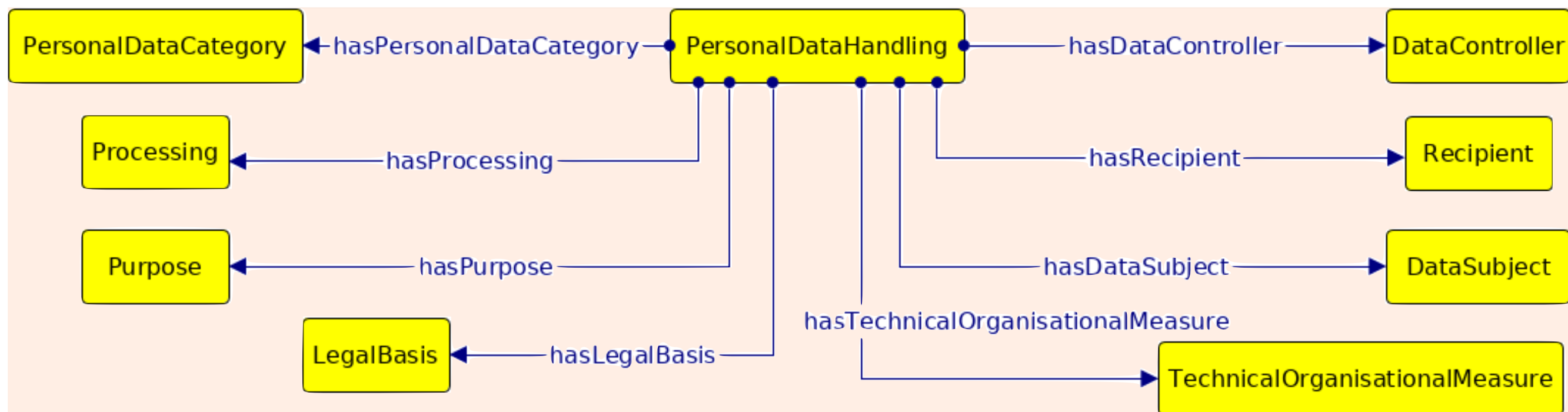
The DPVCG created the Data Privacy Vocabulary to provide an 'ontological model' for personal data being processed, and a 'taxonomical vocabulary' to support its effective use in practical settings.

2019

Ontology

Thing —hasPurpose→ Purpose

Purpose → Context

Purpose → Sector
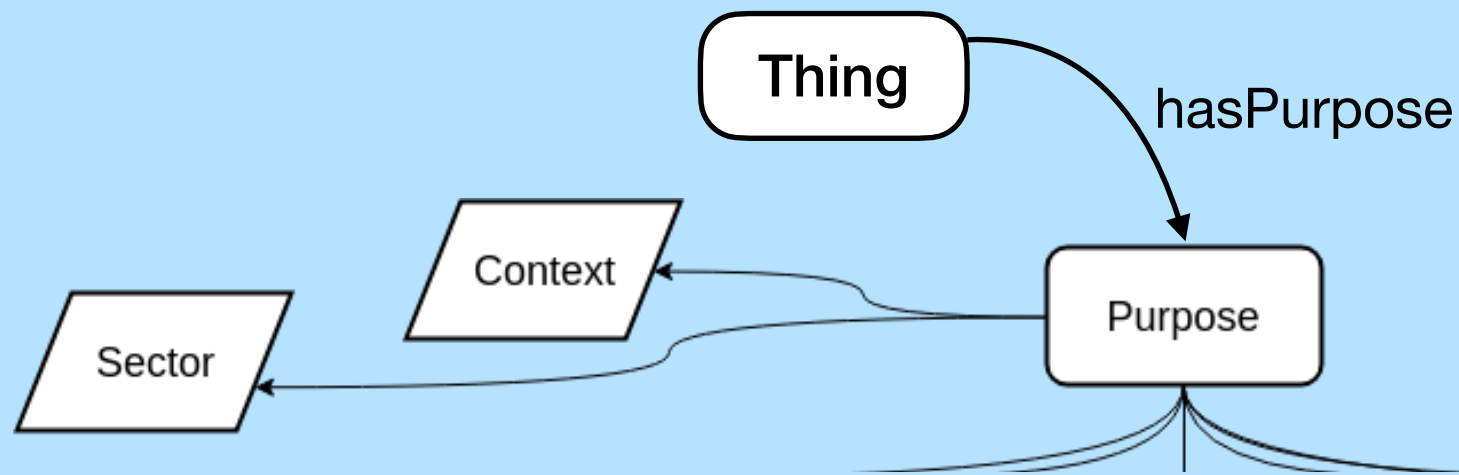
To put it simply, DPV is a two layered approach to create an 'ontology' for modelling the domain (T-box) and providing taxonomies to express use-cases in the domain (A-box)

Taxonomy

Purpose branches to:

- Service Provision → Cusotmer Care / Delivery of Goods
- R & D → Academic / Commercial
- Commercial Interest → Sell to Third Parties, Sell Insights, Analytics
- Security → Access Control Fraud Detection Identity Verification
- Service Optimisation → Optimisation for? Controller Data Subject
- Service Personalisation

2020

DPV provides the 'vocabulary' in a 'privacy language' to express things
(Seems kind of obvious now why its called 'data privacy' and 'vocabulary')
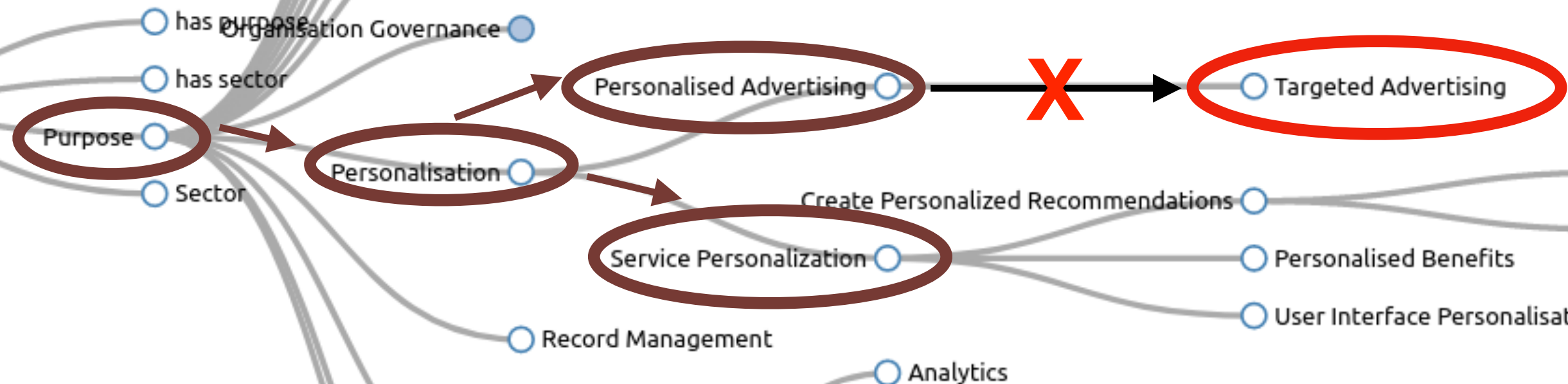
Other languages have many ways of expressing IF *x* THEN *y* or IF *x* THEN *not y*
DPV provides the concepts for *x* and *y* so these languages can be used *practically*

Data Privacy Vocabulary (DPV)

2022

version 1

<u>Final Community Group Report</u> 05 December 2022

This version:
https://www.w3.org/community/reports/dpvcg/CG-FINAL-dpv-20221205/

Latest published version:
https://w3id.org/dpv

Latest editor's draft:
https://w3id.org/dpv/ed/dpv

Editor:
Harshvardhan J. Pandit (ADAPT Centre, Dublin City University)

Former editor:
Axel Polleres (Vienna University of Economics and Business) - Until 31 December 2019

Authors:
Axel Polleres (Vienna University of Economics and Business)
Beatriz Esteves (Universidad Politécnica de Madrid)
Bert Bos (W3C/ERCIM)
Bud Bruegger (Unabhängige Landeszentrum für Datenschutz Schleswig-Holstein)
Elmar Kiesling (Vienna University of Technology)
Eva Schlehahn (Unabhängige Landeszentrum für Datenschutz Schleswig-Holstein)
David Hickey (Dublin City University)
Fajar J. Ekaputra (Vienna University of Technology)
Georg P. Krog (Signatu AS)
Harshvardhan J. Pandit (ADAPT Centre, Dublin City University)
Javier D. Fernández (Vienna University of Economics and Business)
Julian Flake (University of Koblenz-Landau)
Mark Lizar (OpenConsent/Kantara Initiative)
Paul Ryan (Uniphar PLC)
Piero Bonatti (Università di Napoli Federico II)
Ramisa Gachpaz Hamed (Trinity College Dublin)
Rigo Wenning (W3C/ERCIM)
Rob Brennan (University College Dublin)
Simon Steyskal (Siemens)

Approx. ~2100 'hand-crafted' concepts;

DPV-GDPR: GDPR Extension for DPV

version 1

Modelled GDPR specific concepts in a separate namespace

DPV-PD: Extended Personal Data categories for DPV

Extensive taxonomy of personal data categories for use with DPV concepts

DPV-LEGAL: Extension providing Jurisdictions, Laws, and Authorities for DPV

version 0.8.2

DRAFT concepts representing laws, authorities, inter-country agreements

DPV-TECH: Extension providing Technology concepts for DPV

version 0.8.2

DRAFT concepts representing provision methods, stakeholders, networking

2022

BUT… People wanted to use DPV in two completely different ways!

DPV assertions: dpv:hasPurpose rdfs:range dpv:Purpose .

dpv:Marketing a dpv:Purpose

Method 1: Direct use of DPV concepts
ex:Thing dpv:hasPurpose dpv:Marketing .

dpv:Marketing rdfs:subClassOf dpv:Purpose

Method 2: Instantiating DPV concepts
ex:Thing dpv:hasPurpose ex:SpecificMarketing .
ex:SpecificMarketing a dpv:Marketing .

Method 3: Extending DPV concepts
ex:Thing dpv:hasPurpose ex:SpecificMarketing .
ex:SpecificMarketing rdfs:subClassOf dpv:Marketing .

So how do we enable both uses???

We provide DPV with two serialisations:
SKOS+RDFS
and OWL

These are not ontologists! We can't "teach them the correct way"…

SKOS+RDFS Default serialisation / semantics

2022

DPV assertions:
dpv:Marketing, dpv:DirectMarketing a rdfs:Class, dpv:Purpose .
dpv:DirectMarketing skos:broader dpv:Marketing .

✗ **Simple reasoning**

Method 1: Direct use of DPV concepts
ex:Thing dpv:hasPurpose dpv:Marketing .

✓ **Can directly use DPV concepts**

Method 2+3: Instantiating DPV concepts
ex:Thing dpv:hasPurpose ex:Marketing2 .
ex:Marketing2 a dpv:Purpose ;
            skos:broader dpv:Marketing .

✓ **Can create 'instances'**

Use DPV as is

Method 2+3: Extending DPV concepts
ex:Thing dpv:hasPurpose ex:Marketing2 .
ex:Marketing2 a dpv:Purpose ; skos:broader ex:Marketing .

✓ **Can extend in use-case**

Provided under a different namespace **OWL**

Complex reasoning ✓

DPV assertions:
dpv:Marketing a rdfs:Class; rdfs:subClassOf dpv:Purpose .
dpv:DirectMarketing a rdfs:Class; rdfs:subClassOf dpv:Marketing .

Cannot directly use DPV concepts ✗

Method 1: Direct use of DPV concepts
ex:Thing dpv:hasPurpose dpv:Marketing .

Can create 'instances' ✓

Method 2: Instantiating DPV concepts
ex:Thing dpv:hasPurpose ex:Marketing2 .
ex:Marketing2 a dpv:Marketing .

Requires ontology re-engineering

Cannot extend ✗

Method 3: Extending DPV concepts
ex:Thing dpv:hasPurpose ex:Marketing2 .
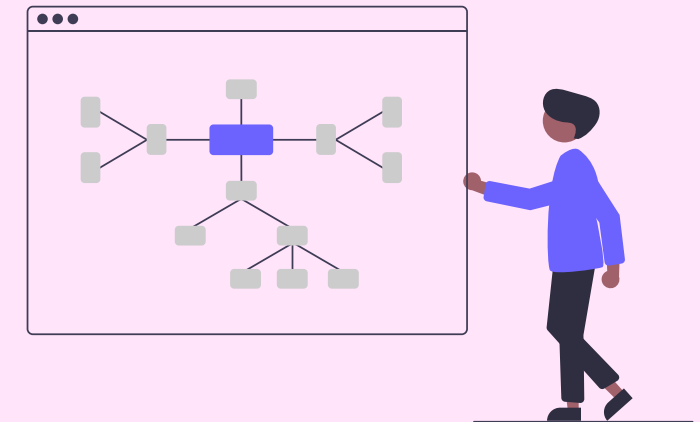ex:Marketing2 rdfs:subClassOf ex:Marketing .

2022
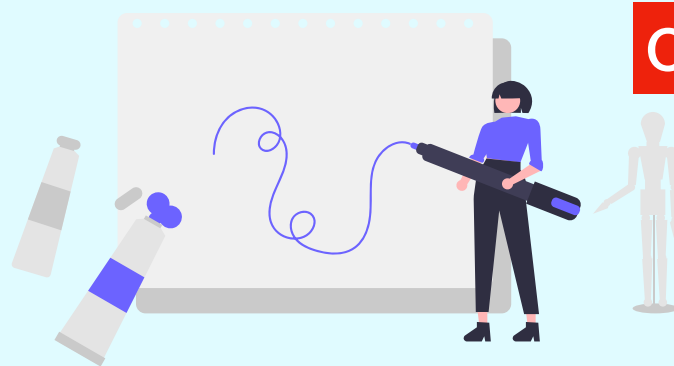
New laws adopted …
How to use DPV for those?

Laws are scoped to a particular jurisdiction. How to model different laws across different jurisdictions?

Some laws address personal and non-personal data. Some address technologies using personal data.

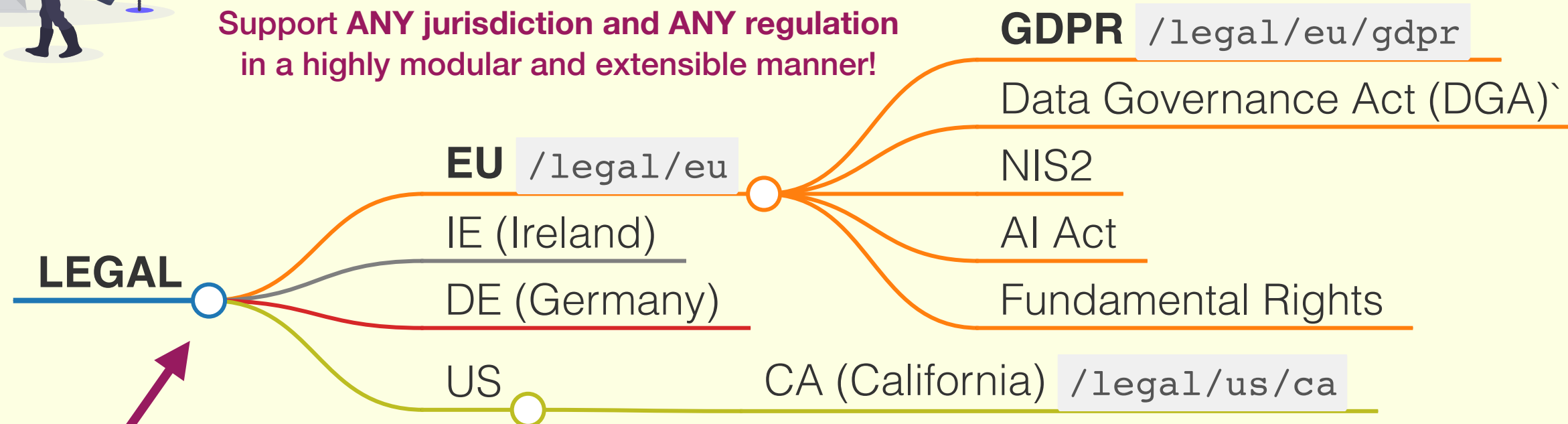Organisations need practical stuff like risk management and security reporting

Need of the hour: Artificial Intelligence (AI) technologies - how do we represent these same concepts like purpose, data, security for AI?

2023

Modular Legal Namespaces

Support ANY jurisdiction and ANY regulation in a highly modular and extensible manner!

GDPR `/legal/eu/gdpr`

Data Governance Act (DGA)`

EU `/legal/eu`

NIS2

IE (Ireland)

AI Act

LEGAL

DE (Germany)

Fundamental Rights

US

CA (California) `/legal/us/ca`

Provide a list of (privacy, data protection, AI) laws in these jurisdictions and the authorities they represent as a knowledge graph

Define ISO 3166-2 location concepts and using these state EU 27 countries / EEA 30 countries

Location concepts

§ 5. Authorities

Authorities are represented as instances of dpv:Authority, and are associated with specific jurisdictions using dpv:hasJurisdiction. The law which the authorities enforce is indicated by dpv:hasApplicableLaw. The webpage for the authority, if available, is indicated using foaf:homepage.

ID	Name	Jurisdictions	Laws	Webpage
legal-de:DPA-DE	The Federal Commissioner for Data Protection and Freedom of Information	Germany	legal-de:law-BDSG legal-eu:law-GDPR	link
legal-de:DPA-DE-BB	The state representative for data protection and the right to inspect files in Brandenburg	Brandenburg, Germany	legal-de:law-BDSG legal-de:law-BE-BbgDSG legal-eu:law-GDPR	link
legal-de:DPA-DE-BE	Berlin Commissioner for Data Protection and Freedom of Information	Berlin, Germany	legal-de:law-BDSG legal-de:law-BE-BlnDSG legal-eu:law-GDPR	link

§ 4. Laws

Laws are represented as instances of dpv:Law, and are associated with their jurisdictions using the [LOC] taxonomy and dpv:hasJurisdiction relation. The webpage for the law, if available, is indicated using foaf:homepage, and the temporal start and end if available, is represented using dct:temporal relation with an instance of time:ProperInterval.

ID	Name	Jurisdictions	Authorities	Webpage	Start/End
legal-de:law-BDSG	Federal Data Protection Act (BDSG)	Germany	legal-de:DPA-DE, legal-de:DPA-DE-BB, legal-de:DPA-DE-BE, legal-de:DPA-DE-BY-non-public, legal-de:DPA-DE-BY-public, legal-de:DPA-DE-HB, legal-de:DPA-DE-HE, legal-de:DPA-DE-HH, legal-de:DPA-DE-MV, legal-de:DPA-DE-NI, legal-de:DPA-DE-NW, legal-de:DPA-DE-RP, legal-de:DPA-DE-SH, legal-de:DPA-DE-SL, legal-de:DPA-DE-SN, legal-de:DPA-DE-ST, legal-de:DPA-DE-TH	link	2019-11-20/ongoing

Risk Management

2023

Statuses for representing incidents
Document security incident, data breaches
Indicate what data was affected, how

7. **risk:Damage**: Impact that acts as or causes
   a. **risk:CorruptionData**: Corruption of Da
   b. **risk:DamageByThirdParty**: Damage
   c. **risk:DataBreach**: Data Breach go to full
   d. **risk:EquipmentFailure**: Equipment Fa
   e. **risk:FinancialLoss**: Financial Loss go
   f. **risk:Harm**: Impact that acts as or caus
   -
      1. **risk:AbusiveContentUtilisation**:
      2. **risk:AttackonPrivateLife**: Attack
      3. **risk:Blackmail**: Blackmail go to ful
      4. **risk:ChildViolence**: Child Violen
      5. **risk:Coercion**: Coercion go to full
      6. **risk:CompromiseAccount**: Com
      7. **risk:CompromiseAccountCrede**
      8. **risk:DangertoCustomers**: Dang
      9. **risk:DangertoPersonnel**: Dange
      10. **risk:Discrimination**: Discriminati

Risk / Impact taxonomy

1. **risk:IncidentStatus**: Status associate
   a. **risk:IncidentConcluded**: The in
      mitigation and with a low likeliho
   b. **risk:IncidentHalted**: The inciden
      go to full definition
   c. **risk:IncidentMitigated**: The inci
      been applied to prevent the sam
   d. **risk:IncidentNearMiss**: The sta
      very close" go to full definition
   e. **risk:IncidentOngoing**: The inci

```
ex:IN12 a risk:Incident ;
  dpv:hasPersonalData pd:EmailAddress ;
  dpv:hasTechnicalMeasure dpv:Encryption
  # Expressing duration using method 1: D
  dpv:hasDuration [
    a dpv:UntilEventDuration ;
    dct:description "Incident continued u
  ];
  # Expressing duration using method 2: D
  dct:temporal "2023-05-24/2023-05-26" ;
dpv:hasStatus risk:IncidentConcluded .
```


1. **risk:RiskMatrix3x3**: A Risk Matrix with
   a. **risk:RM3x3S1L1**: Node in a 3x3 F
      Low go to full definition
   b. **risk:RM3x3S1L2**: Node in a 3x3 F
      Level: Low go to full definition
   c. **risk:RM3x3S1L3**: Node in a 3x3 F
      Moderate go to full definition
   d. **risk:RM3x3S2L1**: Node in a 3x3 F
      Level: Low go to full definition
   e. **risk:RM3x3S2L2**: Node in a 3x3 F
      Risk Level: Moderate go to full definit
   f. **risk:RM3x3S2L3**: Node in a 3x3 F
      Level: High go to full definition
   g. **risk:RM3x3S3L1**: Node in a 3x3 F
      Moderate go to full definition
   h. **risk:RM3x3S3L2**: Node in a 3x3 F
      Level: High go to full definition
   i. **risk:RM3x3S3L3**: Node in a 3x3 F
      High go to full definition

ISO 31000 series risk concepts
Risk Level = Severity x Likelihood

Legal Requirements

AI & AI Act Extension

2023

We increased the scope of our concepts from ***personal data*** to ***any data or technology***

AI Capability

Purpose

Domain

isProvidedBy → AI Provider

AI Operator

AI Technique

hasPurpose

isAppliedWithinDomain

isDeployedBy → AI Deployer

hasCapability

AI Subject

usesTechnique

AI System

hasAISubject

Modality

hasModality

hasAIUser

AI User

Locality Of Use

isUsedWithinLocality

hasComponent

Concepts regarding AI system and its use

AI Component

Area Of Impact

Stakeholder

3.8	Risk Sources
3.8.1	Attack
3.8.2	Adversarial Attack
3.8.3	Data Poisoning
3.8.4	Model Evasion
3.8.5	Model Inversion
3.8.6	Data Risk Source
3.8.7	Error In Data Collection
3.8.8	Error In Data Prepration
3.8.9	Input Data Risk Source
3.8.10	Erroneous Input Data
3.8.11	Test Data Risk Source

3.3	AI Capabilities
3.3.1	Audio Processing
3.3.2	Sound Source Separation
3.3.3	Sound Synthesis
3.3.4	Speaker Recognition
3.3.5	Speech Recognition
3.3.6	Speech Synthesis
3.3.7	Automatic Summarisation
3.3.8	Behaviour Analysis
3.3.9	Biometric Categorisation
3.3.10	Biometric Identification
3.3.11	Computer Vision

Based on
1) "*AIRO: An Ontology for Representing AI Risks Based on the Proposed EU AI Act and ISO Risk Management Standards*" by Golpayegani et al.
2) " *To Be High-Risk, or Not To Be-Semantic Specifications and Implications of the AI Act's High-Risk AI Applications and Harmonised Standards*" by Golpaywgani et a.

Who is using DPV? What are they doing?

2024

Year	Mention	Use	Ext.	Contrib.	Domain	Effort
2020	X				Health	N/A
2020	X				Media	N/A
2020	X					N/A
2020		X				++
2020		X			Health	++
2020			X	X		++
2021	X				Health	N/A
2021	X					N/A
2021		X				++
2021		X			Smart products	+
2021			X	X		+
2022	X					N/A
2022	X			X		N/A
2022			X	X		+
2022	X				Health	N/A
2022		X				+
2022		X			IoT	+
2022		X			Health	+
2023		X			Smart cities	+

Creating a **vocabulary** for **data privacy**: The first-year report of **data privacy vocabularies** and controls community group (DPVCG)
HJ Pandit, A Polleres, B Bos, R Brennan… - On the Move to …, 2019 - Springer
…) was set up to jointly develop such **vocabularies** towards interoperability in the context of **data privacy**. This paper presents the resulting **Data Privacy Vocabulary** (DPV), along with a …
☆ Save ⁇ Cite Cited by 96 Related articles All 9 versions

(At the time of writing the paper we had 81 citations)

We looked at how much *Effort* it would take for these works to integrate the changes made in DPV.

Most works only required minor changes in changed concepts (+ *in table)*, other works required changing IRIs to the new formats as they were using an old pre-v1.0 draft version of the DPV.

🏆 Prototyping an End-User User Interface for the Solid Application Interoperability Specification under GDPR by H Bailly, A Papanna, R Brennan wins **Best In-Use Paper award** at ESWC 2023 🏆

🏆 Using Patterns to Manage Governance of Solid Apps by B. Esteves, H. Pandit wins **Best Paper award** at Workshop on Ontology Patterns (WOP), ISWC 2023 🏆

🏆 Implementing ISO/IEC TS 27560: 2023 Consent Records and Receipts for GDPR and DGA by H. Pandit, J. Lindquist, G. P. Krog wins **Best Paper award** at Annual Privacy Forum 2024 organised by European Union Agency for Cybersecurity (ENISA) and EU Commission 🏆

Are projects* and the industry using DPV? 2024

EU Horizon Projects

SPECIAL   TRAPEZE

MOSAICrOWN

smashHit   FAIRVASC

PROTECT ITN

Industry

Signatu   Inrupt   JLINC

Athumi / Solid Lab

* These are the projects we know about. Most uses of DPV go unreported, and DPV is also commonly referenced in footnotes and inline text which we do not have a way to track. Projects also contained industry partners.

Spotlight: Signatu's use of DPV (Georg P. Krog)

Signatu: Practical Implementation

Signatu provides a practical solution for implementing ISO/IEC TS 27560:2023 using the the Data Privacy Taxonomy for GDPR and DGA compliance. **Signatu's** Legal Compliance and Consent Management Platforms offer tools for creating, managing, and exchanging consent records and receipts in line with the latest standards.

Data Entry Chatbot   Taxonomy Representation

Generate Consent Notice   Generate Consent Records with Events

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

Use of DPV in Standards

2021

DPV's consent modeling and concepts are used to develop ISO/IEC TS 27560:2023 standard

2023

ISO/IEC TS 27560:2023

TECHNICAL SPECIFICATION — ISO/IEC TS 27560:2023

Edition 1 2023-08

Privacy technologies — Consent record information structure
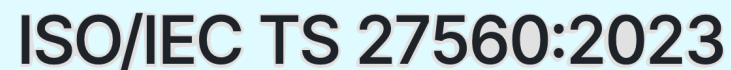
Published (Edition 1, 2023)

DPV is directly used in an ISO standard for representing consent records/receipts as Linked Data.

2024

Consent Records and Receipts as per ISO/IEC TS 27560:2023 using DPV
Final Community Group Report 01 August 2024

DPVCG provides an implementation guide for representing consent records/receipts

https://w3id.org/dpv/guides/consent-27560

EXAMPLE 39: Example of a Consent Record

```
{
    "@id": "https://example.com/a6f58
    "@type": "dpv:ConsentRecord",
    "dct:identifier": "a6f58318-72e6-
    "dct:conformsTo": "https://w3id.c
    "dpv:hasDataSubject": {
        "@id": "0760c9ba",
        "skos:broader": "dpv:Consumer
    },
    "dpv:hasDataController": "ex:Acme
    "dpv:hasDataProcessor": "ex:Beta'
    "dpv:hasJurisdiction": "loc:IE",
    "dpv:hasApplicableLaw": "eu-gdpr
    "dpv:hasProcess": {
        "@type": "dpv:Process",
        "dpv:hasRecipient": ["ex:Acme
        "dpv:hasPurpose": "dpv:Paymer
        "dpv:hasPersonalData": "pd:En
        "dpv:hasStorageCondition": [
            "@type": "dpv:StorageLoca
            "dpv:hasLocation": ["loc
        }, {
            "@type": "dpv:StorageDura
            "dpv:hasDuration": "P6M"
        }, {
            "@type": "dpv:StorageDele
            "dpv:hasDuration": "P1M"
        }]
    },
```

Data Privacy Vocabulary — Version 2 | International Semantic Web Conference (ISWC) | 14 NOV 2024 | Baltimore, Maryland, USA

2024

And then we released DPV version 2.0 on 1 AUG 2024...

DPV 2.0 Release 🎉

The DPVCG is proud to present DPV version 2.0 – a major release that significantly improves and expands the scope and usefulness as compared to DPV 1.0.

The article Data Privacy Vocabulary (DPV) -- Version 2 by Pandit et al. (2024), accepted for presentation at the 23rd International Semantic Web Conference (ISWC 2024), describes DPV 2.0 in terms of its contents, methodology, current adoptions and uses, and future potential. It also describes the relevance and role of DPV in acting as a common vocabulary to support various regulatory (e.g. EU's DGA and AI Act) and community initiatives (e.g. Solid) emerging across the globe. A Search Index of all concepts from DPV and extensions is available.

Summary of Changes

Major Changes

2024
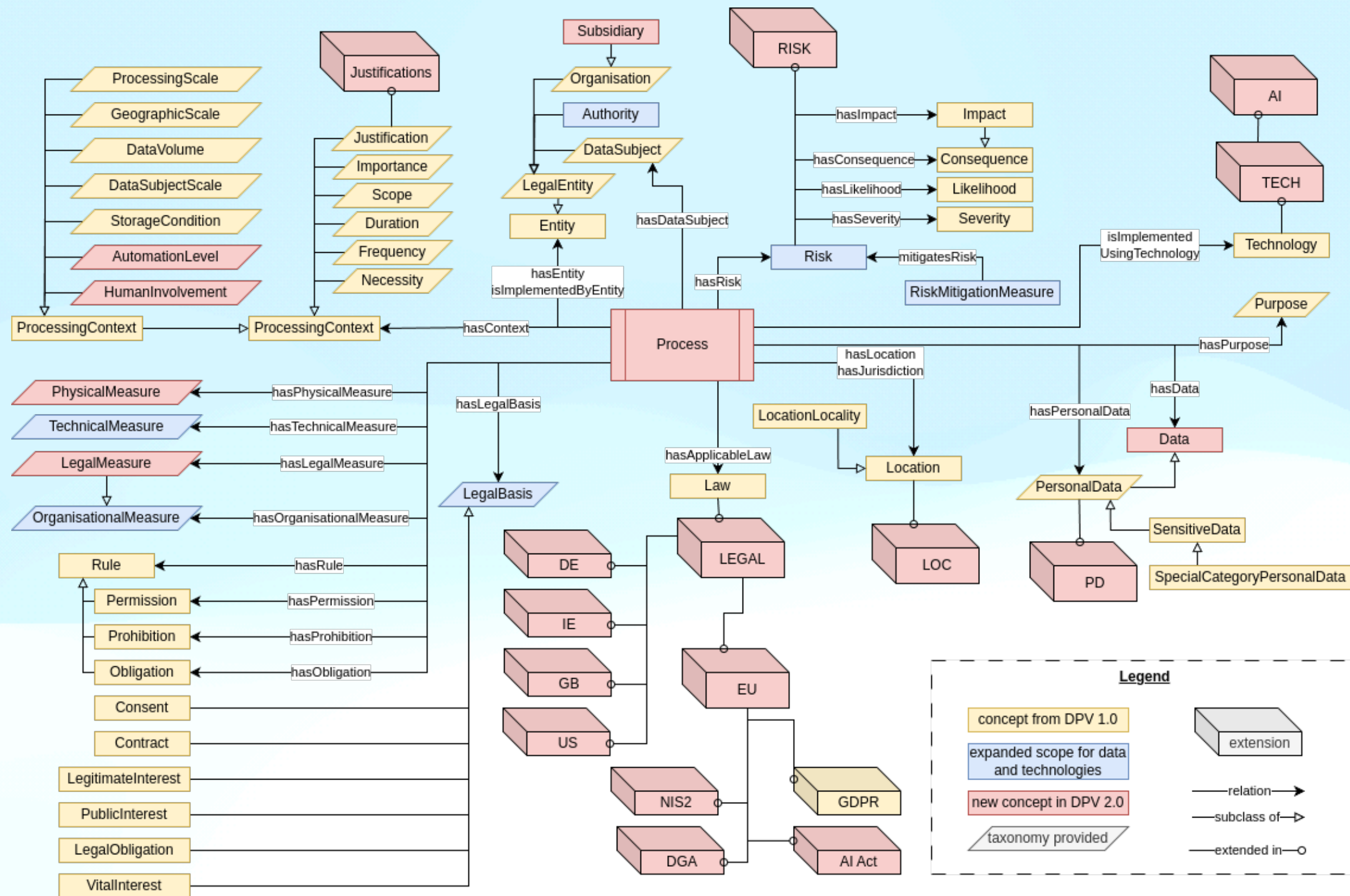
2394 concepts (with 2198 classes and 196 properties)

https://w3id.org/dpv/2.0/changelog

What's coming next?

2025

DPV 2.1 is scheduled for release in JAN'25 with ~6000 concepts - a 3x increase!

Come join us in the W3C Data Privacy Vocabularies and Controls Community Group (DPVCG) !!!

Find a short intro at https://www.dpvcg.org/

Primer

Data Privacy Vocabulary (DPV)

Final Community Group Report 01 August 2024

https://w3id.org/dpv/primer

IEEE SA STANDARDS ASSOCIATION   ◆IEEE

IEEE P7012

IEEE Draft Standard for Machine Readable Personal Privacy Terms

IEEE P7012 is using DPV to express terms

More Jurisdictions
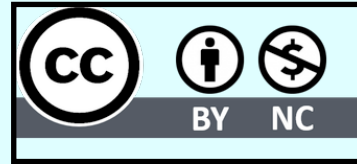
AI taxonomies

ODRL alignment

Risks/Impacts

AI Act compliance

Guidance for using DCAT and PROV

New ISO standard proposed based on DPV modelling of Legal Bases

Machine-Actionable Rights

The W3C Data Privacy Vocabularies and Controls Community Group (DPVCG) proudly presents…

Data Privacy Vocabulary (DPV)

Version 2.0

Bigger!

Better!

Boundless!

https://w3id.org/dpv

Harshvardhan J. Pandit
ADAPT Centre, Dublin City University, Dublin, Ireland me@harshp.com

Beatriz Esteves
IDLab, Ghent University – imec, Ghent, Belgium beatriz.esteves@ugent.be

Georg P. Krog
Signatu AS, Oslo, Norway georg@signatu.com

Paul Ryan
ADAPT Centre, Dublin City University, and Uniphar PLC, Dublin, Ireland paul.ryan76@mail.dcu.ie

Delaram Golpayegani
ADAPT Centre, Trinity College Dublin, Dublin, Ireland delaram.golpayegani@adaptcentre.ie

Julian Flake
University of Koblenz, Koblenz, Germany flake@uni-koblenz.de

Presented at

ISWC
Baltimore, USA