

Scalable Policy-aware Linked Data arChitecture for prlvacy, trAnsparency and complIance (SPECIAL)

Sabrina Kirrane, WU

20 February 2019

National Institute of Informatics Tokyo

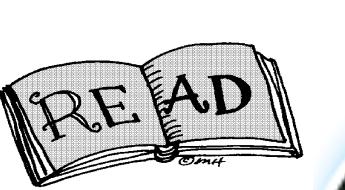


European
Commission

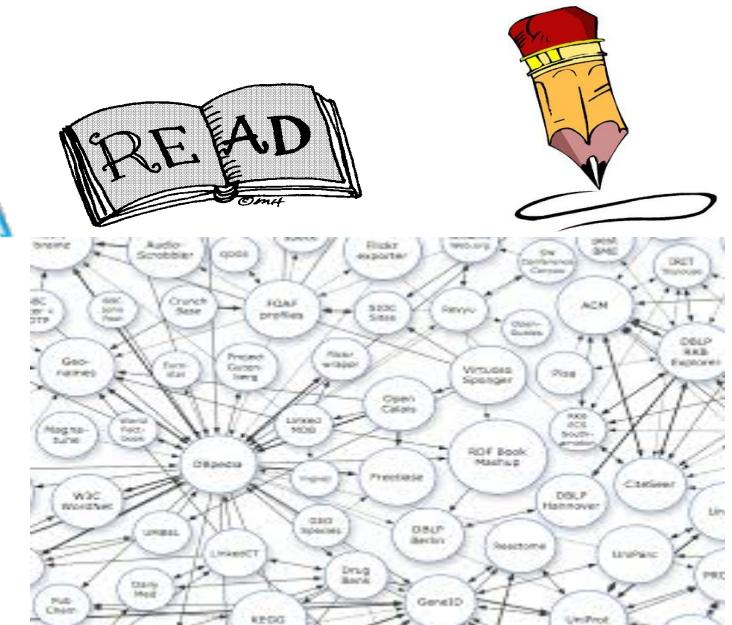
Horizon 2020
European Union funding
for Research & Innovation



Access Control for Linked Data



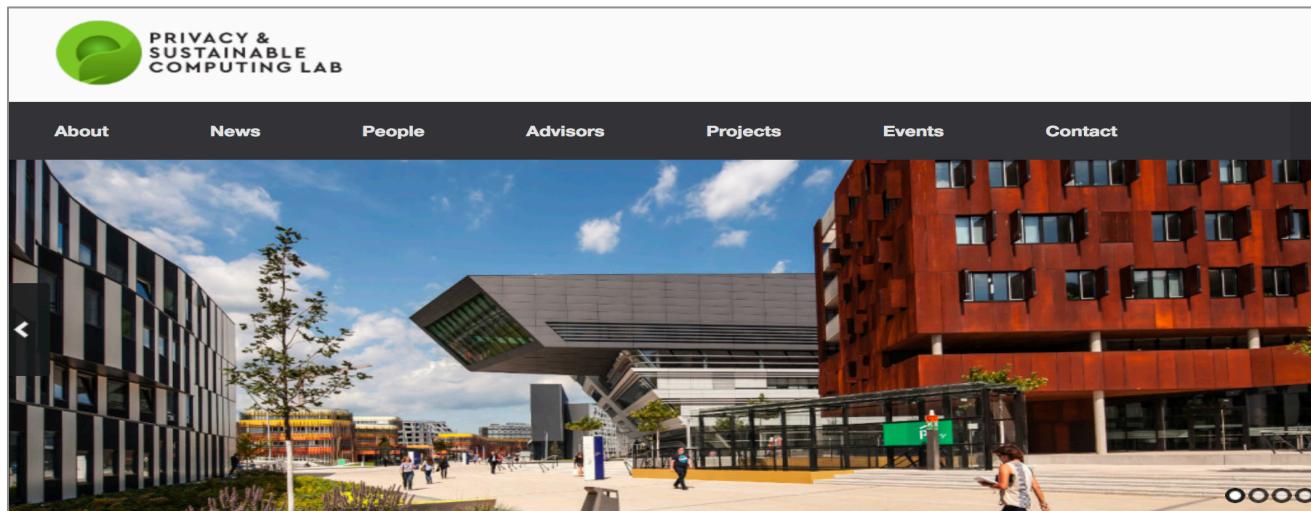
1.1



Privacy & Sustainable Computing

A multidisciplinary perspective...

- Privacy & Sustainable Computing Lab
- Launched September 2016: launch event with various important stakeholders: technologists, standardization, activists...
- Mission: Developing sustainable and privacy-preserving computer systems by bringing together computer science & human-centric behavioral science



Privacy & Sustainable Computing Lab

The increasing ubiquity of IT and data in corporate infrastructures and innovation endeavours, as well as the rising pervasiveness of computing in our daily lives, leads to the question how information systems can be built in a private, secure, ethical and value-driven manner. Current IT systems often fail to deliver on these requirements, commonly characterised as "soft" requirements, while aspects of technical functionality dominate.

The Privacy & Sustainable Computing Lab aims at setting new standards in research, education and practice to address ethical issues in computing.

<http://www.privacylab.at/>



Dr. Ben Wagner
(Director)



Dr. Sabrina Kirrane
(Mgt Board)



Prof. Sarah Spiekermann (Mg
Board)



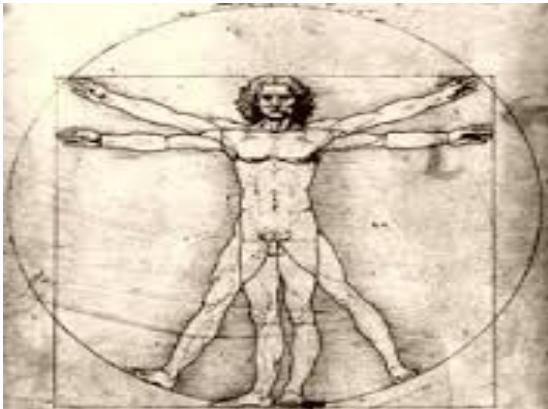
Prof. Axel
Polleres
(Mgt Board)



Privacy & Sustainable Computing

A multidisciplinary perspective...

Online Privacy



Humanities

Net Neutrality

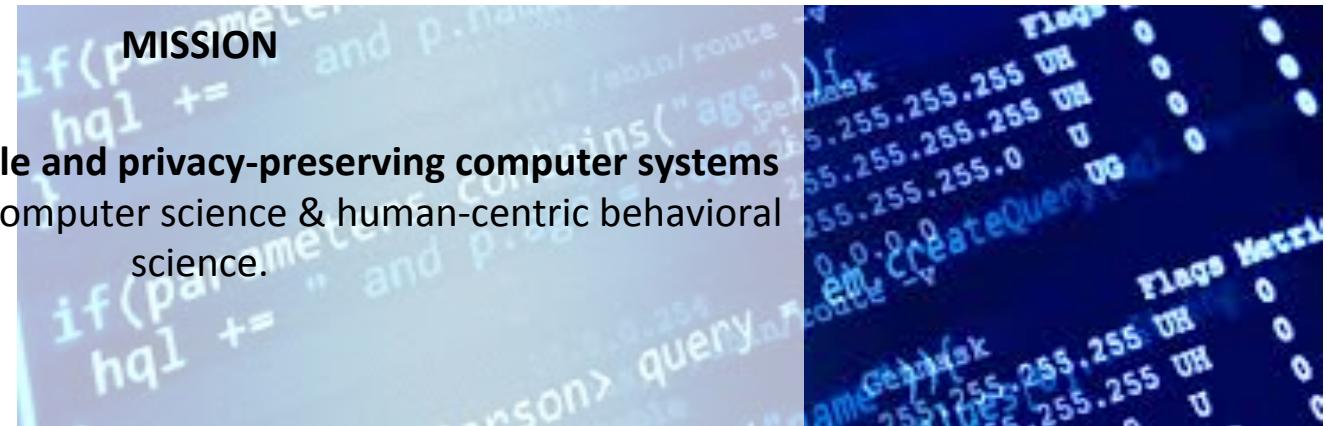
Fragmentation

Open Data

Open Standards

Public Policy and Legislation

Ethical Design

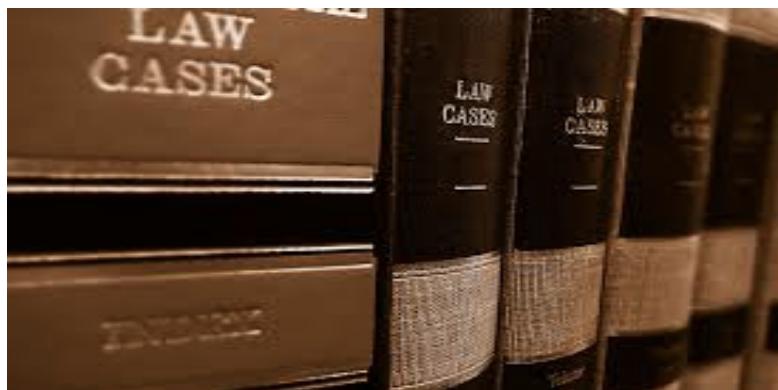


Computer Science

Big Data Analytics

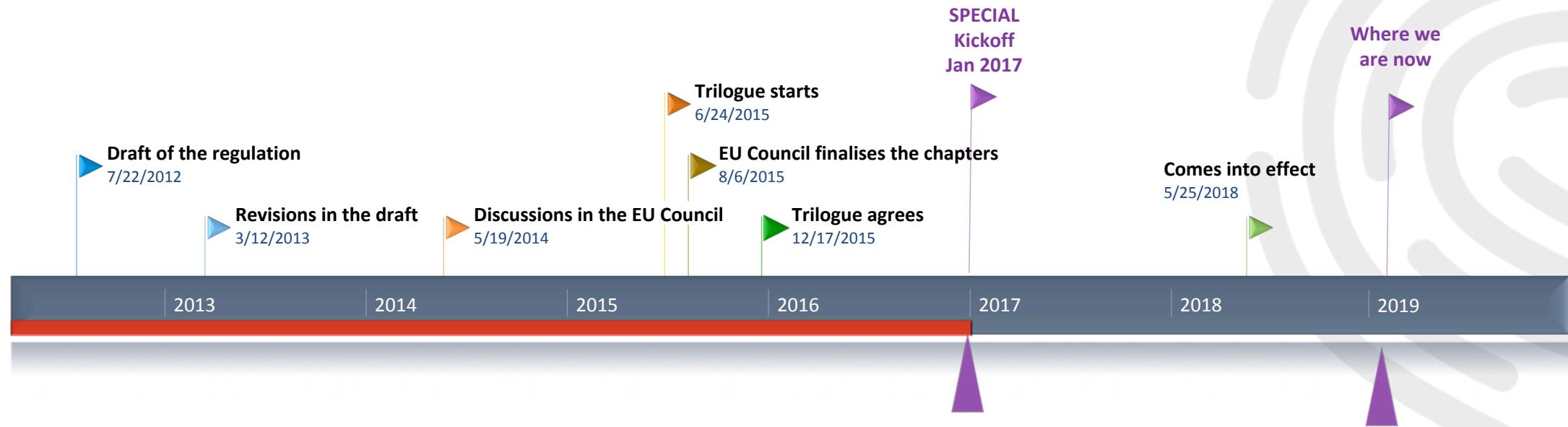
Artificial Intelligence

Data Science



Legal

SPECIAL Aims



Data subjects who would like to declare, monitor and optionally revoke their (often not explicit) preferences on data sharing

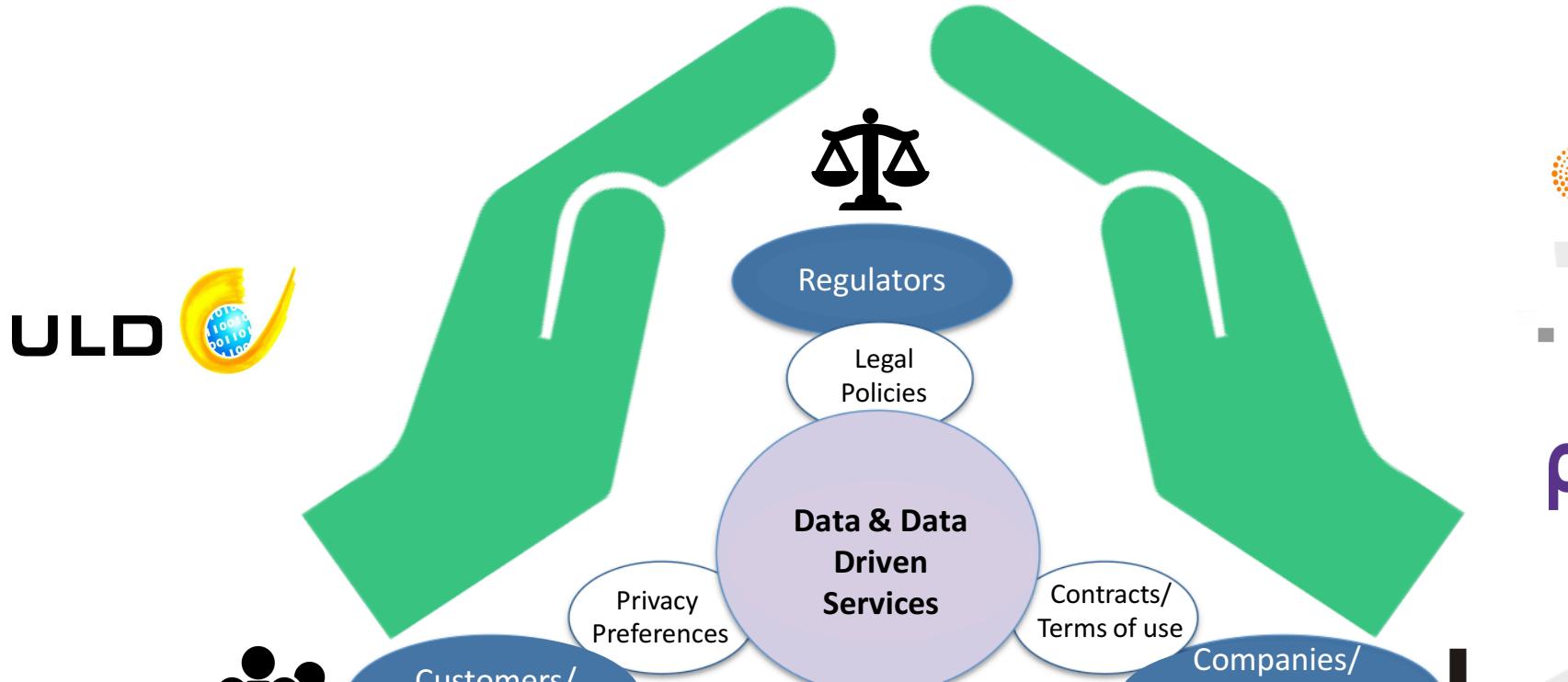


Regulators who can leverage technical means to check compliance with the GDPR



Companies whose business models rely on personal data and for which the GDPR is both a challenge and an opportunity

SPECIAL Aims



SPECIAL Objectives

- Policy management framework
 - ❖ Gives **users control** of their personal data
 - ❖ Represents **usage policies, legislative requirements** and **business policies** in a machine readable format
- Transparency and compliance framework
 - ❖ Provides information on how data is **processed** and with whom it is **shared**
 - ❖ Allows companies to verify that processing is in line with data subject **usage policies** and **legal requirements**
 - ❖ Allows data subjects to take **corrective action**
- Scalable policy-aware Linked Data architecture
 - ❖ Build on top of the Big Data Europe (BDE) platform **scalability and elasticity mechanisms**
 - ❖ Extended BDE with **robust policy, transparency** and **compliance protocols**
 - ❖ Enable personal data **value chains**
- Pilot implementation and evaluation
 - ❖ The architecture will be validated in the context of personal data sharing use cases for the **telecoms** and **financial services** sectors
- Collaboration, Dissemination & Standardisation
 - ❖ Create real-world impact in the form of a **sustainable solution** that we disseminate actively

SPECIAL Use Cases



Events at the Belgian Coast at
your fingertips

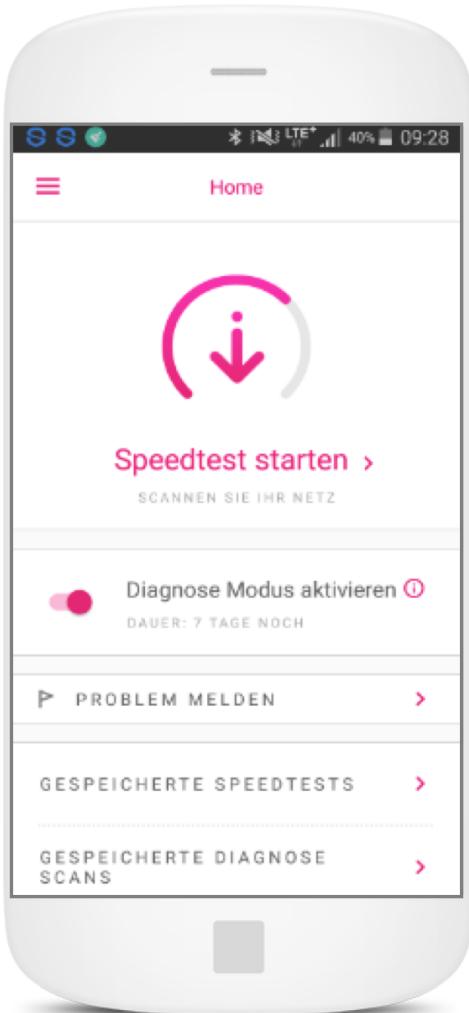
Sign up for free for intelligent tourist event
recommendations tailored to you.

Login

freddy.demeersman@proximus.com

LOGIN

proximus

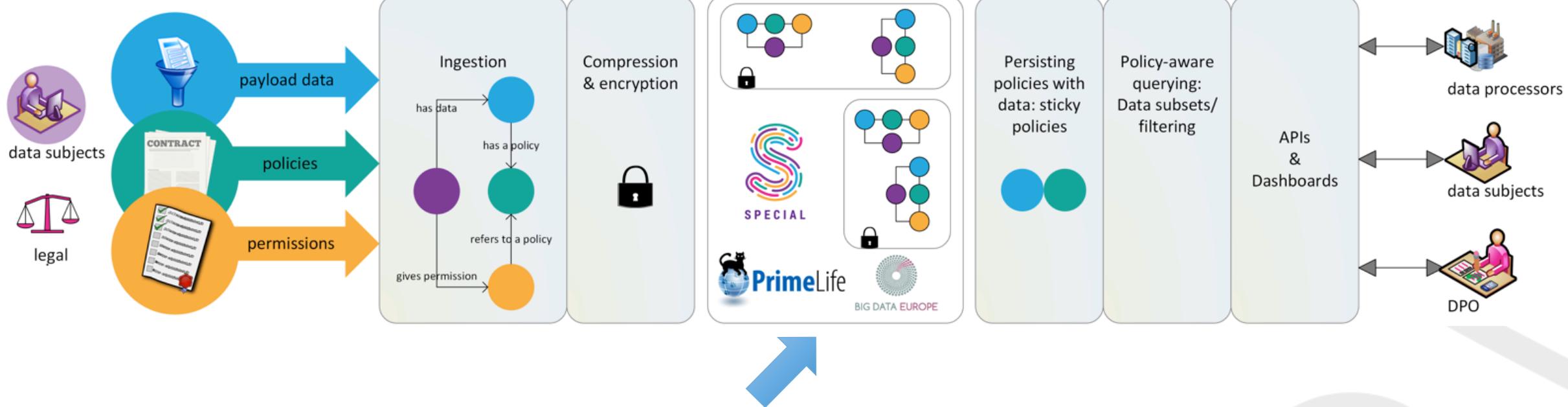


Deutsche
Telekom



SPECIAL Technical Foundations

Big Data and Privacy Foundations

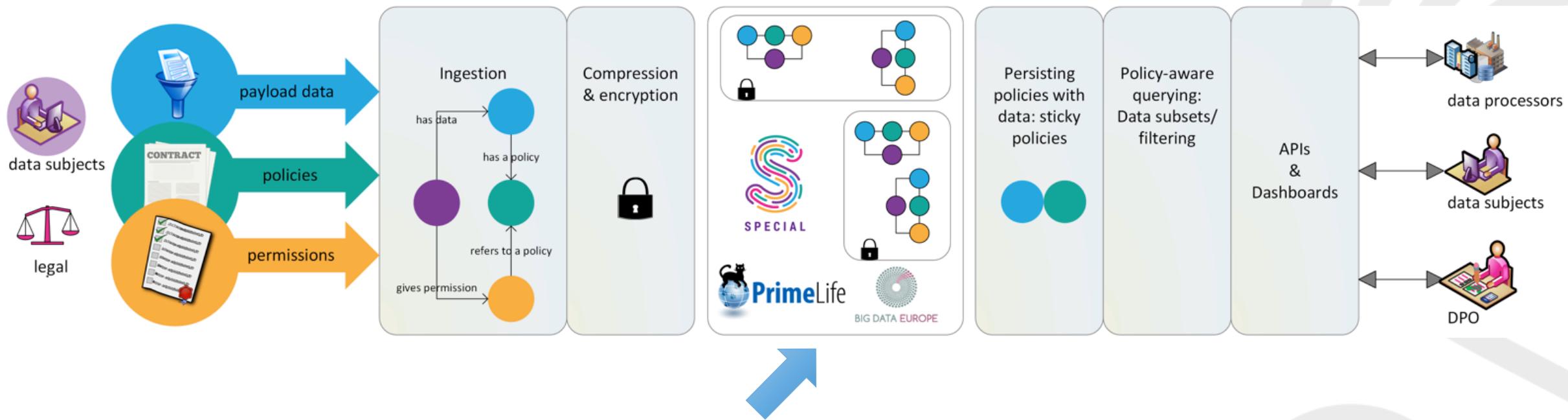


SPECIAL leverages past infrastructure and lessons learned

- ❖ **Big Data Europe** scalability and elasticity
- ❖ **PrimeLife** policy languages, access control policies, release policies and data handling policies
- ❖ The **Platform for Privacy Preferences Project (P3P)** and the **Open Digital Rights Language (ODRL)** vocabularies

SPECIAL Technical Foundations

Linked Data Foundations

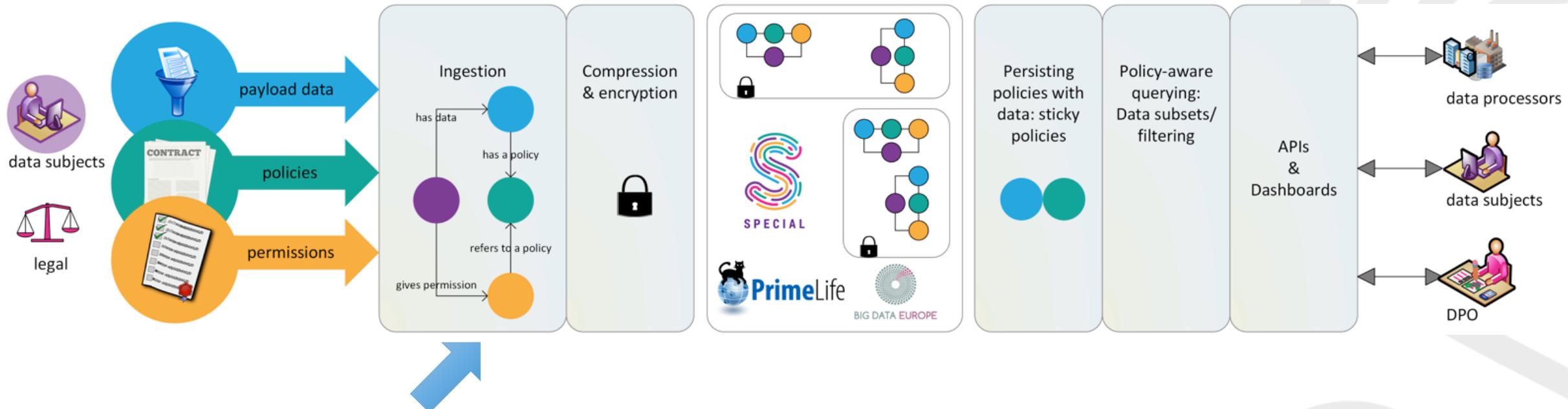


SPECIAL uses the **Linked Data paradigm**

- ❖ All data items are identified by **Internationalised Resource Identifiers (IRI's)**
- ❖ By using HyperText Transfer Protocol (HTTP) IRI's **everything is potentially linkable**
- ❖ IRI's allow SPECIAL to associate **usage constraints** with personal data at different levels of granularity

SPECIAL Technical Foundations

Ingestion

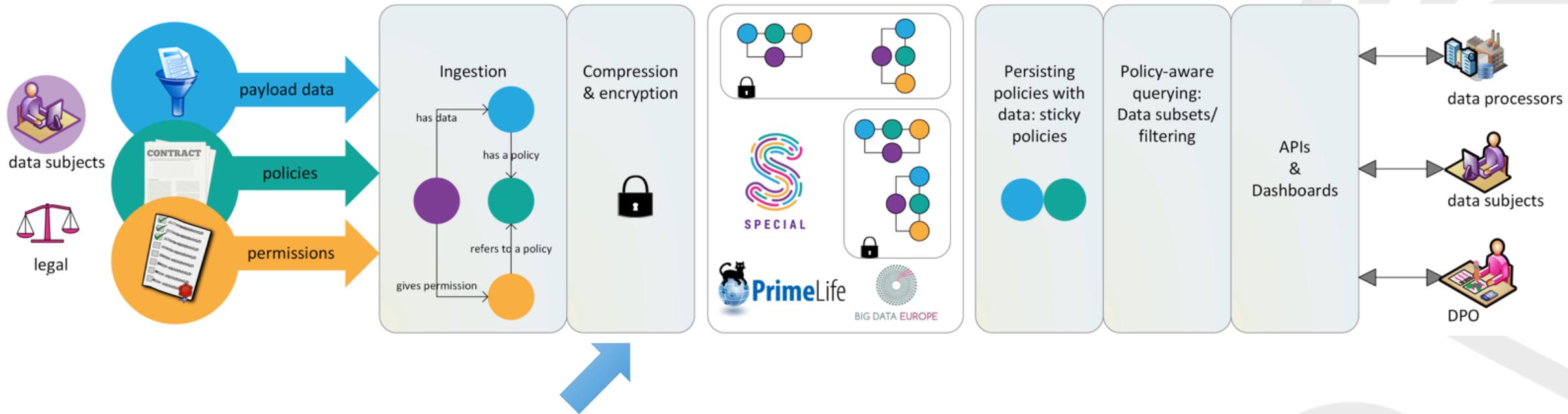


Handle a broad **variety of sources and formats**

- ❖ Integration with Line of Business applications via **transformation middleware**
- ❖ Understanding the personal data that is stored, how it is used, and what constraints are associated with the data needs to be captured in a **personal data processing inventory**
- ❖ **Policy Language** is tightly coupled to the legal process of enquiry (data, processing, purpose, storage and recipients)
- ❖ Allows for the development of an **Intelligent Data Lake**

SPECIAL Technical Foundations

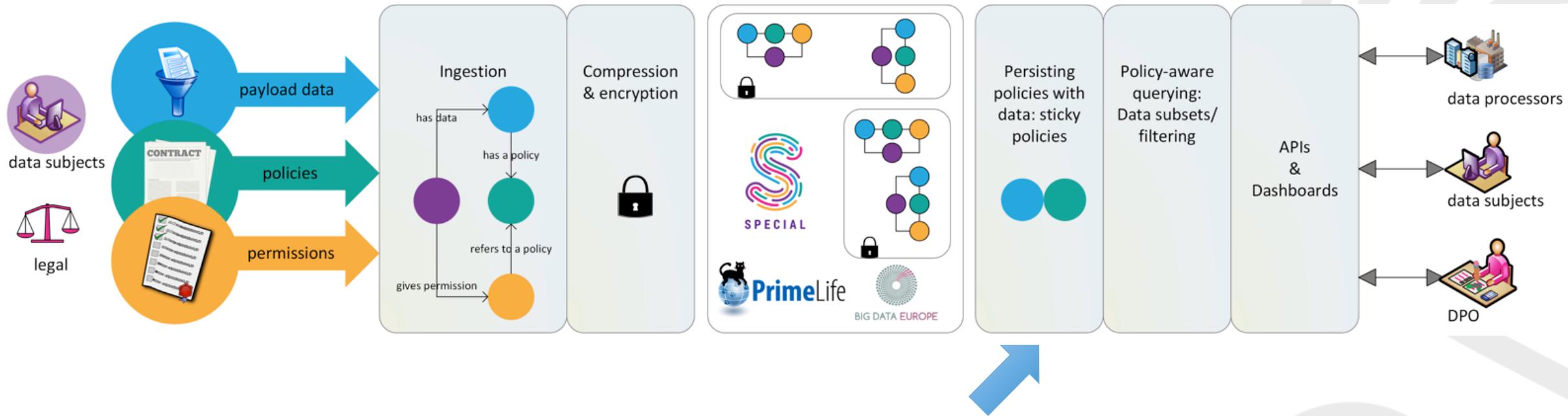
Compression & Encryption



When sharing data or querying results information is **securely stored** and **securely exchanged**

- ❖ Enable efficient **queryable encryption** based on **compressed RDF data**
- ❖ Encryption used for data and policy **integrity**

SPECIAL Technical Foundations Sticky Policies

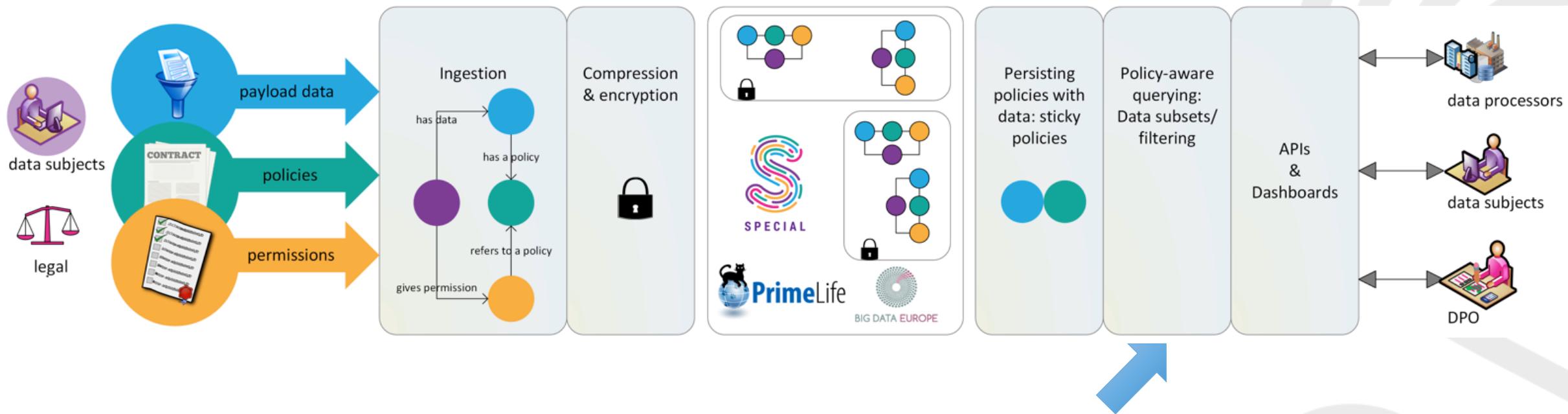


Gluing policy information to the payload data, even across company borders, is called “**sticky policies**”

- ❖ Policy constrained personal data sharing
- ❖ Legal guarantees
- ❖ Integrity and non-repudiation

SPECIAL Technical Foundations

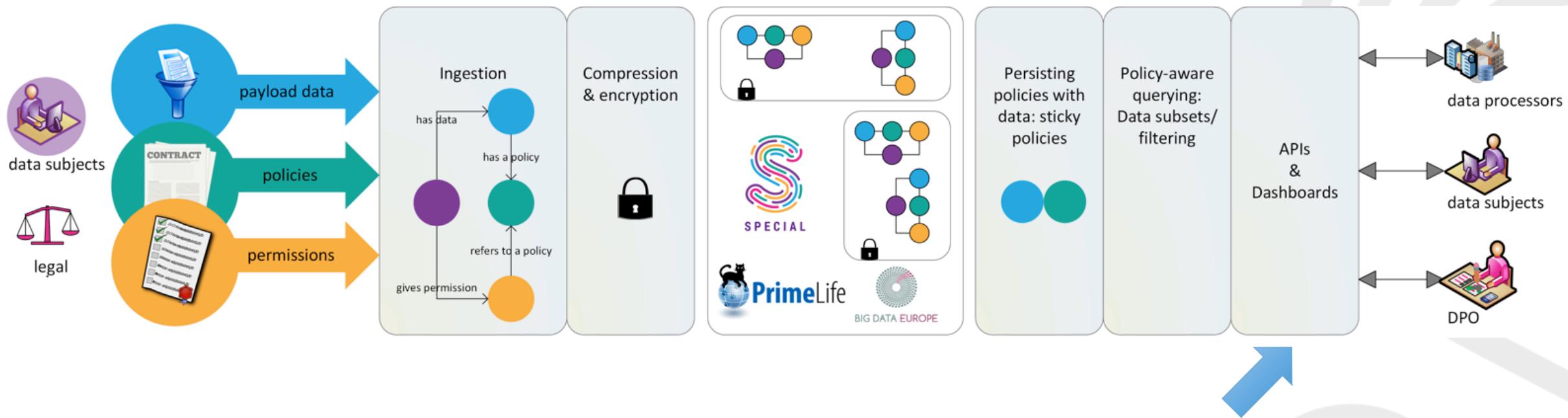
Policy aware querying



Categorise and subdivide data based on fine-grained **usage-policies** and **sensitivity categories/levels**

- ❖ Tackles **consent challenges** via layering, context, transparency and control
- ❖ Retrieve **policies based on data**
- ❖ Policy aware **aggregation** and **anonymisation** techniques

SPECIAL Technical Foundations APIs & Dashboards

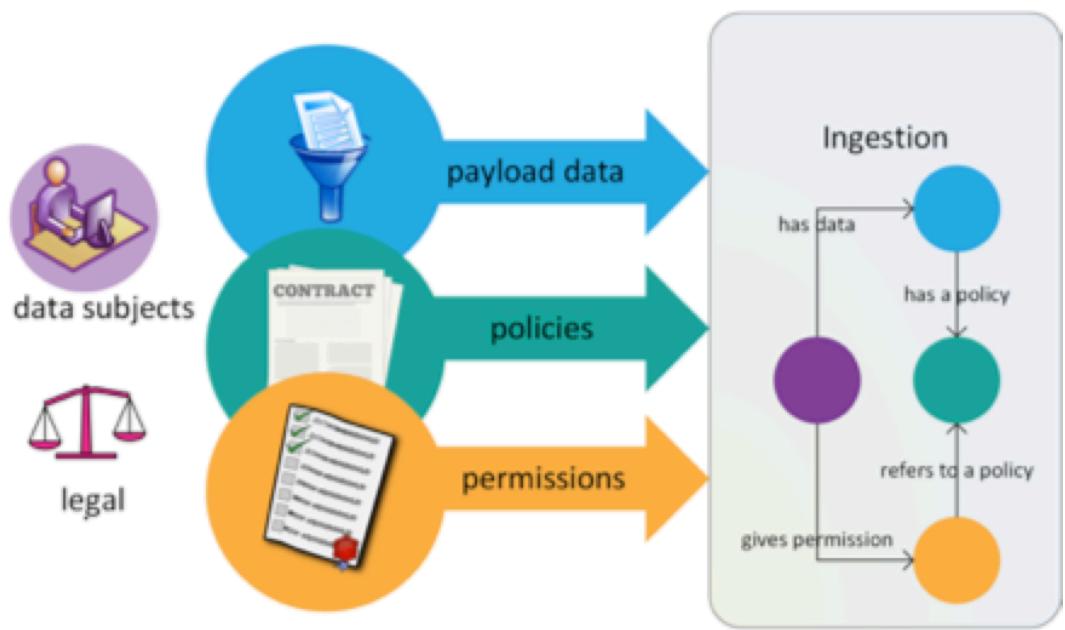


Investigate novel ways of obtaining **consent** and providing **transparency**

- ❖ Consent: display **highly relevant information** to the user based on context
- ❖ Interactive Dashboard: **effective way to represent** data, usage constraints, data processing and data sharing

SPECIAL Technical Foundations

Outline for the rest of the talk



- ❖ Analysing and Modelling the GDPR
- ❖ The SPECIAL usage policy language, vocabularies and compliance checking
- ❖ The SPECIAL transparency and compliance platform
- ❖ SPECIAL Standardisation Activities
- ❖ SPECIAL resources

Analysing and Modelling the GDPR



Analysing & Modelling the GDPR

4.5.2016

EN

Official Journal of the European Union

L 119/1

I

(Legislative acts)

REGULATIONS

REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL
of 27 April 2016
on the protection of natural persons with regard to the processing of personal data and on the free
movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)
(Text with EEA relevance)

THE EUROPEAN PARLIAMENT AND THE COUNCIL OF THE EUROPEAN UNION,

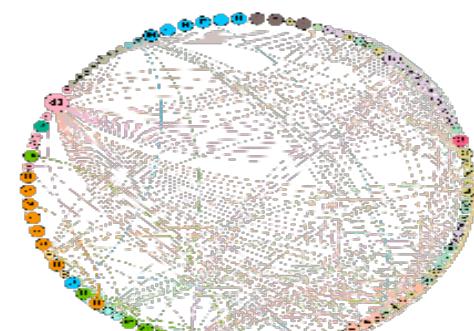
Having regard to the Treaty on the Functioning of the European Union, and in particular Article 16(1) thereof,

Having regard to the proposal from the European Commission,

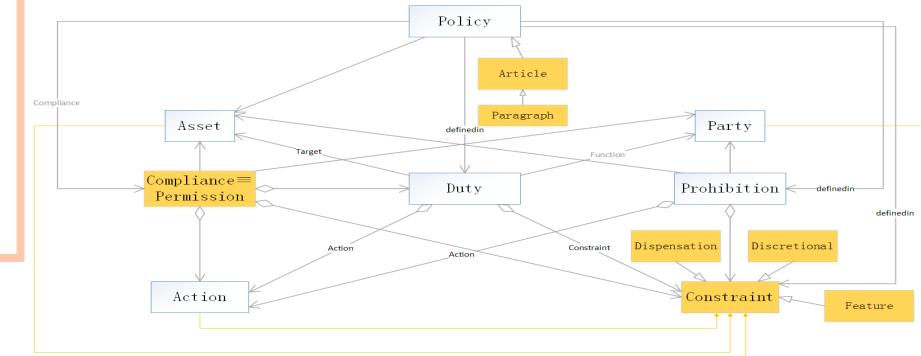
After transmission of the draft legislative act to the national parliaments,

Having regard to the opinion of the European Economic and Social Committee (¹),

Having regard to the opinion of the Committee of the Regions (²),



Relations make it complex for the developers



Actionable instructions

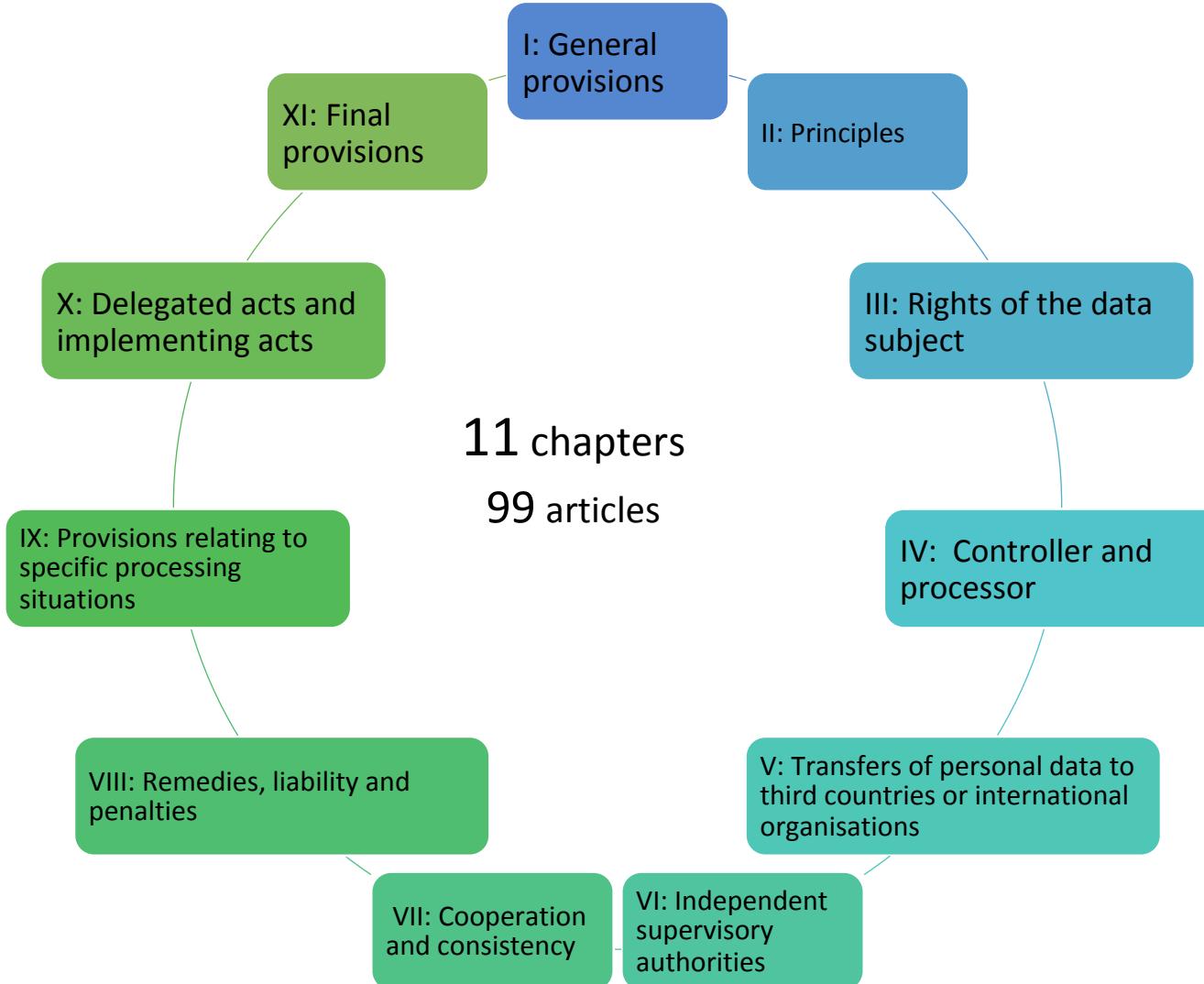


SERAMIS
Sensor-Enabled Real-World Awareness for Management Information Systems



Analysing & Modelling the GDPR

What aspects should we formalise?



Analysing & Modelling the GDPR

What aspects should we formalise?

Chapter III > Section 1 > Article 12

Transparent information, communication and modalities for the exercise of the rights of the data subject

Paragraph 1

The controller shall take appropriate measures to provide any information referred to in **Articles 13 and 14** and any communication under **Articles 15 to 22 and 34** relating to processing to the data subject in a concise, transparent, intelligible and easily accessible form, using clear and plain language

Chapter III > Section 4

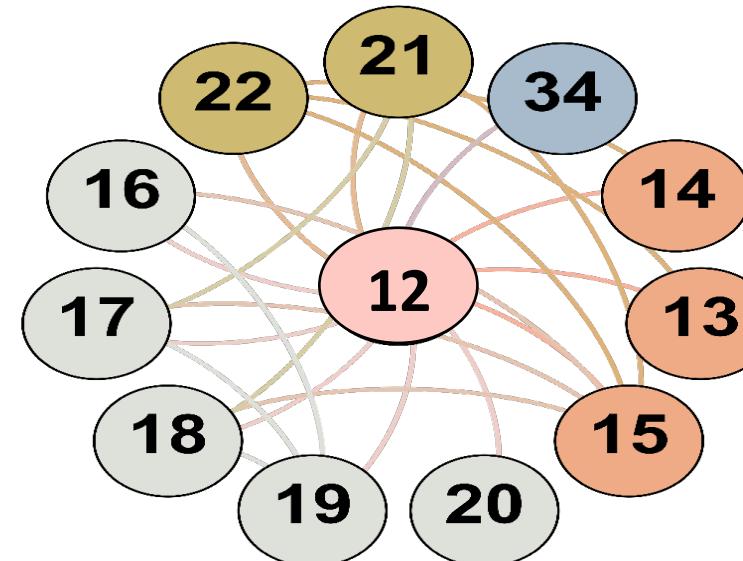
Right to object and automated individual decision-making

Chapter III > Section 3

Rectification and erasure

CHAPTER IV > Section 2

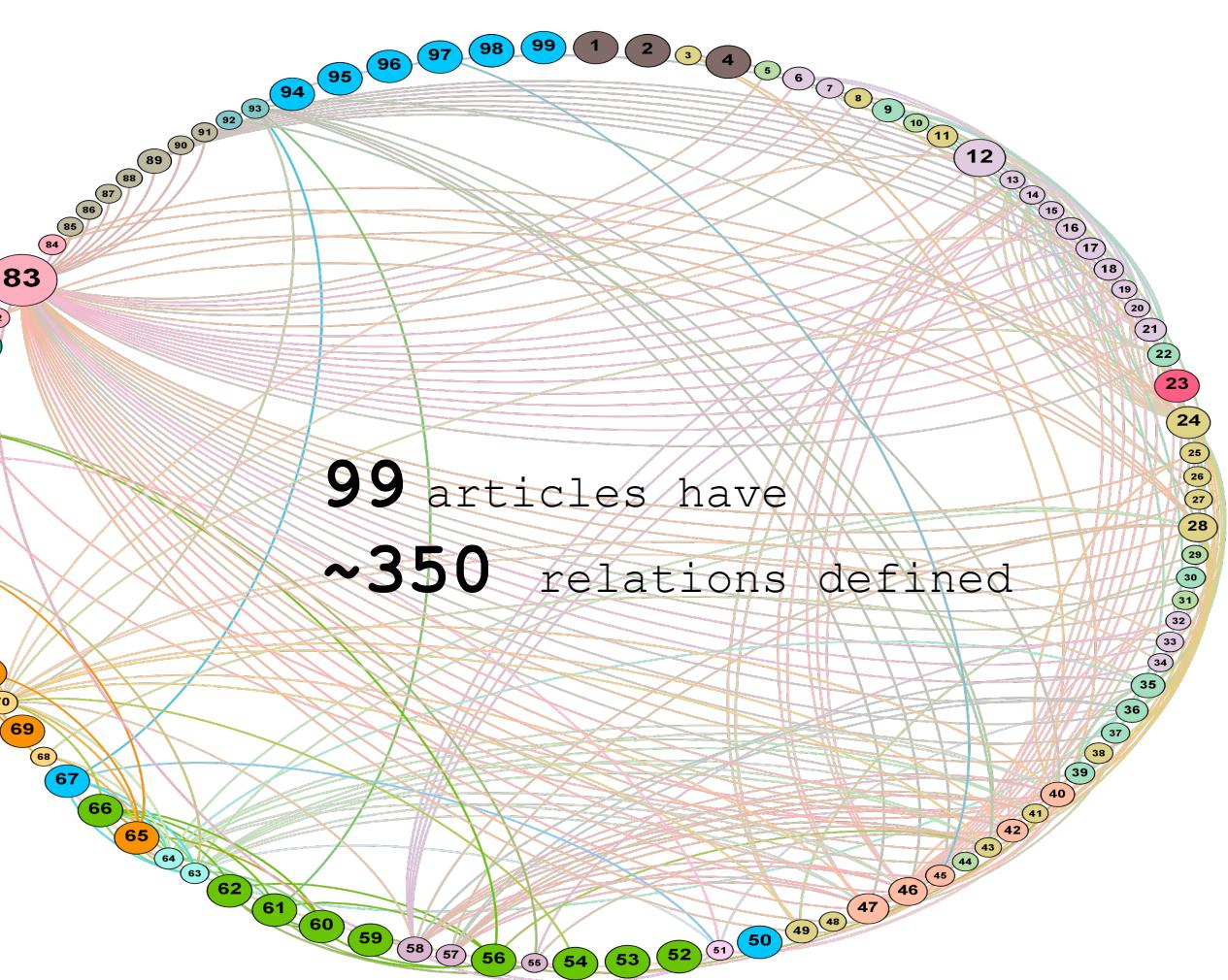
Security of personal data



Chapter III > Section 2
Information and access to personal data

Analysing & Modelling the GDPR

What aspects should we formalise?



Article 83
General conditions for imposing administrative fines

- Other things to be considered
 - Direction of the relationship
 - Article chains
 - Semantics of the relationship
- Model the consequences and fines in the case of non-compliance
- Not all of the 99 articles define obligations (e.g. objectives, definitions, GDPR's entry into force etc.)
- We omit articles defining obligations for other parties like the Supervisory Authorities and European Data Protection Board can be neglected.

Analysing & Modelling the GDPR

What formalism should we use?



COMMUNITY & BUSINESS GROUPS

[Home](#) / [ODRL Community Group](#) / [ODRL 2 Core Model...](#) / ODRL

ODRL Version 2.1 Core Model

Final Specification: 5 March 2015

This Version: <http://www.w3.org/community/odrl/>
Latest Version: <http://www.w3.org/community/odrl/>

Editors:

Renato Iannella, Semantic Identity, ri@semanticidentity.com
Susanne Guth, Vodafone, susanne.guth@vodafone.com
Daniel Paehler, University of Koblenz, tulkas@uni-koblenz.de
Andreas Kasten, University of Koblenz, andreas.kasten@uni-koblenz.de



Permissions & Obligations Expression Working Group Charter

The Web has provided the community with standardized mechanisms for numerous content-management services: publishing, distribution, consumption, describing, and sharing. However, the key area of permissions, obligations and licensing has not been addressed in Web standards to date. Content licenses, rights statements, permissions and obligations express the terms of usage for content. With a standard vocabulary, content owners can express terms and processing systems can determine what permissions and other terms are associated with a given resource or collection of resources.

A permissions and obligations expression system should provide a flexible and interoperable information model that supports transparent and innovative (re)use of digital content across all sectors and communities. The underlying model should support the business models of open, educational, government, and commercial communities through profiles that align with their specific requirements whilst retaining a common semantic layer for wider interoperability. The system should not, however, be the basis of legal compliance or enforcement mechanisms.

A permissions and obligations expression language is composed of detailed terms that are both machine-processable and expressible in a form for human-consumption. Allowable actions, constraints, and requirements are expressed at a level enabling complex and business-specific expressions to be created from a vocabulary with specific semantics. This accommodates a broad range of situations and addresses a different business/user need than systems such as [Creative Commons](#) that provide generic sharing licenses.

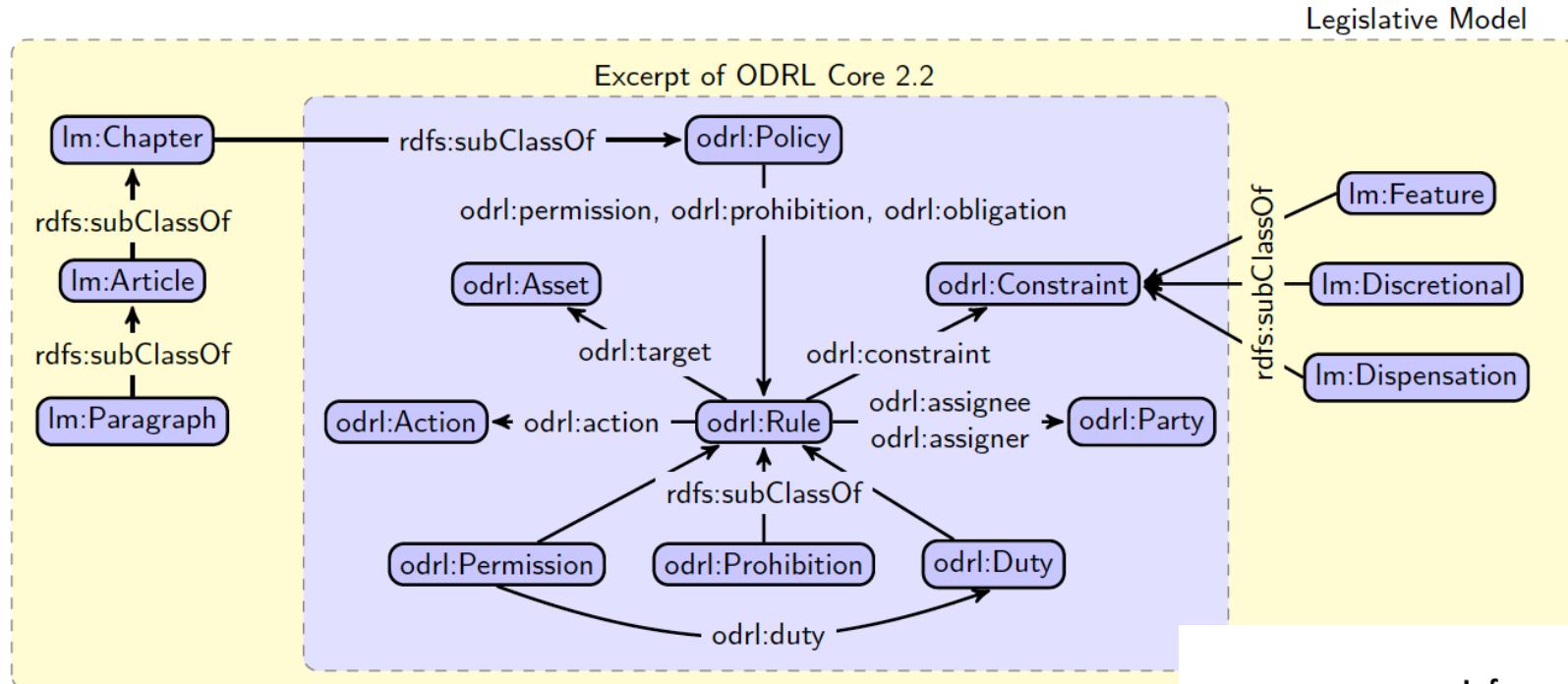
The **mission** of the [Permissions & Obligations Expression Working Group](#) is to define a semantic data model for expressing permissions and obligations statements for digital content, and to define the technical elements to make it deployable across browsers and content systems.

ODRL designed to define rules for the publishing, distribution, and consumption of digital media.

Scope
Deliverables
Dependencies and Liaisons
Participation
Communication
Decision Policy
Patent Policy
About this Charter

Analysing & Modelling the GDPR

What formalism should we use?



Article 13

Information to be provided where personal data
are collected from the data subject

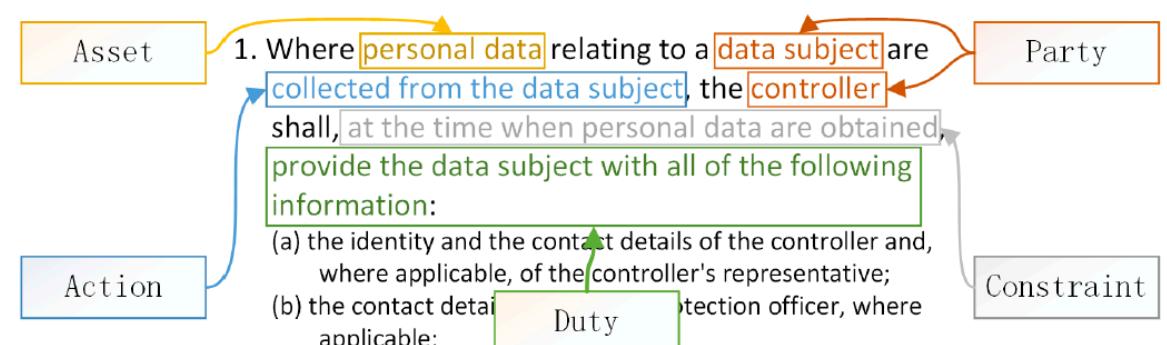


Fig. 3. Breaking down Article 13.1 of the GDPR according to the ODRL model

Analysing & Modelling the GDPR

A regulatory ODRL profile

Listing 1: Snippet of the GDPR instance based on the duty from Article 13.1

```
1 gdpr:P13_1 rdf:type lm:Paragraph .  
2 gdpr:P13_1 odrl:duty gdpr:ProvideInfo .  
3 gdpr:ProvideInfo rdf:type odrl:Duty .  
4 gdpr:ProvideInfo odrl:action gdpr:DirectCollection .  
5 gdpr:ProvideInfo lm:dispensation gdpr:DataSubjecthasInfo .  
6 gdpr:ProvideInfo lm:feature gdpr:Transparency .  
7 gdpr:ProvideInfo lm:feature gdpr:Conciseness .  
8 gdpr:ProvideInfo lm:discretional gdpr:Icons .
```

Analysing & Modelling the GDPR

A regulatory ODRL profile

Listing 2: Snippet of the GDPR instance from Listing 1 with the added question

```
1 gdpr:ProvideInfo rdf:type odrl:Duty .  
2 gdpr:ProvideInfo odrl:action gdpr:DirectCollection .  
3 gdpr:ProvideInfo lm:dispensation gdpr:DataSubjecthasInfo .  
4 gdpr:ProvideInfo lm:feature gdpr:Transparency .  
5 gdpr:ProvideInfo lm:feature gdpr:Conciseness .  
6 gdpr:ProvideInfo lm:discretionary gdpr:Icons .  
7 gdpr:ProvideInfo lm:hasquestion "Does your organisation ensure that the  
8 required information is provided to the data subject?" .
```

Listing 3: Illustration of an *Action* with added question

```
1 gdpr:DirectCollection rdf:type odrl:Action .  
2 gdpr:DirectCollection lm:hasquestion "Does your organisation collect  
3 personal information directly from the data subjects?" .
```

Listing 4: Illustration of a *Feature* related to the duty from Listing 2

```
1 gdpr:Transparency rdf:type lm:Feature .  
2 gdpr:Transparency lm:hasquestion "Does your organisation ensure  
3 transparency for the provided information?" .
```

Analysing & Modelling the GDPR Compliance impact assessment

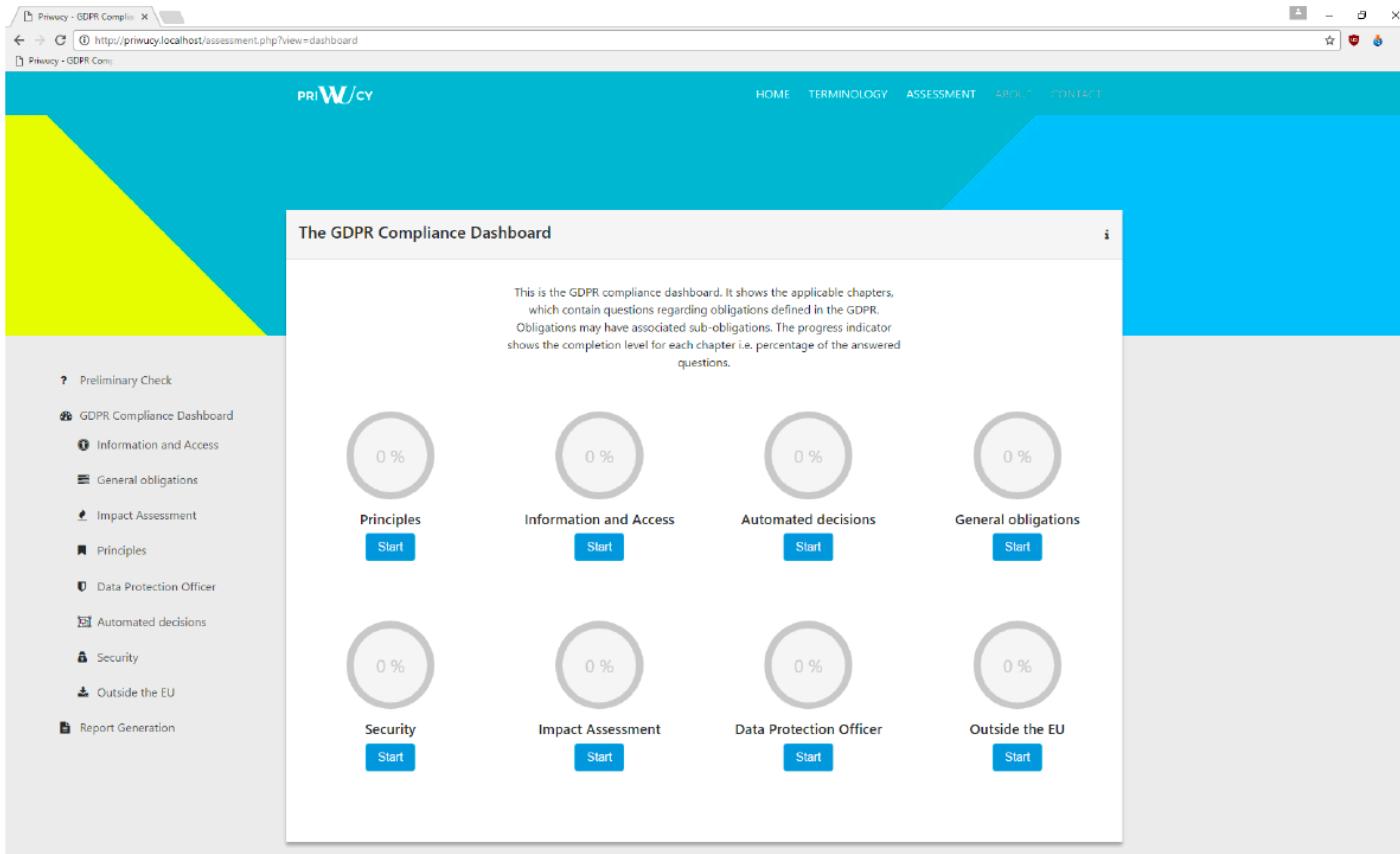


Fig. 7. Dashboard based on the GDPR chapters for the main assessment

Analysing & Modelling the GDPR

Machine readable policies

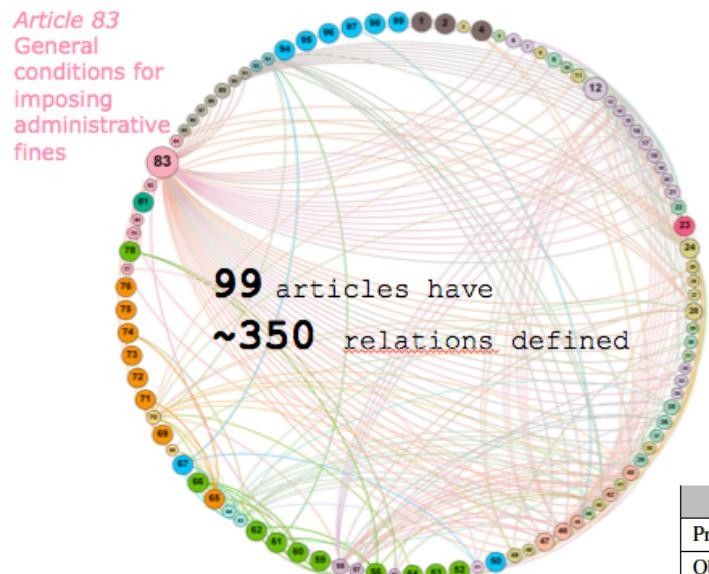


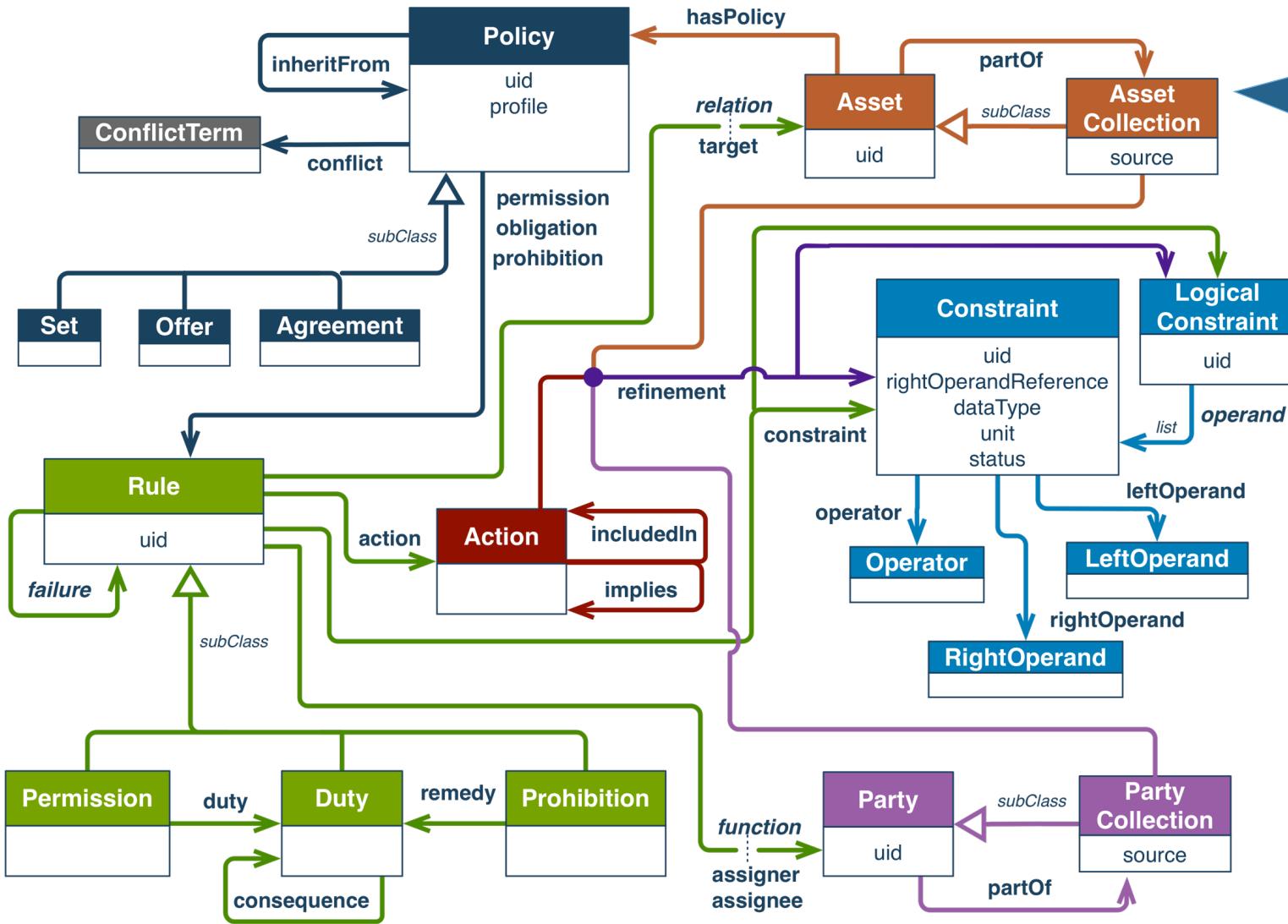
Table 1.1: GDPR Annotation Dictionary		
Type	Annotation	Description
Prohibition	P	you must not (i.e. equivalent to negative obligation)
Obligations	O	you must
Dispensation	D	exemption from the rule (dispensation condition for processing in a legal sense)
Constraints	+C -C	a limitation or restriction (e.g. its allowed if) a limitation or restriction (e.g. its allowed if you don't)
Definitions	Def	explains the meaning of a certain term or defines how an obligation or a constraint must be understood
References	eRef [] tRef []	an article contains an explicit reference (e.g. eRef[Art. 89 (1)]) an article contains a reference related to a certain term (e.g. tRef[consent])
Dispositions	Disp	an example/best practice/suggestion
Opening Clause	OC	indicates a need to consult other legislation (National or European)

In SPECIAL we adopt a systematic approach

- Systematic analysis of the text of the GDPR
- *D2.2 & D2.6 Formal representation of the legislation* detailed the GDPR:
 - ❖ Rule structure
 - ❖ Explicit References
 - ❖ Subjective terms (i.e. single words or parts of a sentence that can be interpreted in various ways)
 - ❖ Implicit knowledge about the law (e.g. the scope of Union Law)
 - ❖ References to other pieces of legislation

Analysing & Modelling the GDPR

Lessons learned



ODRL was heavily guided by licensing use cases.....

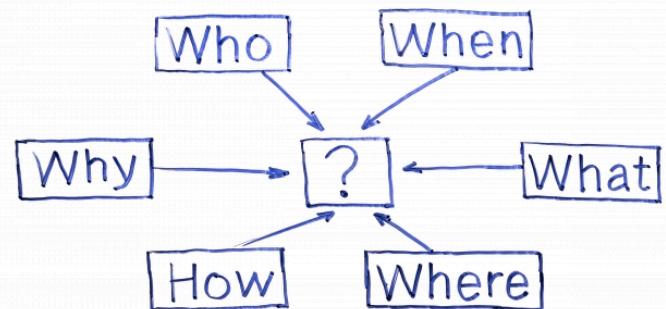
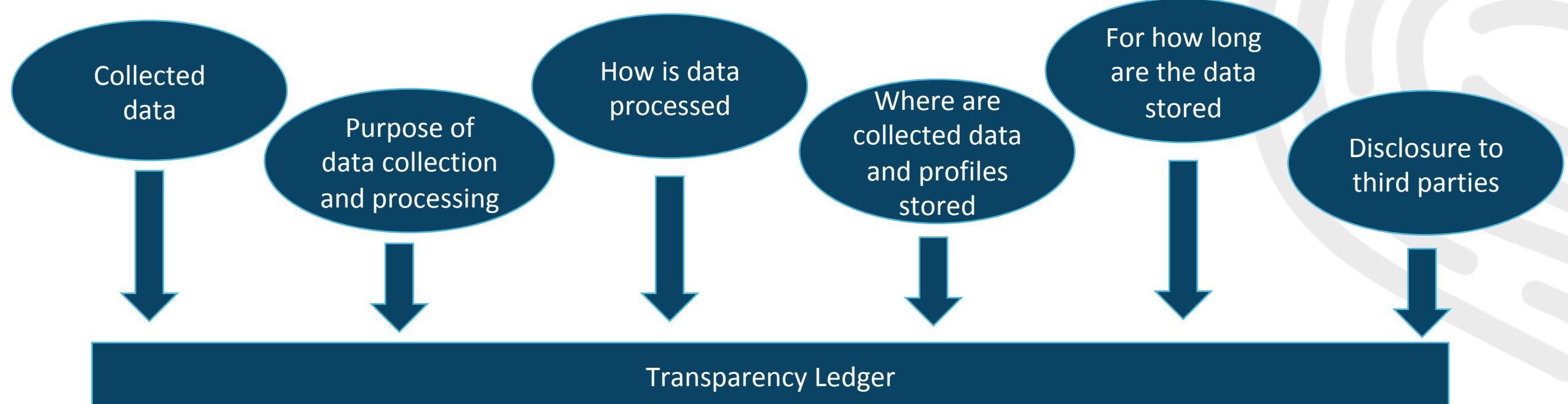
Therefore when modelling regulatory requirements we need to decide how closely we stick to the original model

The SPECIAL usage policy language, vocabularies and compliance checking



Usage policy language

The minimal core model



ORACLE® Symantec

vmware® amazon

Microsoft®

SAP® salesforce.com® Adobe®

Usage policy language

Syntax and expressivity

- Usage policy language, which can be used to express both the data subjects' **consent**, data controllers usage requests, fragments of the **GDPR**, and business policies
- The foundation of the policy language was the **Minimal Core Model (MCM)**
- We propose a new policy language that extensively **re-uses standards** based privacy-related vocabularies
- We are able to leverage existing **Web Ontology Language (OWL)** based reasoners out of the box

Figure 1.1: SPECIAL's Usage Policy Language Grammar

```
UsagePolicy :='ObjectUnionOf' '(' BasicUsagePolicy BasicUsagePolicy { BasicUsagePolicy } ')'
| BasicUsagePolicy

BasicUsagePolicy :='ObjectIntersectionOf' '(' Data Purpose Processing Recipients Storage ')'

Data :='ObjectSomeValueFrom' '(' 'spl:hasData' DataExpression ')'

Purpose :='ObjectSomeValueFrom' '(' 'spl:hasPurpose' PurposeExpression ')'

Processing :='ObjectSomeValueFrom' '(' 'spl:hasProcessing' ProcessingExpression ')'

Recipients :='ObjectSomeValueFrom' '(' 'spl:hasRecipient' RecipientExpression ')'

Storage :='ObjectSomeValueFrom' '(' 'spl:hasStorage' StorageExpression ')'

DataExpression :='spl:AnyData' | DataVocabExpression

PurposeExpression :='spl:AnyPurpose' | PurposeVocabExpression

ProcessingExpression :='spl:AnyProcessing' | ProcessingVocabExpression

RecipientsExpression :='spl:AnyRecipient' | 'spl:Null' | RecipientVocabExpression

StorageExpression :='spl:AnyStorage' | 'spl:Null' |
'ObjectIntersectionOf' '(' Location Duration ')'

Location :='ObjectSomeValueFrom' '(' 'spl:hasLocation' LocationExpression ')'

Duration :='ObjectSomeValueFrom' '(' 'spl:hasDuration' DurationExpression ')'
| 'DataSomeValueFrom' '(' 'spl:durationInDays' IntervalExpression ')'
```

Usage policy language

Syntax and expressivity

Listing 1.1. SPECIAL Namespace Prefixes

```
PREFIX spl: <http://www.specialprivacy.eu/langs/usage-policy#>
PREFIX splog: <http://www.specialprivacy.eu/langs/splog#>
PREFIX svd: <http://www.specialprivacy.eu/vocabs/duration#>
PREFIX svl: <http://www.specialprivacy.eu/vocabs/locations#>.
```

Listing 1.2. Structure of a Usage Control Policy

```
ObjectIntersectionOf(
    ObjectSomeValuesFrom( spl:hasData SomeDataCategory)
    ObjectSomeValuesFrom( spl:hasProcessing SomeProcessing)
    ObjectSomeValuesFrom( spl:hasPurpose SomePurpose)
    ObjectSomeValuesFrom( spl:hasStorage SomeStorage)
    ObjectSomeValuesFrom( spl:hasRecipient SomeRecipient))
```

Usage policy language

Syntax and expressivity

EXAMPLE 4: Example of an ellaborated policy

The following policy P ,

Heart rate and location data are collected and analysed to create a user profile for the purpose of issuing recommendations. Profiles are stored indefinitely in the EU by the data controller and released to third parties.

can be formalised as follows with a factorised general policy:

```
ObjectIntersectionOf(  
    ObjectSomeValueFrom( spl:hasData  
        ObjectUnionOf(  
            ex:HeartRate svd:Location ))  
    ObjectSomeValueFrom( spl:hasProcessing ex:Profiling )  
    ObjectSomeValueFrom( spl:hasPurpose ex:Recommendation )  
    ObjectSomeValueFrom( spl:hasStorage  
        ObjectIntersectionOf(  
            ObjectSomeValuesFrom( spl:hasLocation  
                ObjectIntersectionOf( svl:OurServers svl:EU ))  
            DataSomeValuesFrom(  
                spl:durationInDays  
                DatatypeRestriction( xsd:integer  
                    xsd:mininclusive "0"^^xsd:integer )  
            )  
        )  
    )  
    ObjectSomeValueFrom( spl:hasRecipient spl:AnyRecipient )  
)
```

In this example, the auxiliary vocabularies need to be extended with three new classes: the class `ex:ReartRate` (as a subclass of `svd:Health`), `ex:Profiling` (a subclass of `svpr:Analyze`) and `ex:Recommendation` (a subclass of `svpu:Marketing`).

Usage policy language SPECIAL resources

The SPECIAL Usage Policy Language

version 0.1

Unofficial Draft 06 April 2018



Editor:

Javier D. Fernández (Vienna University of Economics and Business)

Authors:

Piero Bonatti (Università di Napoli Federico II)

Sabrina Kirrane (Vienna University of Economics and

Iliana Mineva Petrova (Università di Napoli Federico I)

Luigi Sauro (Università di Napoli Federico II)

Eva Schleehahn (Unabhängiges Landeszentrum für Da

This document is licensed under a [Creative Commons Attribution 3.0 Li](#)

Abstract

This document specifies usage policy language of SPECIAL, both the data subjects' consent and the data usage policies by a computer, so as to automatically verify that the usage

The ontology defined in this document is publicly available

Vocabulary .../langs/usage-policy#

Bert Bos Last Updated: 17 April 2018

You can [download this ontology as an OWL file.](#)

The following is the formulation in functional syntax of the Usage Policy Language Ontology with identifier

<http://www.specialprivacy.eu/langs/usage-policy#>

The documentation can be found in [Policy Language V1 \(deliverable D2.1\)](#).

```
# NAMESPACE DEFINITIONS
Prefix(spl:=<http://www.specialprivacy.eu/langs/usage-policy#>)
Prefix(owl:=<http://www.w3.org/2002/07/owl#>)
Prefix(rdf:=<http://www.w3.org/1999/02/22-rdf-syntax-ns#>)
Prefix(xml:=<http://www.w3.org/XML/1998/namespace>)
Prefix(xsd:=<http://www.w3.org/2001/XMLSchema#>)
Prefix(rdfs:=<http://www.w3.org/2000/01/rdf-schema#>)

# ONTOLOGY IRI AND ITS VERSION

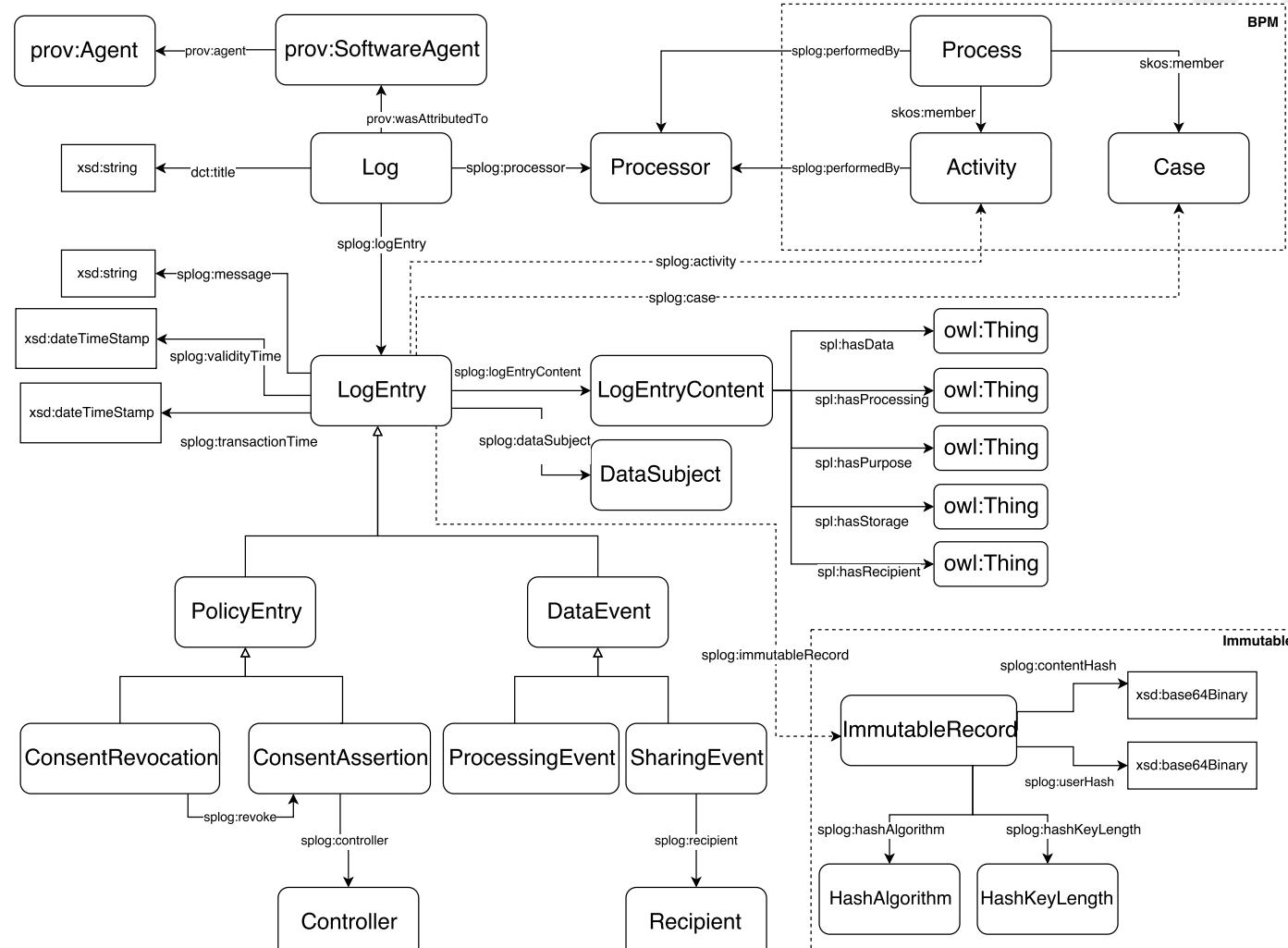
Ontology( <http://www.specialprivacy.eu/langs/usage-policy-ontology>
<http://www.specialprivacy.eu/langs/usage-policy-ontology/1.0>
```

- Detailed in [D2.1 Policy Language V1](#)
- Available for download via the SPECIAL website <https://www.specialprivacy.eu/langs/usage-policy>
- An unofficial *draft specification* has been published online <http://purl.org/specialprivacy/policylanguage>
- Feeds into the standardisation efforts conducted in the [W3C Data Privacy Vocabularies and Controls Community Group](#)

Provenance/event information

The model

- Development of a **log vocabulary** that reuses well-known vocabularies such as **PROV** for representing provenance metadata
- Demonstrate how provenance can be used to support transparency in data value chains



Provenance/event information

Syntax and expressivity

Listing 1.3. A new event for Sue's BeFit device

```
befit:entry3918 a solog:ProcessingEvent ;  
solog:dataSubject befit:Sue ;  
dct:description "Store location in our database in Europe"@en ;  
solog:transactionTime "2018-01-10T13:20:50Z"^^xsd:dateTimeStamp ;  
solog:validityTime "2018-01-10T13:20:00Z"^^xsd:dateTimeStamp ;  
solog:eventContent befit:content3918 ;  
solog:immutableRecord befit:iRec3918 .
```

Listing 1.4. The content of a new event for Sue's BeFit device

```
befit:content3918 a solog:LogEntryContent ;  
spl:hasData svd:Location ;  
spl:hasProcessing befit:SensorGathering ;  
spl:hasPurpose befit:HealthTracking ;  
spl:hasStorage [ spl:haslocation svl:OurServers ] ;  
spl:hasRecipient [ a svr:Ours ] .
```

Provenance/event information

Syntax and expressivity

EXAMPLE 1: Log description

```
eg:log1 a solog:Log;
  dct:title          "Log of Database R2D2"@en;
  dct:description    "This contains a dump of our Database R2D2 used to track BeFit";
  dct:issued         "2018-02-14"^^xsd:dateTimeStamp;
  prov:wasAttributedTo eg:TrackingSystemR2D2 ;
  solog:processor    eg:beFitInc .
```

Provenance/event information

Syntax and expressivity

EXAMPLE 1: Log description

```
eg:log1 a solog:Log;
  dct:title
  dct:description
  dct:issued
  prov:wasAttribute
  solog:processor
```

EXAMPLE 2: An event

```
eg:log1 solog:event eg:logEntry1 .

eg:logEntry1 a solog:ProcessingEvent;

  dct:title          "Collection of new device positions in Database R2D2 on Janua
  solog:dataSubject  eg:user1 ;
  dct:description    "We collected a new position of your BeFit
                      device in our database in Europe"@en;
  solog:transactionTime "2018-01-10T13:20:50Z^^xsd:dateTimeStamp;
  solog:validityTime   "2018-01-10T13:20:00Z^^xsd:dateTimeStamp;
  solog:message        "Tracking position by GPS... collected!" ;
  solog:eventContent   eg:content1 ;
  solog:immutableRecord eg:iRec1 .
```

Provenance/event information

Syntax and expressivity

EXAMPLE 1: Log description

```
eg:log1 a solog:Log;  
  dct:title  
  dct:description  
  dct:issued  
  prov:wasAttribute  
  solog:processor
```

EXAMPLE 2: An event

```
eg:log1 solog:event eg:logEntry1 .  
  
eg:logEntry1 a solog:ProcessingEvent;
```

```
  dct:title  
  solog:dataSubject  
  dct:description  
  
  solog:transactionTime  
  solog:validityTime  
  solog:message  
  solog:eventContent  
  solog:immutableRecord
```

EXAMPLE 3: Event content

```
eg:content1 a solog:logEntryContent;  
  dct:description      "This contains the data item collected by a BeFit device on Janua  
  spl:hasData          svd:Location;  
  spl:hasProcessing    eg:SensorGathering;  
  spl:hasPurpose       eg:HealthTracking;  
  spl:hasStorage        [has:location svl:OurServers];  
  spl:hasRecipient     [a svr:Ours].  
  
eg:SensorGathering rdfs:subClassOf svpr:Collect .  
eg:HealthTracking rdfs:subClassOf svpu:Health .
```

Provenance/event information SPECIAL resources

The SPECIAL Policy Log Vocabulary

A vocabulary for privacy-aware logs, transparency and compliance -
version 0.3



Unofficial Draft 06 April 2018

Editor:

Javier D. Fernández (Vienna University of Economics and Business)

Authors:

Piero Bonatti (Università di Napoli Federico II)

Wouter Dullaert (Tenforce)

Javier D. Fernández (Vienna University of Economics and Business)

Sabrina Kirrane (Vienna University of Economics and Business)

Uros Milosevic (Tenforce)

Axel Polleres (Vienna University of Economics and Business)

This document is licensed under a [Creative Commons Attribution 3.0 License](#).

Abstract

This document specifies *sblog*, a vocabulary to log data processing and sharing a given consent provided by a data subject. We also model the consent actions revocation

Vocabulary .../langs/sblog#

Bert Bos Last Updated: 17 April 2018

(You can [download this ontology as an OWL file](#).)

This is the SPECIAL Policy Log Vocabulary, with identifier

<http://www.specialprivacy.eu/langs/sblog#>

For the documentation, see the upcoming [Deliverable D2.3](#).

```
@prefix : <http://www.specialprivacy.eu/langs/sblog#> .  
@prefix dct: <http://purl.org/dc/terms/> .  
@prefix owl: <http://www.w3.org/2002/07/owl#> .  
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .  
@prefix spl: <http://www.specialprivacy.eu/langs/usage-policy#> .  
@prefix xml: <http://www.w3.org/XML/1998/namespace> .  
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .  
@prefix prov: <http://www.w3.org/ns/prov#> .  
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .  
  
<http://www.specialprivacy.eu/langs/sblog#> a owl:Ontology ;  
    rdfs:seeAlso "https://aic.ai.wu.ac.at/qadlod/policyLog/" ;  
    owl:versionInfo "0.3"@en .
```

- Detailed in *D2.3 Transparency Framework V1*
- Available for download via the SPECIAL website
<https://www.specialprivacy.eu/langs/sblog>
- An unofficial *draft specification* has been published online
<http://purl.org/specialprivacy/sblog>
- Feeds into the standardisation efforts conducted in the *W3C Data Privacy Vocabularies and Controls Community Group*

Transparency and compliance checking

Subsumption Algorithm

- The development of a compliance checking algorithm for the SPECIAL policy language devised in T2.1
- A company's policy can be checked for compliance with data subjects' consent and with part of the GDPR by means of **subsumption queries**
- We provide a **complete and tractable structural subsumption algorithm** for compliance checking
- Detailed in *D2.4 & D2.8 Transparency and Compliance Algorithms*

Algorithm 1: $\text{STS}(\mathcal{K}, C \sqsubseteq D)$

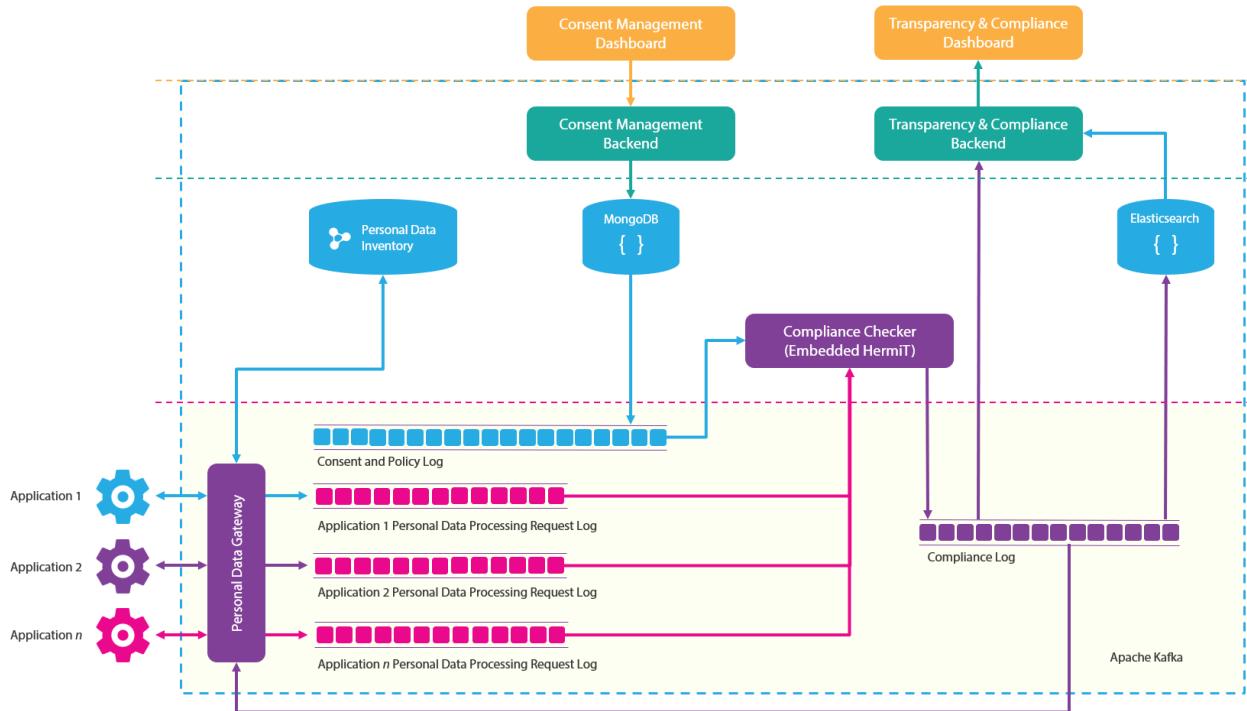
Input: \mathcal{K} and an elementary $C \sqsubseteq D$ where C is normalized
Output: true if $\mathcal{K} \models C \sqsubseteq D$, false otherwise
Note: Below, by $C = C' \sqcap C''$ we mean that either $C = C'$ or C' is a conjunct of C (possibly not the first one)

```
1 begin
2   if  $C = \perp$  then return true
3   if  $D = A$ ,  $C = A' \sqcap C'$  and  $A' \sqsubseteq^* A$  then return true
4   if  $D = [l, u](f)$  and  $C = [l', u'](f) \sqcap C'$  and  $l \leq l'$  and
       $u' \leq u$  then return true
5   if  $D = \exists R.D'$ ,  $C = (\exists R.C') \sqcap C''$  and
       $\text{STS}(\mathcal{K}, C' \sqsubseteq D')$  then return true
6   if  $D = D' \sqcap D''$ ,  $\text{STS}(\mathcal{K}, C \sqsubseteq D')$ , and
       $\text{STS}(\mathcal{K}, C \sqsubseteq D'')$  then return true
7   else return false
8 end
```

The SPECIAL transparency and compliance platform

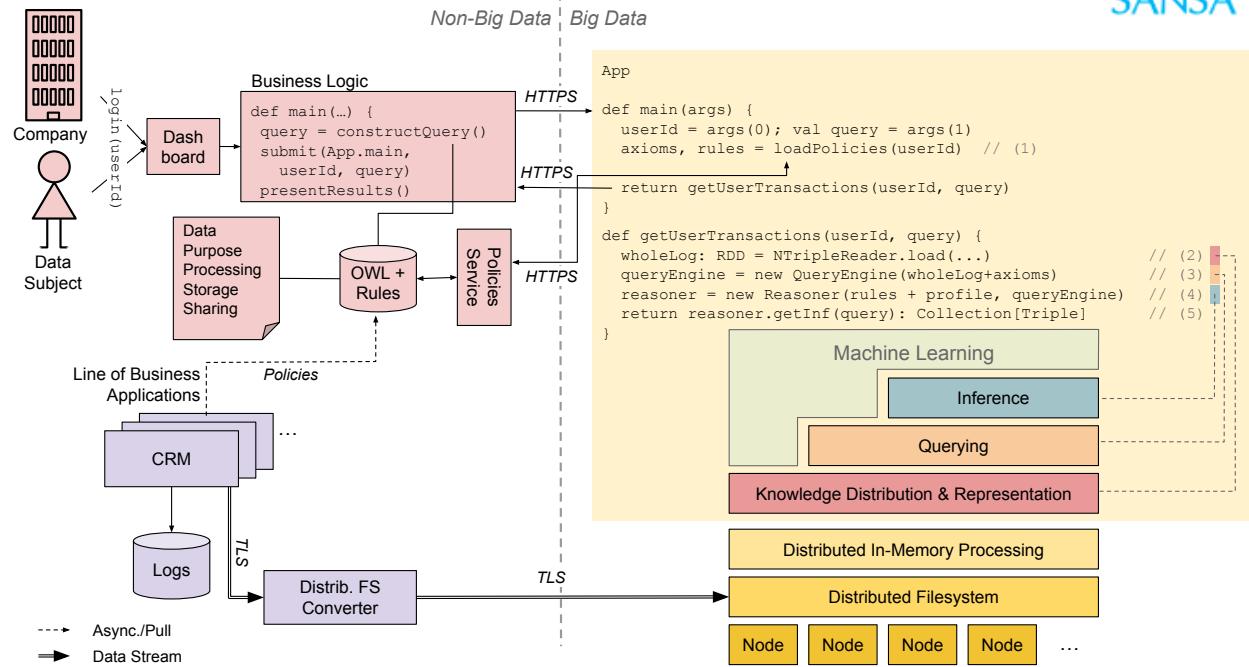


Transparency and compliance checking platforms



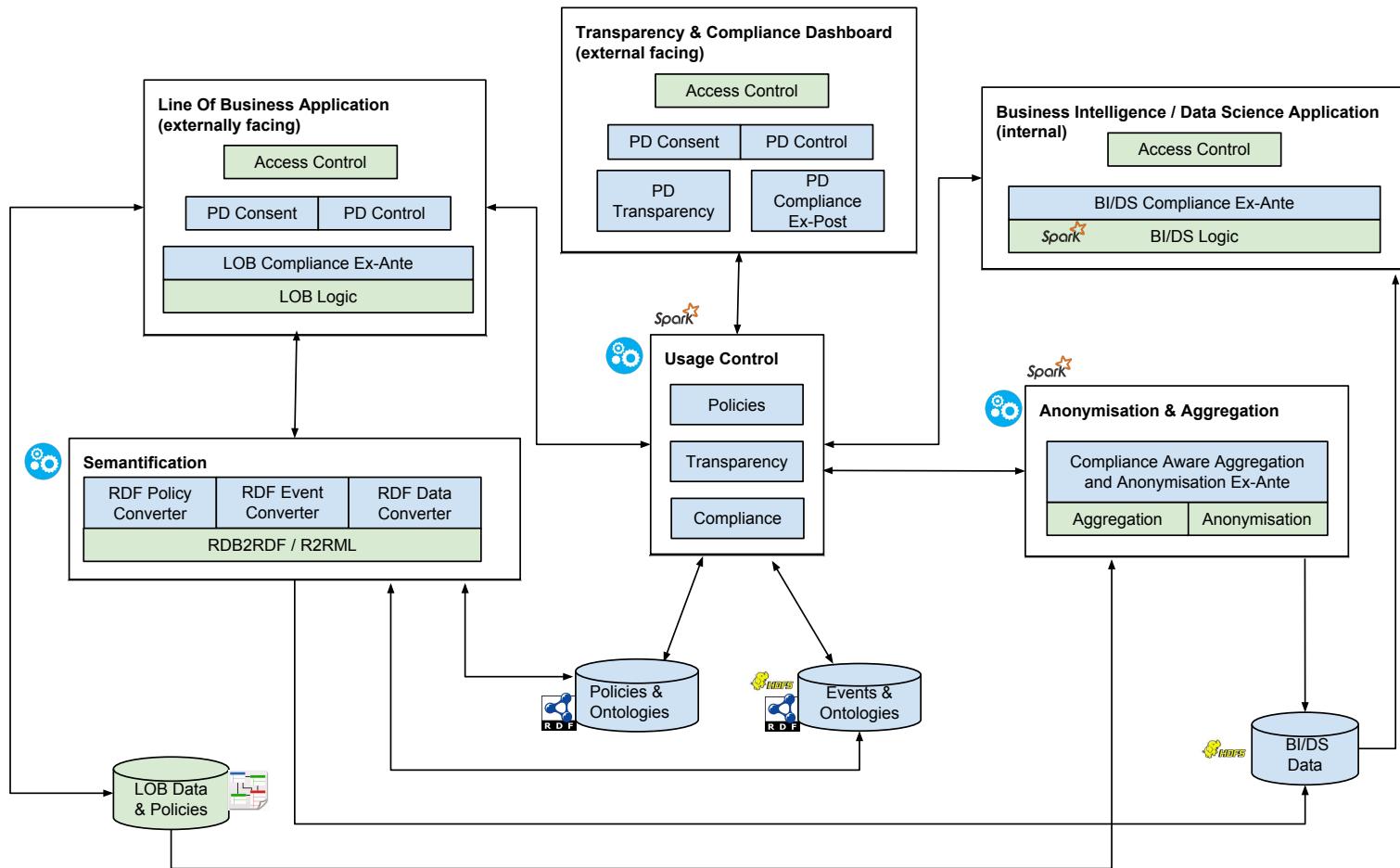
- Data processing and sharing event logs are stored in the **Kafka** distributed streaming platform, which in turn relies on Zookeeper for configuration, naming, synchronization, and providing group services.
- We assume that consent updates are infrequent and as such usage policies and the respective vocabularies are represented in a **Virtuoso triple store**.
- The compliance checker, which includes an embedded **HermiT reasoner**, uses the consent saved in Virtuoso together with the application logs provided by Kafka to check that data processing and sharing complies with the relevant usage control policies.
- As logs can be serialized using JSON-LD, it is possible to benefit from the faceting browsing capabilities of **Elasticsearch** and the out of the box visualization capabilities provided by **Kibana**.

Transparency and compliance checking platforms



- SANSA is an open source semantic data processing stack that supports distributed computations on large-scale RDF data
- SANSA is built on top of the two prevalent distributed in-memory big data processing frameworks Apache Spark and Apache Flink
- Demonstrating how SANSA can be used for personal data processing compliance checking

Provenance/event information



- In D2.3 we frame the SPECIAL policy, transparency and compliance components within the wider scope of a general **Enterprise setting**
- SPECIAL can be used in conjunction with existing **Line of Business (LOB)** and in **Business Intelligence (BI) / Data Science (DS)** settings
- Key role of the **Personal Data processing Inventory**

SPECIAL Standardisation Activities

Data Privacy, Vocabularies and Controls Community Group (DPVCG)

- ❖ Launched on the 25th of May 2018
- ❖ Presentation at MyData on the 31st of August-2018
- ❖ F2F in Vienna on the 3rd and 4th of December
- ❖ The current goal is to agree on first public drafts of minimal sets of vocabularies with first stable working drafts being reached latest on **25 May 2019**.

The screenshot shows the homepage of the W3C Data Privacy Vocabularies and Controls Community Group (DPVCG). The header features the W3C logo and the text "COMMUNITY GROUP". The main content area is titled "DATA PRIVACY VOCABULARIES AND CONTROLS COMMUNITY GROUP". It includes a mission statement about developing a taxonomy of privacy terms related to GDPR. Below this, it mentions the start date of 25th May 2018 and a link to the "Workshop on Data Privacy Controls and Vocabularies". On the right side, there are sections for "Tools for this group" (Mailing List, Wiki, IRC), "Chairs" (Bert Bos, Axel Polleres), and "Participants (52)" (a grid of small profile pictures). A handwritten note "Chairs" with an arrow points to the chairpersons' section.

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

Data Privacy, Vocabularies and Controls Community Group (DPVCG)

- ❖ Consent Receipt Specification (Kantara)
- ❖ GDPRText (Trinity College Dublin)
- ❖ DECODE project
- ❖ CitySPIN project
- ❖ Expedite project
- ❖ Pret-a-LLOD.eu project

N.B. More Industry involvement needed

The screenshot shows the homepage of the W3C Data Privacy Vocabularies and Controls Community Group (DPVCG). The header features the W3C logo and the text "COMMUNITY GROUP". The main content area is titled "DATA PRIVACY VOCABULARIES AND CONTROLS COMMUNITY GROUP". It includes a mission statement about developing a taxonomy of privacy terms from the GDPR. Below this, it mentions the official start date of May 25, 2018, resulting from a workshop in Vienna. On the right side, there are sections for "Tools for this group" (Mailing List, Wiki, IRC), "Chairs" (Bert Bos, Axel Polleres), and "Participants (52)" (a grid of small profile pictures). The URL <https://www.w3.org/community/dpvcg/> is provided at the bottom.

W3C COMMUNITY GROUP

Home / Data Privacy Vocabularies...

DATA PRIVACY VOCABULARIES AND CONTROLS COMMUNITY GROUP

The mission of the W3C Data Privacy Vocabularies and Controls CG (DPVCG) is to develop a taxonomy of privacy terms, which include in particular terms from the new European General Data Protection Regulation (GDPR), such as a taxonomy of personal data as well as a classification of purposes (i.e., purposes for data collection), and events of disclosures, consent, and processing such personal data.

The Community Group shall officially start on 25th of May 2018, the official date of the GDPR coming into force, as a result of the W3C [Workshop on Data Privacy Controls and Vocabularies](#) in Vienna earlier this year.

Tools for this group [i](#)

- [Mailing List](#)
- [Wiki](#)
- [IRC](#)

Chairs

Bert Bos

Axel Polleres

Participants (52)

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

SPECIAL Resources



Achievements: Exploitable Results

- Resources
 - ❖ The SPECIAL Usage Policy Language
<http://purl.org/specialprivacy/policylanguage>
 - ❖ The SPECIAL Vocabularies
<https://www.specialprivacy.eu/vocabs>
 - ❖ The SPECIAL Policy Log Vocabulary
<http://purl.org/specialprivacy/splog>
- *SPECIAL Ex-Post Compliance Checking*
 - ❖ Demonstrates how usage policies together with event logs can be used to perform ex-post compliance checking
- *SPECIAL Consent and Transparency Interfaces*
 - ❖ Various consent user interfaces and the transparency dashboard
 - ❖ Guidelines for legally compliant consent retrieval

The SPECIAL Policy Log Vocabulary

A vocabulary for privacy-aware logs, transparency and controls version 0.3

Unofficial Draft 06 April 2018

Editor:

Javier D. Fernández (Vienna University of Economics and Business)

Authors:

Piero Bonatti (Università di Napoli Federico II)

Wouter Dullaert (Tenforce)

Javier D. Fernández (Vienna University of Economics and Business)

Sabrina Kirrane (Vienna University of Economics and Business)

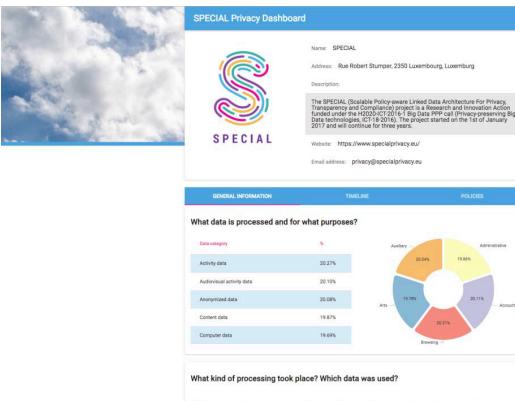
Uros Milosevic (Tenforce)

Axel Polleres (Vienna University of Economics and Business)

This document is licensed under a Creative Commons Attribution 3.0 License.

Abstract

This document specifies *splog*, a vocabulary to log data processing and sharing events given a consent provided by a data subject. We also model the consent actions related to revocation



The screenshot shows the SPECIAL Privacy Dashboard interface. It includes sections for General Information (Name: SPECIAL, Address: Rue Robert Stumper, 2350 Luxembourg, Luxembourg), Timeline, Policies, and Data Processing. The Data Processing section displays a pie chart showing the distribution of data categories: Auditory (20.27%), Administrative (19.86%), Accounting (20.11%), Billing (19.47%), and Computer (19.49%). Below the chart, a table lists data categories and their percentages: Activity data (20.27%), Audience activity data (20.11%), Anonymized data (20.08%), Content data (19.47%), and Computer data (19.49%).

Vocabulary .../langs/splog#

Bert Bos Last Updated: 17 April 2018

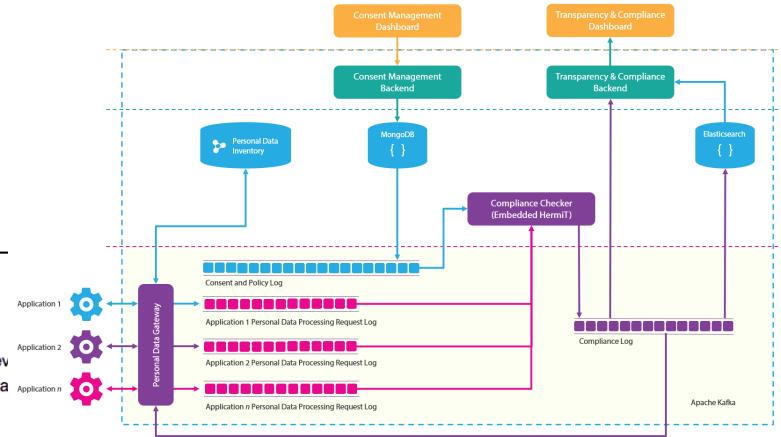
(You can download this ontology as an OWL file.)

This is the SPECIAL Policy Log Vocabulary, with identifier

<http://www.specialprivacy.eu/langs/splog#>

For the documentation, see the upcoming Deliverable D2.3.

```
@prefix : <http://www.specialprivacy.eu/langs/splog#> .  
@prefix dct: <http://purl.org/dc/terms/> .  
@prefix owl: <http://www.w3.org/2002/07/owl#> .
```



The screenshot shows the W3C Data Privacy Vocabularies Community Group page. It features the W3C logo and a banner with the text "COMMUNITY & BUSINESS GROUPS". The page includes links for "Home" and "Data Privacy Vocabularies...".

DATA PRIVACY VOCABULARIES AND CONTROLS COMMUNITY GROUP

The mission of the W3C Data Privacy Vocabularies and Controls CG (DPVCG) is to develop a taxonomy of privacy terms, which include in particular terms from the new European General Data Protection Regulation (GDPR), such as a taxonomy of personal data as well as a classification of purposes (i.e., purposes for data collection), and events of disclosures, consent, and processing such personal data.

The Community Group shall officially start on 25th of May 2018, the official date of the GDPR coming into force, as a result of the W3C [Workshop on Data Privacy Controls and Vocabularies](#) in Vienna earlier this year.

- Tools for this group
- Mailing List
- Wiki
- IRC
- Tracker
- RSS
- Contact This Group

Any Questions?



Events at the Belgian Coast at your fingertips

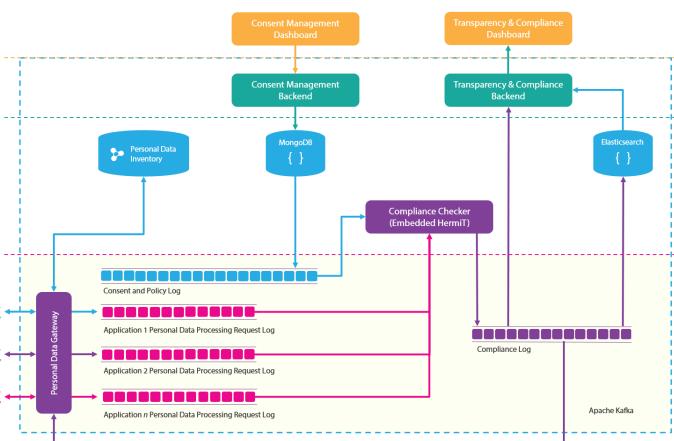
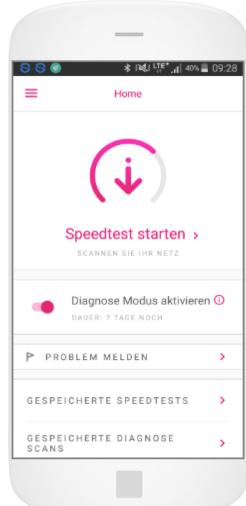
Sign up for free for intelligent tourist event recommendations tailored to you.

Login

freddy.demeersman@proximus.com

LOGIN

pro~~ximus~~



The SPECIAL Usage Policy Language

version 0.1



Unofficial Draft 06 April 2018

Editor:

Javier D. Fernández (Vienna University of Economics and Business)

Authors:

Piero Bonatti (Università di Napoli Federico II)
Sabrina Kirrane (Vienna University of Economics and Business)
Iliana Mineva Petrova (Università di Napoli Federico II)
Luigi Sauro (Università di Napoli Federico II)
Eva Schleihahn (Unabhängiges Landeszentrum für Datenschutz (ULD))

This document is licensed under a Creative Commons Attribution 3.0 License.

Abstract

This document specifies usage policy language of SPECIAL. The usage policy language is meant to express both the data subjects' consent and the data usage policies of data controllers in formal terms, understandable by a computer, so as to automatically verify that the usage of personal data complies with data subjects' consent.

The ontology defined in this document is publicly available at <http://www.specialprivacy.eu/langs/usage-policy>.

SPECIAL Privacy Dashboard

Name: SPECIAL
Address: Rue Robert Stumper, 2300 Luxembourg, Luxembourg
Description: The SPECIAL (Scalable Policy-aware Linked Data Architecture For Privacy Transparency and Compliance) project is a Research and Innovation Action funded by the European Commission. The project started on the 1st of January 2017 and will continue for three years.
Website: <http://www.specialprivacy.eu/>
Email address: privacy@specialprivacy.eu

GENERAL INFORMATION

Data category	%
Activity data	20.27%
Audited activity data	20.15%
Anonymized data	20.05%
Content data	19.87%
Computer data	19.65%

TIMELINE

Timeline chart showing project progress over time.

POLICIES

Donut chart showing the distribution of purposes (e.g., Marketing, Advertising, Accounting, Billing).

What kind of processing took place? Which data was used?

The SPECIAL Policy Log Vocabulary

A vocabulary for privacy-aware logs, transparency and compliance - version 0.3



Unofficial Draft 06 April 2018

Editor:

Javier D. Fernández (Vienna University of Economics and Business)

Authors:

Piero Bonatti (Università di Napoli Federico II)
Wouter Duillaert (Tenseforce)
Javier D. Fernández (Vienna University of Economics and Business)
Sabrina Kirrane (Vienna University of Economics and Business)
Uros Milosevic (Tenseforce)
Axel Polleres (Vienna University of Economics and Business)

This document is licensed under a Creative Commons Attribution 3.0 License.

Abstract

This documents specifies splog, a vocabulary to log data processing and sharing events that should comply with a given consent provided by a data subject. We also model the consent actions related to consent giving and revocation

COMMUNITY & BUSINESS GROUPS

W3C

Community & Business Groups

DATA PRIVACY VOCABULARIES AND CONTROLS

Community Group

The mission of the W3C Data Privacy Vocabularies and Controls CG (DPVCG) is to develop a taxonomy of privacy terms, which include in particular terms from the new European General Data Protection Regulation (GDPR), such as a taxonomy of personal data as well as a classification of purposes (i.e., purposes for data collection), and events of disclosures, consent, and processing such personal data.

The Community Group shall officially start on 25th of May 2018, the official date of the GDPR coming into force, as a result of the W3C [Workshop on Data Privacy Controls and Vocabularies](#) in Vienna earlier this year.

Tools for this group

- Mailing List
- Wiki
- IRC
- Tracker
- RSS
- Contact This Group