Privacy: Standards and Vocabularies for Transparency & Interoperability

Axel Polleres Joint work with: Piero Bonatti, Bert Bos, Stefan Decker, Javier D. Fernández, Sabrina Kirrane, Vassilios Peristeras, Rigo Wenning, Martin Kurze, Ben Whittam Smith & the members of the **W3C Data Privacy Vocabularies and Controls CG**

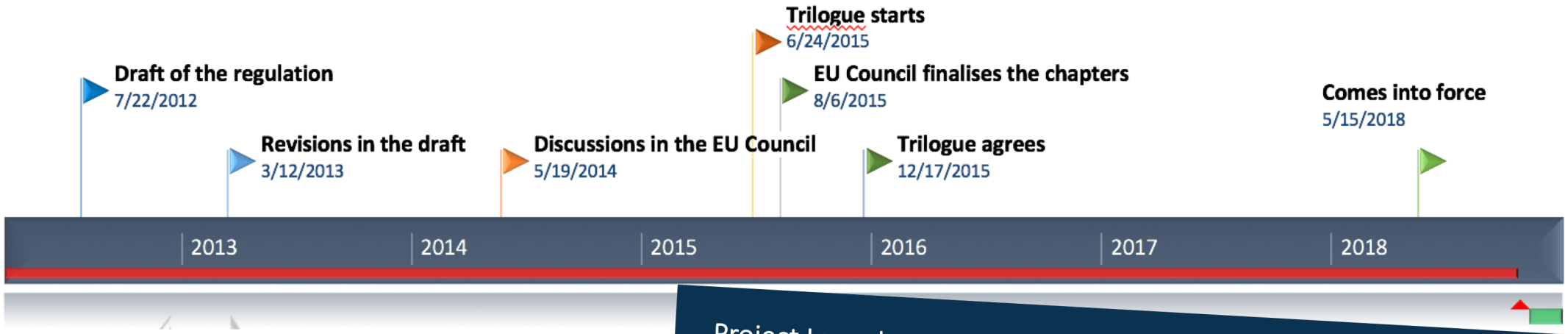
Horizon 2020
European Union funding
for Research & Innovation

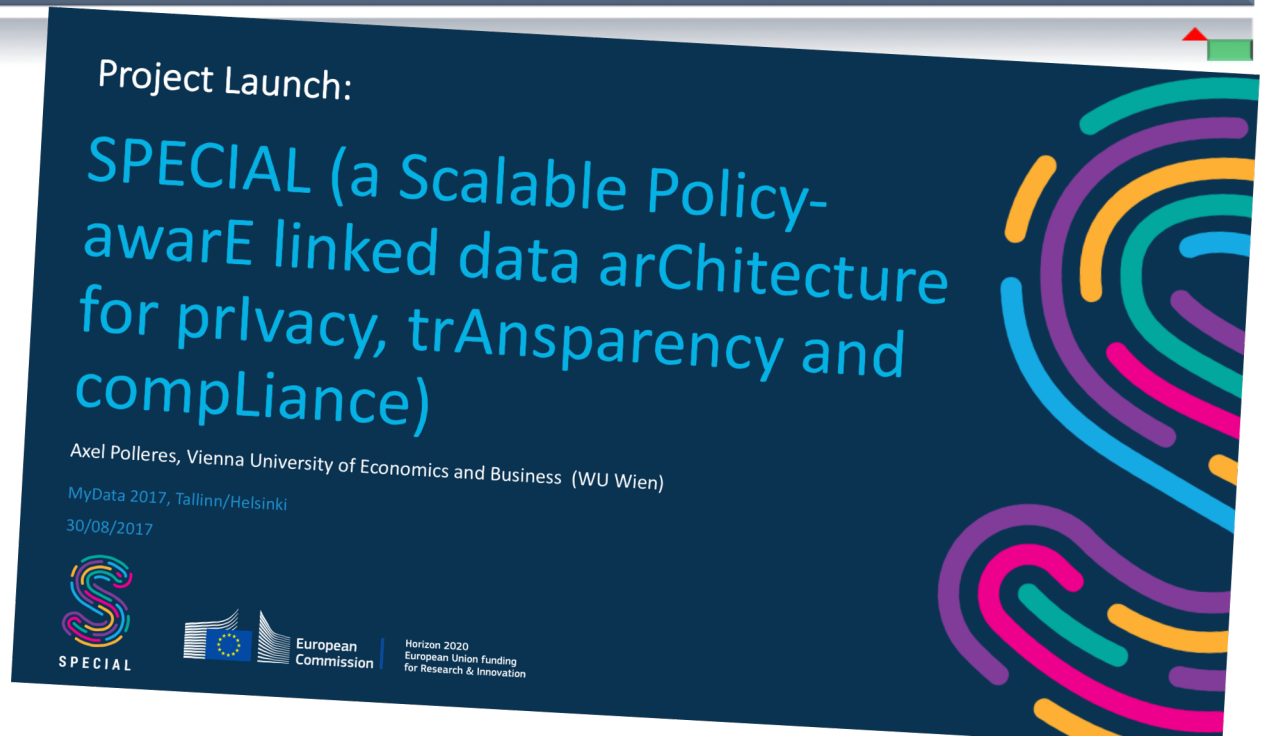
FFG
Promoting Innovation.

Federal Ministry
Republic of Austria
Transport, Innovatic
and Technology

Federal Ministry
Republic of Austria
Digital and
Economic Affairs

Background…

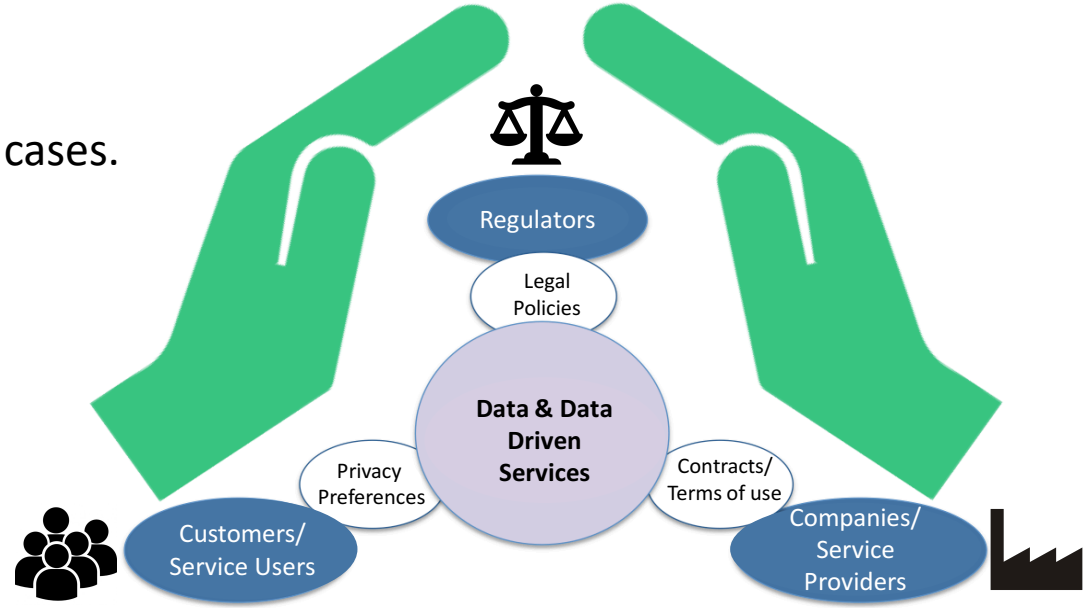

Trilogue starts
6/24/2015
Draft of the regulation
7/22/2012
EU Council finalises the chapters
8/6/2015
Comes into force
5/15/2018
Revisions in the draft
3/12/2013
Discussions in the EU Council
5/19/2014
Trilogue agrees
12/17/2015
2013
2014
2015
2016
2017
2018


Project Launch:
SPECIAL (a Scalable Policy-awarE linked data arChitecture for prIvacy, trAnsparency and compLiance)
Axel Polleres, Vienna University of Economics and Business (WU Wien)
MyData 2017, Tallinn/Helsinki
30/08/2017
SPECIAL
European Commission
Horizon 2020
European Union funding for Research & Innovation

Use Cases for Transparency and Interoperability in Privacy:

- Companies: Ensuring Regulatory Compliance for Companies
- Regulators: Checking and enforcing GDPR
- *Data Subjects: Personal Data Markets: from "Data Collection" to "Data Donations"*

Different roles have different use cases.



Components of Personal Data Processing (not exhaustive…)

Rules/Policies

- Consent
- Regulations

Purpose

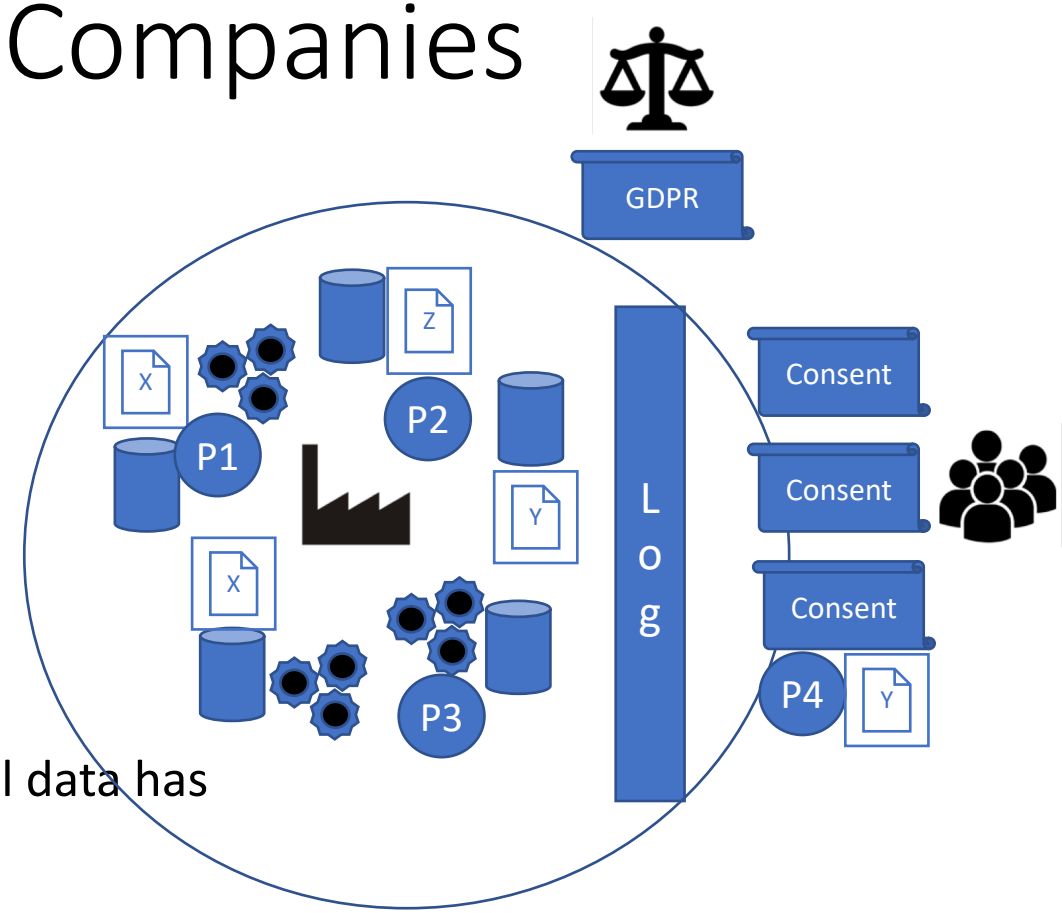
Processing

Storage

Personal Data (categories, formats)

Regulatory Compliance for Big Companies

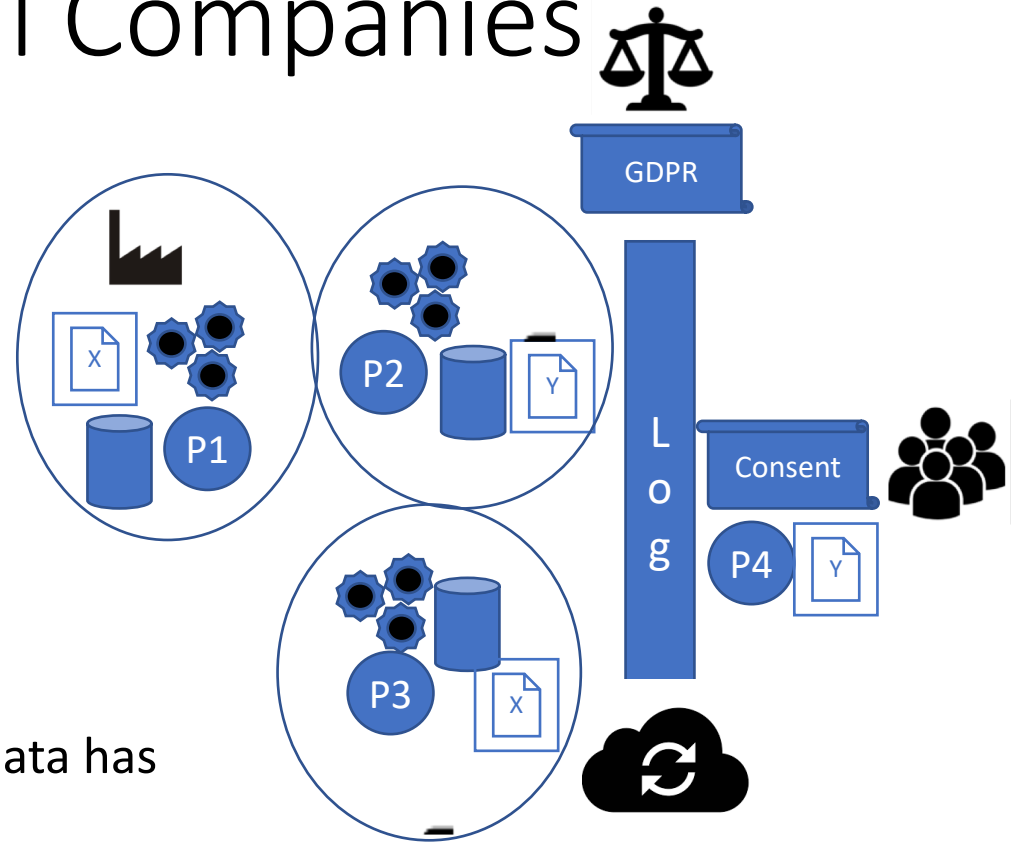
- Many heterogeneous systems that process personal data
- Potentially many different places that store and hold consent
- How to deal with GDPR data requests at scale?
- How to prove to the regulator and to the customer that personal data has been handled in compliance to consent only?



$$P4 \stackrel{?}{\subseteq} P2$$

Regulatory Compliance for Small Companies

- No resources to build their own compliance infrastructure
- How to deal with GDPR data requests at scale?
- How to prove to the regulator and to the customer that personal data has been handled in compliance to consent only?



$$P4 \stackrel{?}{\subseteq} P2$$

Semantic Interoperability boils down to:

- What is a common core to address these use cases?
- How do we benefit them all at the same time?

W3C

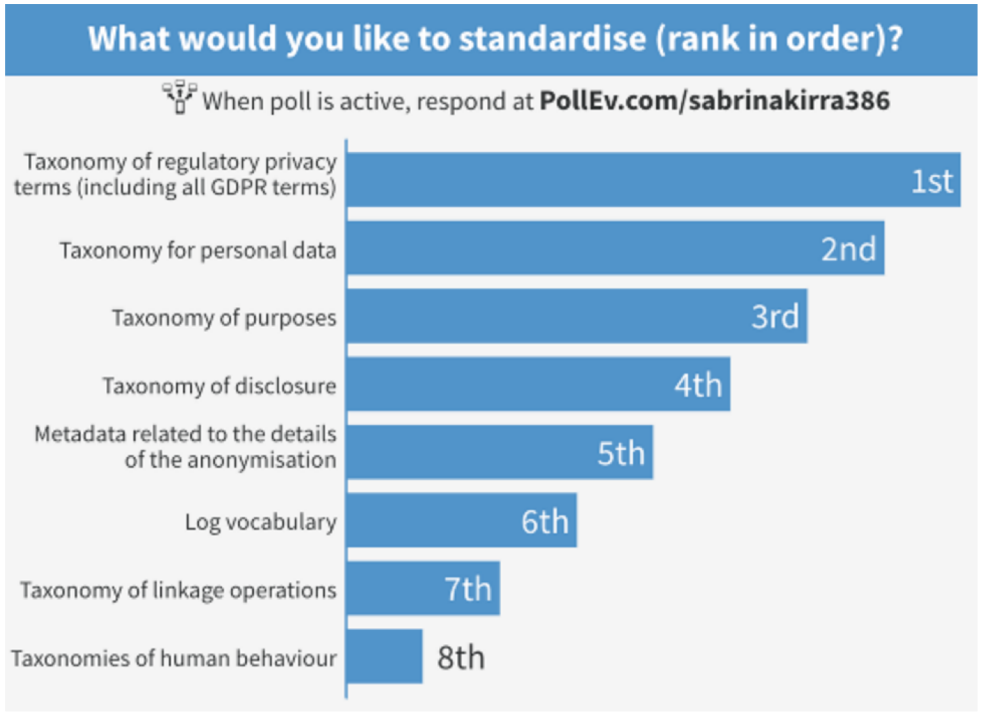
Data Privacy Controls and Vocabularies

A W3C Workshop on Privacy and Linked Data

17–18 April 2018, WU Vienna, Vienna, Austria, Europe https://www.w3.org/2018/vocabws/report.html

Semantic Interoperability boils down to:
- What is a common core to address these use cases?
- How do we benefit them all at the same time?

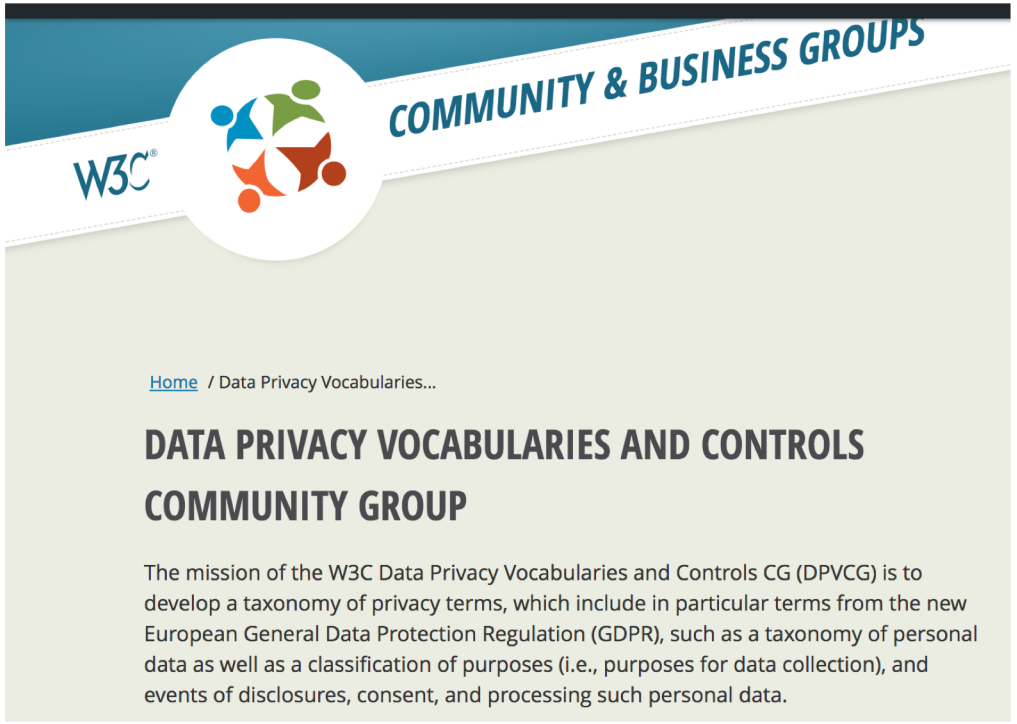
Rough workshop outcome / scoping:



1. Taxonomy of regulatory privacy terms (including all GDPR terms).
2. Taxonomy for personal data.
3. Taxonomy of purposes.
4. Taxonomy of disclosure/processing.
5. Metadata (e.g. related to processing details of anonymization)
6. Log vocabulary.
7. Taxonomy of linkage operations.
8. Taxonomies of human behavior.

Semantic Interoperability boils down to:
- What is a common core to address these use cases?
- How do we benefit them all at the same time?

→ **Foundation of a W3C Community Group (25th May 2018)**



→ Collect concrete **Use cases**
→ Collect **Existing Vocabularies**

→ **Align Core Vocabularies**

1. Taxonomy of regulatory privacy terms (including all GDPR terms).
2. Taxonomy for personal data.
3. Taxonomy of purposes.
4. Taxonomy of disclosure/processing.
5. Metadata (e.g. related to processing details of anonymization)
6. Log vocabulary.
7. Taxonomy of linkage operations.
8. Taxonomies of human behavior.

Semantic Interoperability boils down to:
- What is a common core to address these use cases?
- How do we benefit them all at the same time?

→ Foundation of a W3C Community Group (25th May 2018)

→ Collect concrete **Use cases**
→ Collect **Existing Vocabularies**

→ **Align Core Vocabularies**

Use-Cases [edit]

- SPECIAL/Proximus use case - personalized touristic recommendations
- SPECIAL/DT use case - mobile network quality measurements
- SPECIAL/TR use case - 'Know Your Customer' (finance, anti-money-laundering)
- DECODE/DEC01 use case - Online voting system with privacy
- DECODE/DEC02 use case - Rental Register
- DECODE/DEC03 use case - Sharing sensor data

⋮

We need your input!
→Join DPVCG!

Vocabularies [edit]

- CDMM Consent Ontology
- COEL
- Data Protection Ontology by Bartolini et. al
- GDPRov
- GDPRtEXT
- IEEE 7012
- ODRL
- P3P
- P-Plan
- Privacy Preference Ontology
- PROV-O
- SPECIAL Usage Policy
- SPECIAL Policy Log

https://www.w3.org/community/dpvcg/wiki/Use-Cases,_Requirements,_Vocabularies

Starting Point: Use Cases/Vocabularies from SPECIAL

SPECIAL

European Commission

Horizon 2020
European Union funding
for Research & Innovation

Three Distinct Use Cases:

THOMSON REUTERS

Know-Your-Customer services for the banking industry

proximus

Recommendation engine for subscribers

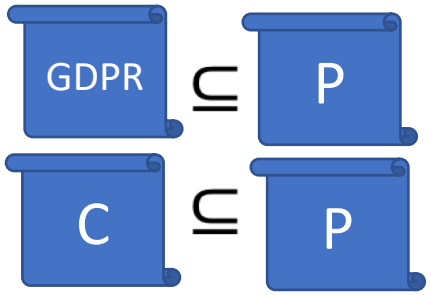
Service quality monitoring

One Compliance Solution:

Processing requires **PERMISSIONING**

Permissions must be compliant with the **GDPR**

Permissions must be compliant with **Consent**



i.e., COMPLIANCE is a logical operation

What to Standardise:

Core Logic

Core Vocabularies

Compliance Services

Against What Criteria:

Completeness and Correctness: $\subseteq$

Market adoption

SPECIAL's view on Core Interoperability Components:

Rules/Policies: ← W3C ODRL/POE

SPECIAL Usage Policy Language (SPL)

Log/Transparency

SPECIAL Policy Log Vocabulary(SPLOG)

Purposes ← W3C P3P

Processing ← W3C ODRL

Storage

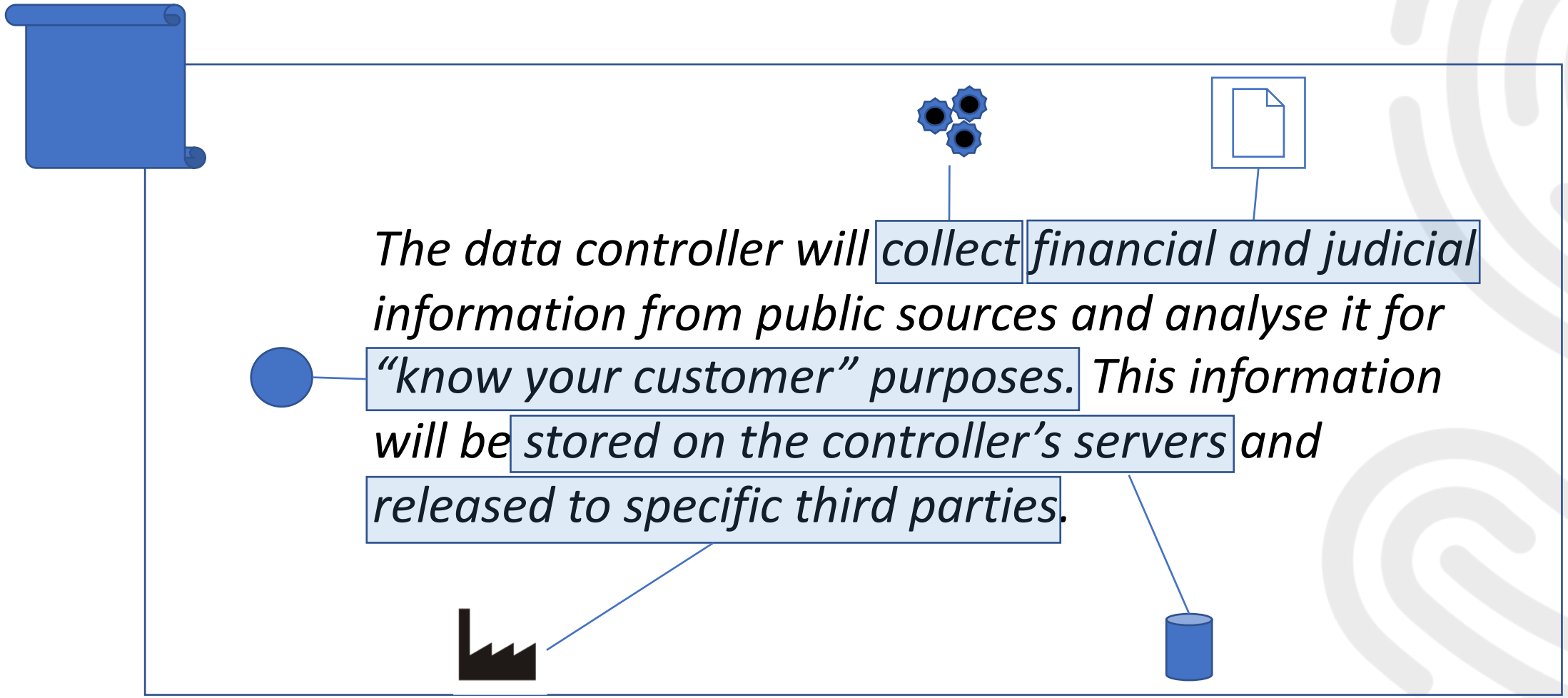
Data ← W3C P3P, (OASIS COEL?)

Recipients ← W3C P3P

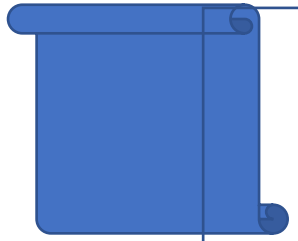
SPECIAL namespaces:

@prefix spl: <http://www.specialprivacy.eu/langs/usage-policy#>.
@prefix svpu: <http://www.specialprivacy.eu/vocabs/purposes#>.
@prefix svpr: <http://www.specialprivacy.eu/vocabs/processing#>.
@prefix svd: <http://www.specialprivacy.eu/vocabs/data#>.
@prefix svr: <http://www.specialprivacy.eu/vocabs/recipients#>.
@prefix splog: <http://www.specialprivacy.eu/langs/splog#>.
…

Use Cases/Vocabularies from SPECIAL: Example



Use Cases/Vocabularies from SPECIAL: Example (OWL)



ObjectIntersectionOf(

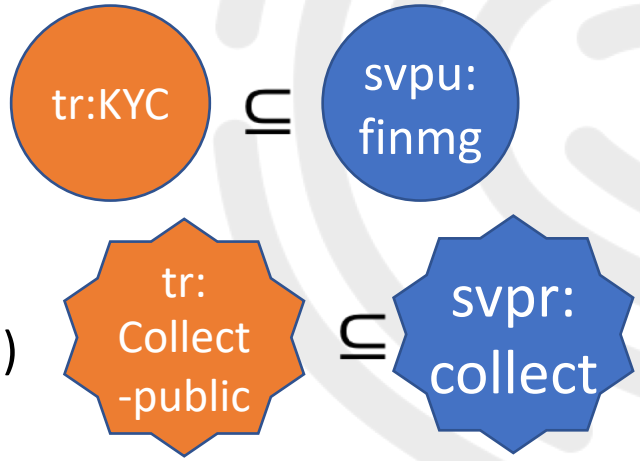
ObjectSomeValueFrom(**spl:hasData**
ObjectUnionOf(svd:Financial svd:Judicial))

ObjectSomeValueFrom(**spl:hasProcessing**
ObjectUnionOf(tr:Collect-public svpr:Analyze))

ObjectSomeValueFrom(**spl:hasPurpose** tr:KYC)

ObjectSomeValueFrom(**spl:hasStorage**
ObjectIntersectionOf(
ObjectSomeValueFrom(**spl:hasLocation** spl:ControllerServers)
DataSomeValuesFrom(**spl:durationInDays**
DatatypeRestriction(xsd:integer **xsd:mininclusive** "0"^^xsd:integer))
))

ObjectSomeValueFrom(**spl:hasRecipient** svr:AnyRecipient)
)



Discussion…
How to structure those taxonomies?
What are the important use cases to cover?

Components of Personal Data Processing (not exhaustive…)

Rules/Policies

- Consent
- Regulations

Purpose

Processing

Storage

Personal Data (categories, formats)

Personal Data (categories, formats)

the MyData model

The MyData model could be integrated on the “high level” it presents data as

- human-centrically grouped into “areas of interest”

as well as how it’s

- processed
- and used

Quelle: MyData

Personal Data (categories, formats)

the Worls Economic Forum model

Figure 10: A taxonomy of personal data by origin

TYPE	EXAMPLE		
Individually provided	• Photos • Blogs	• Emails • Tweets	• Online transaction data • Registration forms & job applications
Observed	• Internet browsing preferences	• Surveillance video • Location data	• Call detail records
Inferred	• Credit scores • Consumer profiles	• Predictive traffic flows • Patterns in the spread of infectious diseases	• Targeted advertisement

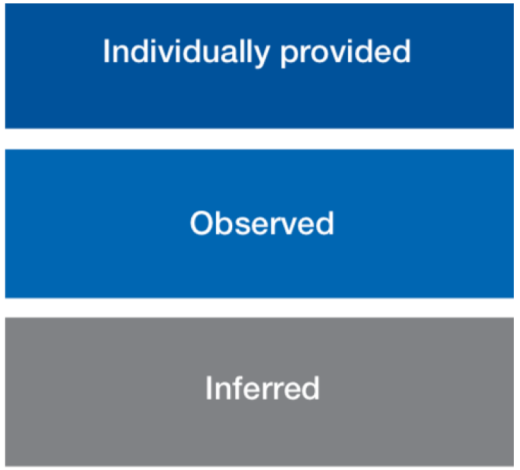
Source: Information Accountability Foundation, World Economic Forum, Marc E. Davis

Quelle: MyData
Quelle: http://reports.weforum.org/rethinking-personal-data/?doing_wp_cron=1538384793.1468729972839355468750

Personal Data (categories, formats)

How should we structure our taxonomy?

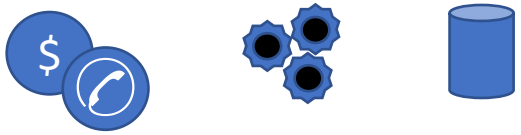
vs.



Source: Information Accountability Foundation, Worl

- Which one is more fit for purpose?
- Which one covers 80:20 use cases?

How to structure our other taxonomies?



What expressivity do we need (conditionals, etc.)?

Call for Action: Join DPVCG!

- More use cases matter!
- Existing efforts for interoperability/vocabularies matter!

Joining is easy!

→The group is Open to everyone!

→Just create a W3C account

CLICK HERE ↓

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

- Looking forward to discussions…