



European Union
European Social Fund
Operational Programme Employment

Semantic Government Vocabulary

Petr Křemen

Project: Implementace strategií v oblasti otevřených dat II
Reg. number: CZ.03.4.74/0.0/0.0/15_025/0004172

Keyword ambiguity



OTEVŘENÁ DATA

Datasets Publishers Keywords More ▾

Publisher (6)

Český úřad zeměměřický a katastrální (6 259)

HLAVNÍ MĚSTO PRAHA (17)

Český statistický úřad (3)

Statutární město Plzeň (2)

Ministerstvo financí (1)

Statutární město Ostrava (1)

Keyword (21)

Buildings

Budovy (6 259)

INSPIRE (6 259)

Query:

budovy

Show advanced filters

Clear filters

Title Ascending ▾

6 283 datasets found with query: "budovy"

3D budovy centrum Plzně

Statutární město Plzeň

3D model budov centra města Plzně z roku 2005 byl využit pro projekt Turistické olympiády IVV 2005 v Plzni. Více na <http://mapy.plzen.eu/aplikace-a-mapy/tematicke-kategorie/uzemni-cleneni/3d-model-plzne.aspx>

ZIP

Bloková mapa

Statutární město Ostrava

Bloková mapa - generalizovaná mapa, která vzniká zevšeobecněním obsahu technické a katastrální mapy tak, že veškeré prvky mapy jsou začleněny do kategorií (budovy, komunikace, voda, les, zeleň, prům. areály, hřbitovy) a průběh jejich hranic se zjednoduší.

ZIP

Keyword ambiguity



Building is a construction

- both above and below ground
- spatially compact
- with walls and roof
- with heating

*Act 406/2000
Coll., on Energy
Management*

Building is a construction

- above ground
- with solid foundations
- spatially compact
- with walls and roof

*Act 256/2013 Coll,
Cadastral Law*

Bus stops ?



Lack of semantic relationships between keywords

- How to find semantically interrelated datasets ?

Registry of candidate lists for Parliamentary elections 2017

Political Parties in the Chamber of Deputies

Votes in the Chamber of Deputies

Donations for elections to the Chamber of Deputies

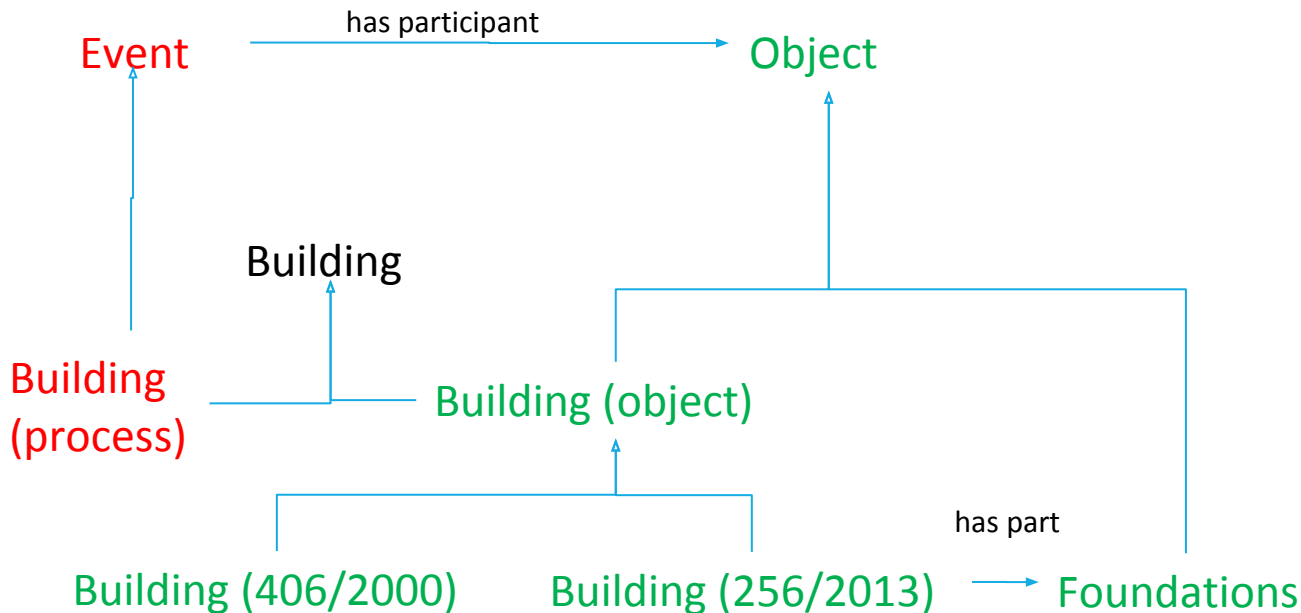
Registry of candidates for Parliamentary elections 2017

Banking accounts of political parties

Members of the Chamber of Deputies

Votes of members of the Chamber of Deputies

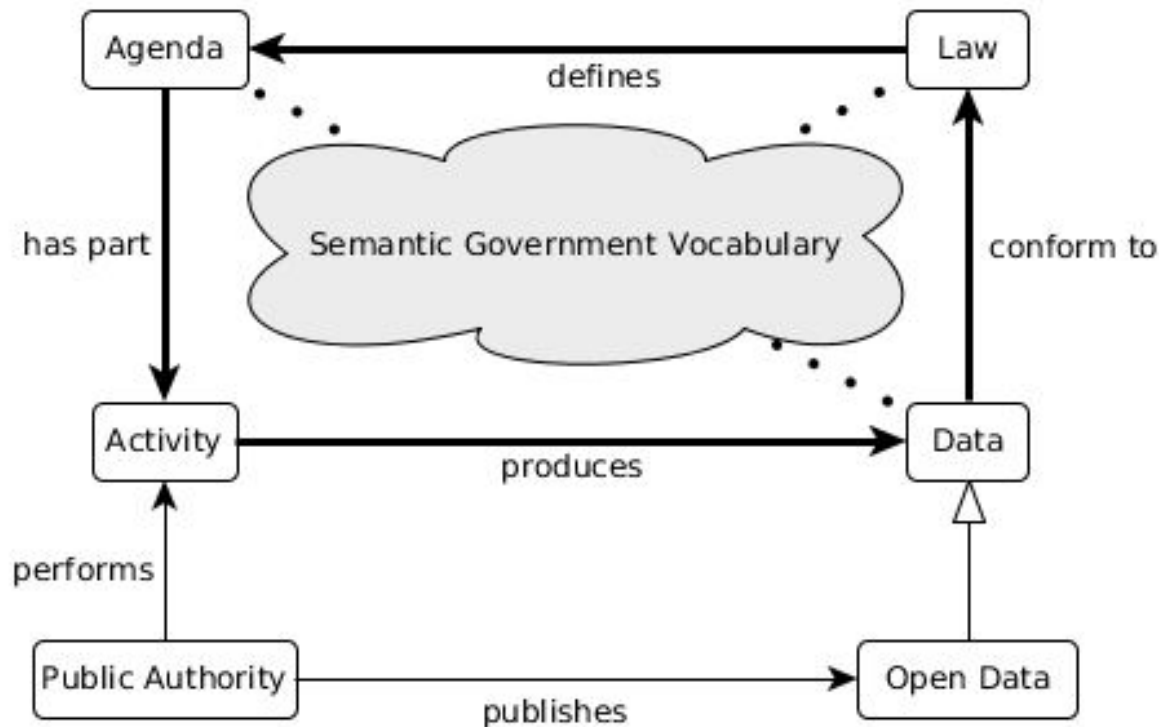
Discovering related datasets



Semantic Government Vocabulary (SGoV)



European Union
European Social Fund
Operational Programme Employment

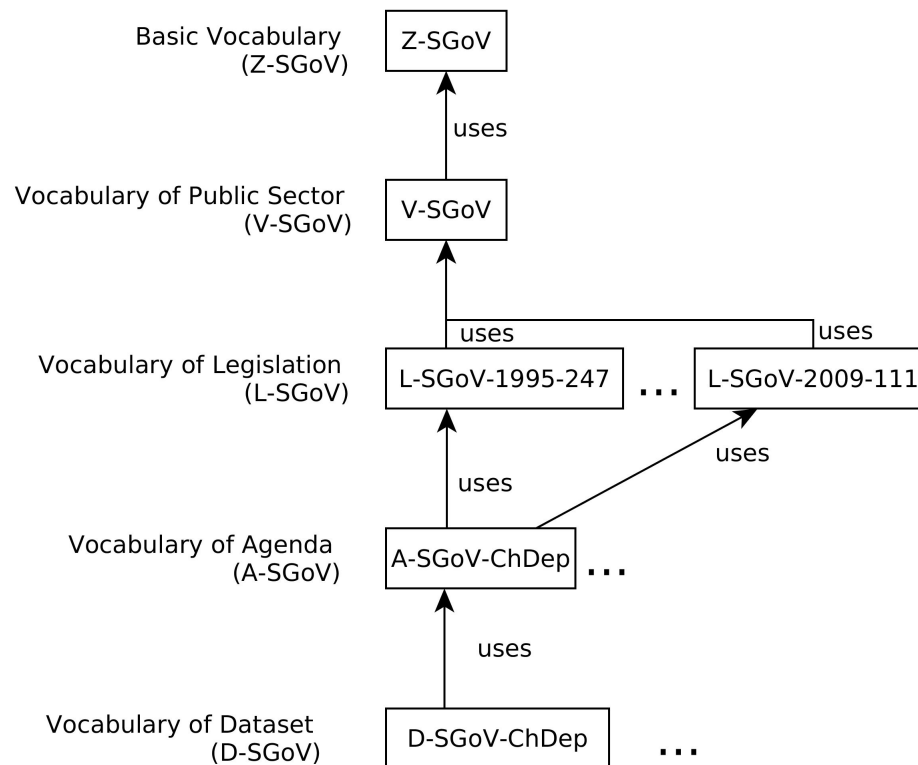


SGoV Architecture

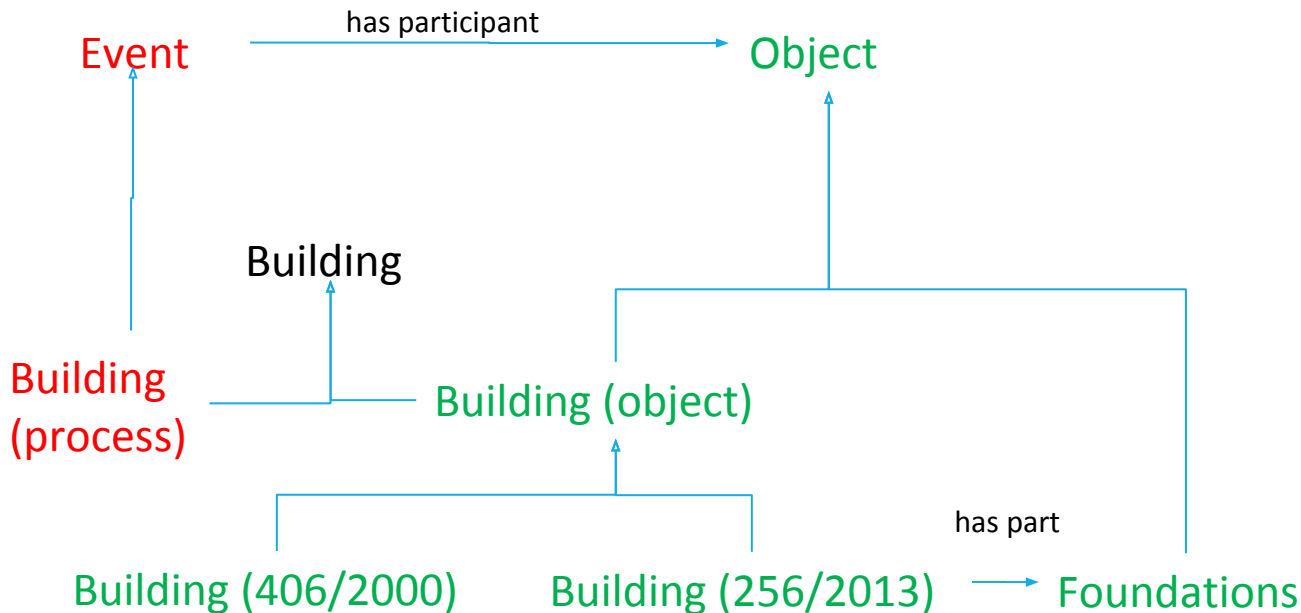


European Union
European Social Fund
Operational Programme Employment

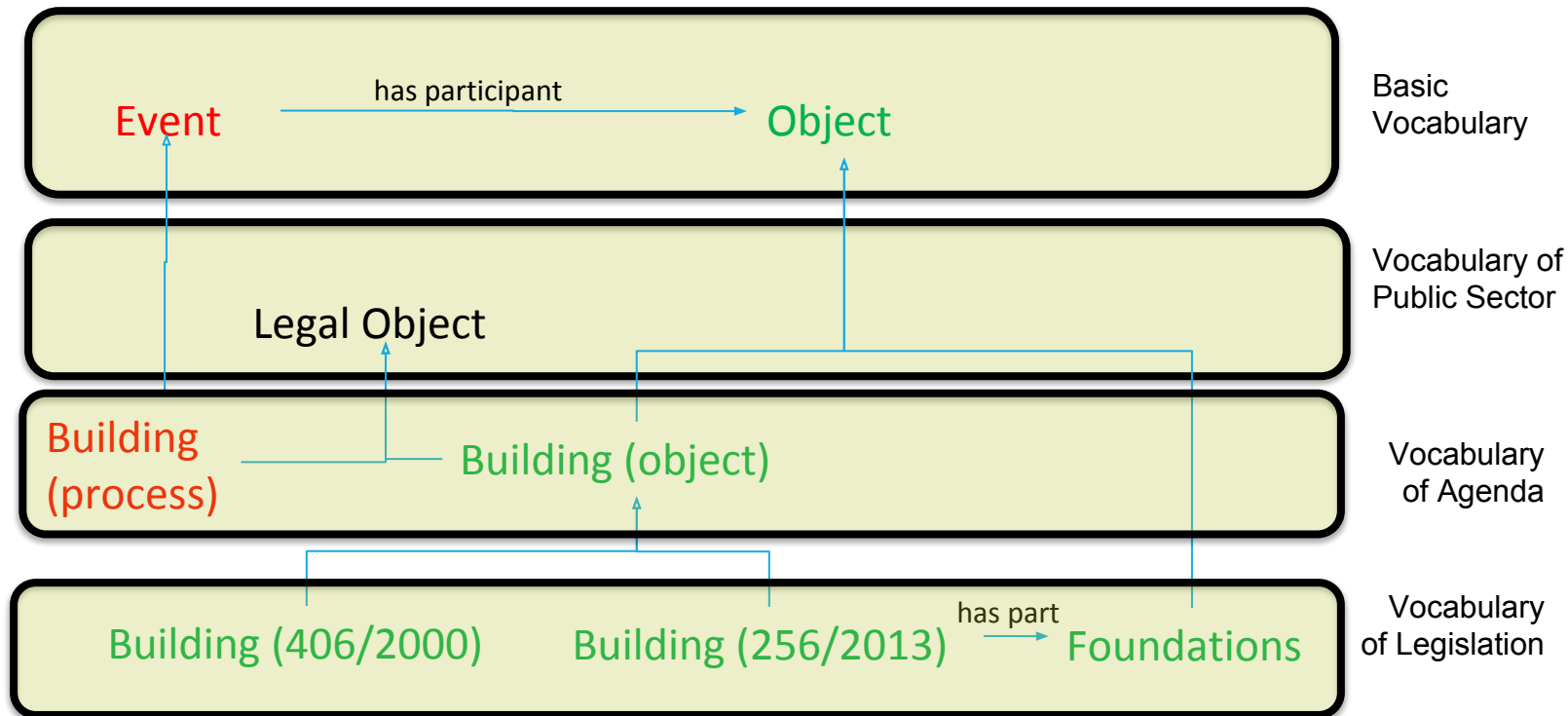
- Basic Vocabulary
 - *E.g. Event, Kind, Relator*
 - Based on the **Unified Foundational Ontology**
- Vocabulary of Public Sector
 - *E.g. Legal Subject, Organization*
 - Mapped to **ISA Core**
- Vocabularies of Legislation
 - *E.g. Building (Act 406/2000 Coll., on Energy Management)*
 - *1 vocabulary per Act*
- Vocabularies of Agenda
 - *1 vocabulary per Agenda*
- Vocabularies of Datasets
 - *1 vocabulary per Data Schema (to support data series)*



Discovering related datasets



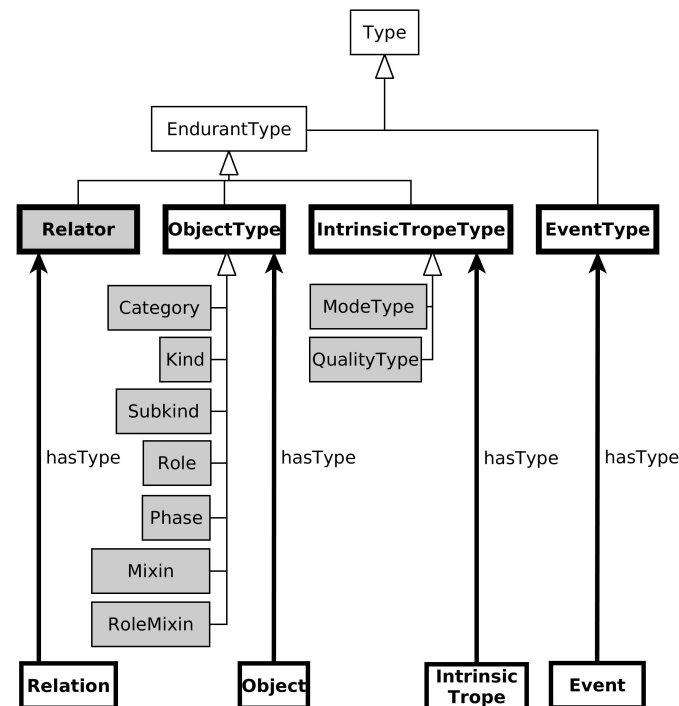
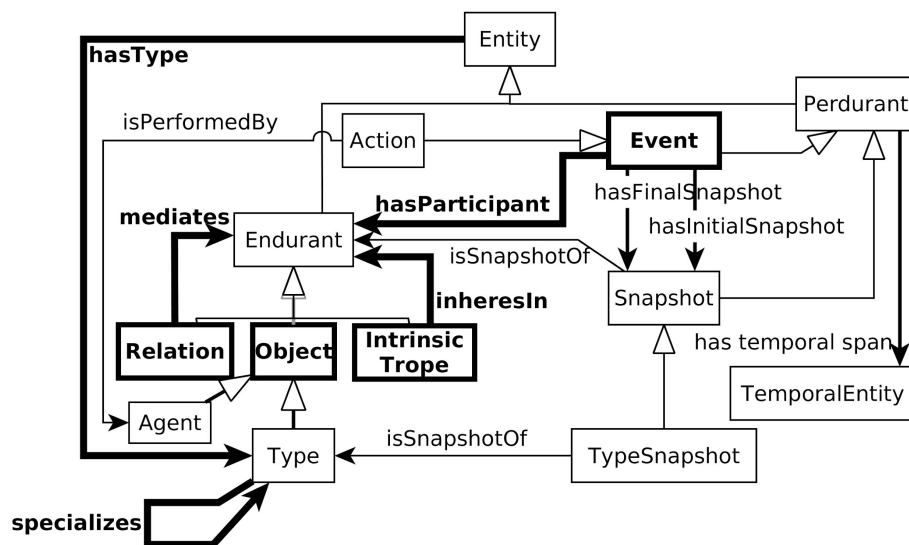
Discovering related datasets



Unified Foundational Ontology



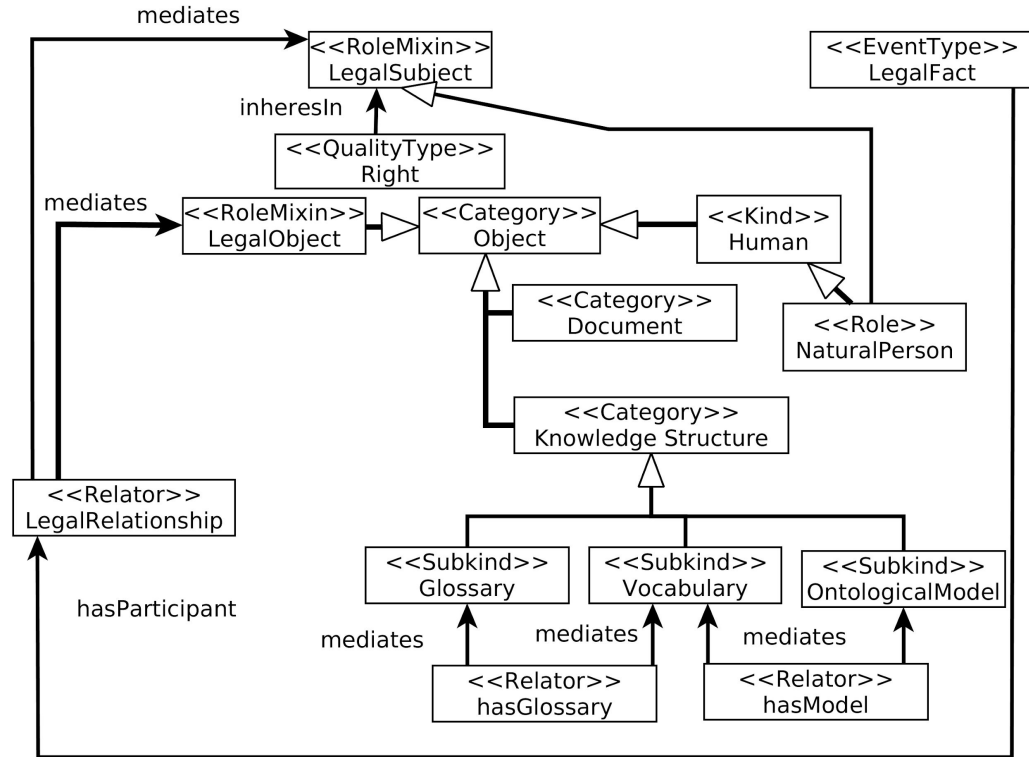
European Union
European Social Fund
Operational Programme Employment



Vocabulary of Public Sector



- Modeled using OntoUML
 - Inspired by OntoClean
- Aimed at describing common-sense notions in the public Sector
- To be reused by multiple vocabularies



Leg./ Agenda Vocabularies



European Union
European Social Fund
Operational Programme Employment

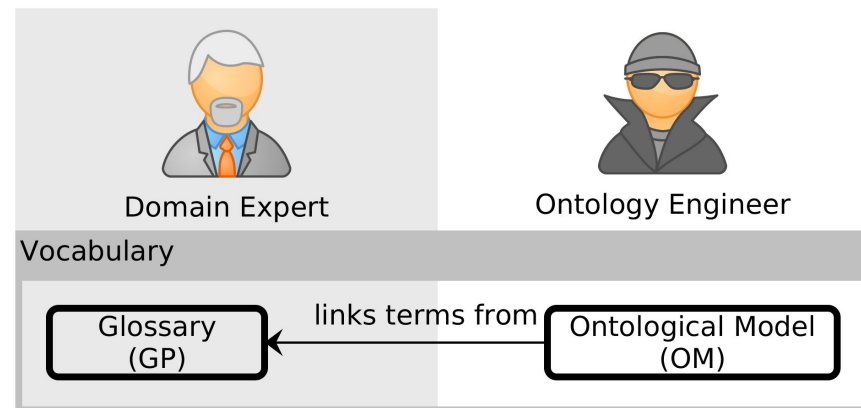
- Modeled using the Vocabulary of Public Sector
- Aimed at describing legal/agenda terms **necessary to describe the meaning of a dataset**

SGoV Vocabulary Internals



European Union
European Social Fund
Operational Programme Employment

- Each vocabulary is in OWL2-DL (expressiveness **ALCNI**)
 - Glossary
 - SKOS
 - Ontological model
 - OWL2-DL



Current Vocabulary Usage



European Union
European Social Fund
Operational Programme Employment

- Open data of the Register of rights and responsibilities of public authorities
- Experiment on the domain of Parliamentary Elections in the Czech Republic
 - 3 people - 4 hours long training for them
 - Using only basic language (not OntoUML)
 - Object Type
 - Event Type
 - Relator Type
 - Intrinsic Trope Type
 - + links according to UFO
 - Partial vocabularies for 40 Acts
 - Over 1000 terms

SGoV Resources



European Union
European Social Fund
Operational Programme Employment

- <https://slovník.gov.cz/sparql>
 - SPARQL endpoint over SGoV
- <https://slovník.gov.cz/prohlížeč>
 - simple faceted browser of SGoV
- <https://github.com/opendata-mvcr/ssp>

Summary



- Main ideas
 - Layered architecture
 - To have the vocabulary lifecycle manageable
 - Some well-founded top-level/public sector ontology
 - To define metrics for dataset similarity
 - Legislation-founded vocabularies
 - To be used for (not only open) data description
 - Broader use
 - Share meaning of terms used in eGovernment
 - Share meaning of datasets (rich metadata)