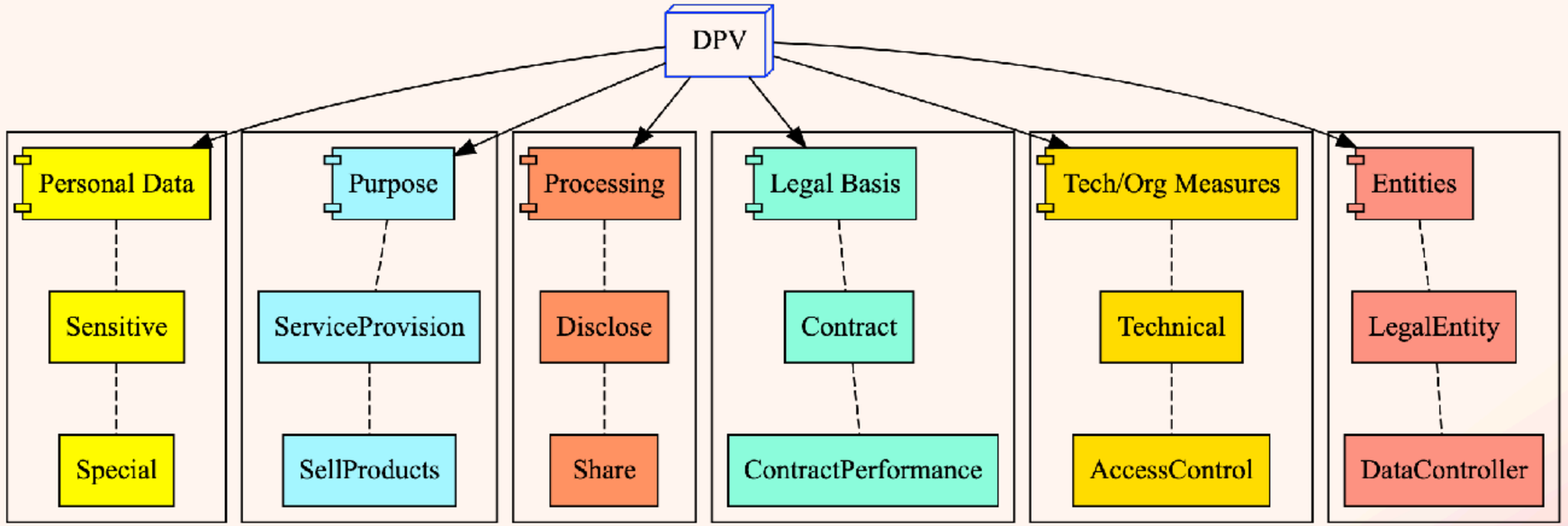


https://w3id.org/dpv

Data Privacy Vocabulary (DPV)

Semantics for GDPR and Personal Data Processing

Harsh(vardhan J. Pandit)

Research Fellow @ ADAPT Centre, Trinity College Dublin

Email: pandith@tcd.ie | dpv@harshp.com | Twitter: @coolharsh55

https://harshp.com/research/presentations

Presentation for SEMANTiCS 2022

From Compliance Information to Compliance Automation

GDPR Requirement := Maintain Information

Actors := Departments, Authorities, Data Subject, Processor, Controller

Information := Personal Data, Purpose of processing, Implementation details, Storage, Recipients, Technical or Organisational Measures to protect data

1) automation possible only if information is machine-readable

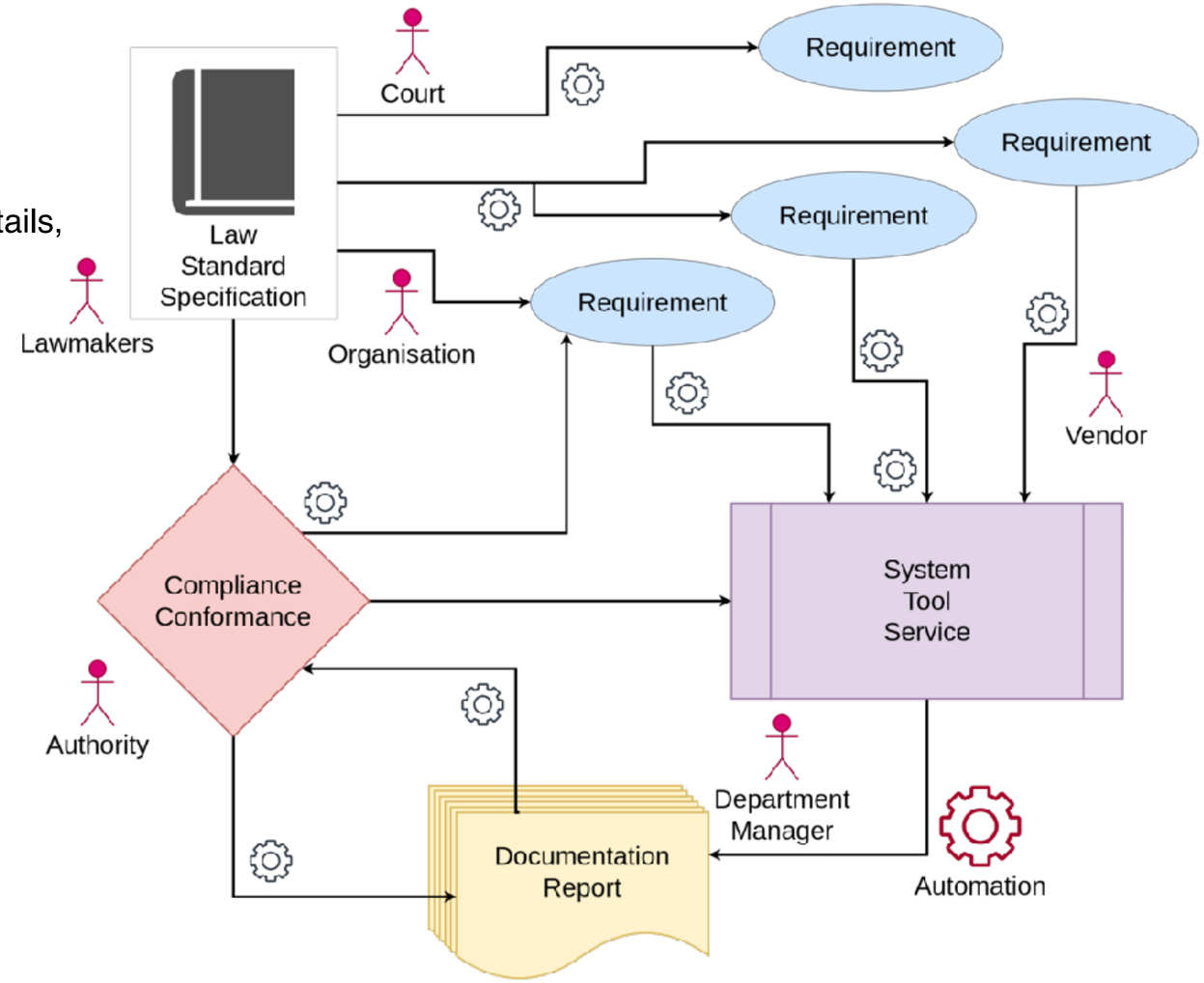
Tasks := Ensure correctness of information, Evaluate compliance

2) solutions can be shared only if information is interoperable

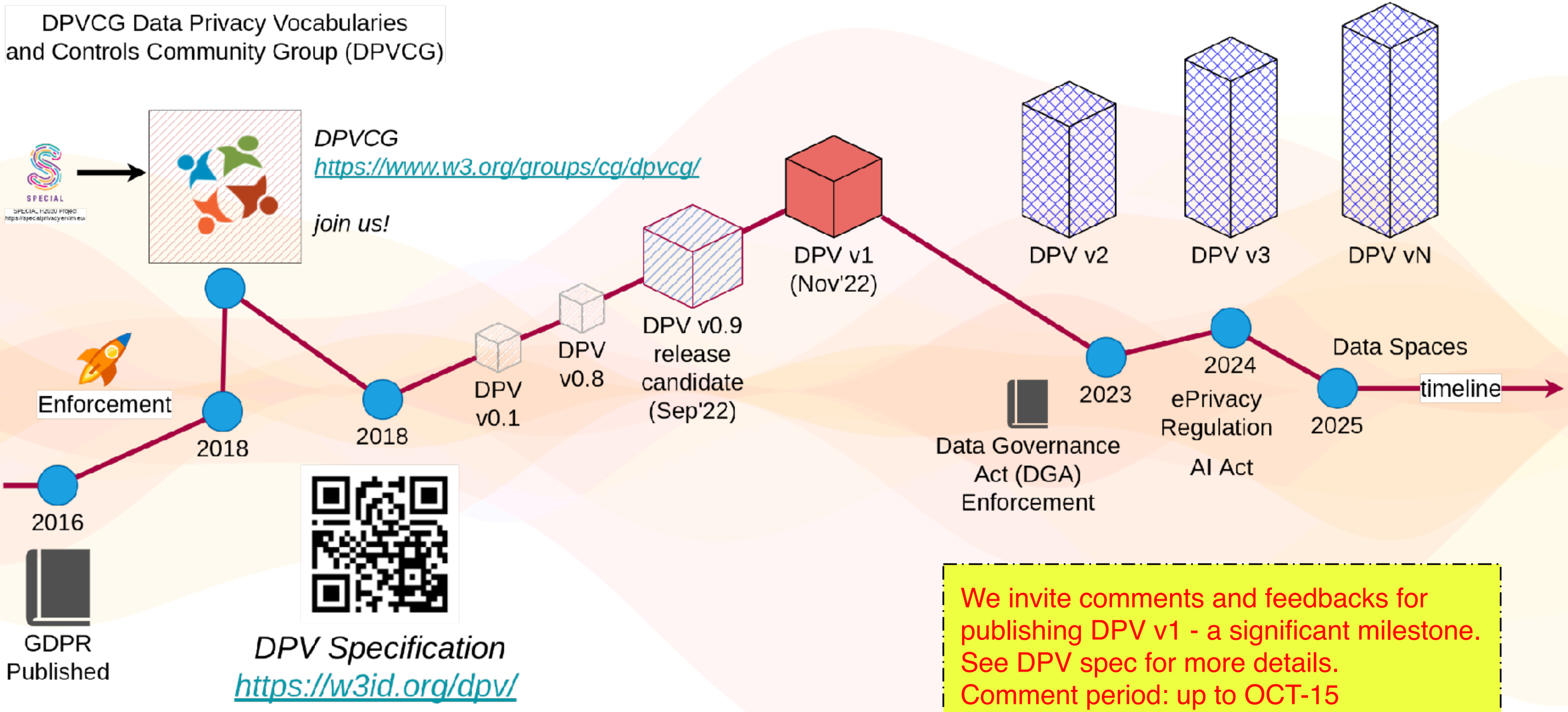
Outputs := Internal Reports, Register of Processing Activities (ROPA), Privacy Notices, Controller-Processor Communications

3) solutions need to be consistent for uniform compliance

Our solution:
The **Data Privacy Vocabulary (DPV)** https://w3id.org/dpv/



DPVCG Data Privacy Vocabularies and Controls Community Group (DPVCG)



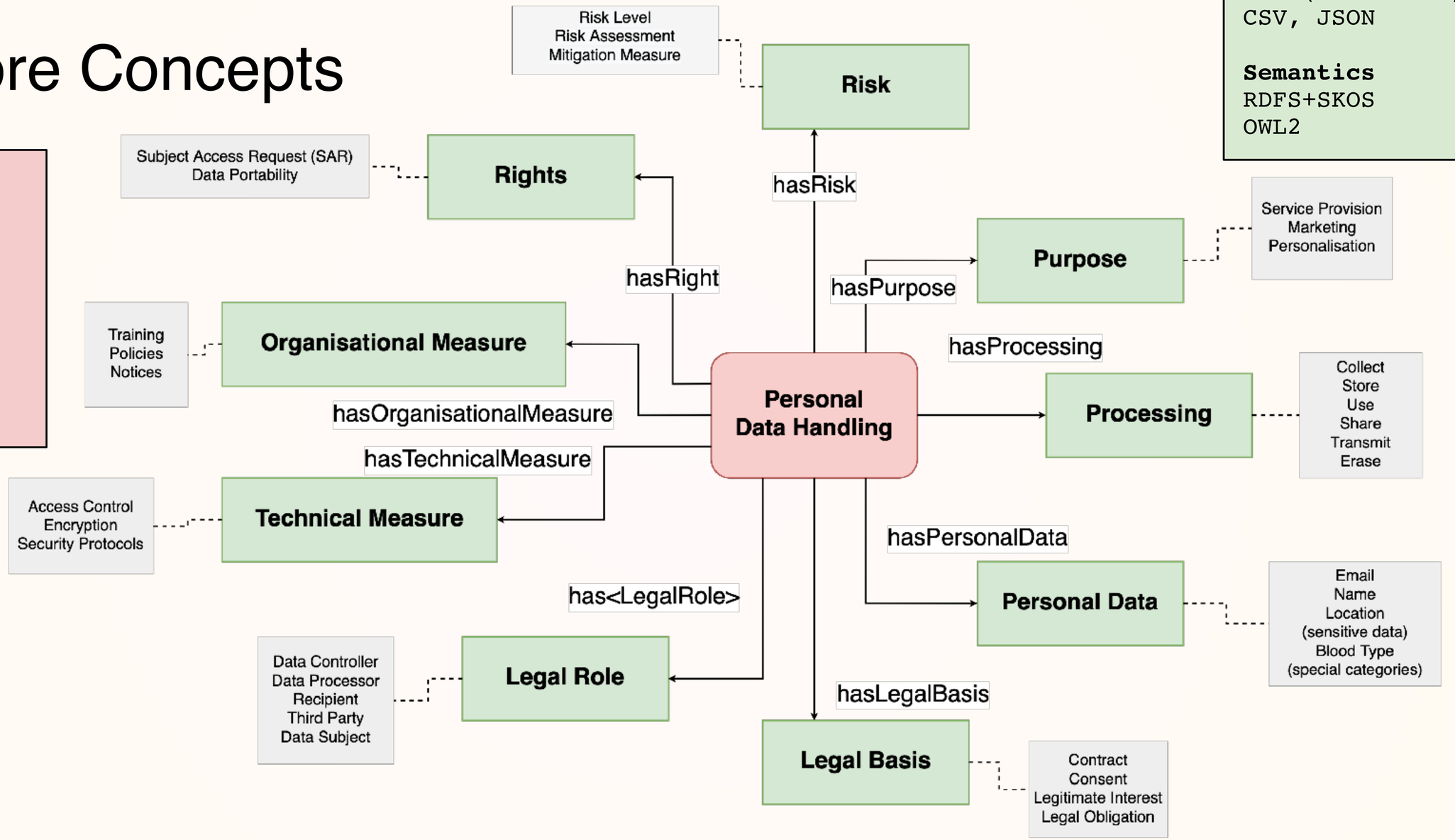
We invite comments and feedbacks for publishing DPV v1 - a significant milestone.
See DPV spec for more details.
Comment period: up to OCT-15

DPV Core Concepts

Purpose
Personal Data
Processing
Legal Basis
Legal Role
Tech/Org Measures
Risk
Rights

Serialisations
RDF (canonical)
CSV, JSON

Semantics
RDFS+SKOS
OWL2



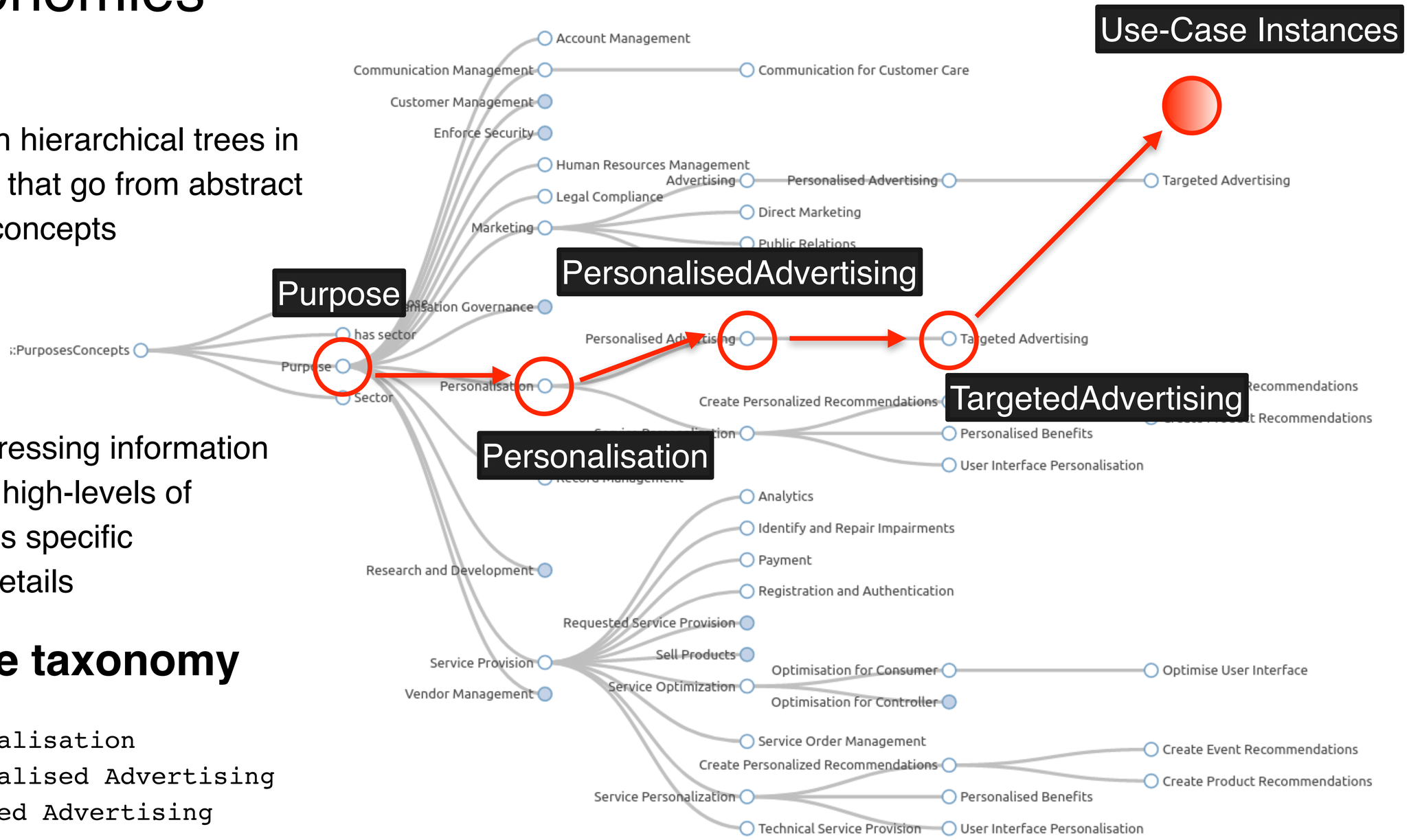
DPV Taxonomies

DPV provides rich hierarchical trees in top-down fashion that go from abstract to more specific concepts

This enables expressing information and rules at both high-levels of abstraction and as specific implementation details

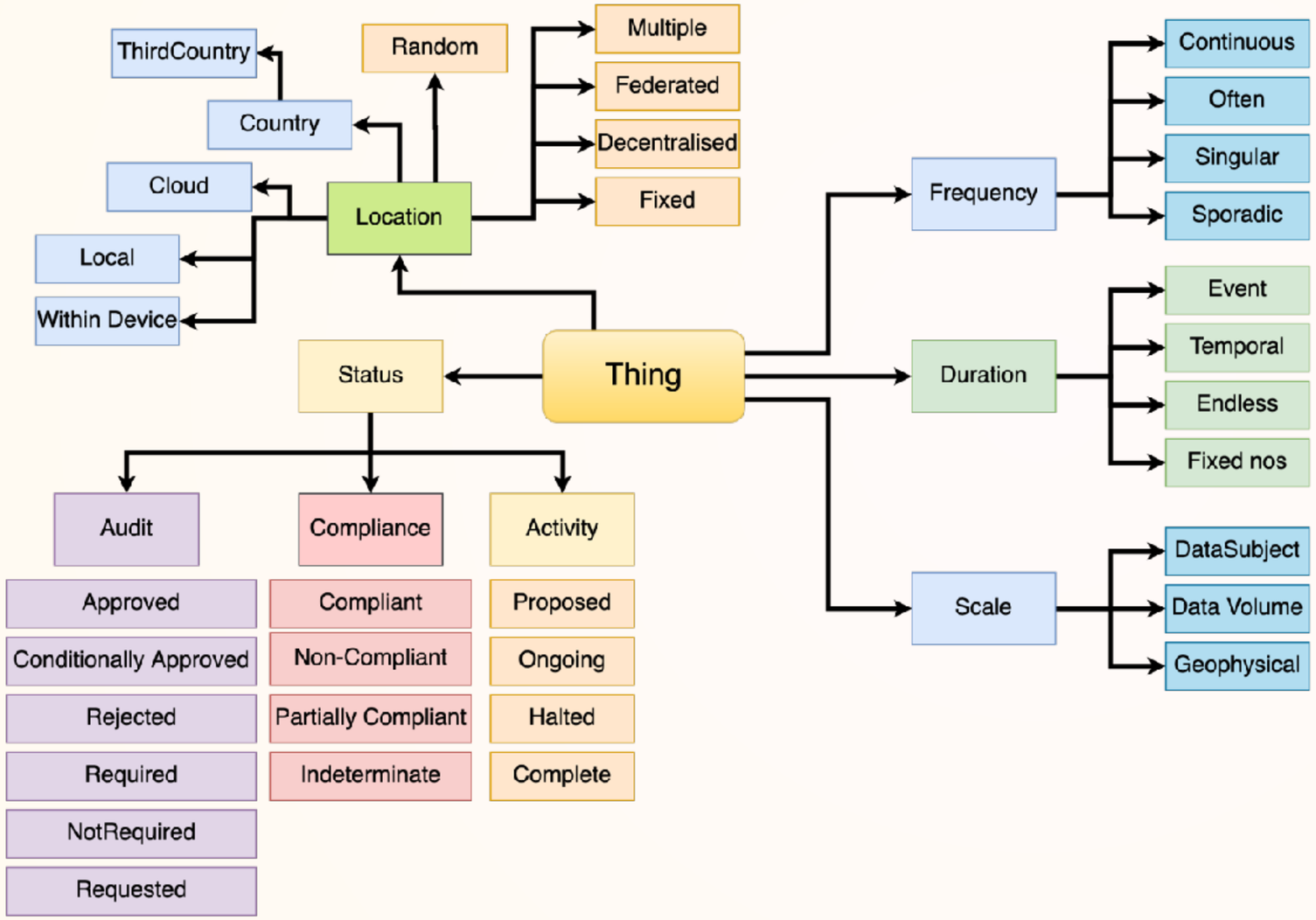
E.g. Purpose taxonomy

```
Purpose → Personalisation
        → Personalised Advertising
        → Targeted Advertising
```



generated by SKOSPlay https://skos-play.sparna.fr/

Contextual Information



Duration

Scale

Frequency

Status

Audits

Compliance

Location

Jurisdiction

Adequacy Decisions

Data Source

Automation

Storage conditions

Decision Making

Human Involvement

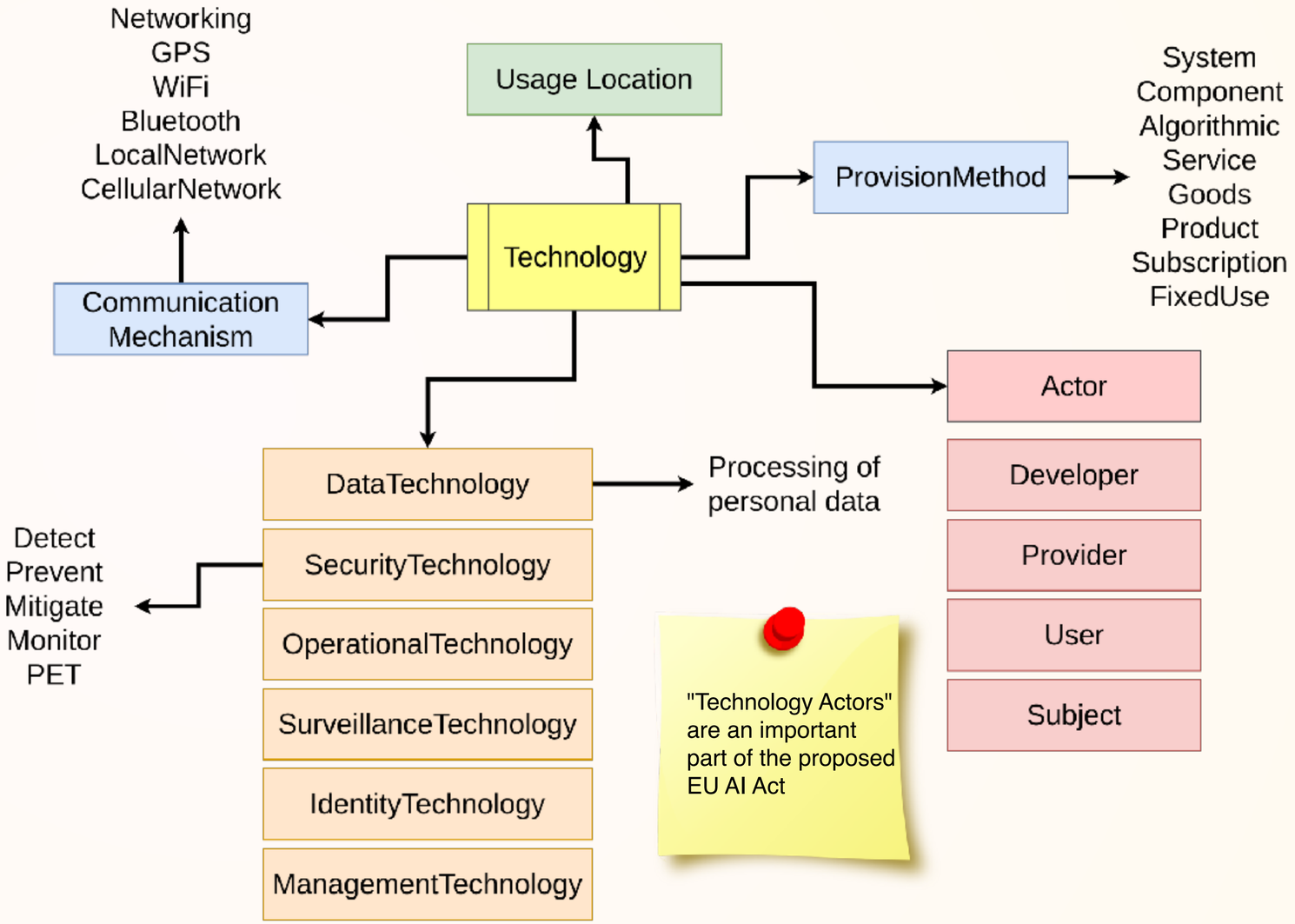
Consequences & Impacts

Technologies and Implementations

"Technology" is how you implement a process.

It can be a service, tool, or a system you develop, reuse, or purchase from a vendor.

Representing them is important to provide *practical knowledge* about personal data and its processing.



DPV Extensions

"*Extensions*" are additional vocabularies that provide enrichments to concepts defined within DPV (as the primary vocabulary).

Extensions can be rapidly changing concepts, opinionated models, domain or jurisdiction specific taxonomies, or simply provided for convenience.

dpv-pd	Personal Data concepts
dpv-legal	Countries, EU Membership, Laws, Authorities, Adequacy Decision
dpv-tech	Technologies and its provision and functioning
risk	Risk Assessment and Risk Management
rights	Exercising rights, communicating about rights, EU fundamental rights
dpv-gdpr	GDPR specific concepts such as legal bases, data transfer tools, rights

DPV Applications

current work

1. Register of Processing Activities (ROPA)
2. Consent Records
3. Compliance Checking
4. Impact Assessments (PIA / DPIA)
5. Data Input/Output Assistance
6. Annotating code / documents
7. Expressing and Evaluating Rules

work in progress

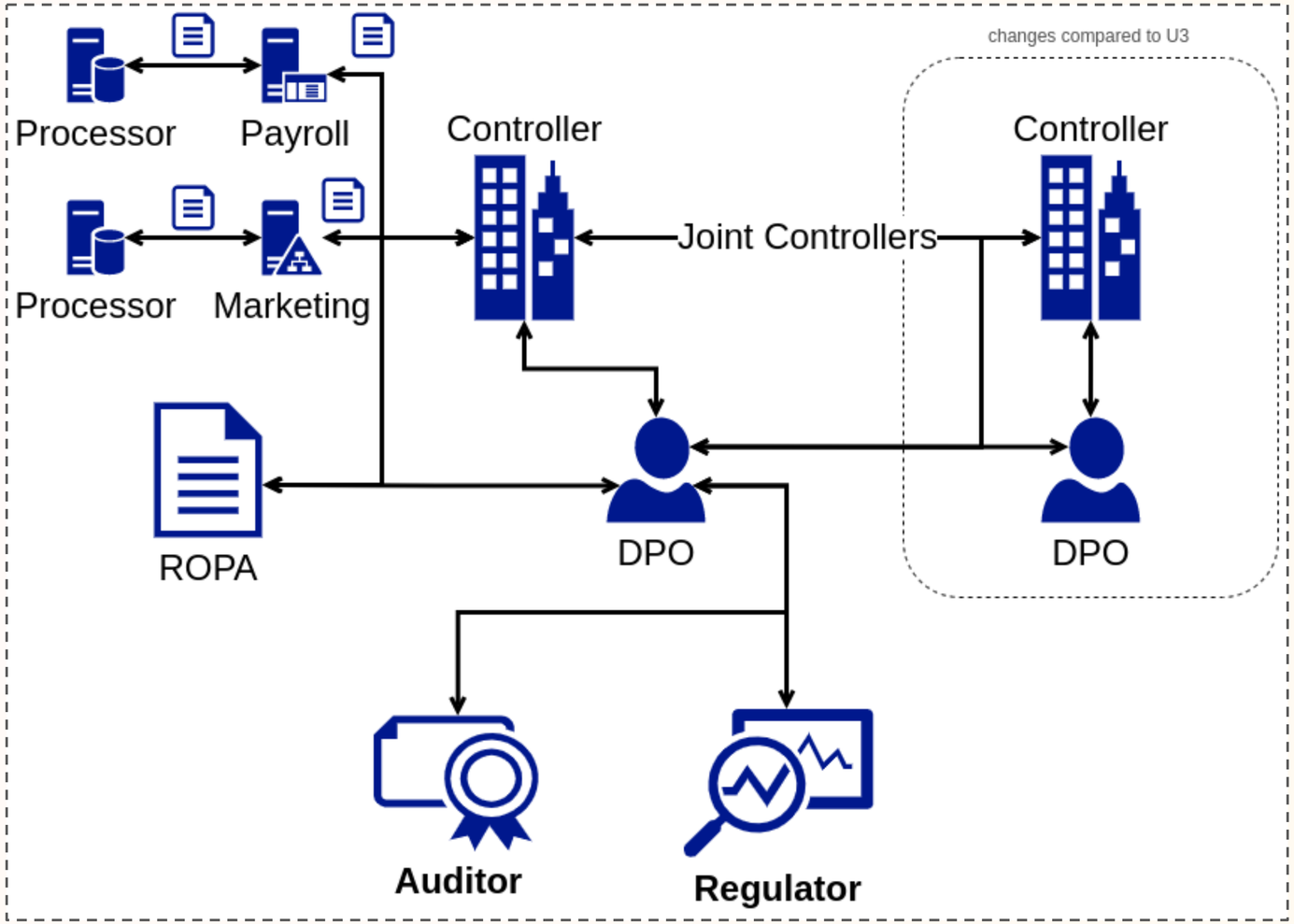
1. Risk Management
2. Data Breach Records
3. Subject Access Request
4. Data Portability
5. Data Transfers
6. Privacy Policies
7. Standards & Guidelines

Register of Processing Activities (ROPA)

Information flows between:
- Organisational units
- DPOs (internal/external)
- Controller
- Joint-controller
- Processor
- Auditors
- Authorities

DPV enables **interoperable solutions** for:
- intra-organisation data flows
- inter-organisation data flows

DPV can be used as:
- common language to communicate
- internal language to build and use tools



DPCat: Data Processing Catalogue Specification

ROPA as 'information'

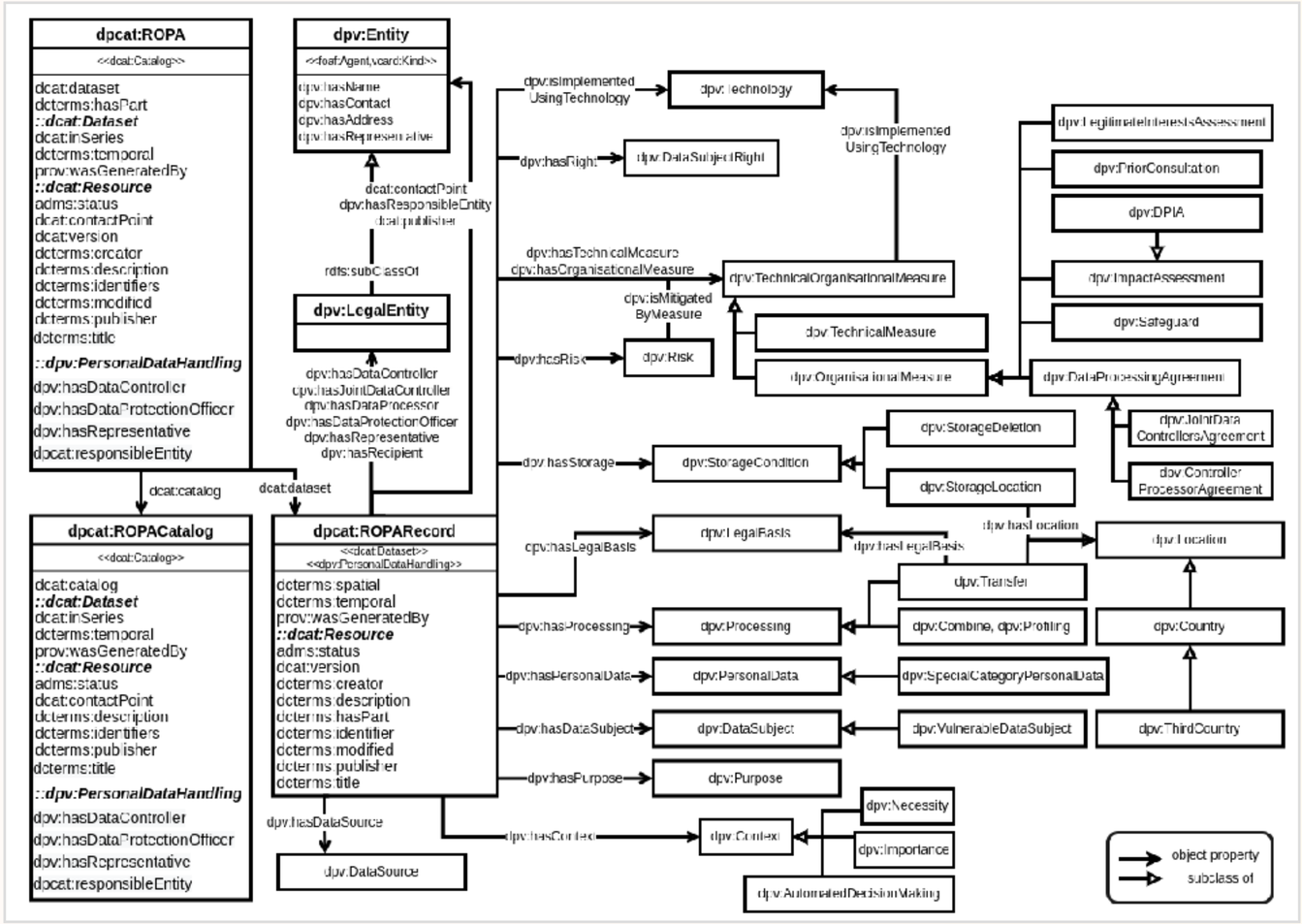
Use DPV to represent concepts collated from EU DPAs about personal data processing based on GDPR

ROPA as 'information record'

Extend DCAT to use ROPA as a dataset that can be packaged and shared and tracked with provenance

ROPA as 'information process'

Identify data governance and information flows for GDPR's ROPA based on actors such as controllers, processors, units, DPOs, authorities and use DPV to build solutions for internal use or interoperability



DPCat: Specification for an Interoperable and Machine-Readable Data Processing Catalogue Based on GDPR by *P. Ryan, R. Brennan, and H. J. Pandit* https://doi.org/10.3390/info13050244

Consent, and Consent Records

```
:63ded36f-4acd-4f3c-991e-6cb636698523 a dpv:ConsentRecord ;
    dct:hasVersion "27560-WD5" ;
    dpv:hasIdentifier "63ded36f-4acd-4f3c-991e-6cb636698523" ;
    dpv:hasDataSubject "96121fde-199f-4848-8942-4436e270513a" ;
    dpv:hasNotice "https://example.com/privacy-notice"^^xsd:anyURI ;
    dpv:hasPersonalDataHandling [
      a dpv:PersonalDataHandling ;
      dct:title "Send Newsletters with Seasonal Offers"@en ;
      dpv:hasPurpose dpv:Marketing ;
      dpv:hasLegalBasis dpv:Consent ;
      dpv:hasPersonalData dpv-pd:Email ;
      dpv:hasDataController ex:Acme ;
      dpv:hasProcessing dpv:Collect, dpv:Store ;
      dpv:hasStorageCondition [
          dpv:hasLocation dpv-legal:IE ;
          dpv:hasDuration 63115200 ;
          ] ;
      dpv:hasJurisdiction dpv-legal:EU ;
      dpv:hasRecipient ex:Beta, ex:Epsilon ;
    ] ;
    dpv:hasConsentStatus dpv:ConsentGiven ;
    dct:hasPart [
        a dpv:ConsentGiven, dpv:ExplicitlyExpressedConsent ;
        dpv:isIndicatedAtTime "2021-05-28T12:24:00"^^xsd:dateTime ;
        dpv:hasDuration 63115200 ;
        dpv:hasEntity "96121fde-199f-4848-8942-4436e270513a"
    ] .
```

One of the most common topics!

DPV (v0.1) initially modelled a simple and straightforward consent model based on common requirements

DPV v0.8+ has comprehensive 'Consent Records' to enable tracking consent forms, events, and use its state in implementations

Intended conformance with standards such as ISO/IEC TR 27560

Rule-based compliance checking

Creation of rules requires domain experts (legal, logic, computing)

Use of rules requires real-world concepts - which DPV's taxonomies provide

Compliance checking can be combination of concepts:

```
implement_process = personal data ^ purpose ^ processing ^ recipient ^ legal basis
```

Compliance evaluation can be validation of constraints:

```
[ a sh:NodeShape ; sh:property [ sh:path dpv:hasLegalBasis ; sh:minCount 1 ; ] ]
```

EU H2020 projects have produced several rule checking mechanisms that can be (re-)used, e.g. see SPECIAL, TRAPEZE, MIREL, DAPRECO, BPR4GDPR

Pan-Jurisdictional KG

DPV seems to incorporate several concepts from GDPR, but it is not *exclusive* to it.

DPV concepts are intended to be jurisdiction agnostic, with GDPR specific concepts declared in an extension: **DPV-GDPR**

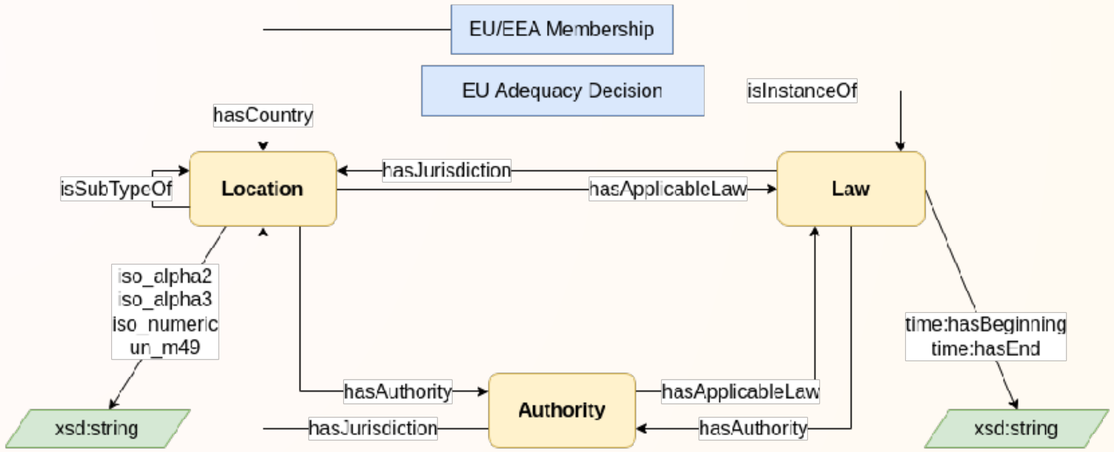
Similarly, other extensions can model different jurisdictions and domains by *extending* from DPV as a base vocabulary.

For specifying instances for different jurisdictions within the same graph, location can be utilised, e.g.

```
  <Product>  hasJurisdiction B .
<Controller> hasJurisdiction C .
```

Examples of what DPV can be extended with:

dpv-ccpa: CCPA/US concepts and requirements
dpv-iso: Aligning DPV concepts with ISO terminology
dpv-dga: Modelling Data Governance Act (DGA) as an extension of GDPR (via DPV)



The **`dpv-legal`** extension helps with pan-jurisdictional information by providing a graph of laws, authorities, etc;

Adoption and Use

SPECIAL H2020 Project *https://specialprivacy.ercim.eu/* - policy based compliance and provenance for GDPR

Signatu AS *https://signatu.com/* - automated tools and uniform machine-readable compliance

TRAPEZE H2020 Project *https://trapeze-project.eu/* - policies and compliance mechanisms for GDPR

smashHit H2020 Project *https://smashhit.eu/* - represent consent and contract as a legal KG

MOSAICrOWN H2020 Project *https://mosaicrown.eu/* - model data processing as a legal KG

FAIRVASC H2020 Project *https://fairvasc.eu/* - represent patient consent information/forms, and DPIAs

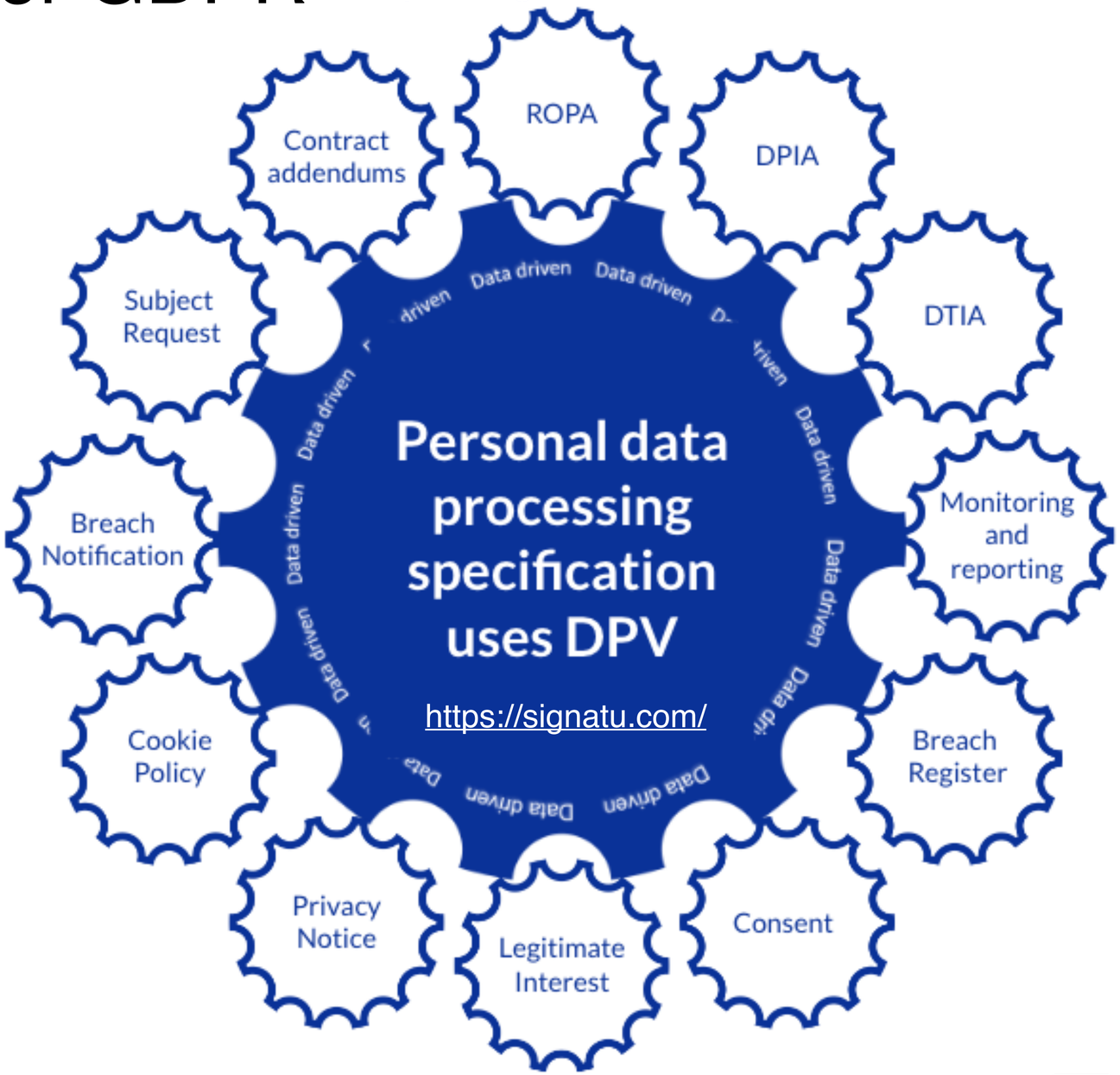
GDPR data transfer compliance framework *https://doi.org/10.3233/FAIA210332* by D. Hickey & R. Brennan

>>> for more, see *https://www.w3.org/community/dpvcg/wiki/Adoption_of_DPVCG*

Signatu AS - A SaaS Platform for GDPR https://signatu.com/

Signatu delivers INNOVATIVE Legal Compliance and Data Governance as SaaS or Managed Service:

- Data model & Knowledge Graph
- Controlled Vocabularies for Use-Cases
- Uniform and Reliable Application
- Automation for Documentation
- Data/Compliance is Machine-readable
- Data/Compliance is Interoperable
- Data/Compliance can be Queried
- Custom Views as Required
- Going beyond data lakes/warehouses/silos



TRAPEZE H2020 Project

https://trapeze-project.eu/

Transparency, Privacy & Security for European Citizens

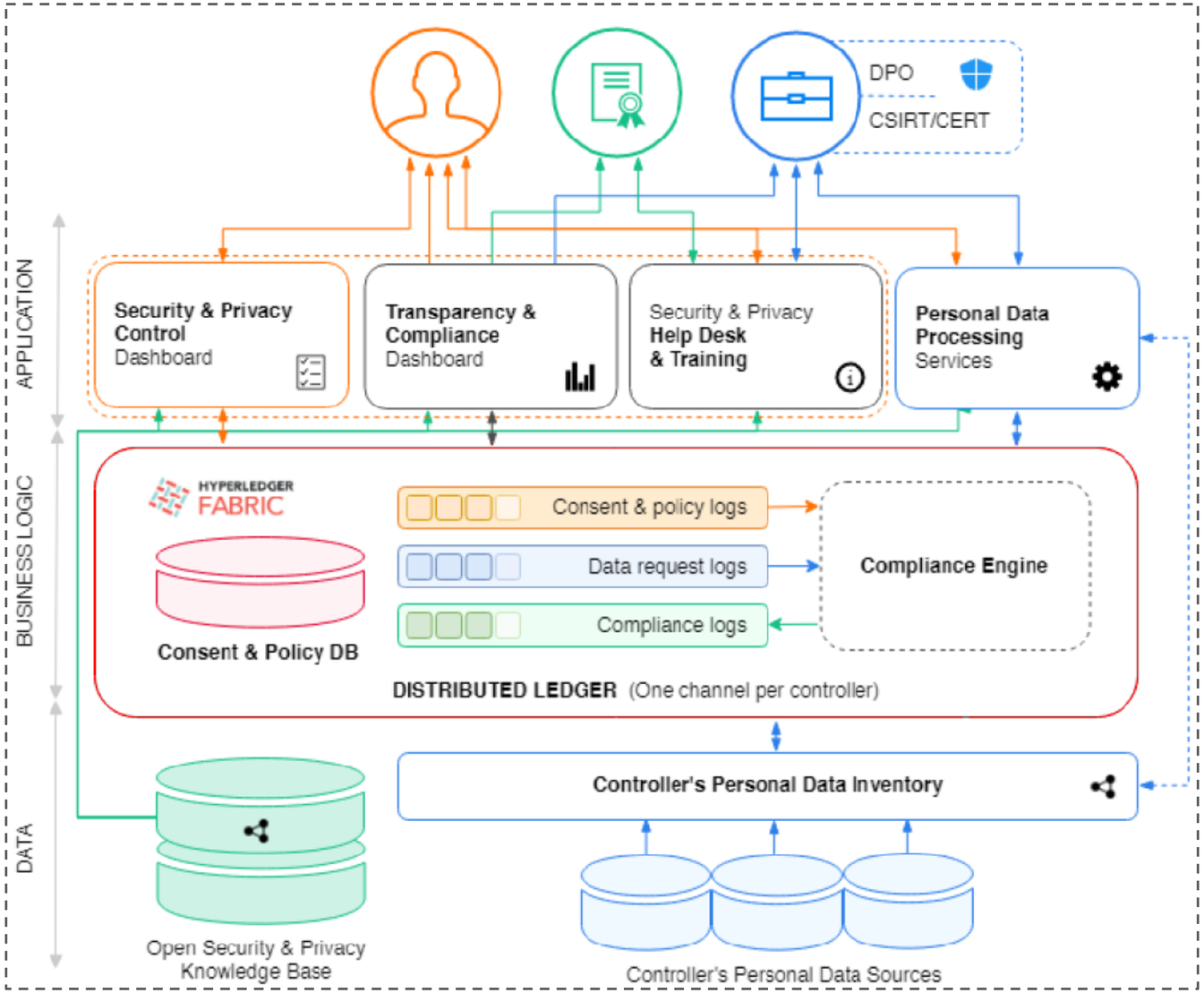
Uses DPV for creating a *semantic policy language* that uniformly represents: Privacy policies, data subjects' consent, data protection regulations

Machine understandable + automated compliance checking **using an OWL2 profile** that uses formal semantics to provide provably correct compliance (*no false positives nor negatives*)

Real time compliance checking: *~200 µsec/check*

Usability: The employees of TRAPEZE's industrial partners can write their privacy policies

Facilitating adoption: JSON serialisation available (equivalent to OWL2 version)

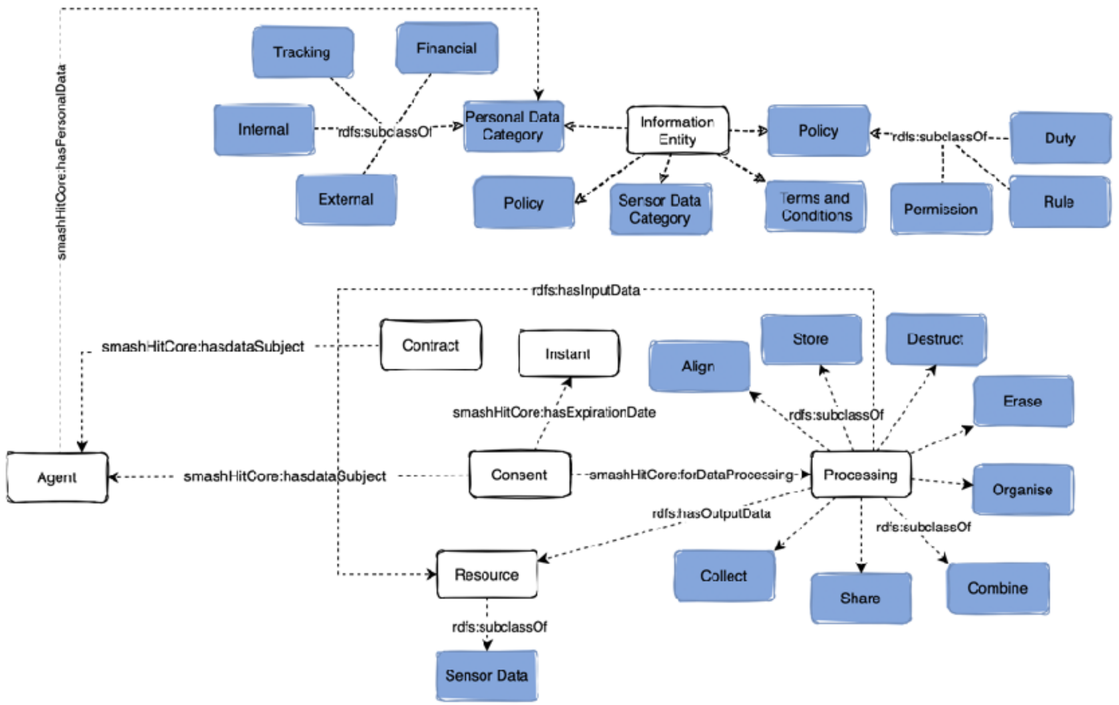


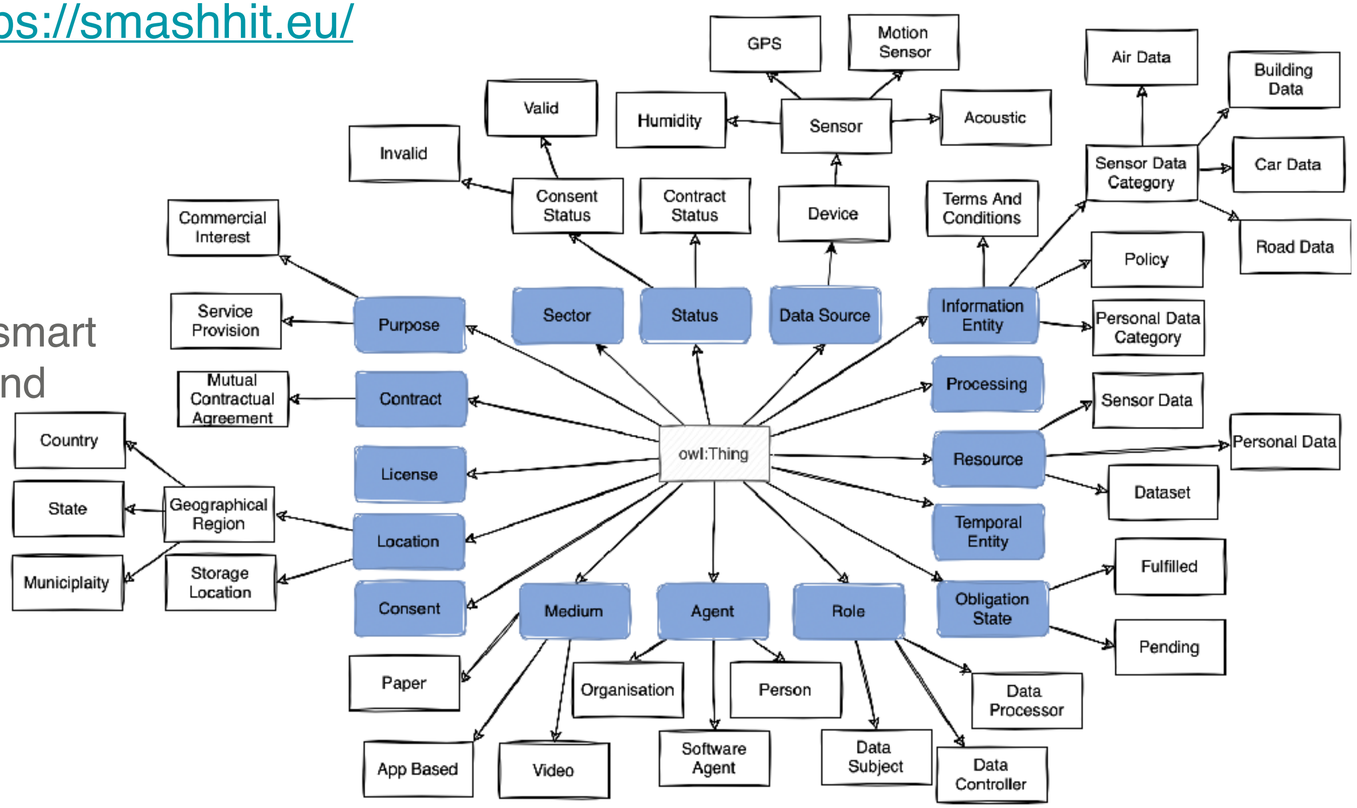
TRAPEZE Architecture Overview

smashHit H2020 project

https://smashhit.eu/

Ontology for GDPR-compliant sensor data sharing in the smart cities and insurance domains that models both ***consent*** and ***contracts*** as GDPR legal basis for data processing.

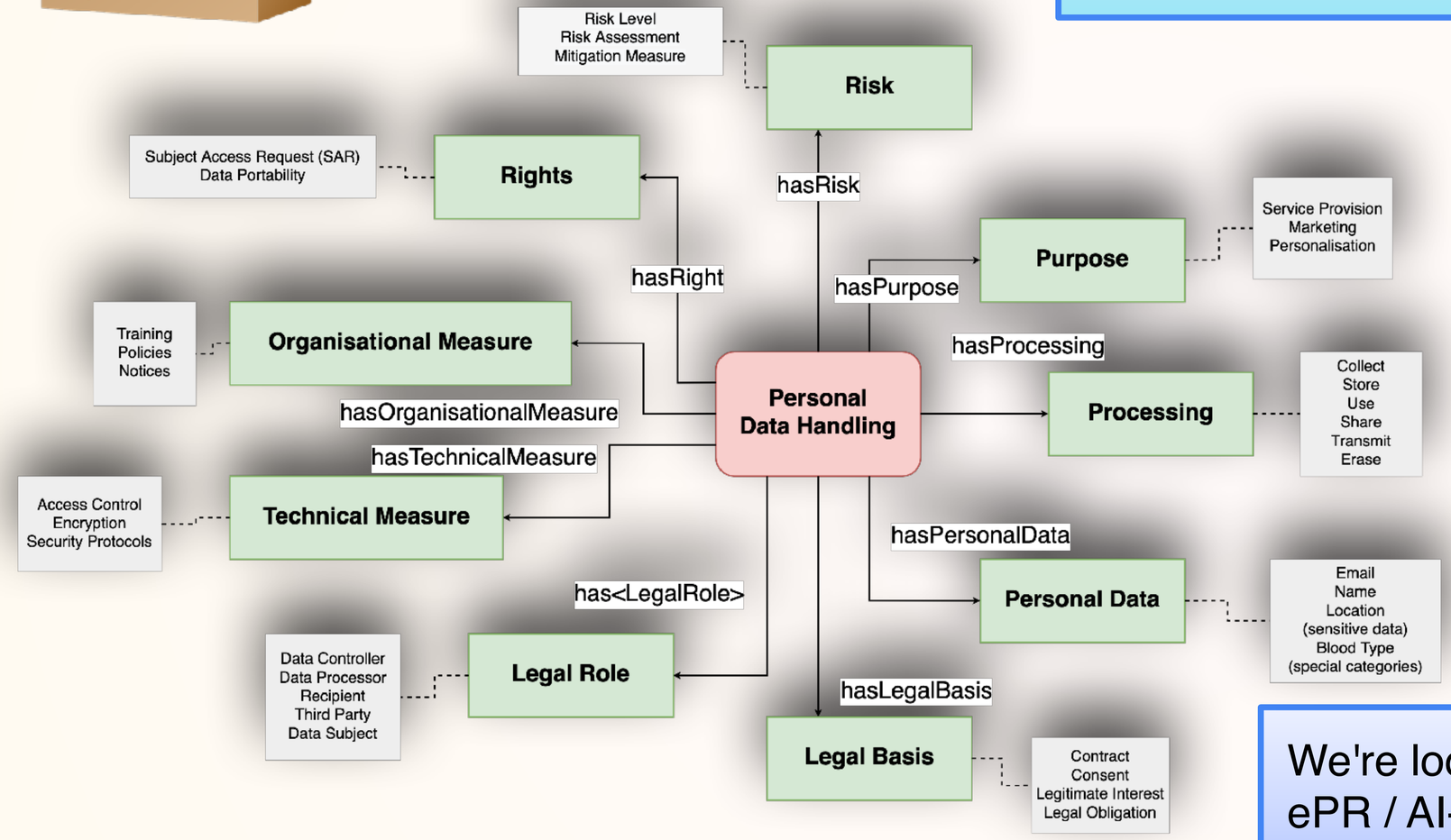




Reuses concepts from *Data Privacy Vocabulary (DPV)* to model personal data, purposes, consent, etc.

smashHitCore is available at https://gitlab.atb-bremen.de/smashhit/semantic-model

The Data Privacy Vocabulary (DPV) reflects ~5 years of efforts in creating an open resource providing concepts related to personal data processing, privacy, data protection, and GDPR



DPV's taxonomies provide semantic interoperability, which enables new, innovative, smart, and automated solutions

Demonstrated usefulness for important use-cases, e.g. ROPA, consent, compliance checking

We're looking to the future! DGA / ePR / AI-Act / Data Spaces

DPV Specification
https://w3id.org/dpv
Github
https://github.com/w3c/dpv/
Joining DPVCG
https://www.w3.org/community/dpvcg/

interested? questions? contact at:

twitter@coolharsh55

pandith@tcd.ie

dpv@harshp.com

public-dpvcg@w3.org

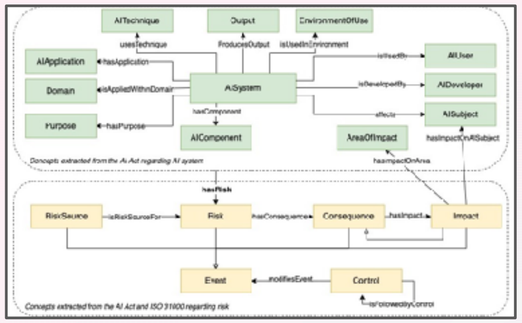
Data Privacy Vocabulary (DPV)

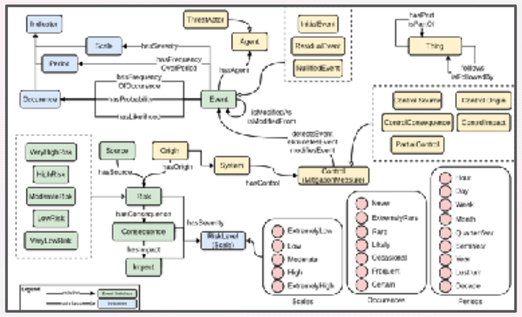
version 0.8

Draft Community Group Report 26 August 2022

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

We invite comments and feedbacks for publishing DPV v1 - a significant milestone. See DPV spec for more details. Comment period: up to OCT-15





AI Risk Management based on AI Act
Session 2.3
WED 12:45
Lohengrin (3)

DPV for DPIA & Risk Management
Session 7.3
THU 15:40
Lohengrin (3)

more @ SEMANTiCS