

Enhancing Transparency in the Big Data and AI Landscape

Sabrina Kirrane, Vienna University of Economics and Business

Beyond Privacy: Learning Data Ethics

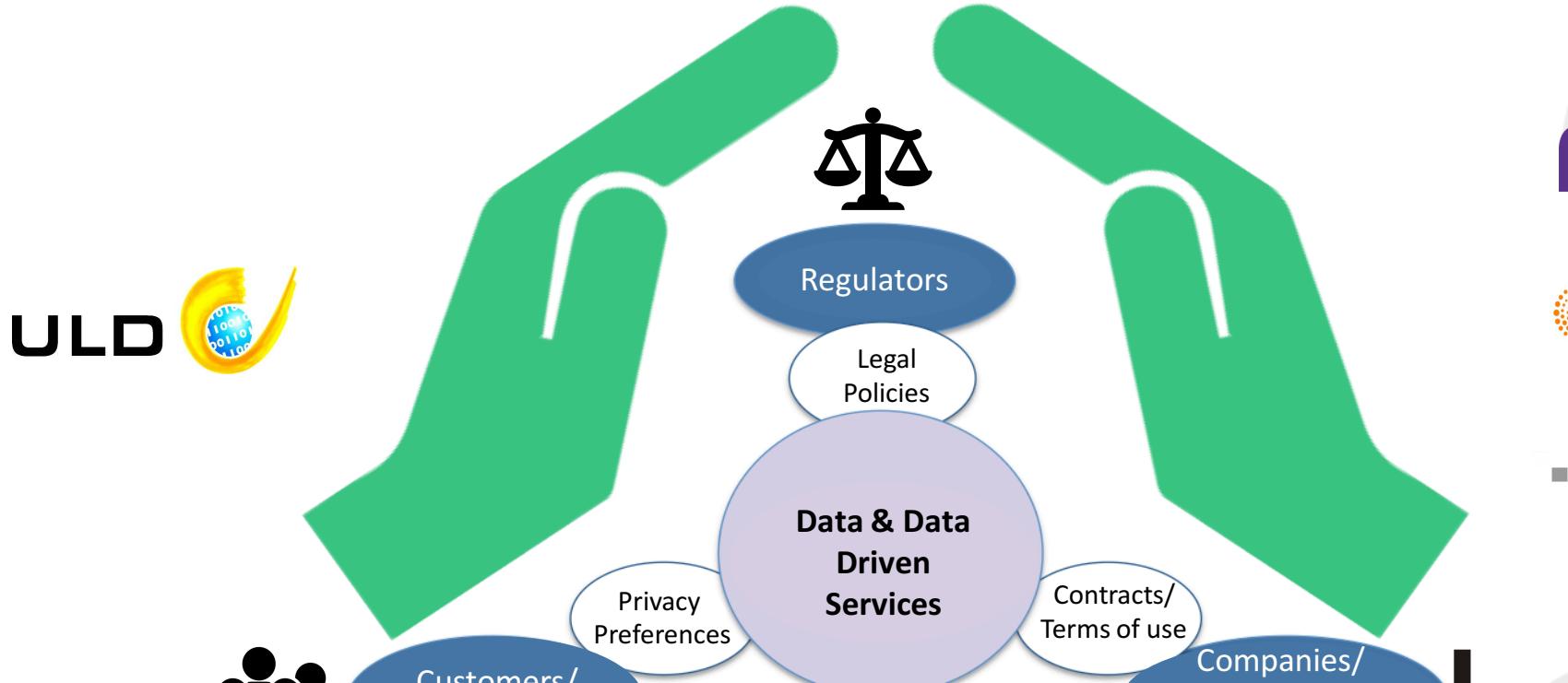
13th of November 2019



Horizon 2020
European Union funding
for Research & Innovation



SPECIAL Aims



proximus

the answer company™
THOMSON REUTERS®

Deutsche
Telekom

The SPECIAL Usage policy language

Syntax and expressivity

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 Schlehahn (Unabhängiges Landeszentrum für Da

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



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 & D2.5 Policy Language V2*
- Available for download via the SPECIAL website:
<https://www.specialprivacy.eu/publications/public-deliverables>
- An unofficial *draft specification* has been published online
<https://www.specialprivacy.eu/platform/ontologies-and-vocabularies>

The SPECIAL Policy Log Vocabulary

Syntax and expressivity

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 documents 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* delivered in M14
- Available for download via the SPECIAL website
<https://www.specialprivacy.eu/langs/sblog>
- An unofficial *draft specification* has been published online
<https://www.specialprivacy.eu/platform/ontologies-and-vocabularies>

SPECIAL ODRL Regulatory Compliance Profile

Syntax and expressivity

ODRL Regulatory Compliance Profile

version 0.1

Unofficial Draft 29 May 2019

Editor:

[Sabrina Kirrane](#) (Vienna University of Economics and Business)

Authors:

[Sabrina Kirrane](#) (Vienna University of Economics and Business)

[Marina De Vos](#) (University of Bath)

[Julian Padget](#) (University of Bath)

This document is licensed under a [Creative Commons Attribution 4.0 License](#).

Abstract

The Open Digital Rights Language (ODRL) is a policy expression language that provides a flexible and interoperable information model, vocabulary, and encoding mechanisms for representing statements about the usage of content and services.

This document constitutes an ODRL profile that adapts the ODRL Core Model and Vocabulary with concepts and terms to support regulatory compliance checking of business policies.

In essence, ODRL Regulatory Compliance Profile policies are used to represent regulatory permissions, prohibitions, obligations, and dispensations, which may be limited by constraints (e.g., temporal, spatial).

- Preliminary Analysis Detailed in *D2.2 Formal Representation of the legislation V1* & *D2.6 Formal Representation of the legislation V2*
- Available for download via the SPECIAL website:
<https://www.specialprivacy.eu/publications/public-deliverables>
- An unofficial *draft specification* has been published online
<https://www.specialprivacy.eu/platform/ontologies-and-vocabularies>

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*

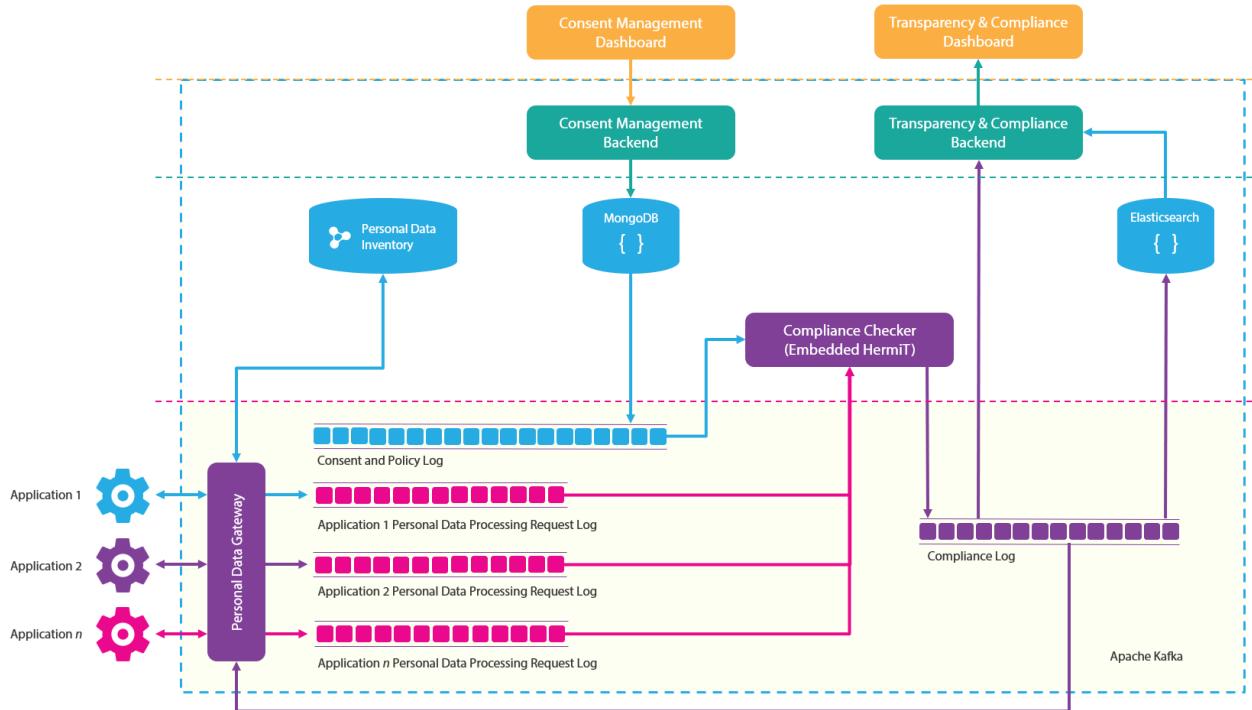
Algorithm 1: $\text{STS}(\mathcal{K}, C \sqsubseteq D)$

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
```

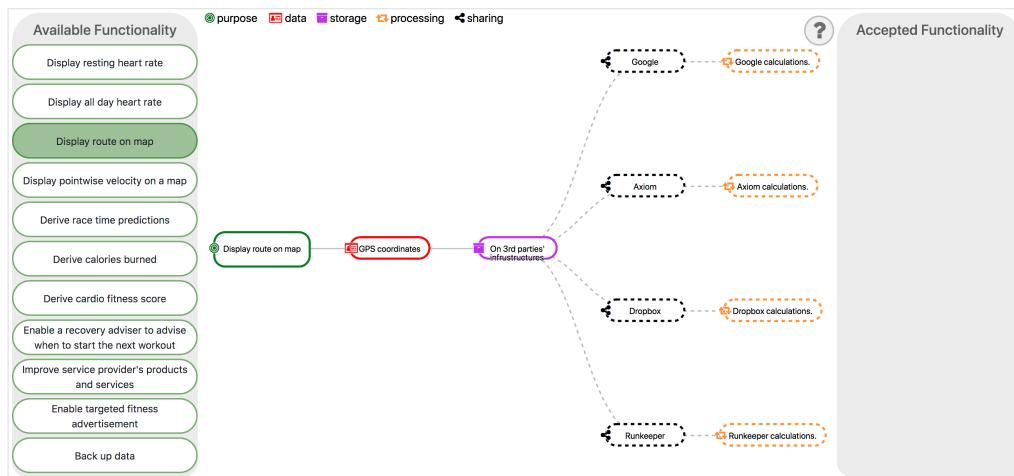
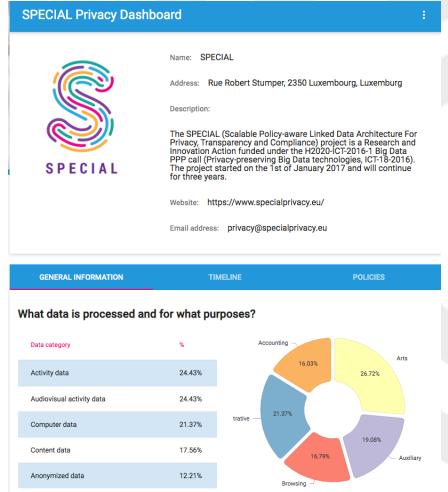
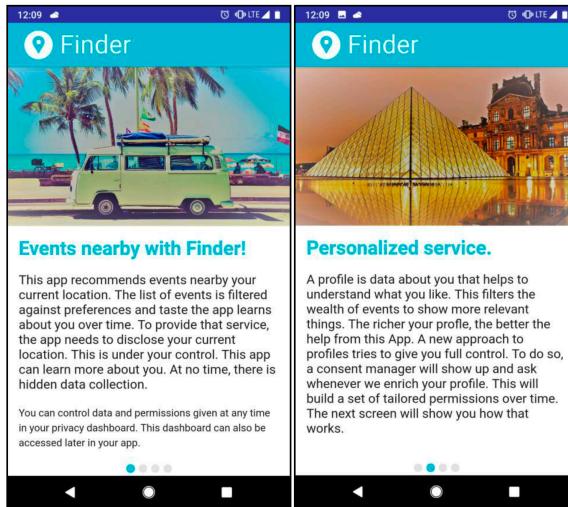
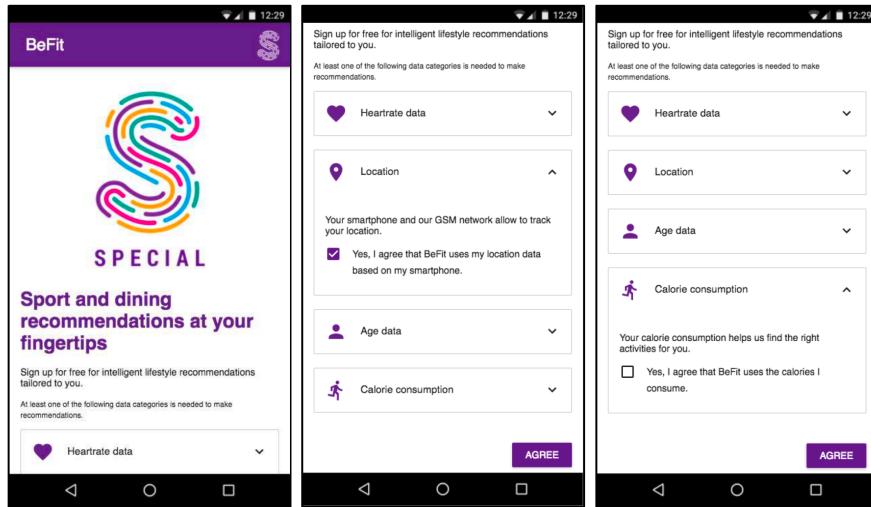
Transparency and compliance checking

Stream processing platform



- 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**.

The SPECIAL Mobile Consent UI



Open Challenges & Opportunities

- Standardisation of vocabularies (data, processing, purpose, storage, sharing) is difficult
- There are cognitive limitations in terms of understanding consent and transparency
- GDPR Compliance is only the tip of the iceberg, from a usage control perspective we also need to consider other regulations, licenses, social norms, cultural differences
- We need to embrace distributed and decentralised systems, which complicates things further
- Ensuring such systems are well behaved is a crucial to success (i.e., all usage constraints are adhered to and the system as a whole works as expected)

Any Questions?



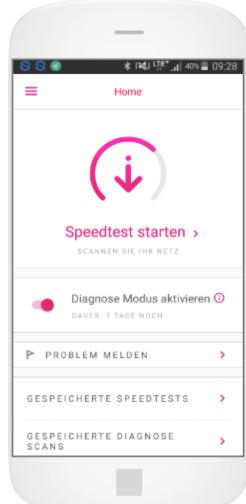
Events at the Belgian Coast at your fingertips

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

Login

freddy.demeersman@proximus.com

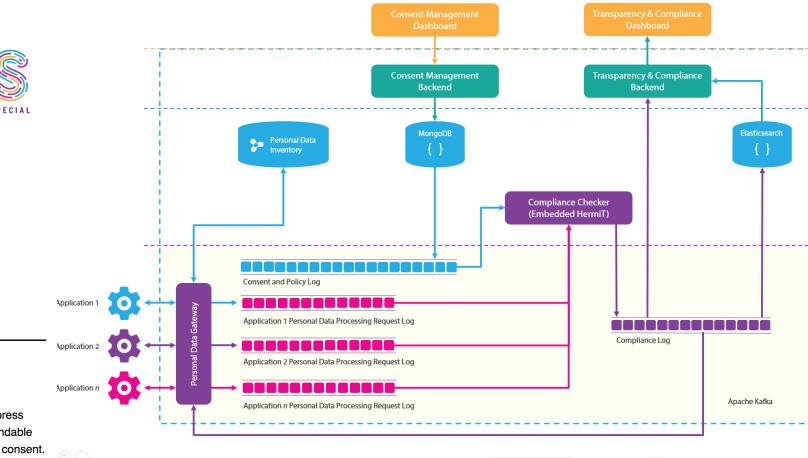
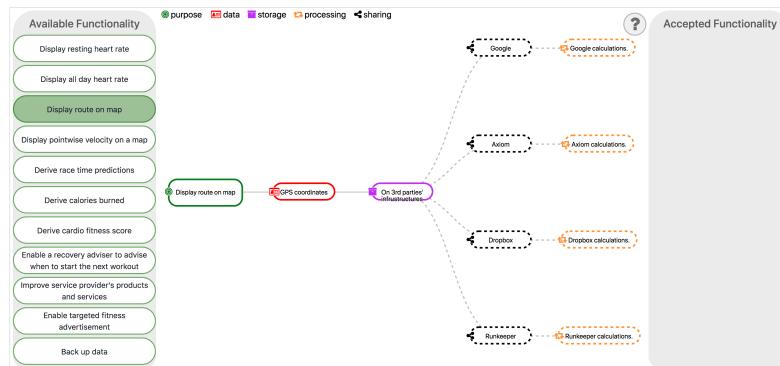
LOGIN



the answer company™
THOMSON REUTERS®

proximus

Deutsche Telekom



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>.

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

Contact Details



Horizon 2020
European Union funding
for Research & Innovation

The project SPECIAL (Scalable Policy-aware linked data arChitecture for privacy, transparency and compliance) has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 731601 as part of the ICT-18-2016 topic Big data PPP: privacy-preserving big data technologies.

Technical/Scientific contact

Sabrina Kirrane

Vienna University of Economics and Business

sabrina.kirrane@wu.ac.at