

# Workshop on Privacy & Linked Data

Rigo Wenning

Dagstuhl Seminar 18233 on Applied Machine Learning

04-06. Juni 2018

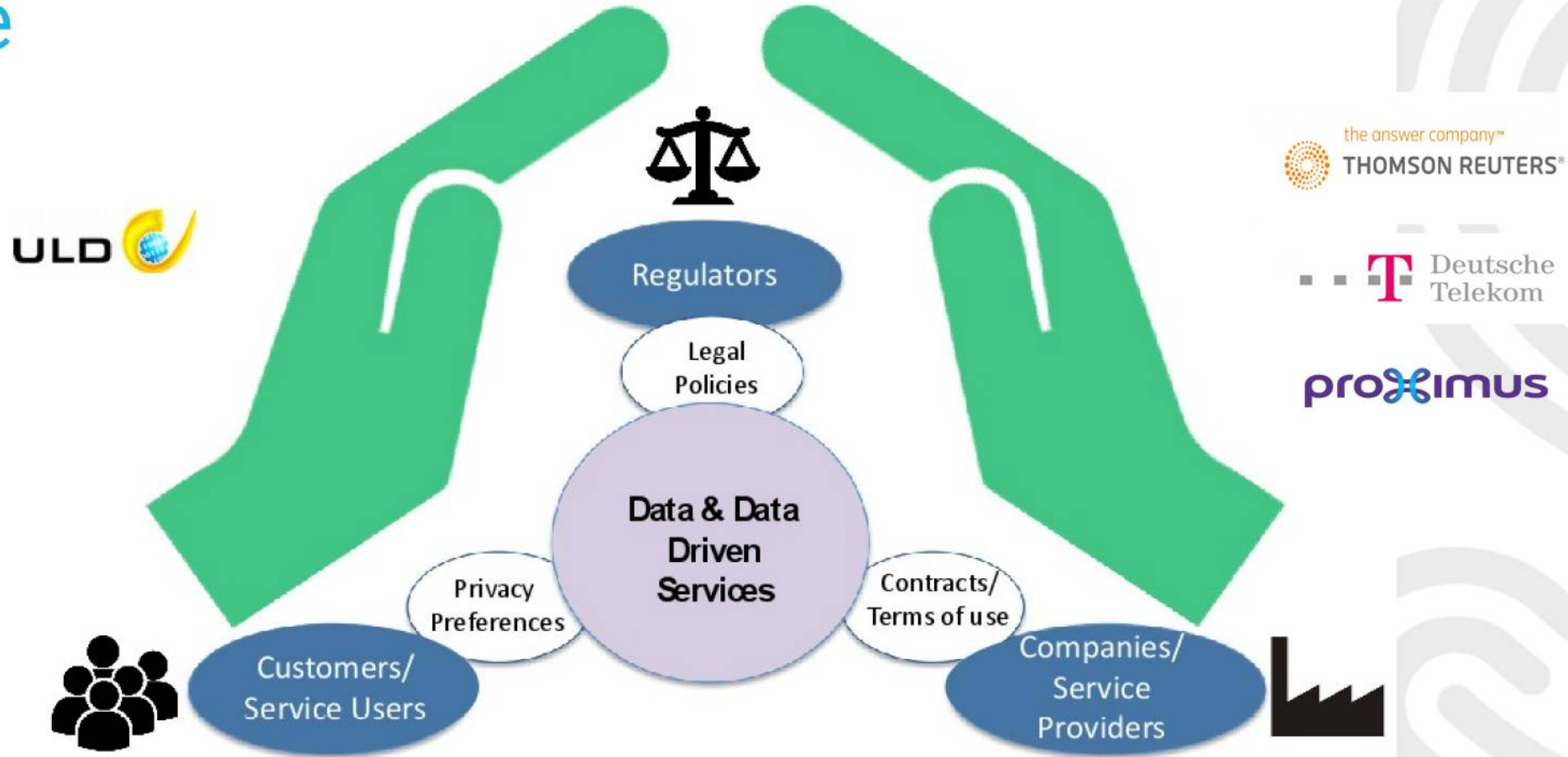


Horizon 2020  
European Union funding  
for Research & Innovation



# SPECIAL Überblick

## Ziele

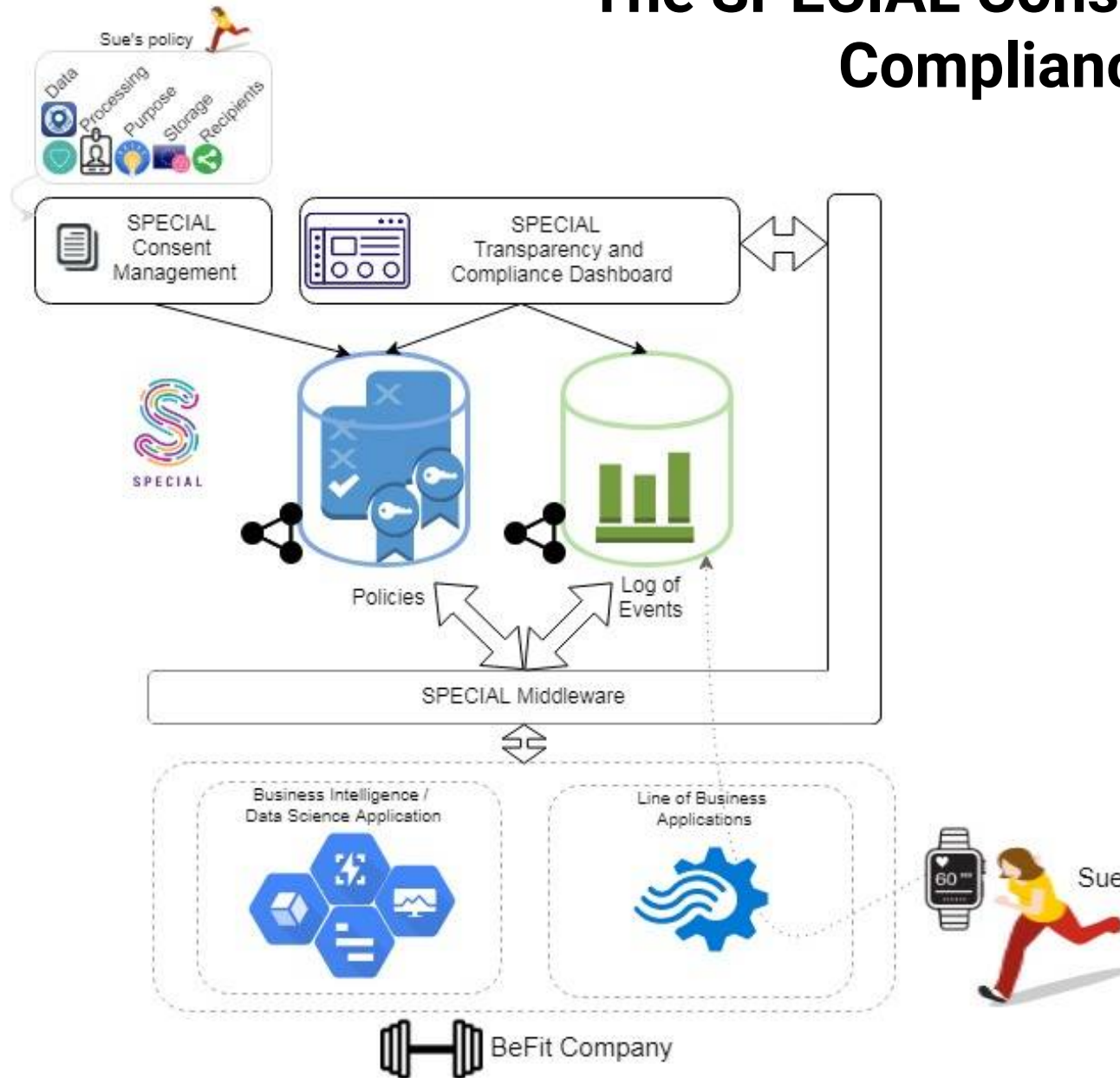


# SPECIAL Überblick

## Ziele

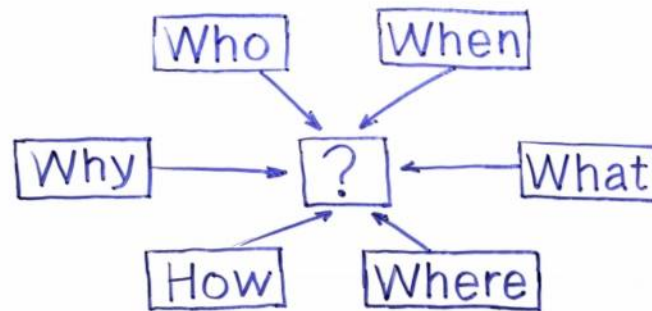
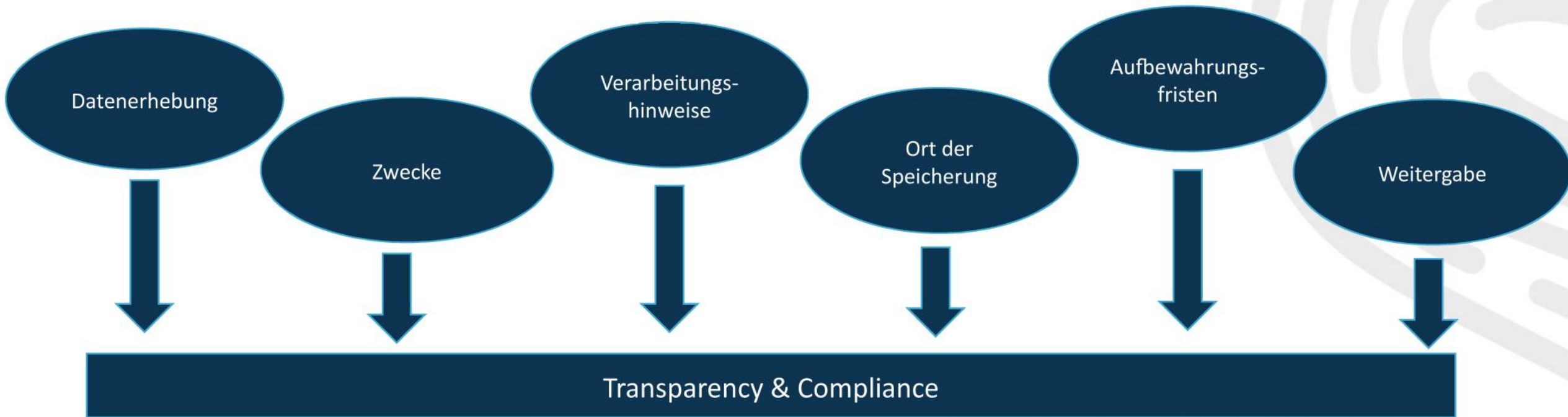
- Policy management framework
  - ❖ **Kontrolle des Nutzers** über seine Daten
  - ❖ Repräsentation von **Zugangs- und Nutzungsregel** and **lgesetzlichen Anforderungen** in **maschinenlesbarem Format**
- Transparency and compliance framework
  - ❖ Liefert Information zur **Verarbeitung** und **Weiterleitung** an Dritte
  - ❖ Ermöglicht Eingriff und **Korrektur** durch den Nutzer
- Scalable policy-aware Linked Data architecture
  - ❖ Auf Basis der Big Data Europe (BDE) Plattform **für Skalierbarkeit**
  - ❖ Erweitert BDE mit **policy, transparency** and **compliance Protokollen**

# The SPECIAL Consent, Transparency and Compliance framework



# The SPECIAL Policy Language

## Starting Points



# Die SPECIAL Policy Language

## Bestehende Vokabulare

- **FOAF**, **vCard** und **schema.org** bieten Vokabulare zur Modellierung von statischen Personendaten
- **DICOM** für Gesundheitsdaten mit vielen Attributen zur Fitness
- **NeoGeo** Vokabular, das **GeoSPARQL** Vokabular oder das **WGS84 Geo Positioning** Vokabular können für Ortsdaten genutzt werden.
- **P3P** WG hatte ein RDF Vokabular veröffentlicht mit Zwecken
- **ODRL** trägt ein Modell bei, das den Ausdruck von actions, prohibitions, und obligations erlaubt um Erklärungen und Consent zu beschreiben
- **OWL Time** kann Zeit und Dauer der Verarbeitung ausdrücken, sowie Aufbewahrungszeiten
- **PROV** ist ein guter Startpunkt für die Modellierung von Herkunft und Verarbeitung von Einwilligungen

# The SPECIAL Policy Language

## More Details

Unofficial Draft

### TABLE OF CONTENTS

1. Introduction
    - 1.1 What is a Usage Policy?
    - 1.2 Audience and scope
    - 1.3 Document conventions and namespaces
  2. Conformance
  3. The SPECIAL's Usage Policy Language
    - 3.1 Basic Usage Policies
    - 3.2 General Usage Policies
    - 3.3 Storage Expressions
  4. Compliance Checking
  5. SPECIAL's Usage Policy Language Grammar
  6. Acknowledgements
- A. References**
- A.1 Normative references
  - A.2 Informative references

## 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 Schliehahn (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>.

ReSpec



## Project deliverables

Linh Nguyen Last Updated: 26 March 2018

### Deliverable 6.1 - SPECIAL website set up (M3)

[Deliverable 6.1](#)

This document provides an overview of the initial SPECIAL project website as it stands at Project Month 3.

### Deliverable 1.2 - Legal requirements for a privacy enhancing Big Data V1 (M6)

[Deliverable 1.2](#)

This is the report providing details of all legal and ethical considerations, as a main input for the SPECIAL privacyaware platform.

### Deliverable 3.1 - Initial setup of policy aware Linked Data architecture and engine (M6)



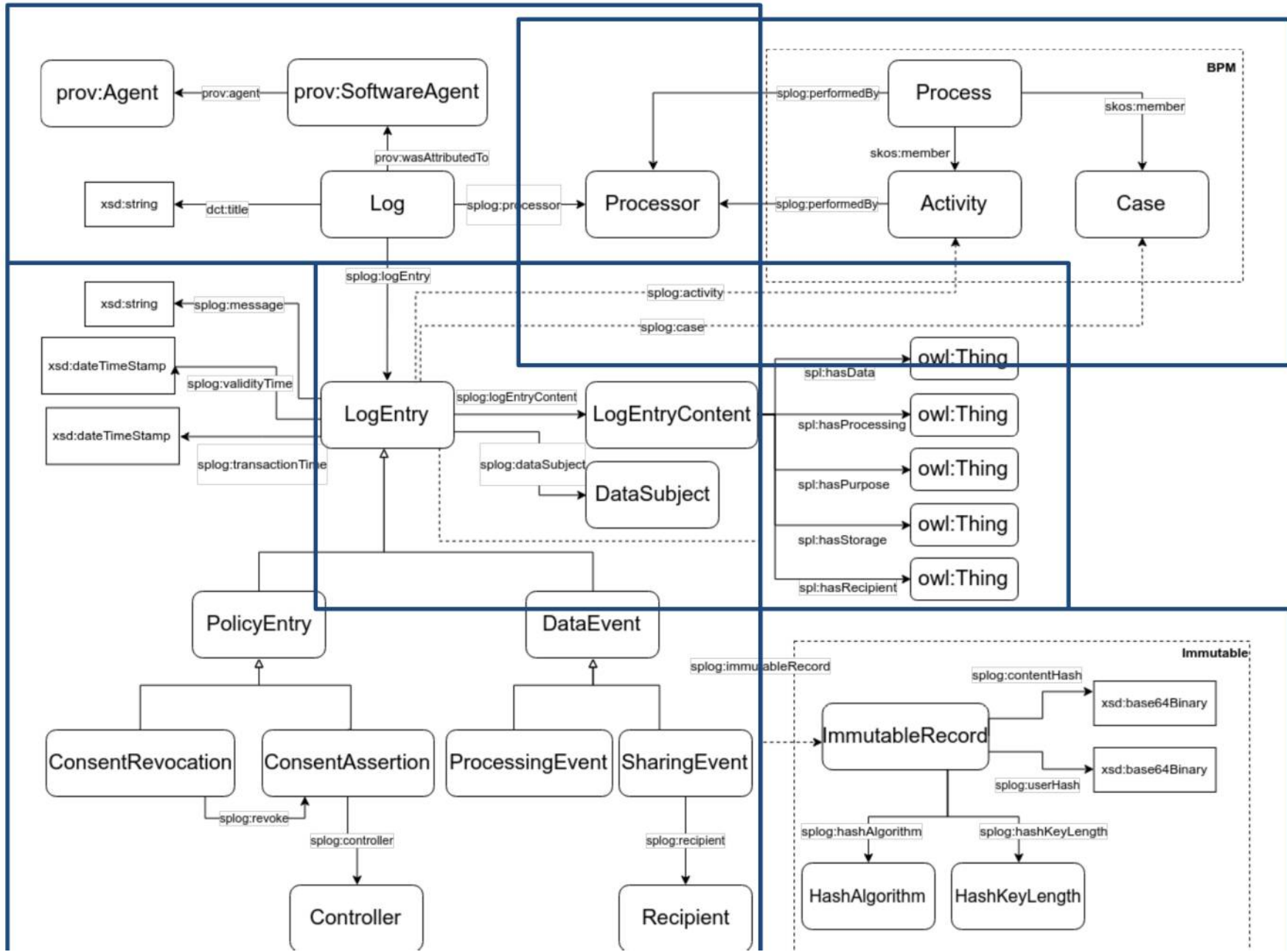
Project deliverables

Papers & presentations

Press releases and newsletters

Studies & essays

Dissemination material





# Shortcomings we try to overcome

- lack of specific and standard **vocabularies** for representing **privacy-related events**
- It should be possible to describe the event content at different **granularities**
  - from a simple **taxonomy** (stating the type of data, processing, etc.), to the most fine-grained description of the **actual data** associated to the event;
- **Interoperability/standard APIs**
  - different systems within a company and/or different companies
  - Use case example: cyber-physical social systems

<http://cityspin.net/>



CITYSPIN

# Anonymisation

- Protokoll zur Anonymisierung von Daten
- Best Practices
- Anforderungen
- Meta-Information zu Anonymisierung



# OASIS COEL TC

COEL is a privacy-by-design approach to organising data about people.

It is based on the coding of events (micro-structured data called 'atoms') using a hierarchical taxonomy.

This taxonomy includes codes for consent & notice as well as over 5,000 other behaviours.

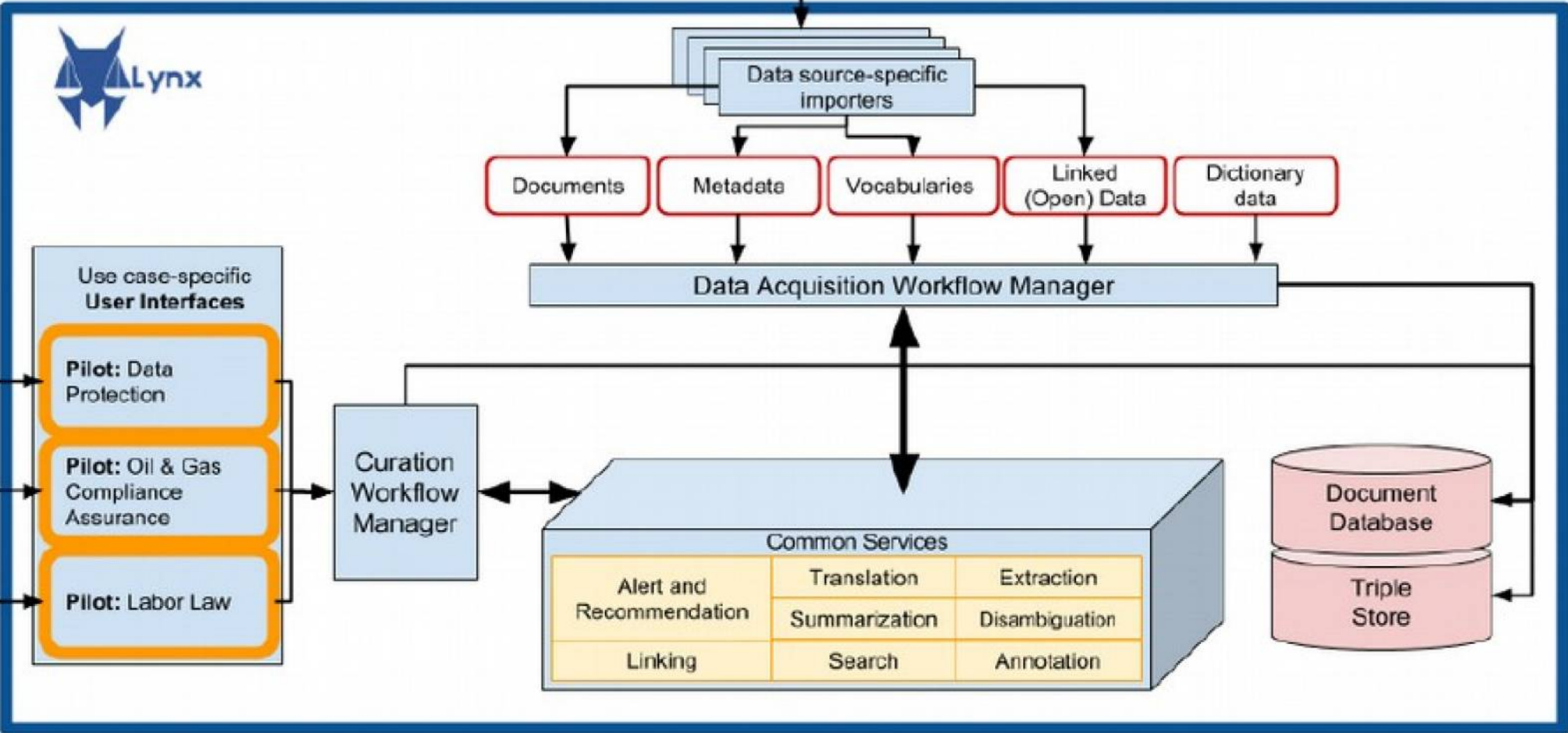
## JSON Behavioural Atom

|         |                                 |
|---------|---------------------------------|
| Who     | pseudonymised key               |
| When    | Unix time                       |
| What    | taxonomy code                   |
| Where   | actual, relative or descriptive |
| How     | data source                     |
| Context | social & environmental          |
| Consent | consent for use                 |

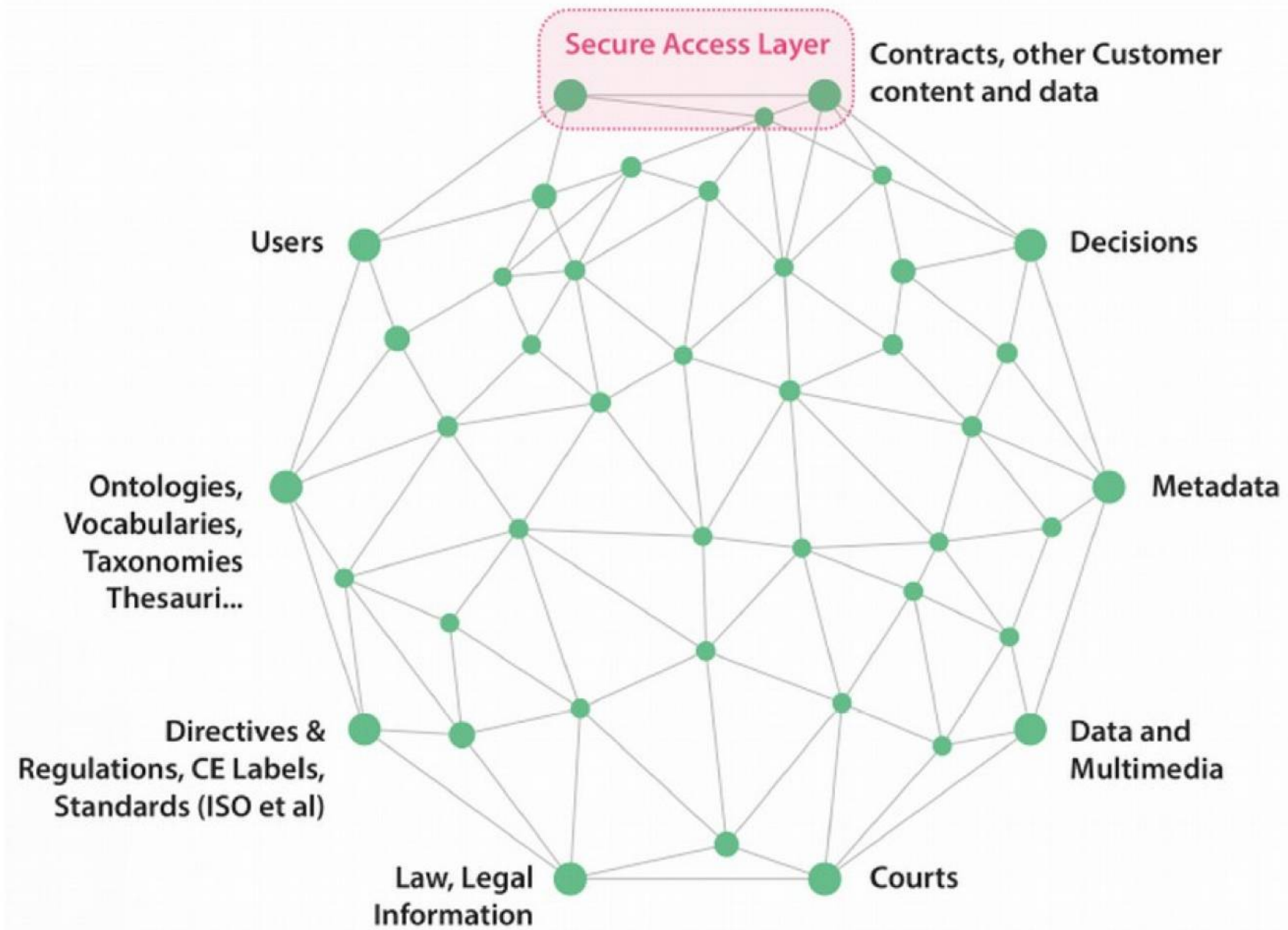
(syntactic layer)



Data & Documents



# The Legal Knowledge Graph



---

# ODRL-based Usage Control

## **Establishing a Community of Practice**

Position Statement – W3C Workshop on Privacy and Linked Data

---

Jaroslav Pullmann, Fraunhofer FIT  
Christian Mader, Fraunhofer IAIS  
Andreas Eitel, Fraunhofer IESE

# Decentralized IDentifiers (DIDs)

- Developed at Rebooting-the-Web-of-Trust workshop and W3C Credentials CG
- Persistent, dereference-able, cryptographically verifiable identifiers
- Registered in a blockchain or other decentralized network
- **did:sov:3k9dg356wdcj5gf2k9bw8kfg7a**
- Modular specification using “methods”:
- **did:sov, did:btcr, did:v1, did:uport, ...**
- Can be pairwise unique for each relationship
- Resolution: DID → DID Document
  - Set of public keys
  - Set of service endpoints

| Method     | DID Prefix |
|------------|------------|
| Sovrin     | did:sov:   |
| Bitcoin    | did:btcr:  |
| uPort      | did:uport: |
| VeresOne   | did:v1:    |
| IPFS       | did:ipid:  |
| IPDB       | did:ipdb:  |
| Blockstack | did:stack  |



# What would you like to standardise (rank in order) V2?

Respond at [PollEv.com/sabrinakirra386](https://pollev.com/sabrinakirra386)



# Data Privacy Vocabularies & Controls CG

- harmonize related efforts
- bring together stakeholders
- develop respective vocabularies to enable semantic interoperability and interchange of transparency logs
- exact scope of use cases related to making personal data processing interoperable
- <https://www.w3.org/community/dpvcg/>

# Deliverables

- 1) Use cases and requirements
- 2) Alignment of vocabularies and identification of overlaps
- 3) Glossary of GDPR terms
- 4) Vocabularies based on the heterogeneity or homogeneity of the agreed upon use cases

