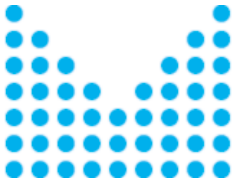
European Union
European Social Fund
Operational Programme Employment

Guaranteed and semantically interoperable public government data published and shared as Linked Open Data

Jakub Klímek, Martin Nečaský, Petr Křemen



MINISTRY OF THE INTERIOR
OF THE CZECH REPUBLIC

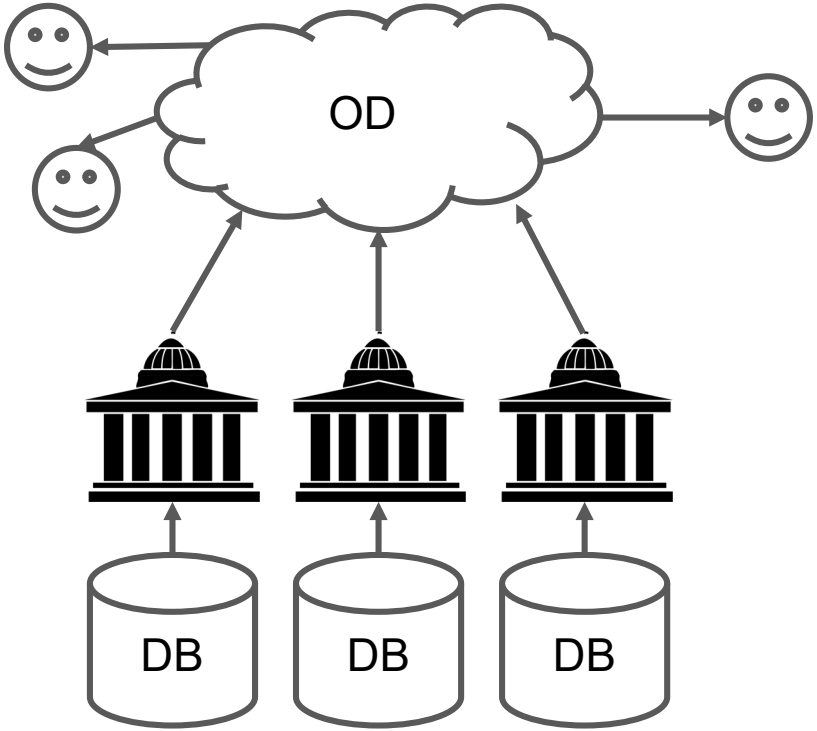
FACULTY
OF MATHEMATICS
AND PHYSICS
Charles University

FACULTY
OF ELECTRICAL
ENGINEERING
CTU IN PRAGUE

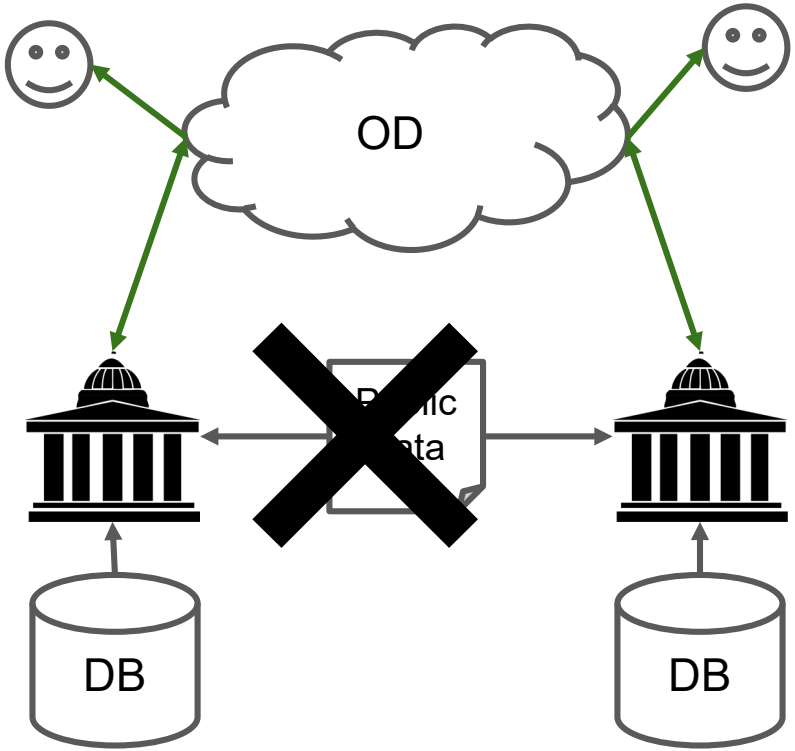
Project: Development of data policies in the field of improving the quality and interoperability of public administration data. Reg. number: CZ.03.4.74/0.0/0.0/15_025/0013983

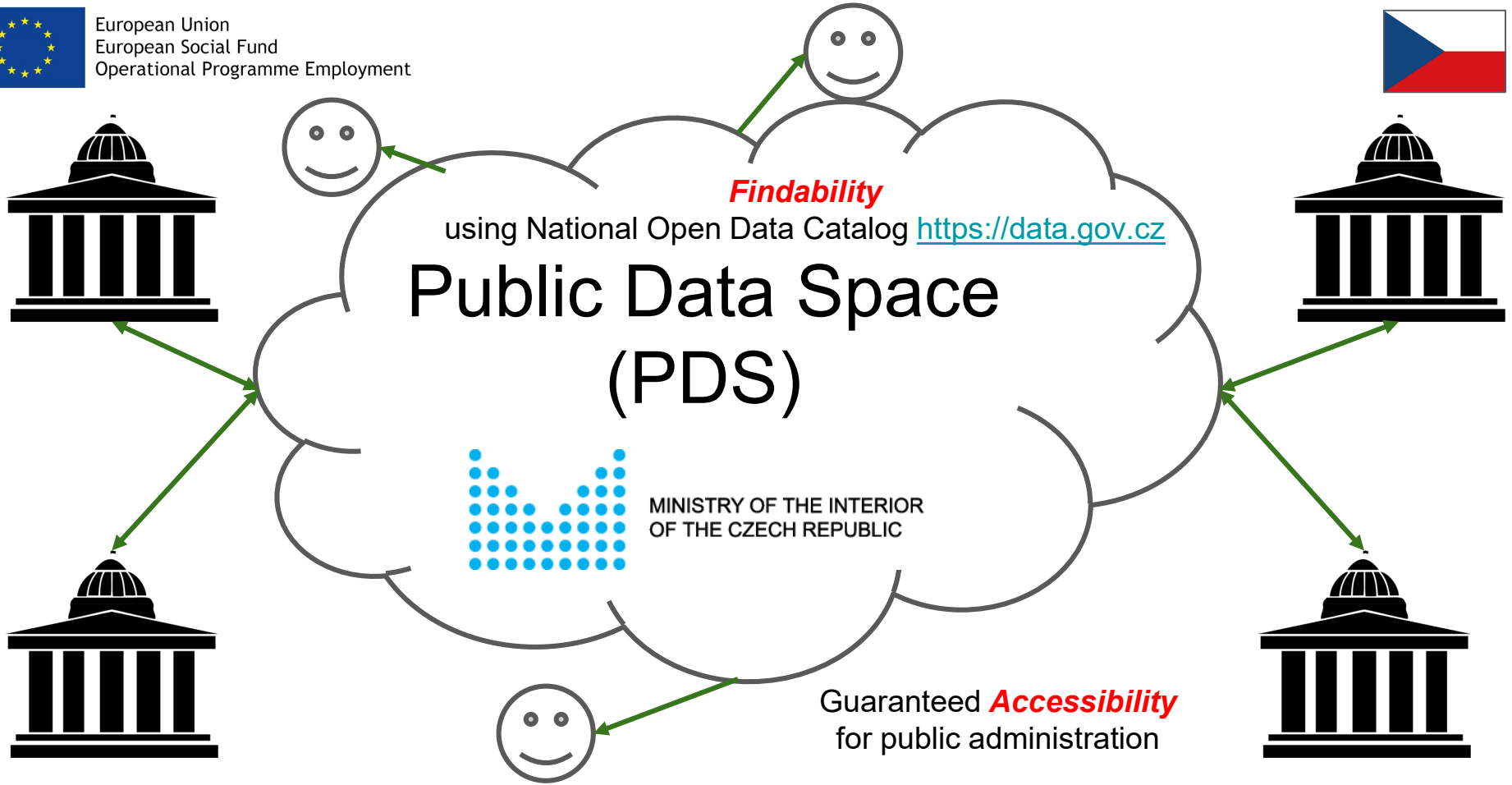
2021 legislation update in Czechia

Obligation to **publish** data as open data by default

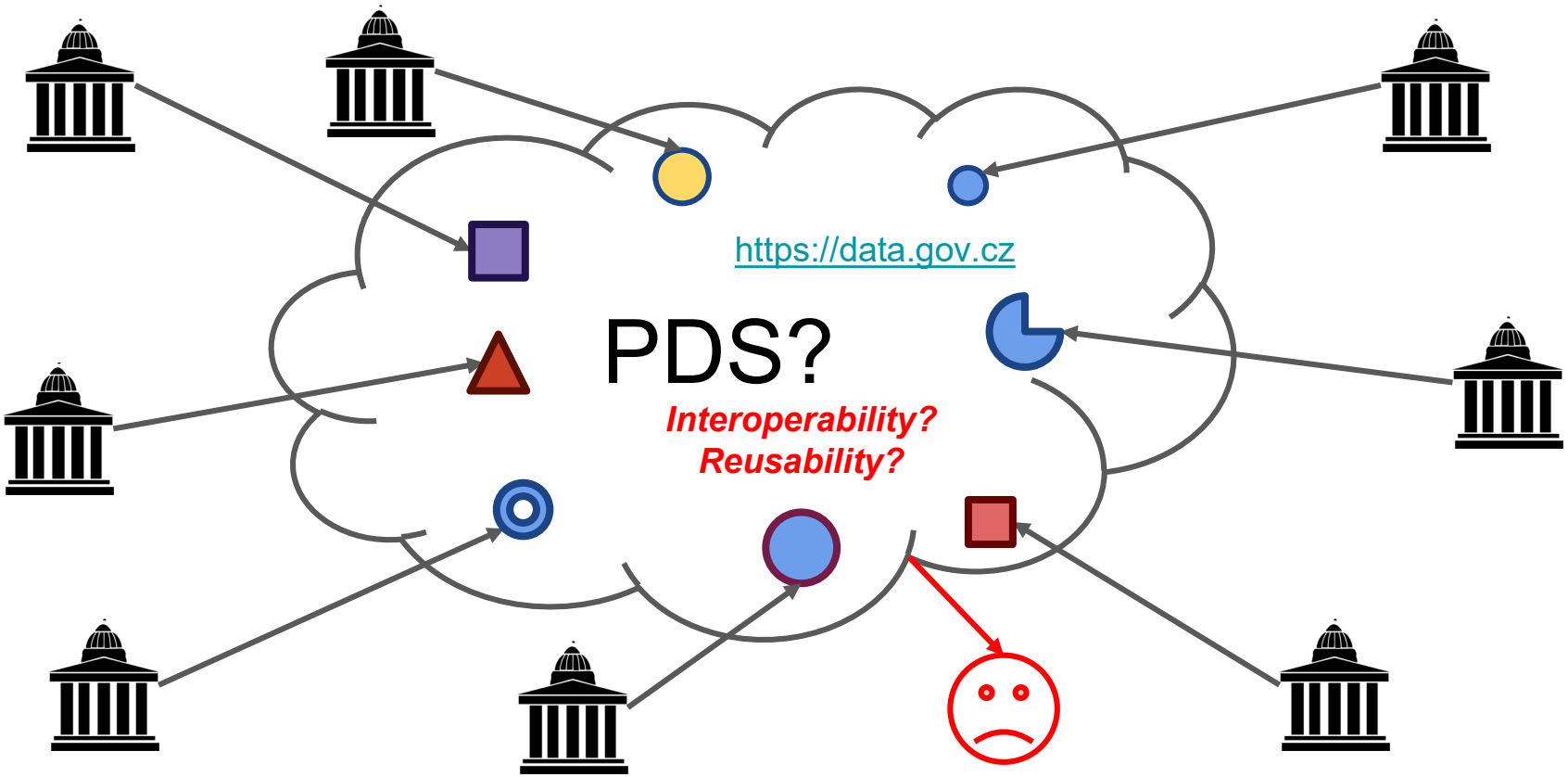


Obligation to **consume** open data




Findability
using National Open Data Catalog https://data.gov.cz
Public Data Space
(PDS)
MINISTRY OF THE INTERIOR
OF THE CZECH REPUBLIC
Guaranteed Accessibility
for public administration

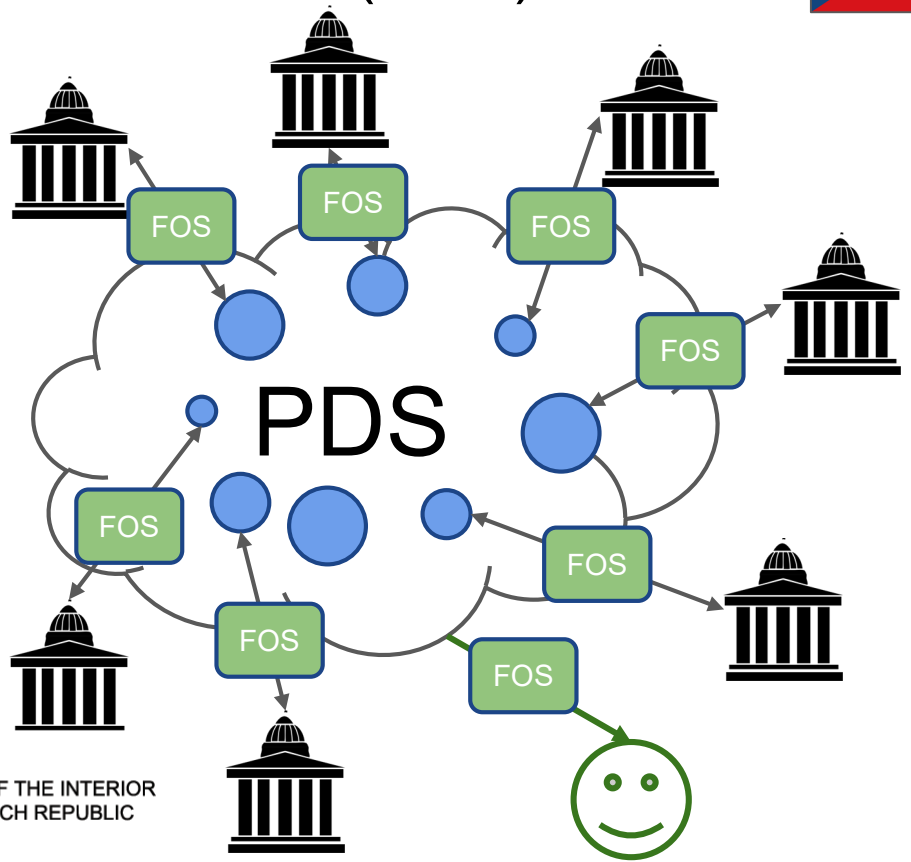
PDS: regular open data not enough



Formal open standard (FOS)

EU DIRECTIVE 2019/1024
Article 2 (15)

'formal open standard' means a standard which has been laid down in written form, detailing specifications for the requirements on how to ensure software interoperability

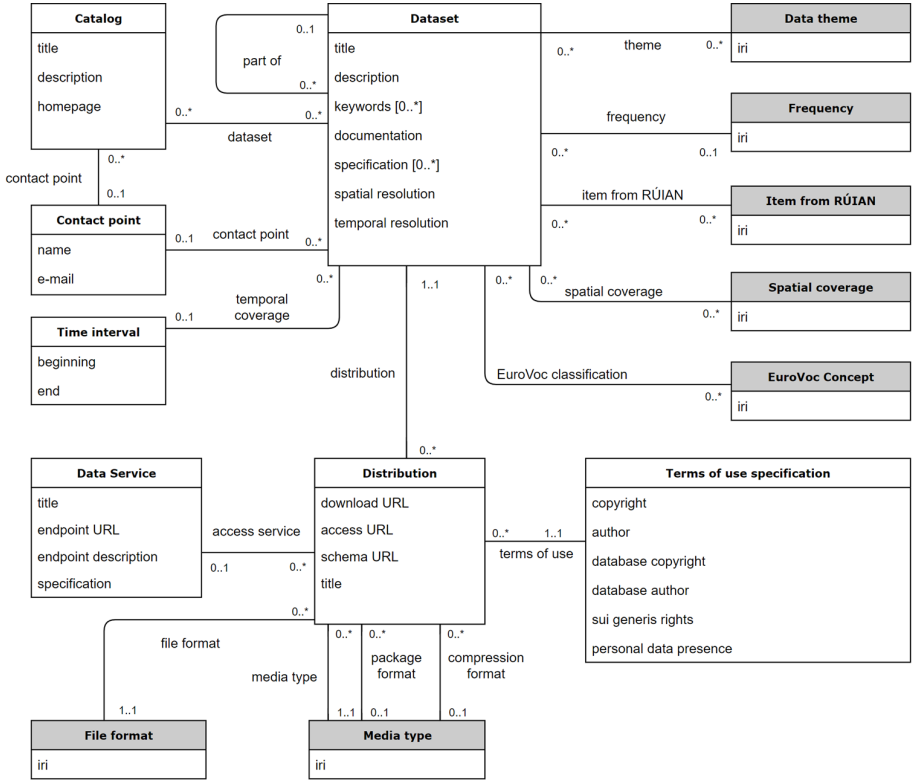


In Czech, FOSes are called
Otevřené formální normy (OFN)
and they are coordinated by

MINISTRY OF THE INTERIOR
OF THE CZECH REPUBLIC

FOS - what it contains

1. Conceptual data model
   - data format independent



FOS - what it contains

1. Conceptual data model
   - data format independent
2. Specification
   - detailed description of classes, attributes and relations
   - reused code-lists

RDF Class:	dcat:Catalog
Definition:	A curated collection of metadata about resources (e.g., datasets and data services in the context of a data catalog)
Sub-class of:	dcat:Dataset
Usage note:	A Web-based data catalog is typically represented as a single instance of this class.
See also:	§ 6.5 Class: Catalog Record, § 6.6 Class: Dataset

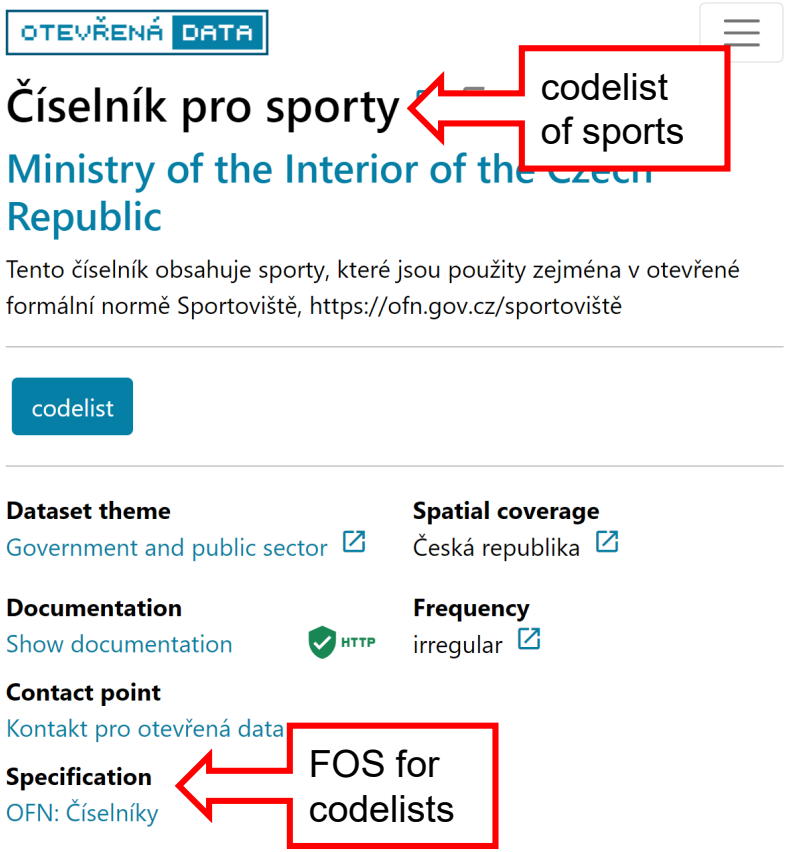
FOS - what it contains

1. Conceptual data model
   - data format independent
2. Specification
   - detailed description of classes, attributes and relations
   - reused code-lists
3. Examples and data schemas
   - JSON(-LD) + JSON Schema
   - possibly CSV + CSV on the Web
   - possibly XML + XML Schema


```
{
    "@context": "https://ofn.gov.cz/rozhraní-katalogů-otevřených-dat/2021-01-11/kontexty/rozhraní-katalogů-otevřených-dat.jsonld",
    "iri": "https://data.gov.cz/lkod/mdcr/datové-sady/vld",
    "typ": "Datová sada",
    "název": {
        "cs": "Jízdní řády veřejné linkové dopravy",
        "en": "Public transport timetables"
    },
    "popis": {
        "cs": "Obsahem datové sady jsou schválené a aktuálně platné jízdní řády veřejné linkové dopravy postoupené do Celostátního informačního systému o jízdních řádech ve strojově zpracovatelném formátu.",
        "en": "This dataset contains approved timetables and timetables in effect for public transport entered into the state-wide timetable information system."
    },
    "poskytovatel": "https://rpp-opendata.egon.gov.cz/odrpp/zdroj/orgán-veřejné-moci/66003008",
    "téma": [
        "http://publications.europa.eu/resource/authority/data-theme/TRAN"
    ],
    "periodicita_aktualizace": "http://publications.europa.eu/resource/authority/frequency/WEEKLY_3",
    "klíčové_slovo": {
        "cs": [
            "jízdní řády",
            "veřejná linková doprava",
            "autobus"
        ],
        "en": [
            "timetable",
            "public transport",
            "bus"
        ]
    },
```

FOS - what it contains

1. Conceptual data model
   - data format independent
2. Specification
   - detailed description of classes, attributes and relations
   - reused code-lists
3. Examples and data schemas
   - JSON(-LD) + JSON Schema
   - possibly CSV + CSV on the Web
   - possibly XML + XML Schema
4. Predefined metadata
   - DCAT-AP
   - for National Open Data Catalog

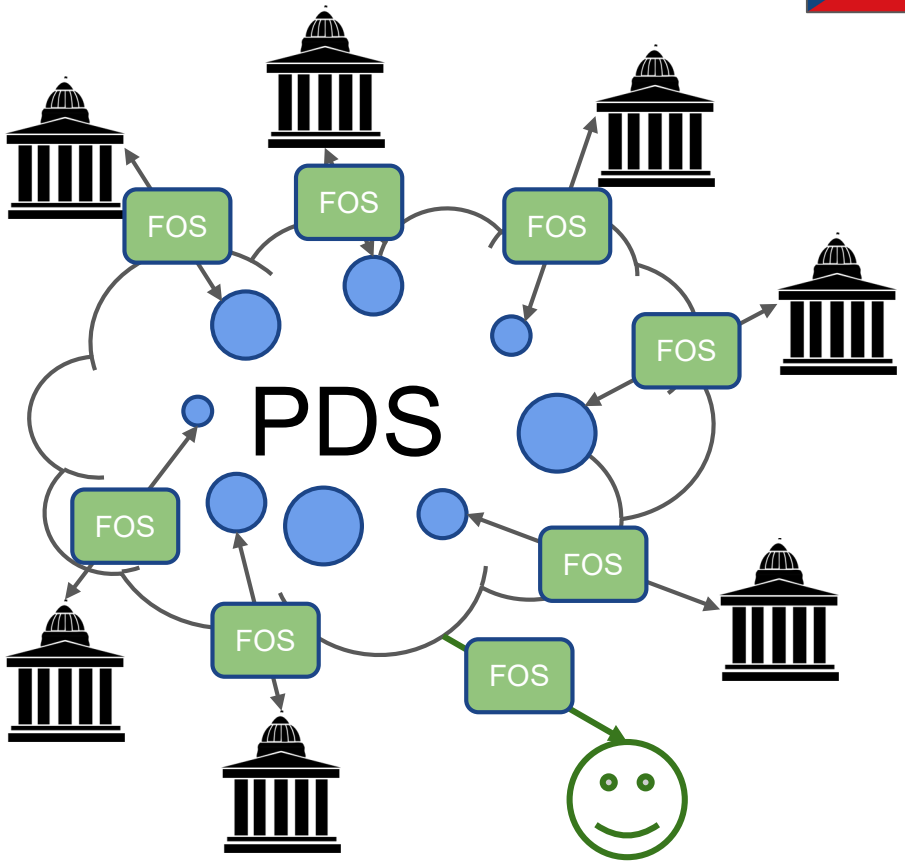


FOS - why is it useful

Formal open standard (FOS) function

- Obviously: *syntactic* interoperability among various publishers
- But also: protection of interoperability
  - When system supplier changes
  - When system operator changes
- Moreover: faster Linked Data adoption
  - JSON in fact JSON-LD => LD
  - CSV + CSVW => LD
- Compliance with FOS => LD

Interoperability
Reusability

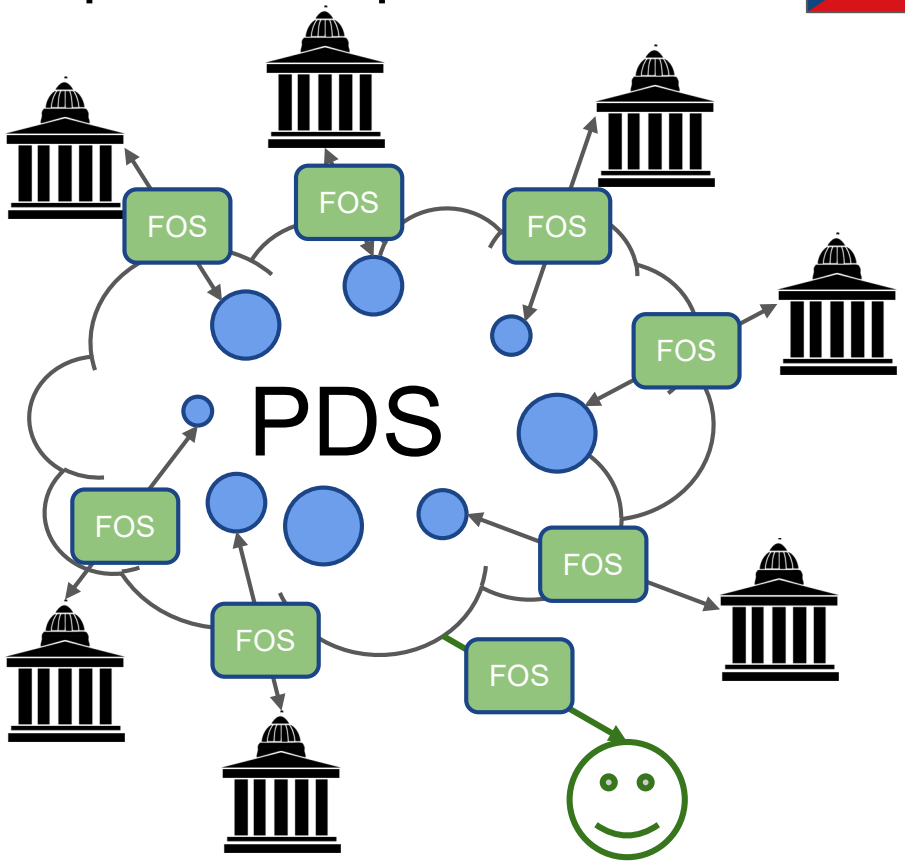


Public Data Space recap

What we have so far in Public Data Space (PDS)

- Open data
- ***Findability*** in National Open Data Catalog
  - and therefore data.europa.eu
- Guaranteed ***accessibility*** for public administration
- ***Syntactic interoperability*** + ***reusability***
  - JSON-LD => Linked Data
  - Formal Open Standards (FOSes)

What about ***semantic interoperability***?



Semantic mapping of FOSes

Data structures in FOSes can be mapped to multiple semantic definitions

```
{
    "@context": "https://ofn.gov.cz/turistické-cíle/2020-07-01/kontexty/turistický-cíl.jsonld",
    "@context": "https://ofn.gov.cz/turistické-cíle/2020-07-01/kontexty/turistický-cíl.ssp.jsonld",
    "typ": "Turistický cíl",
    "iri": "https://www.spilberk.cz/",
    "název": {
        "cs": "Hrad Špilberk",
        "en": "Špilberk Castle"
    },
```

Context 1: Common Linked Data vocabularies

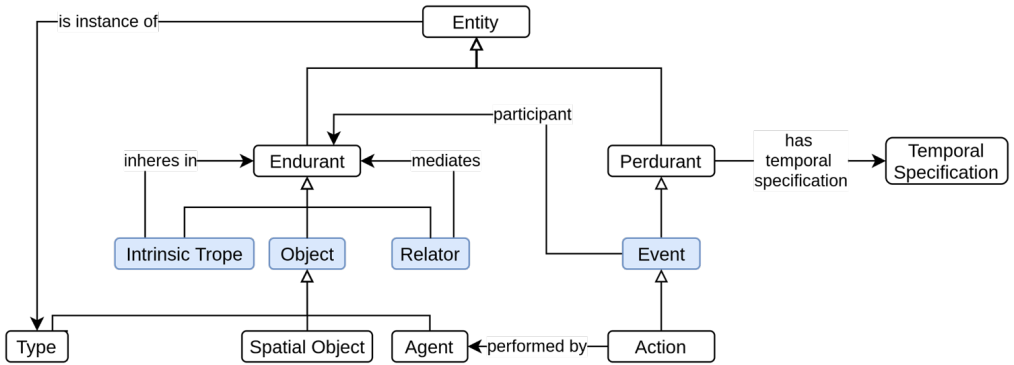
```
@prefix schema: <http://schema.org/> .
@prefix dcterms: <http://purl.org/dc/terms/> .
<https://www.spilberk.cz/> a schema:TouristAttraction ,
    dcterms:title "Hrad Špilberk"@cs, "Špilberk Castle"@en .
```

Context 2: SGoV - semantic government vocabulary

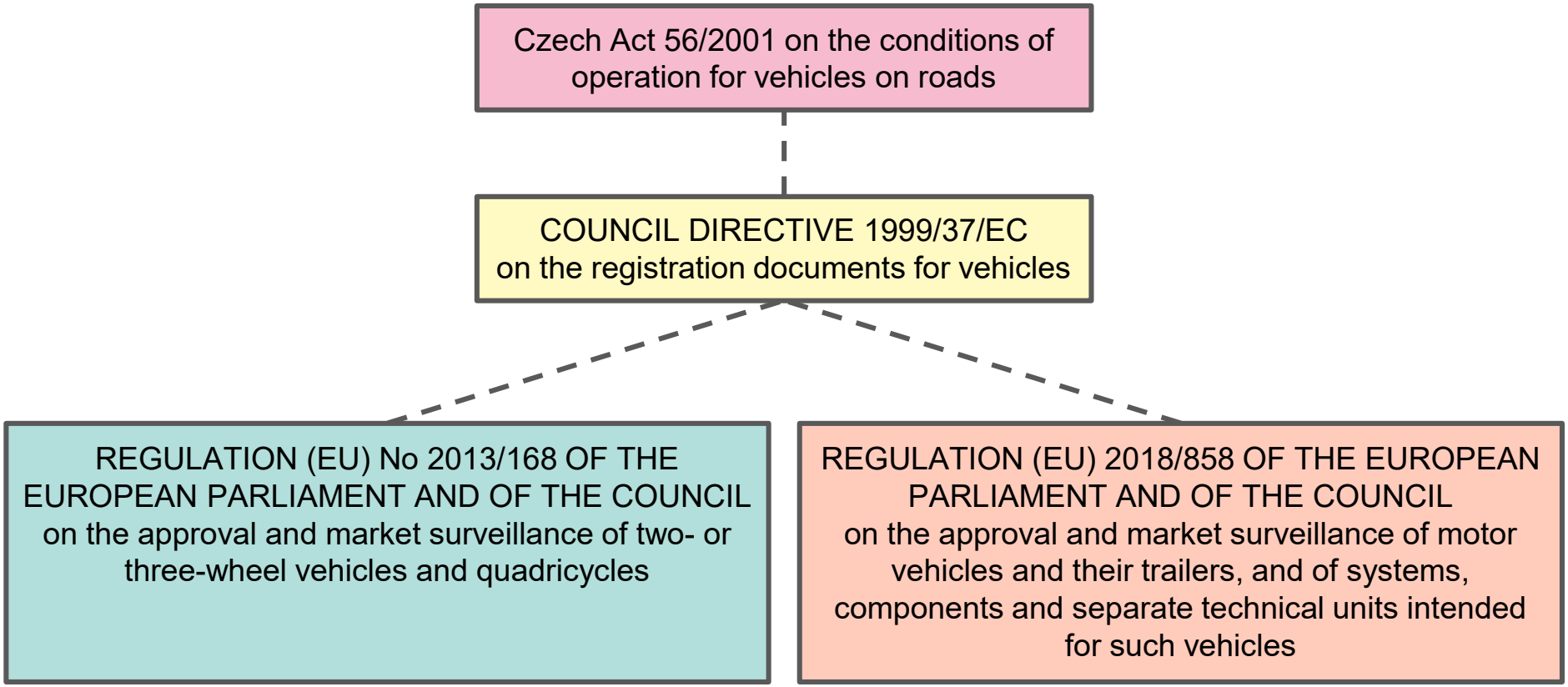
```
@prefix cíle: <https://slovník.gov.cz/datový/turistické-cíle/pojem/> .
@prefix věci: <https://slovník.gov.cz/generický/věci/pojem/> .
<https://www.spilberk.cz/> a cíle:turistický-cíl ;
    věci:název "Hrad Špilberk"@cs, "Špilberk Castle"@en .
```

Semantic government vocabulary (SGoV)

- ecosystem of layered linked ontologies **based on** Czech **legislation**
  - e.g. one per Act
  - e.g. one per specific domain
  - ...
- semantic definitions of
  - concepts
  - relationships
- mappings among concepts from different domains
  - including ISA² Core Vocabularies and other LD vocabularies
- based on UFO-A, UFO-B, UFO-MLT
  - + OntoUML as modeling language
  - inspired by Giancarlo Guizzardi

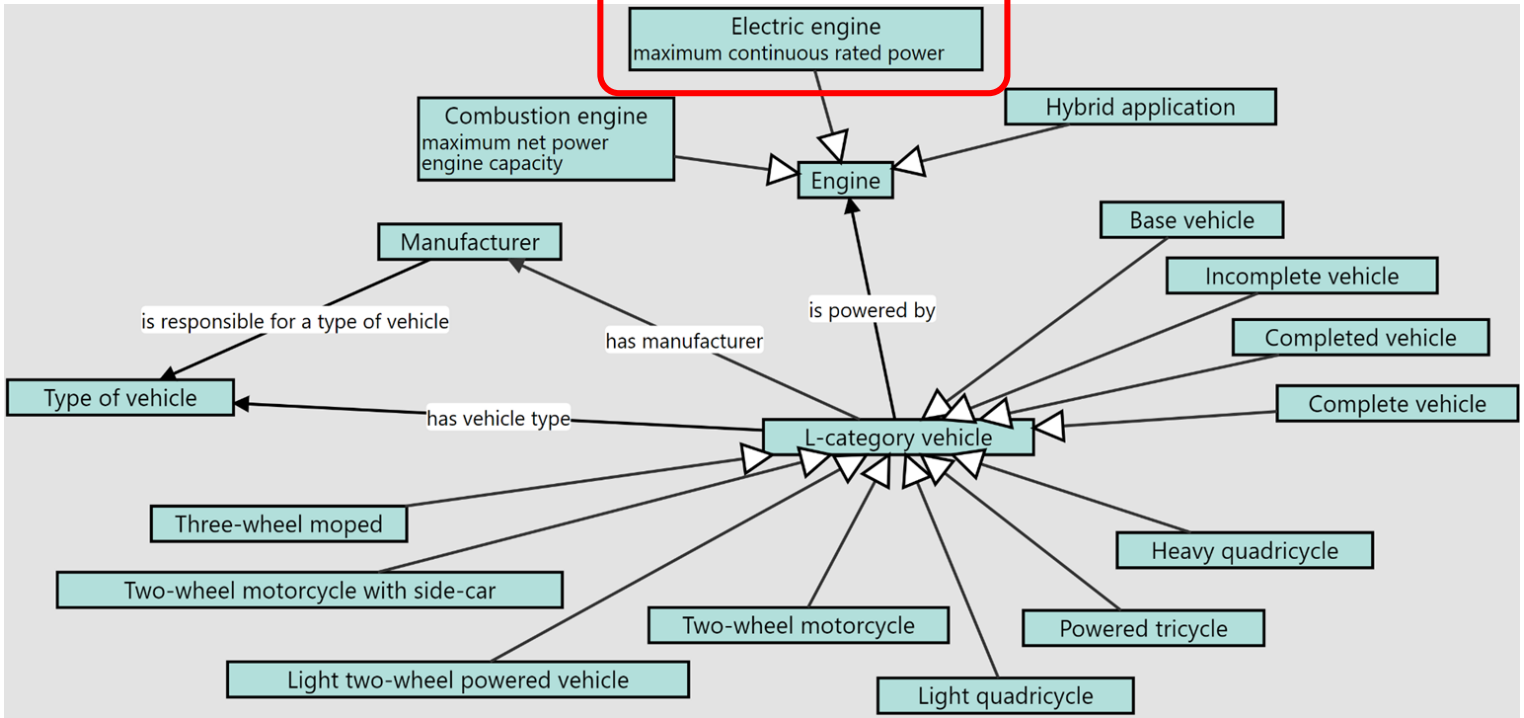


SGoV example: Vehicles



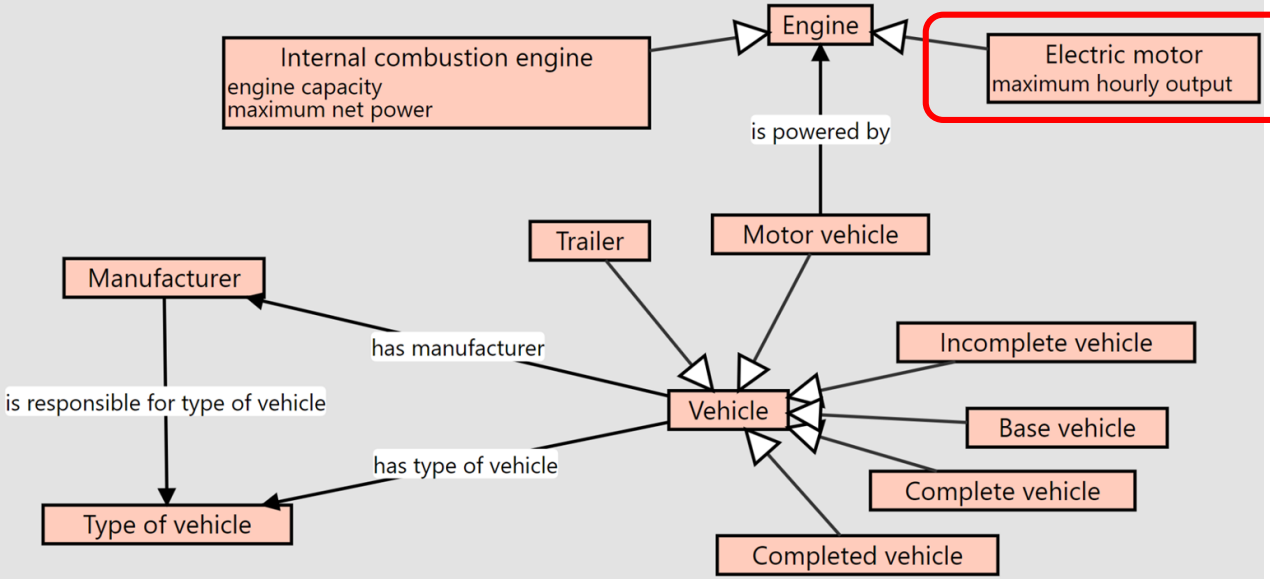
SGoV example: Vehicles: EU reg. 1

REGULATION (EU) No 2013/168 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on the approval and market surveillance of two- or three-wheel vehicles and quadricycles



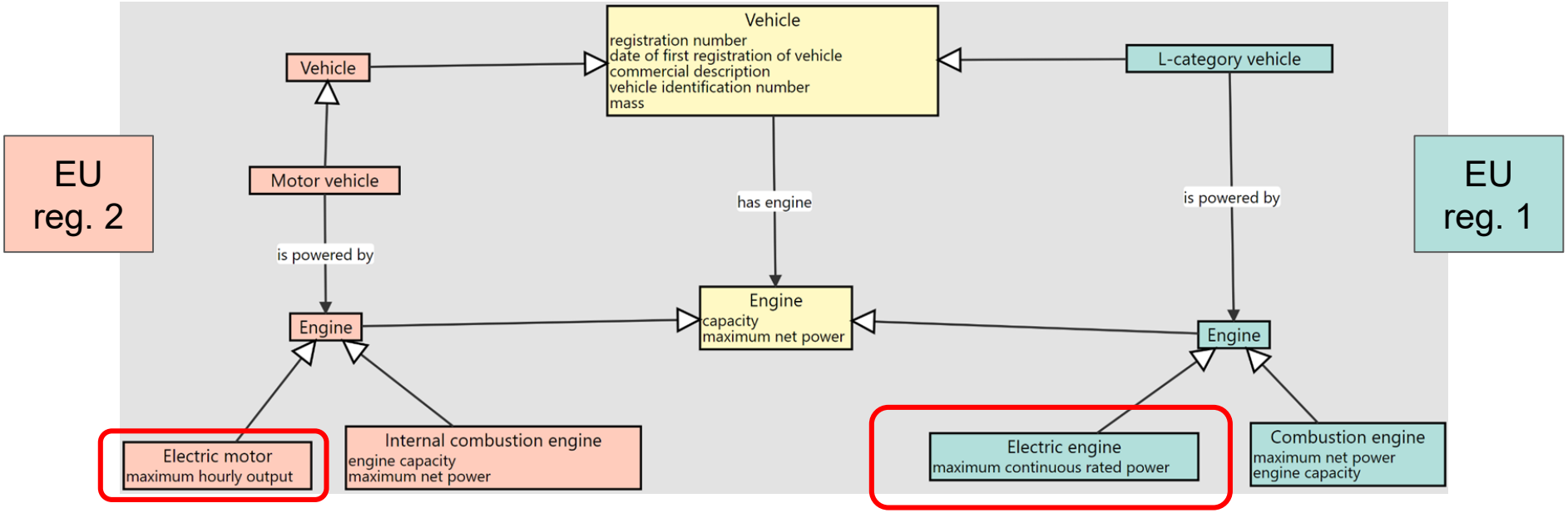
SGoV example: Vehicles: EU reg. 2

REGULATION (EU) 2018/858 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL on the approval and market surveillance of motor vehicles and their trailers, and of systems, components and separate technical units intended for such vehicles



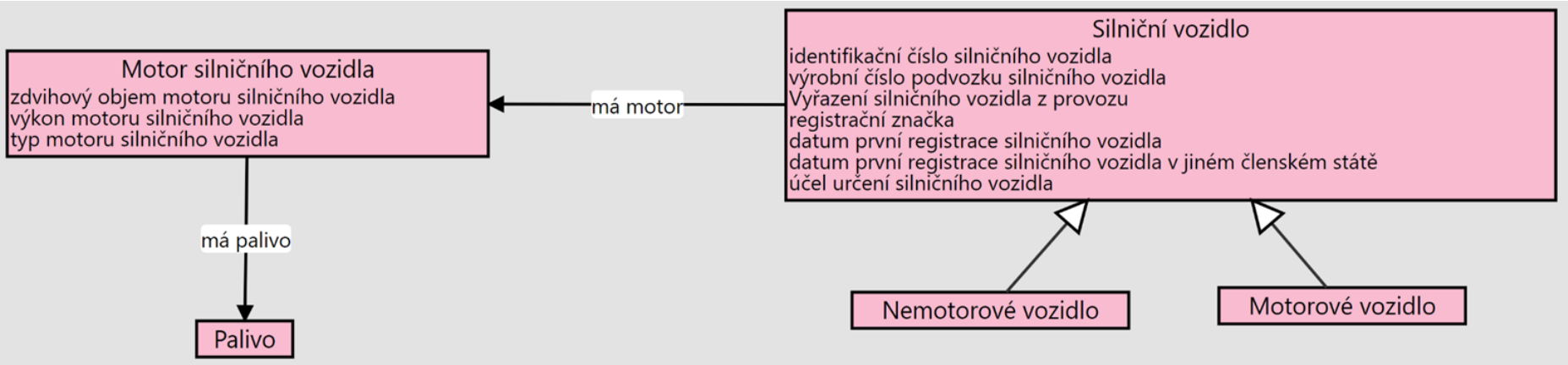
SGoV example: Vehicles: EU directive

COUNCIL DIRECTIVE 1999/37/EC on the registration documents for vehicles



SGoV example: Vehicles: Czech act

Czech Act 56/2001 on the conditions of operation for vehicles on roads



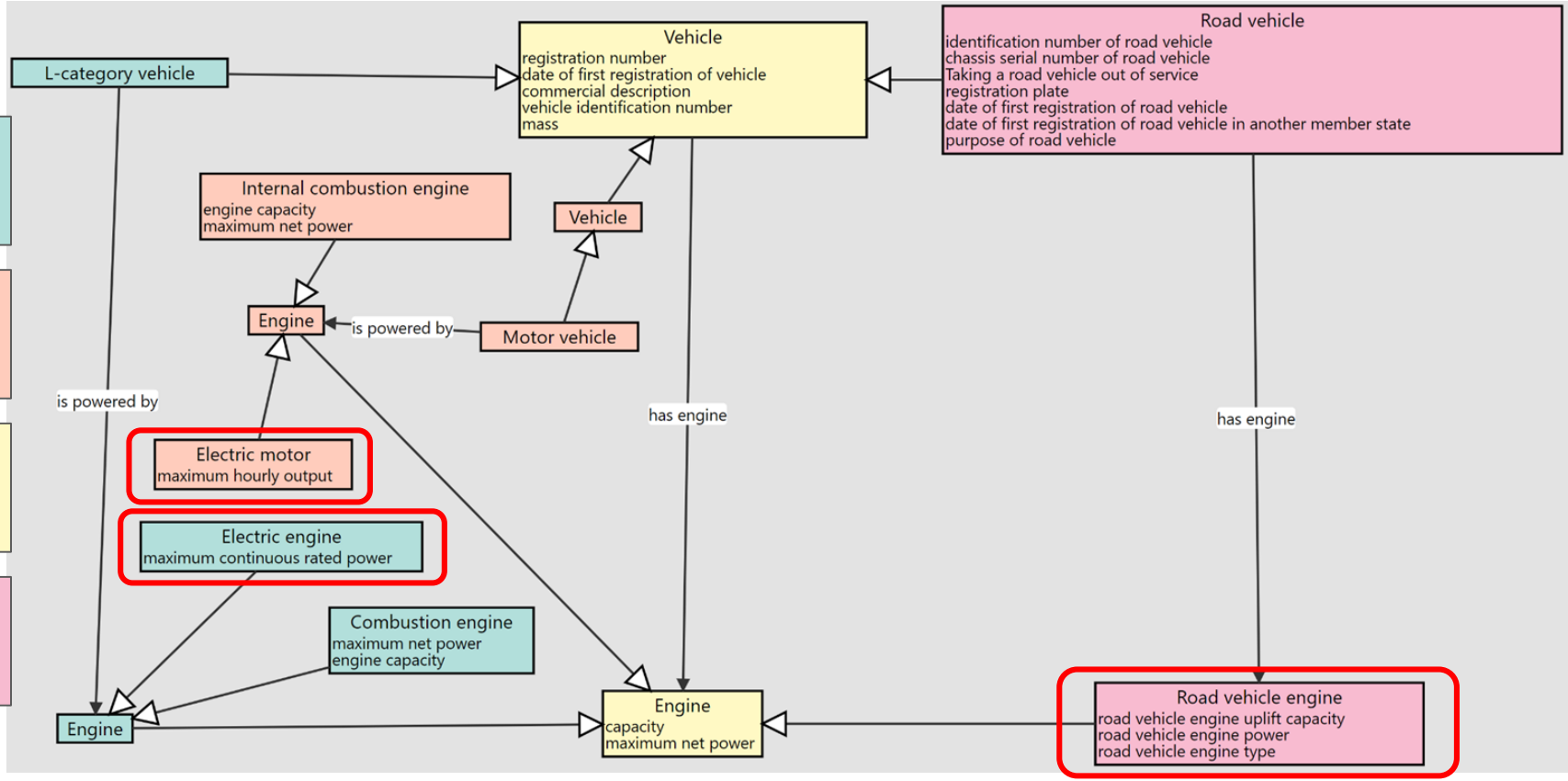
CZ <-> EU legislation mapping (en)

EU reg. 1

EU reg. 2

Council directive

Czech act



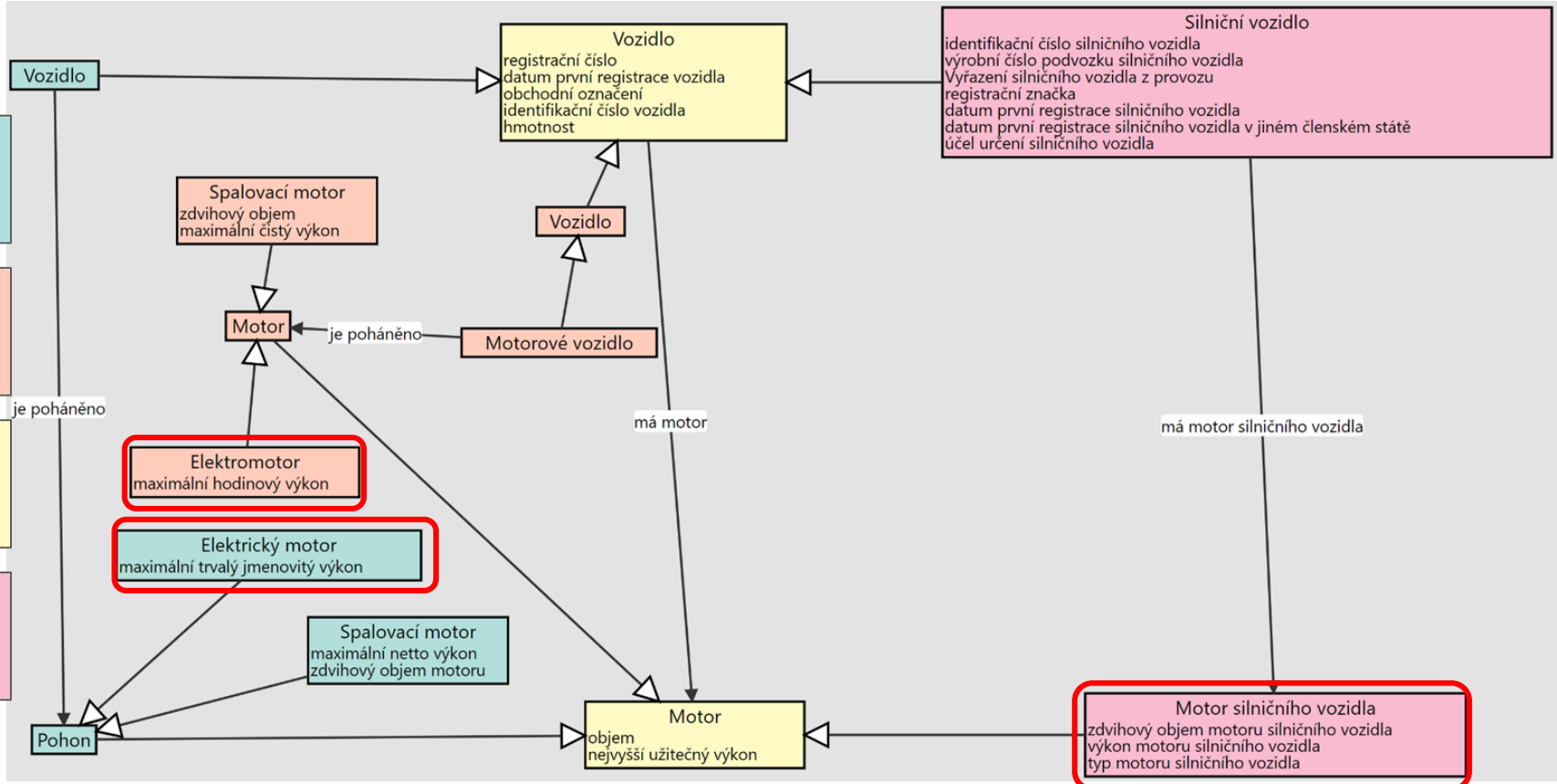
CZ <-> EU legislation mapping (cs)

EU reg. 1

EU reg. 2

Council directive

Czech act

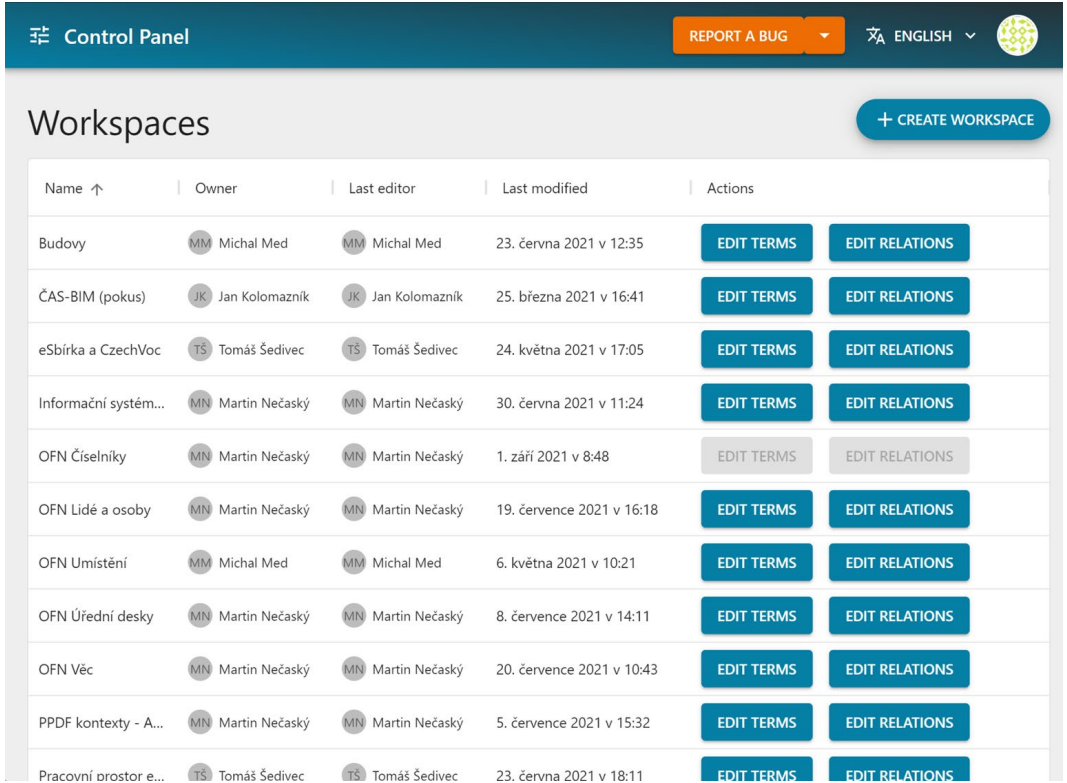


SGoV tools and workflow

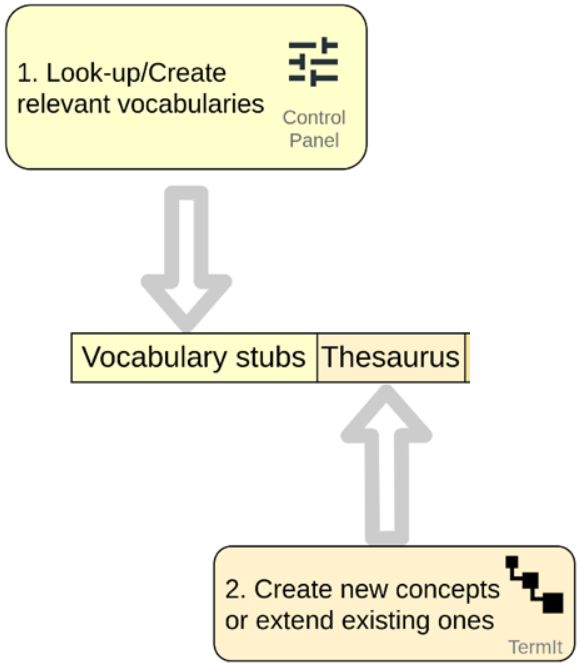
1. Look-up/Create relevant vocabularies

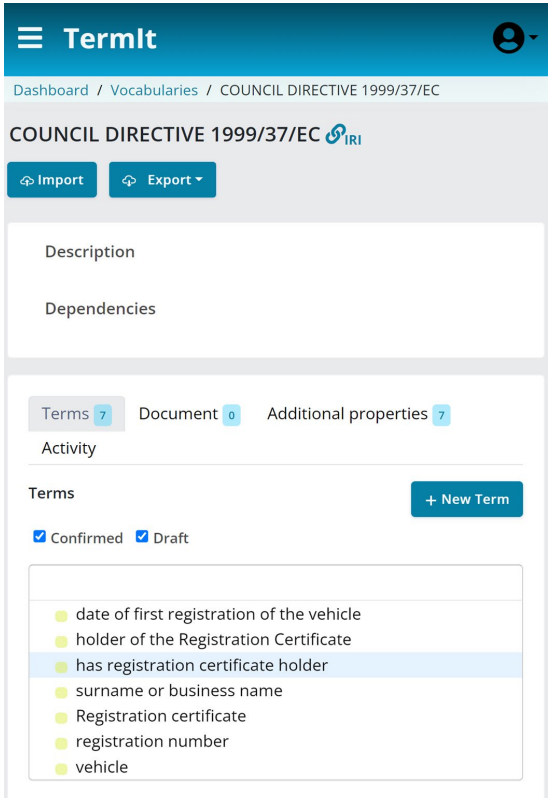
Control Panel

Vocabulary stubs

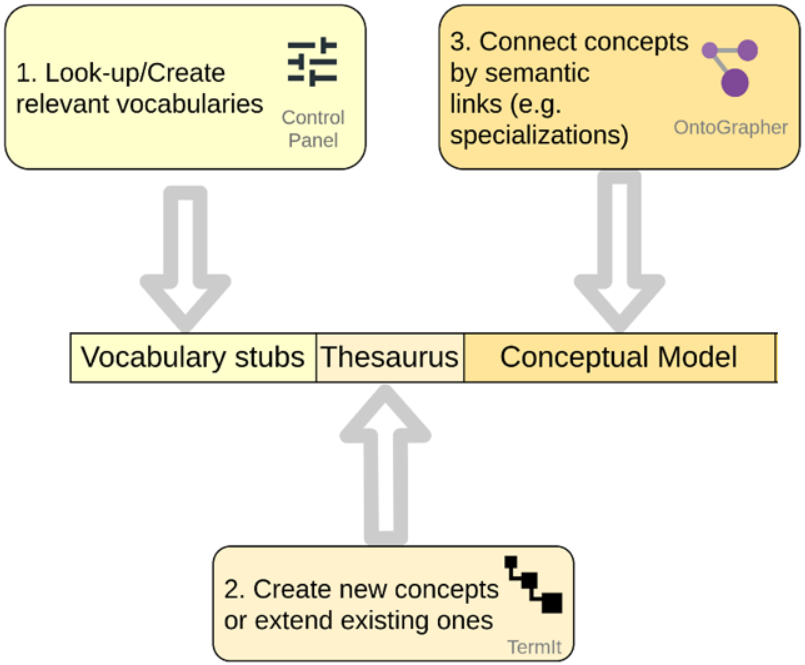


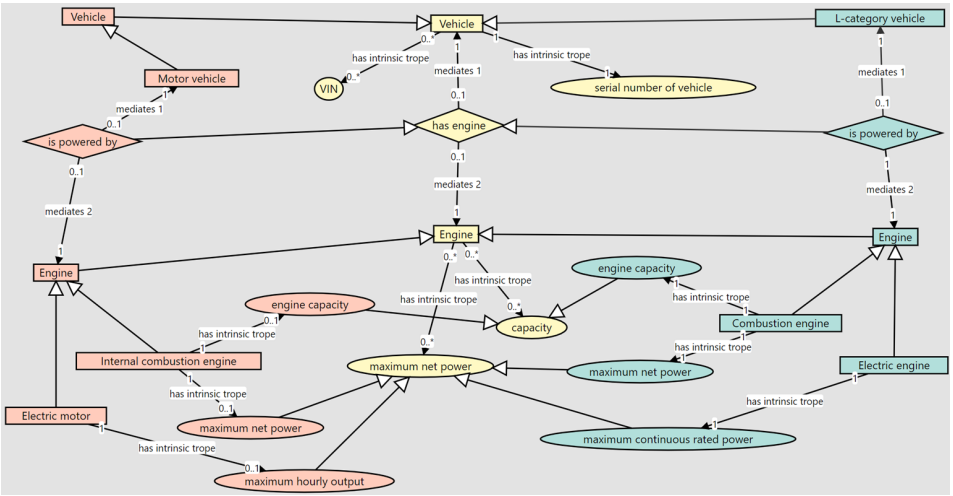
SGoV tools and workflow



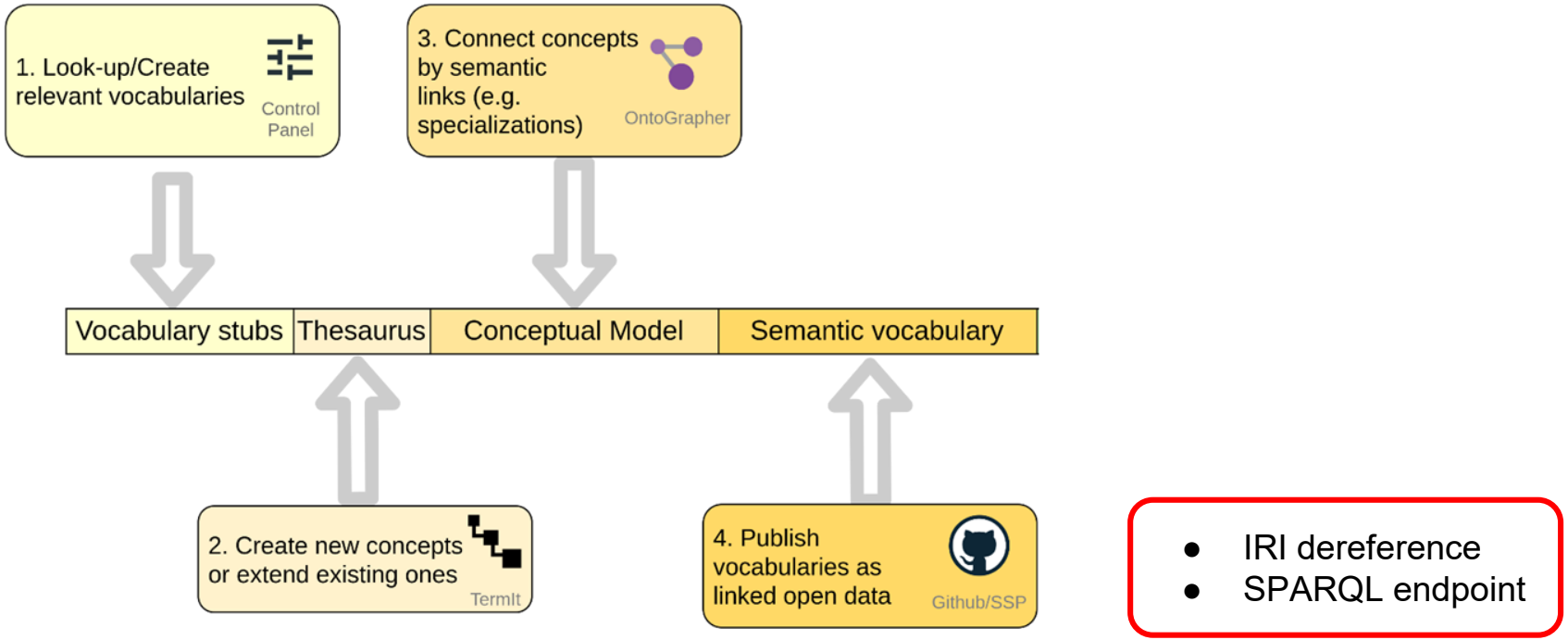


SGoV tools and workflow



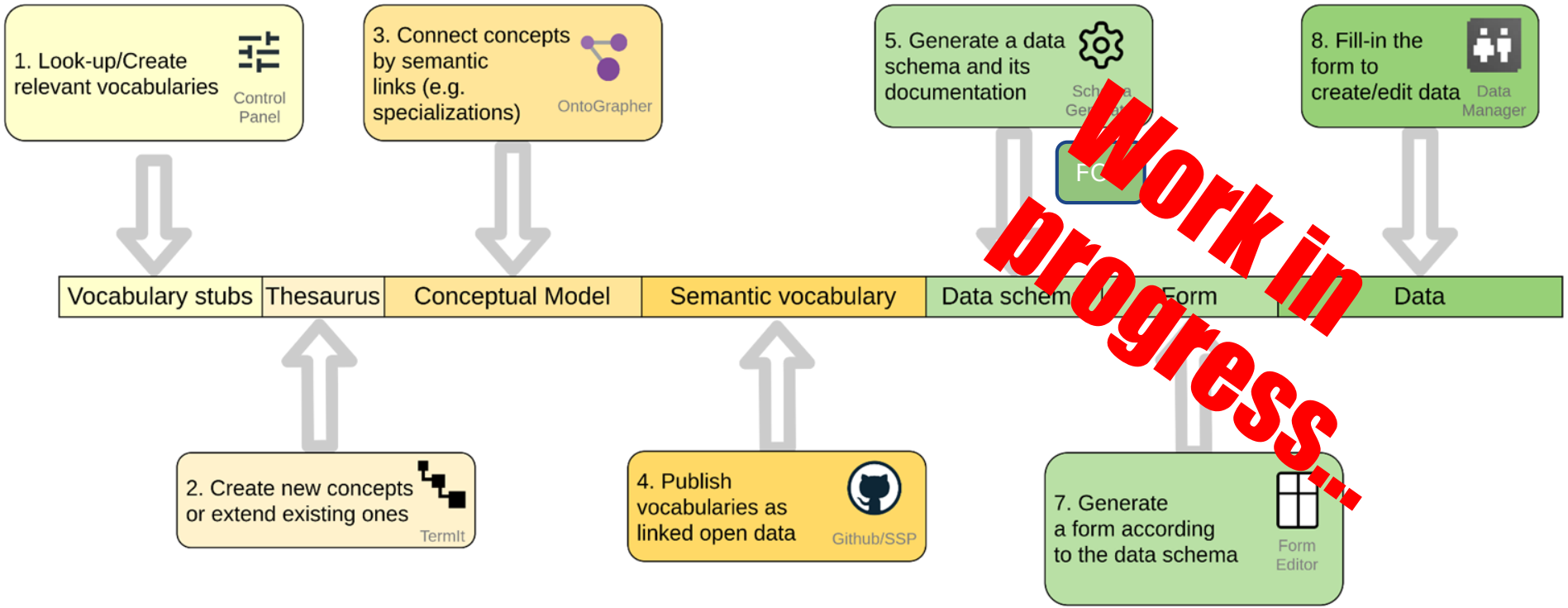


SGoV tools and workflow

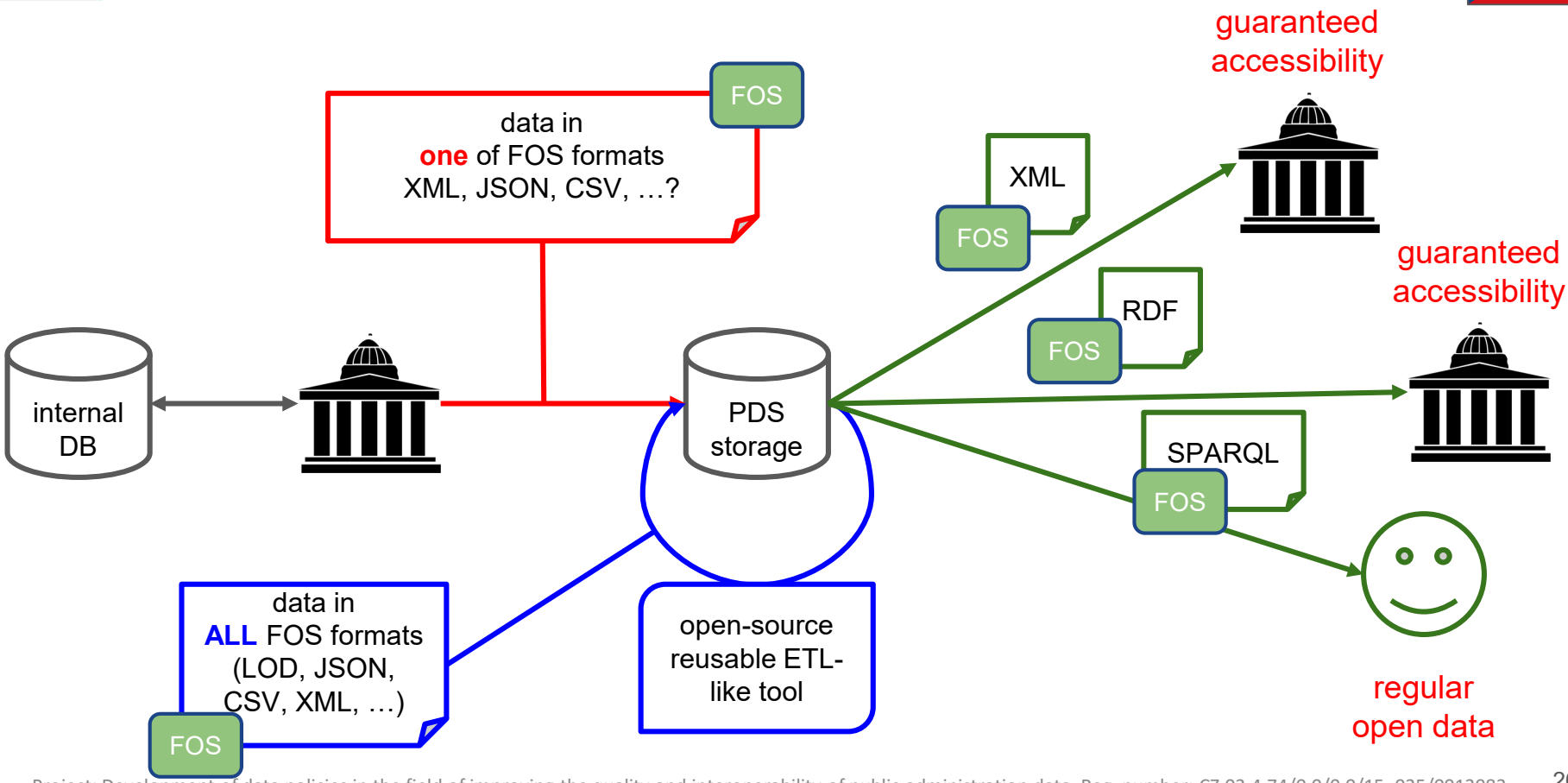


- IRI dereference
- SPARQL endpoint

SGoV tools and workflow



Public Data Space (PDS) architecture



Current state of PDS

- Public Data Space (PDS) architecture is done
- Allocated finances from NextGenerationEU
- Ongoing preparations for implementation

- First datasets to populate PDS will be codelists
- FOS for codelists based on SKOS

MINISTRY OF THE INTERIOR
OF THE CZECH REPUBLIC

Vision of extension to EU context

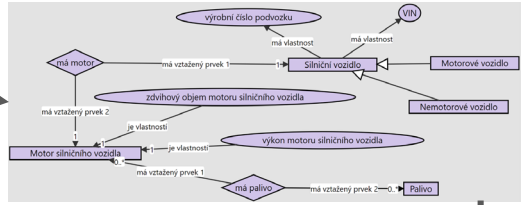
Czech data about vehicle registrations

```
{
    "@context": "https://ofn.gov.cz/vozidla/2022-12-24/kontexty/vozidla.jsonld",
    "typ": "Vozidlo",
    "iri": "https://vozidla.cz/vin/YBHNASD123458791",
    "název": {
        "cs": "Moje nové elektroauto",
        "en": "My new electric vehicle"
    },
```

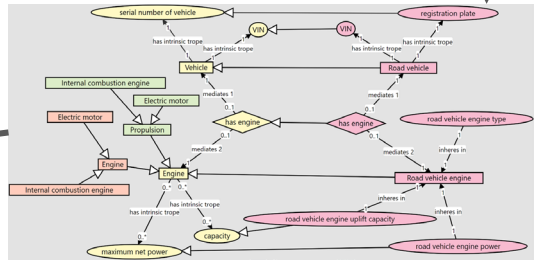
FOS

automatic, semantically interoperable transformations

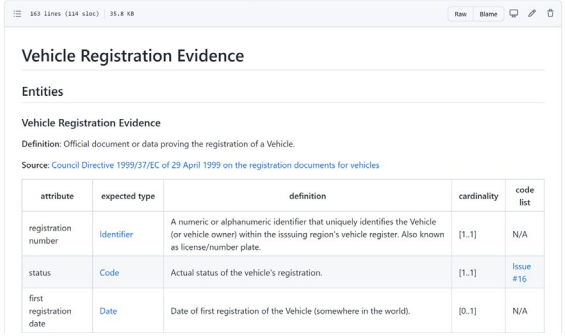
Czech vehicles act (SGoV)



CZ <-> EU legislation mapping (SGoV)



SEMIC Vehicle registration evidence (SDG)



Vehicle Registration Evidence

Entities

Vehicle Registration Evidence

Definition: Official document or data proving the registration of a Vehicle.

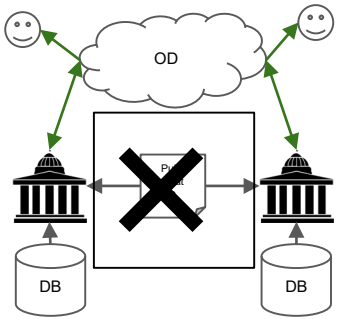
Source: Council Directive 1999/37/EC of 29 April 1999 on the registration documents for vehicles

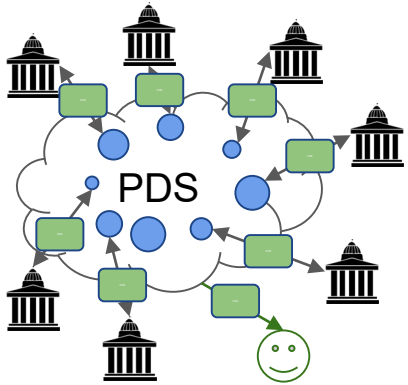
attribute	expected type	definition	cardinality	code list
registration number	Identifier	A numeric or alphanumeric identifier that uniquely identifies the Vehicle (or vehicle owner) within the isssuing region's vehicle register. Also known as license/number plate.	[1..1]	N/A
status	Code	Actual status of the vehicle's registration.	[1..1]	Issue #16
first registration date	Date	Date of first registration of the Vehicle (somewhere in the world).	[0..1]	N/A

EU legislation based SEMIC Vocabulary mapping

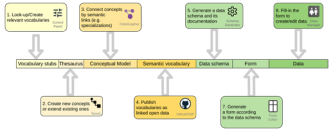
European Union
European Social Fund
Operational Programme Employment

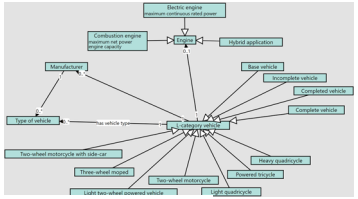
Guaranteed and semantically interoperable public government data published and shared as Linked Open Data





Questions?





Jakub Klímek, Martin Nečaský, Petr Křemen

MINISTRY OF THE INTERIOR OF THE CZECH REPUBLIC

FACULTY OF MATHEMATICS AND PHYSICS Charles University

FACULTY OF ELECTRICAL ENGINEERING CTU IN PRAGUE

Project: Development of data policies in the field of improving the quality and interoperability of public administration data. Reg. number: CZ.03.4.74/0.0/0.0/15_025/0013983