



European Union
European Social Fund
Operational Programme Employment

Linked Open Data in The Czech Republic

Martin Nečaský

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



Existing LOD



- Czech Social Security Administration
 - Statistical data about pensions
 - RDF dumps, SPARQL endpoint, CSV dumps
 - <https://data.cssz.cz/web/otevrena-data/katalog-otevrenych-dat>
 - <https://data.cssz.cz/web/otevrena-data/-/duchodci-v-cr-krajich-okresech>
- General Financial Directorate
 - Central Register of Subsidies
 - RDF dumps, SPARQL endpoint, CSV dumps

Problems of LOD



European Union
European Social Fund
Operational Programme Employment

Consumers
Publishers } do not understand LOD benefits

Benefits of LOD





Code lists

Problems with codelists



- Different data models, structures and formats
- Duplicated by publishers
- Not shared by publishers
- No notifications about changes
- No propagation of changes

LOD solution for codelists

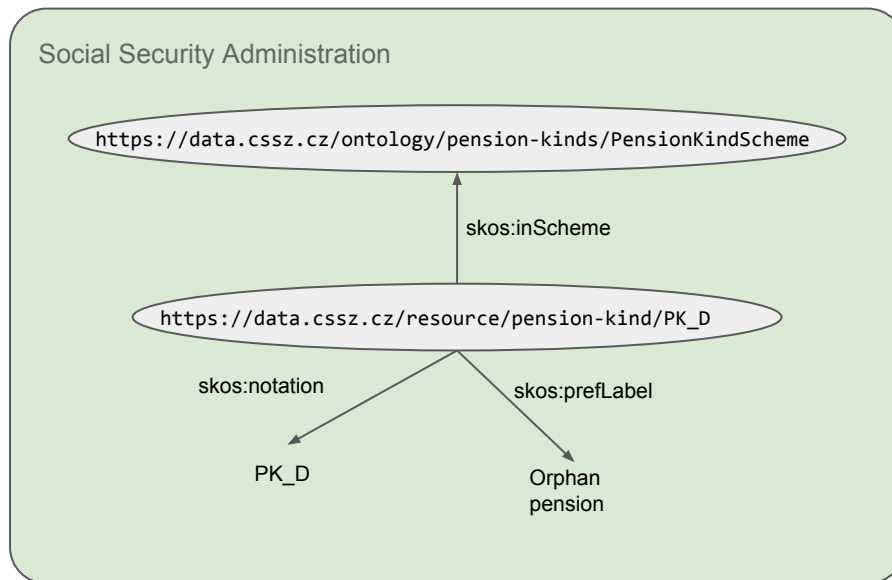


Unified data model, structure and formats :

RDF

SKOS

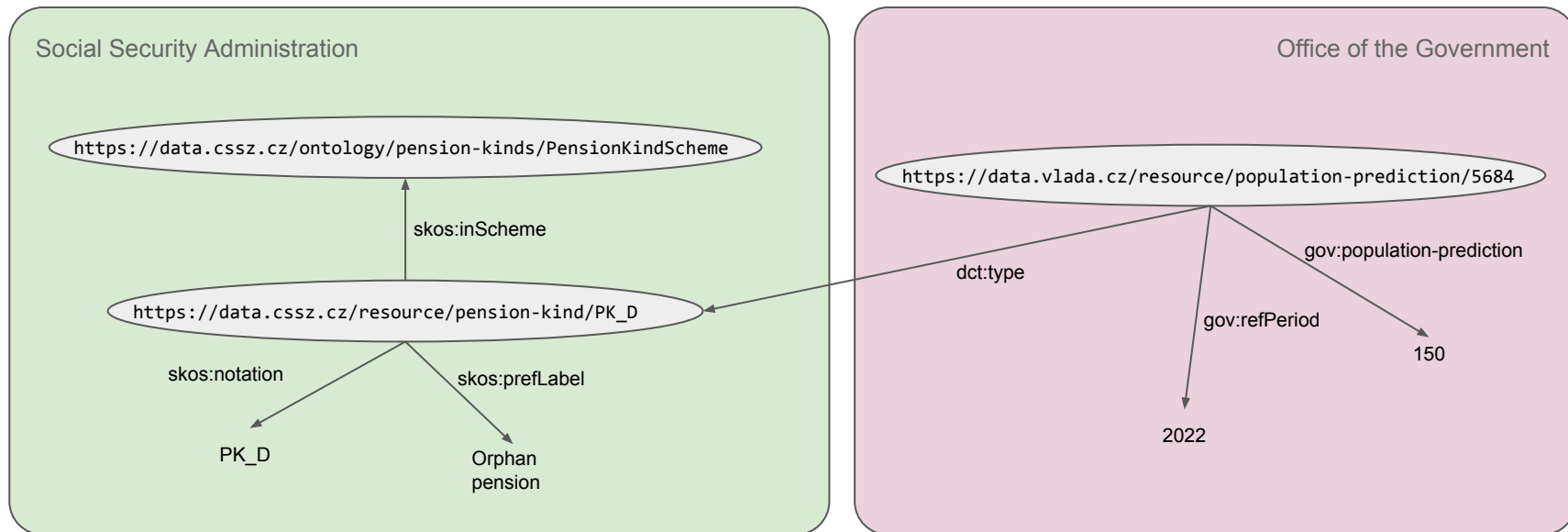
IRI patterns



LOD solution for codelists



Minimizing duplicities with Linking



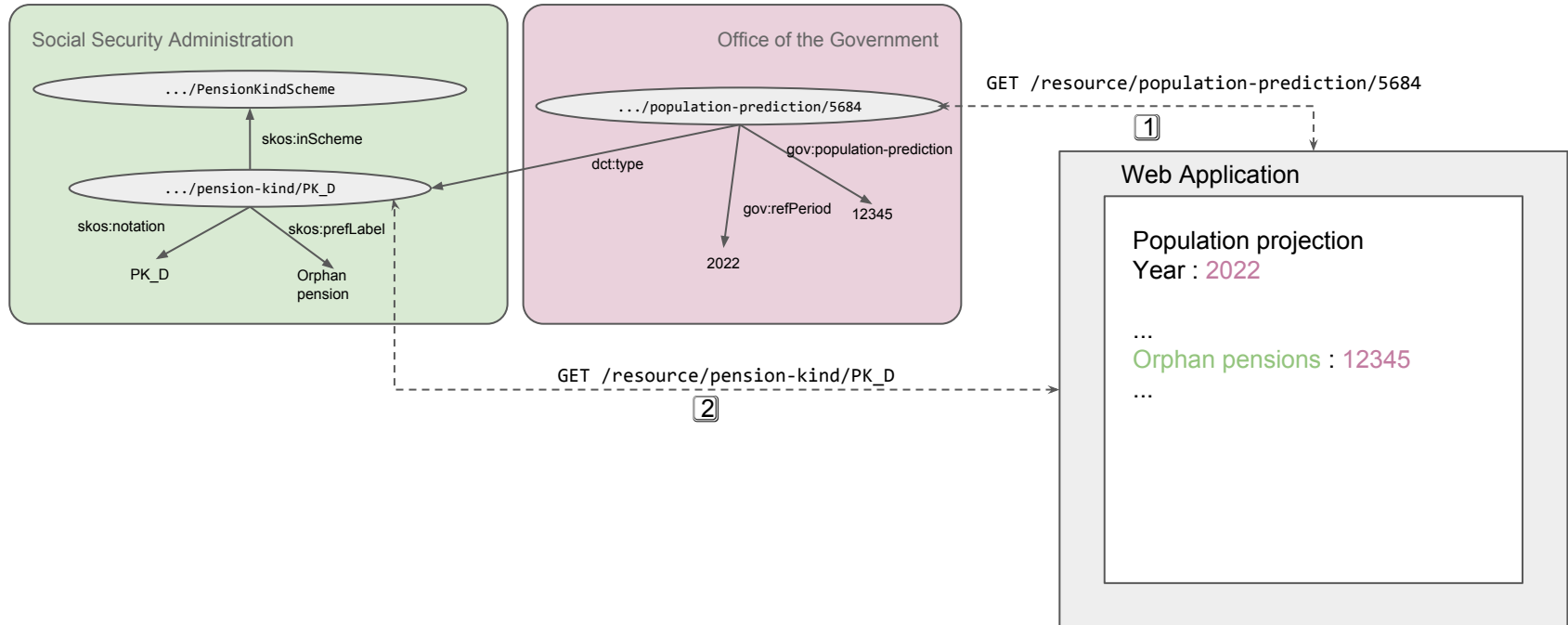
LOD solution for codelists



Sharing codelists and their items :

HTTPS

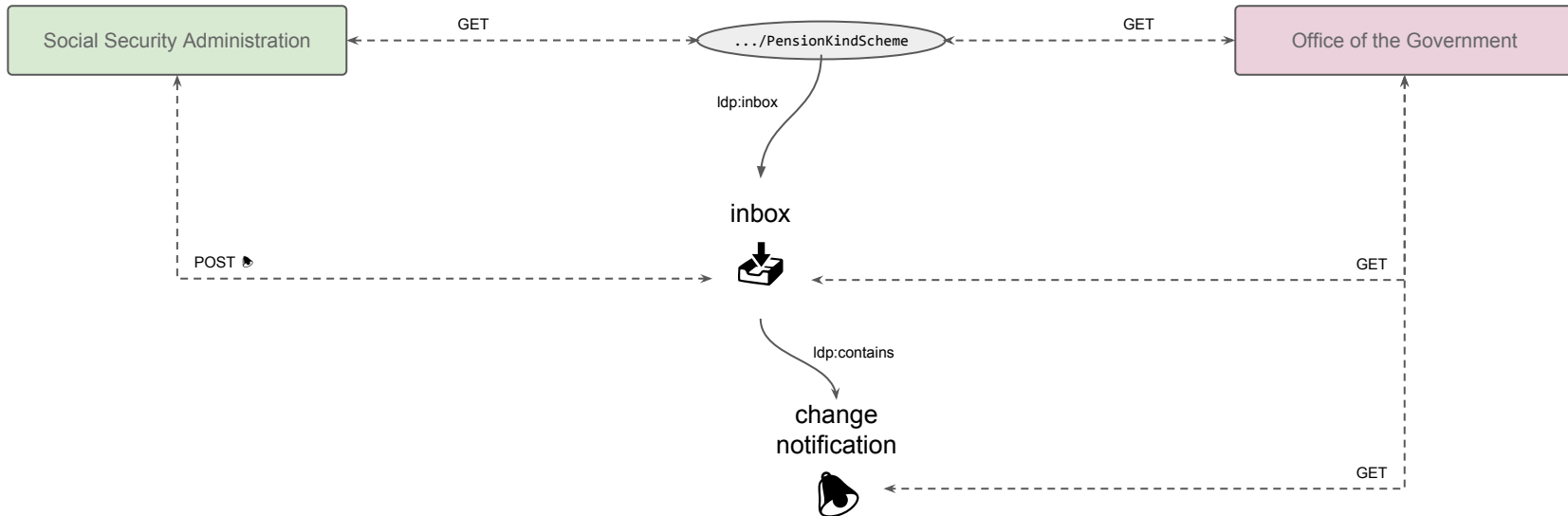
IRI dereferencing



LOD solution for codelists



Change notification and propagation based on LD Notifications



Our next steps towards LOD codelists



- National standard for publishing codelists as LOD
- Mandatory publishing of important codelists as OD
- Publishing codelists managed by Czech Statistical Office
 - 100s of codelists used by many public bodies
- Alignment of codelists
 - Czech Statistical Office
 - Czech Social Security Administration
 - EU Vocabularies (Publications Office of the EU)



Statistical data

Problems with statistical data



- Different data models, structures and formats
- Hard to explore
- Hard to integrate

LOD solution for statistical data



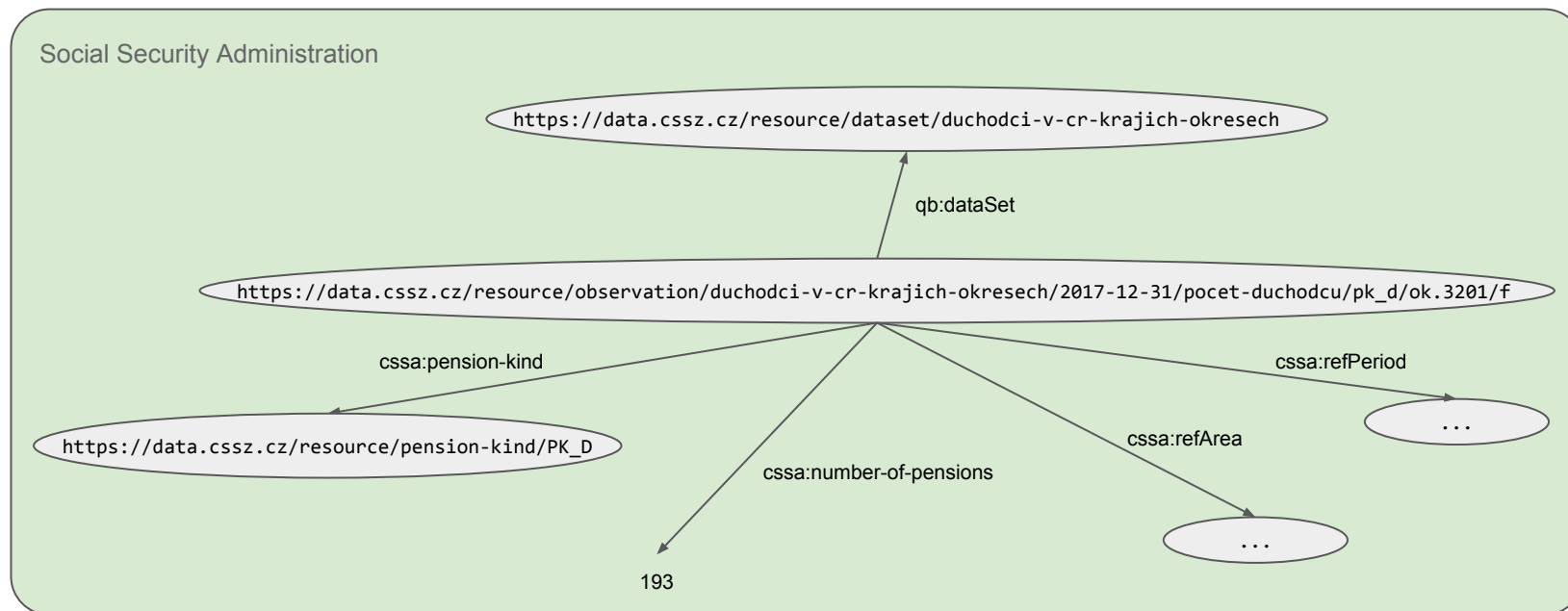
Unified data model, structure and formats :

RDF

Data Cube Voc.

SDMX

IRI patterns



LOD solution for statistical data



European Union
European Social Fund
Operational Programme Employment

Easier exploration

Which statistical measures do we
measure in Czech regions?

Who publishes these measures in
which datasets?

LOD solution for statistical data



Easier exploration based on

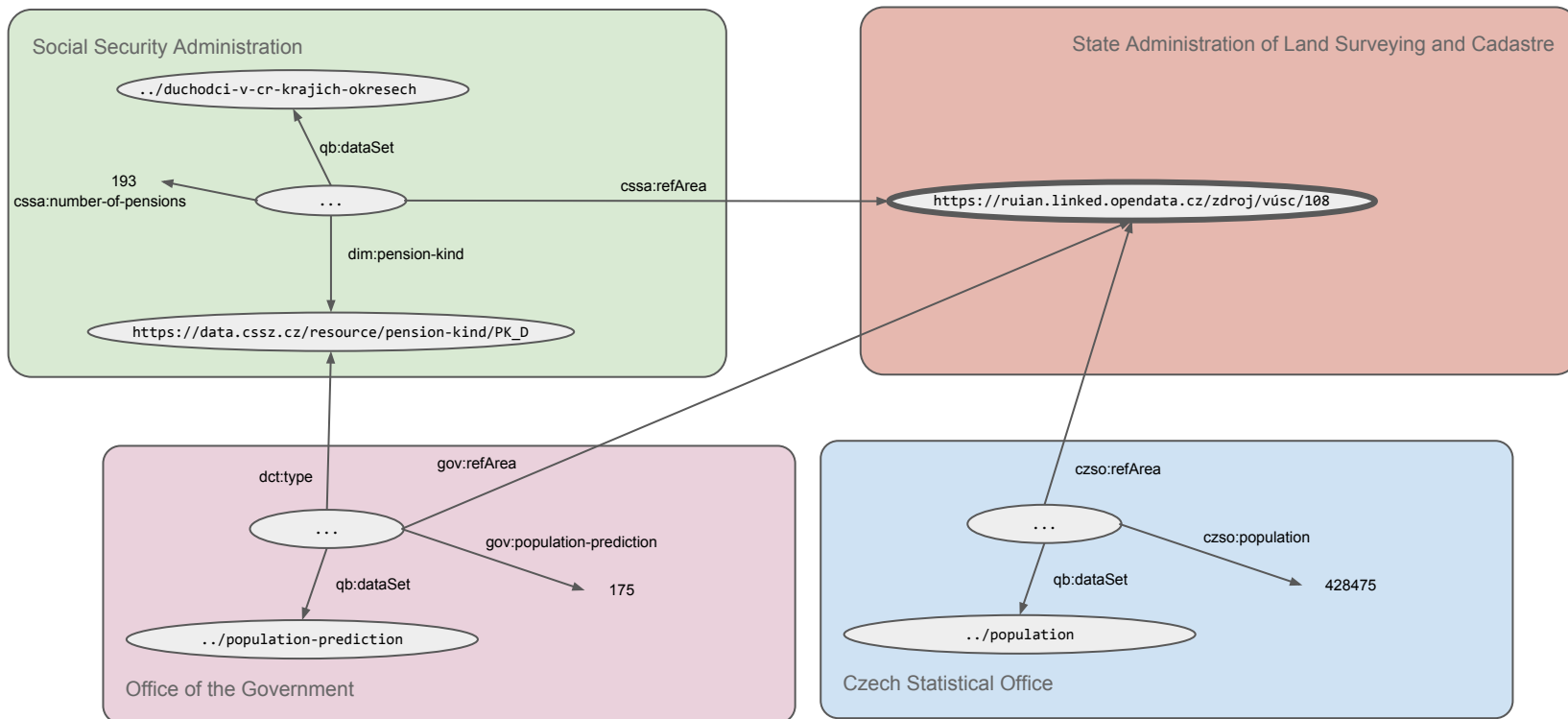
RDF

Data Cube Voc.

SDMX

Codelists reuse

SPARQL



LOD solution for statistical data



Easier exploration based on

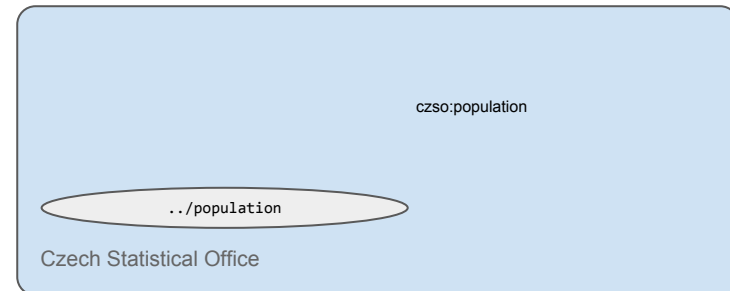
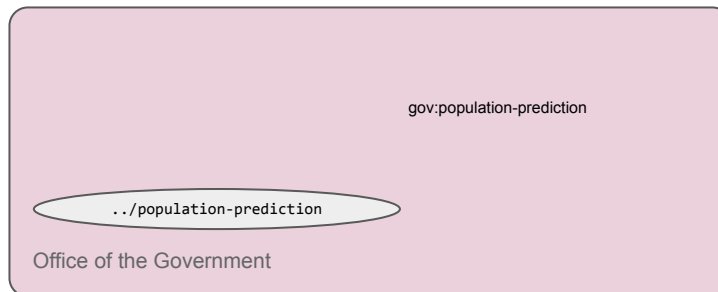
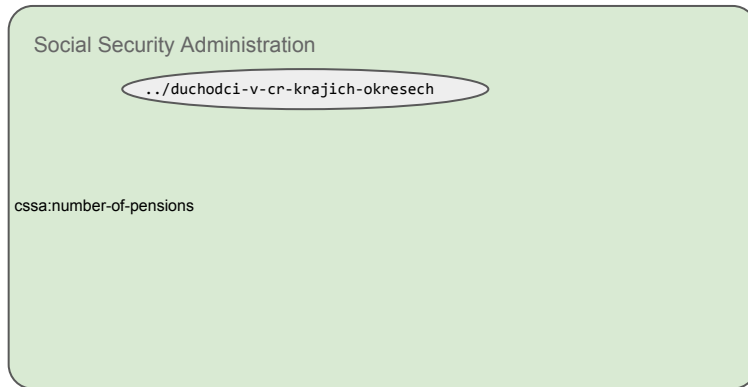
RDF

Data Cube Voc.

SDMX

Codelists reuse

SPARQL



LOD solution for statistical data



Straightforward integration thanks to

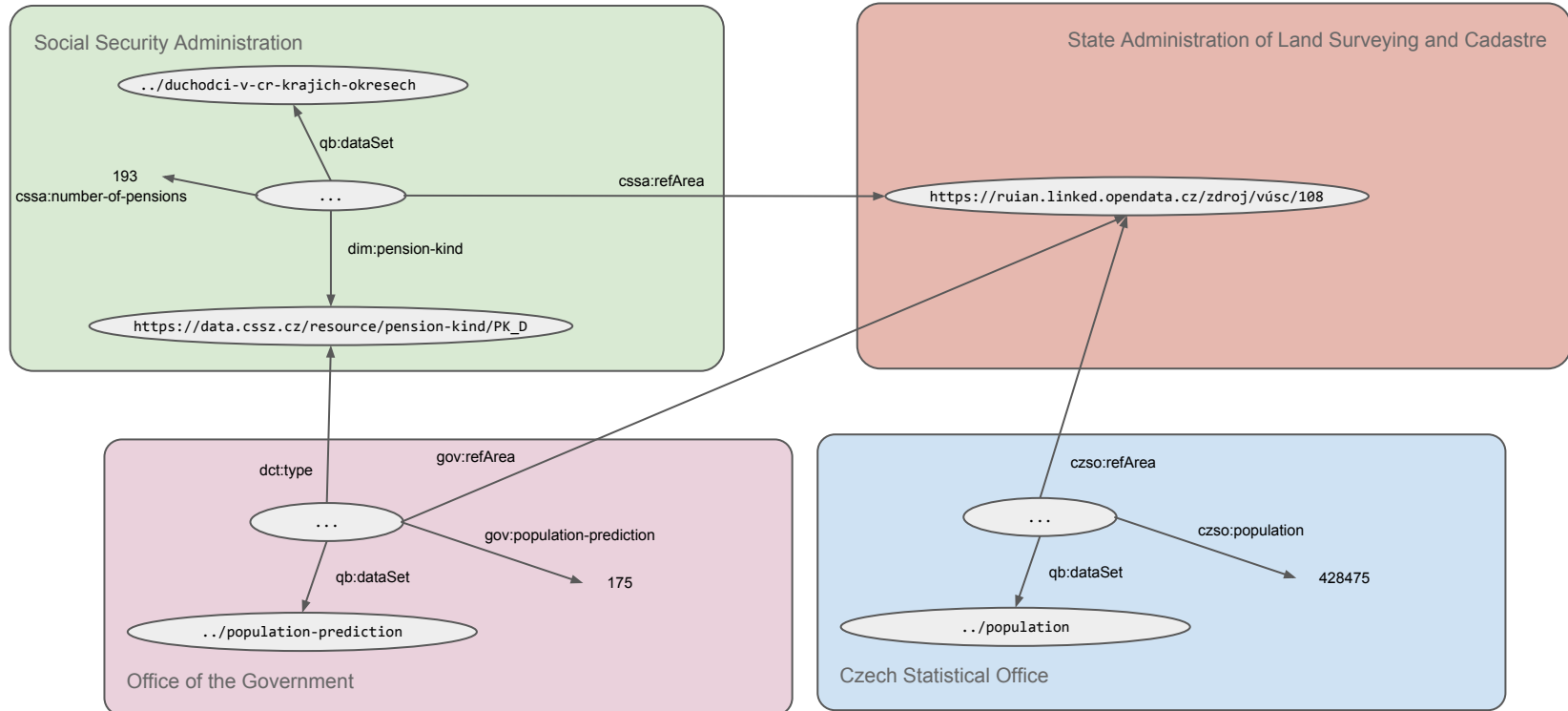
RDF

Data Cube Voc.

SDMX

Codelists reuse

SPARQL



LOD solution for statistical data



Straightforward integration thanks to

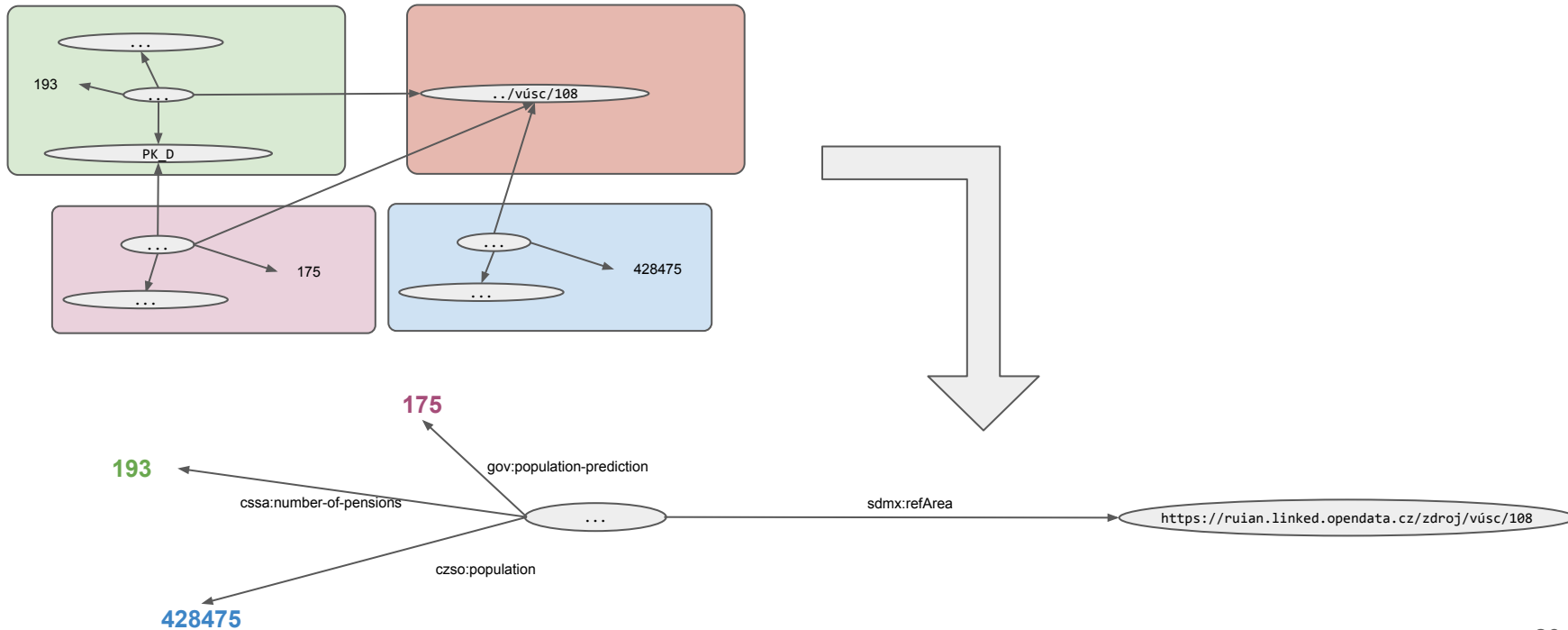
RDF

Data Cube Voc.

SDMX

Codelists reuse

SPARQL



LOD solution for statistical data



European Union
European Social Fund
Operational Programme Employment

Straightforward integration thanks to

RDF

Data Cube Voc.

SDMX

Codelists reuse

SPARQL

<https://etl.opendata.cz/#/pipelines/edit/canvas?pipeline=https:%2F%2Fetl.opendata.cz%2Fresources%2Fpipelines%2F1544018816319>

Our next steps towards LOD statistical data



- National standard for publishing statistical data as LOD
- Publishing statistical data
 - Czech Statistical Office
 - Czech Social Security Administration
 - (
 - Technology Agency of the Czech Republic
 - Institute of Health Information and Statistics of the Czech Republic
 - ...
 -)
- Alignment with codelists
- Alignment with Register of territorial identification, addresses and real estates



Base registries

Czech base registries



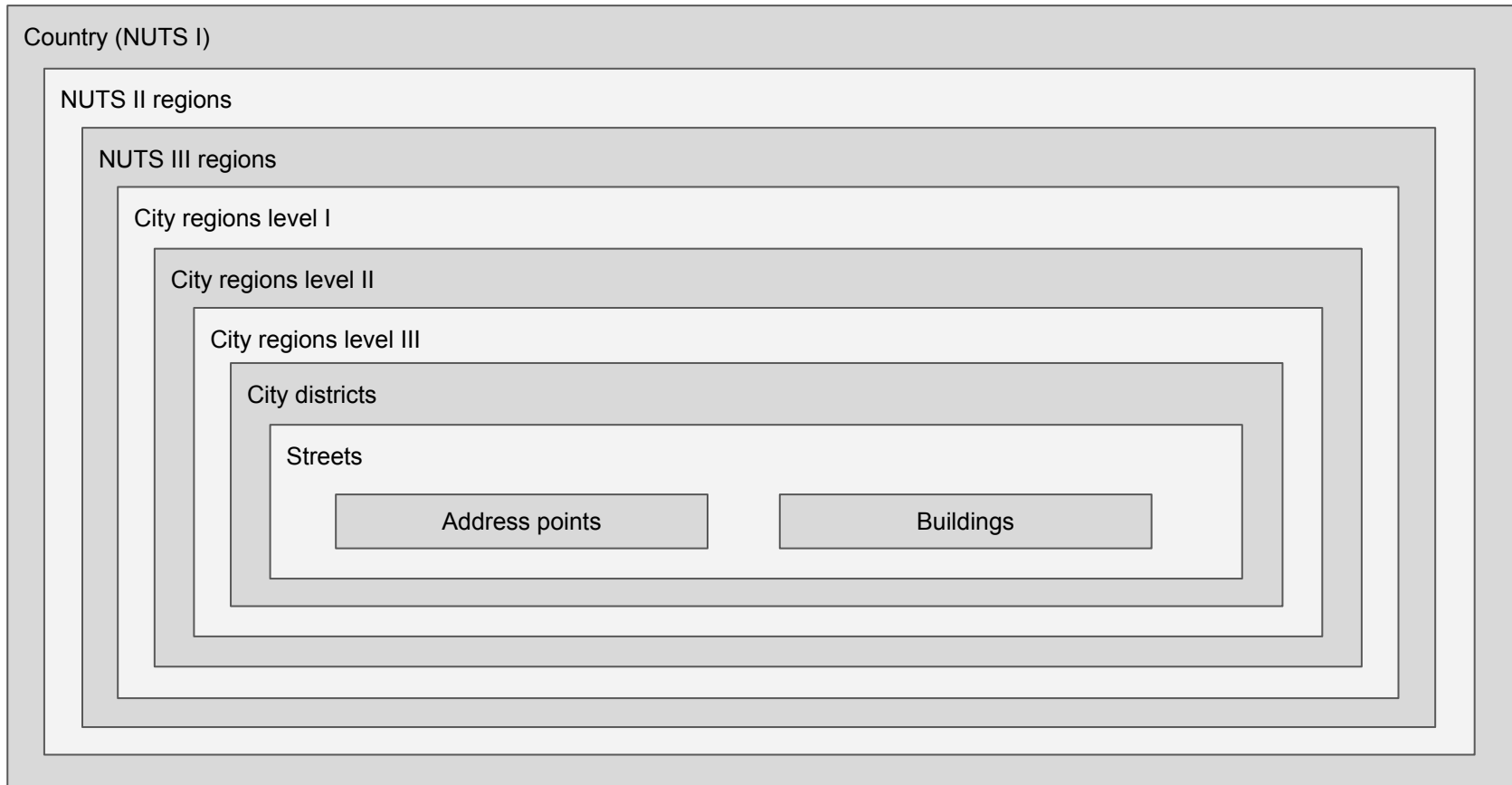
“Base registries, important components of European public services, are a trusted and authoritative source of information. They provide basic information on data items such as people, companies, vehicles, licences, buildings, locations and roads.” -- ISA2

- Register of inhabitants
- Register of persons (companies)
- Register of territorial identification, addresses and real estates
- Register of rights and responsibilities of public authorities
- Collection of Laws

Register of territorial identification



European Union
European Social Fund
Operational Programme Employment



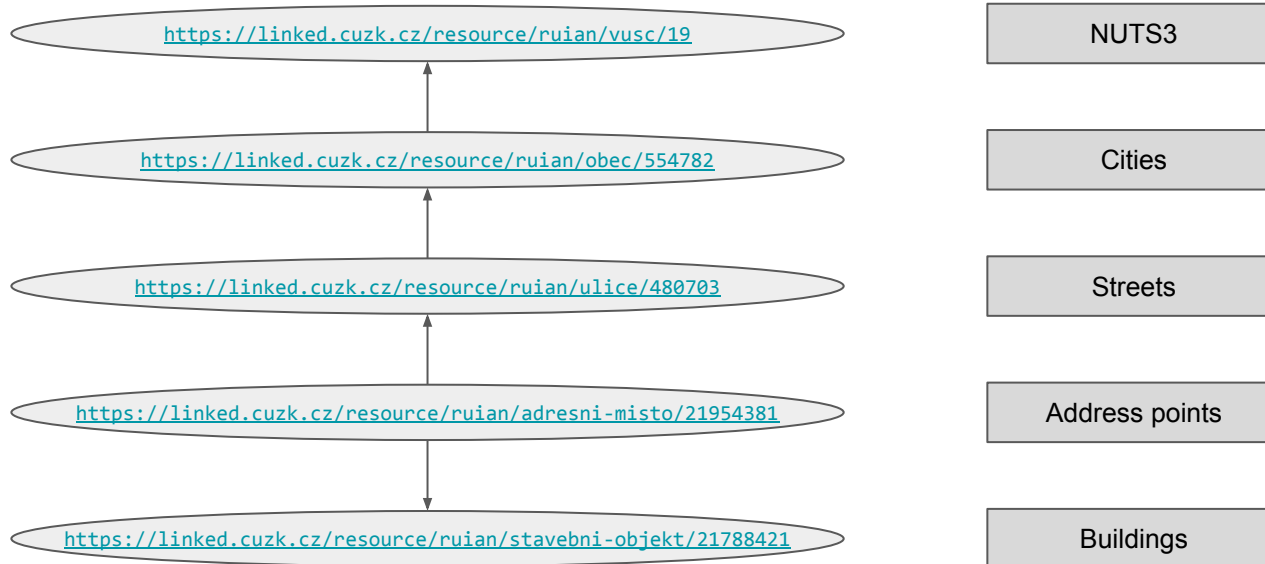
Register of territorial identification



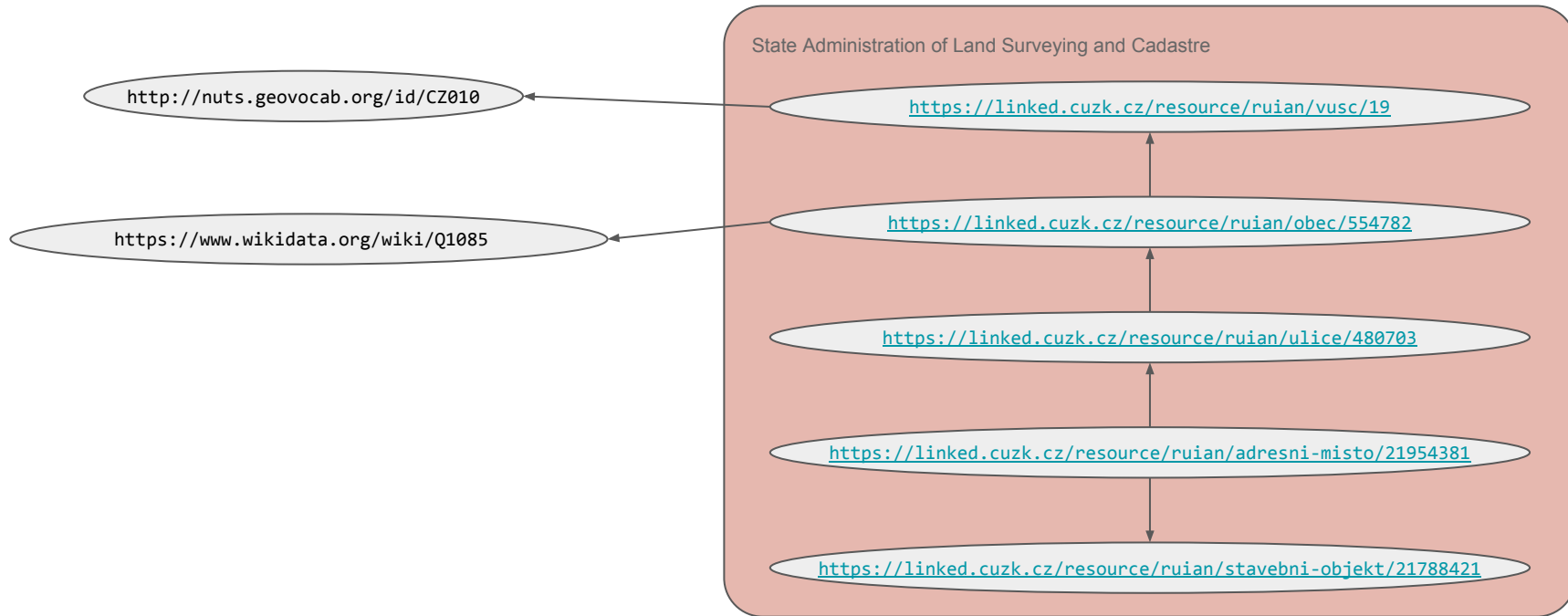
European Union
European Social Fund
Operational Programme Employment

- Primary and the only territorial identities
- Territorial knowledge graph
 - names, shapes, relationships, ...
- Problems
 - Identities technically expressed as integers which are hard to link and reuse on the Web
 - Knowledge base published in XML which is hard to integrate with other data sources

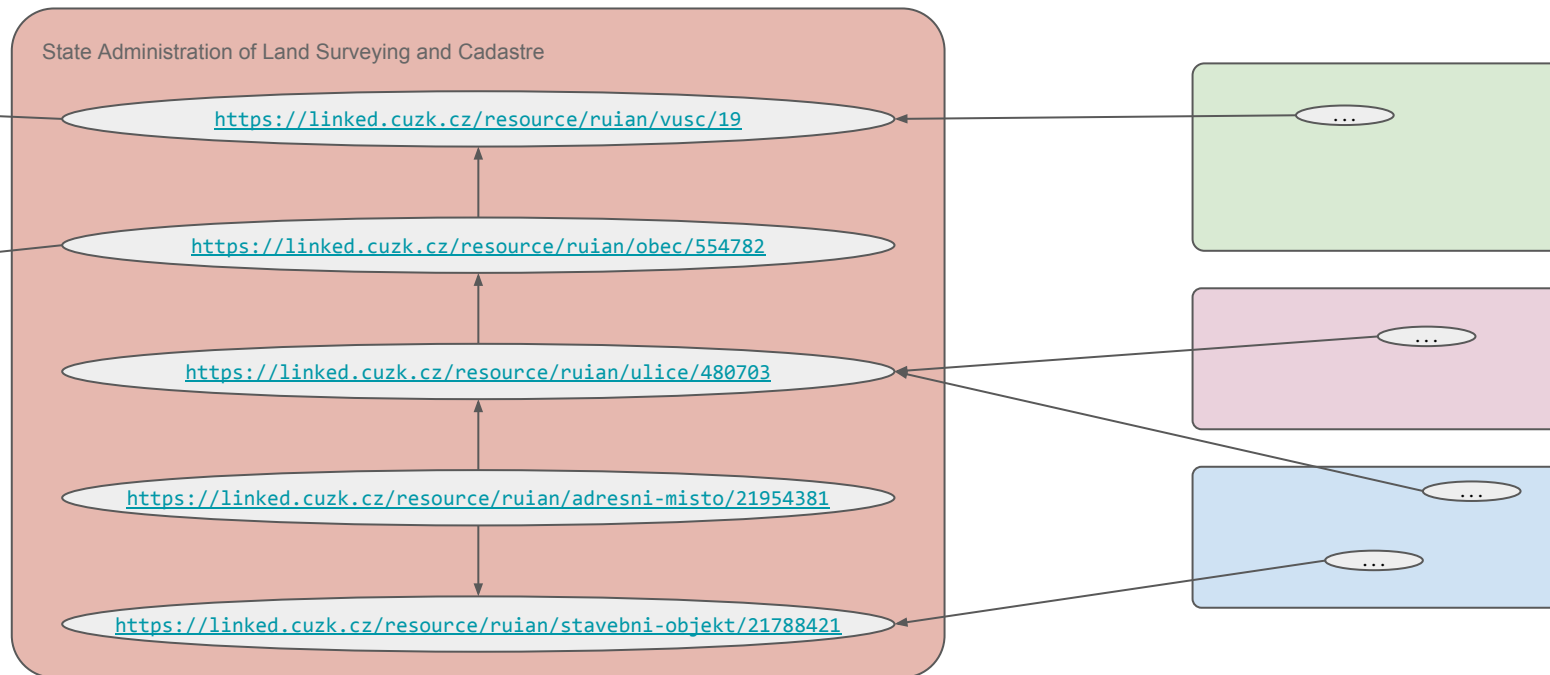
LOD solution for Register of territorial identification



LOD solution for Register of territorial identification



LOD solution for Register of territorial identification



LOD solution for Register of territorial identification



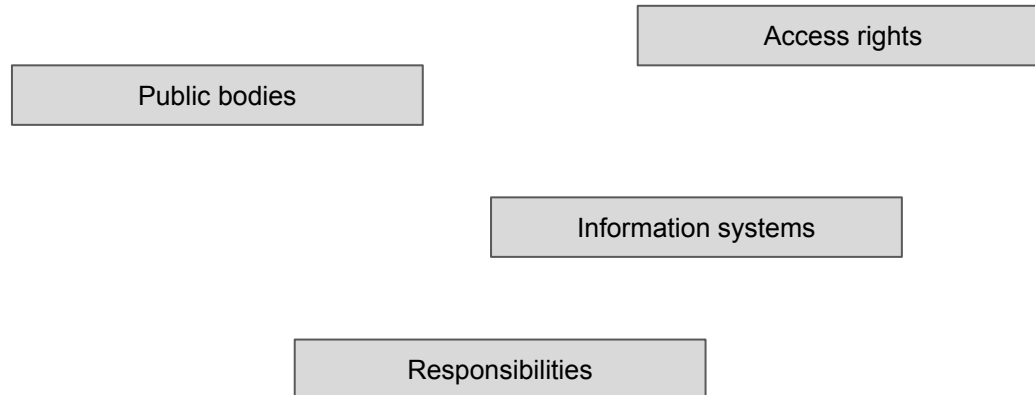
European Union
European Social Fund
Operational Programme Employment

- Experimental representation <https://linked.cuzk.cz/opendata.cz/sparql>
- IRIs alive, but currently only HTML representation available
- Discussions with State Administration of Land Surveying and Cadastre

Register of rights and responsibilities



European Union
European Social Fund
Operational Programme Employment



Register of rights and responsibilities



European Union
European Social Fund
Operational Programme Employment

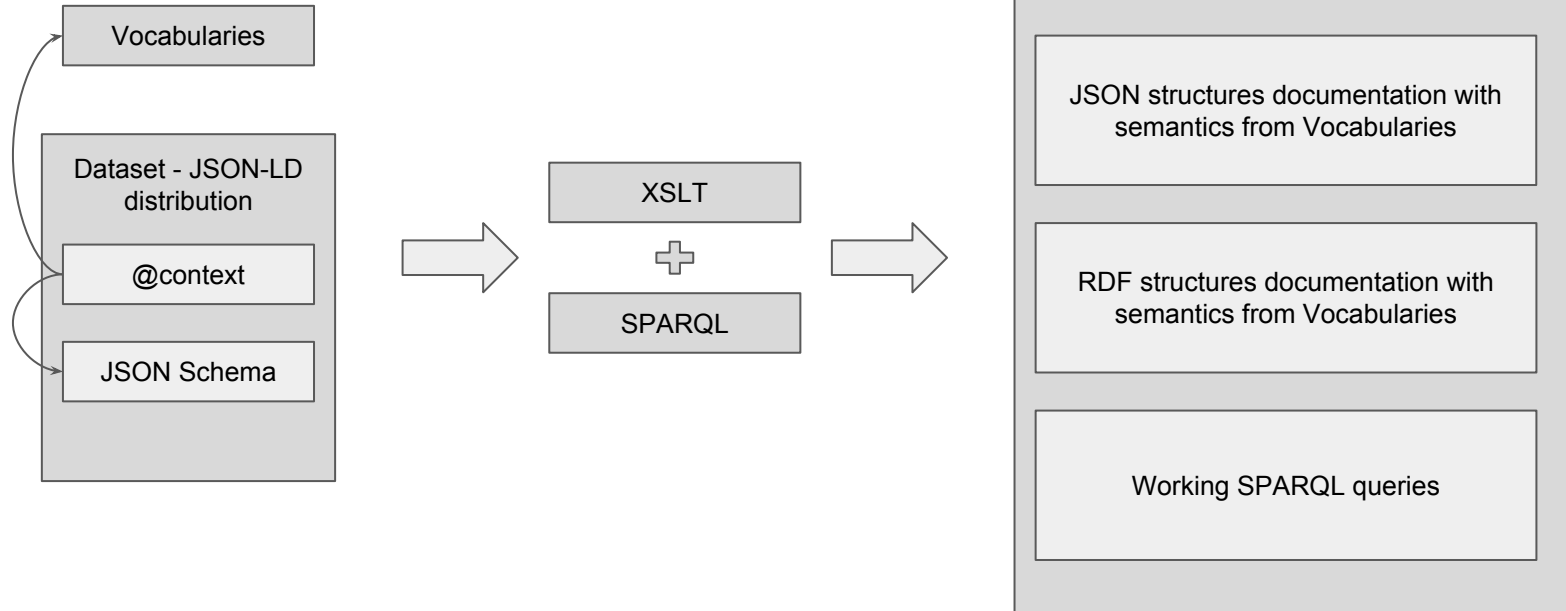
- Primary and the only identities of public bodies, their responsibilities, etc.
- Governmental knowledge graph
 - Relationships between public bodies, responsibilities, access rights, information systems, etc.
 - Links to related legislation
- Problems
 - Identities technically expressed as integers which are hard to link and reuse on the Web
 - Knowledge base published in XML which is hard to integrate with other data sources

LOD solution for Register of rights and responsibilities

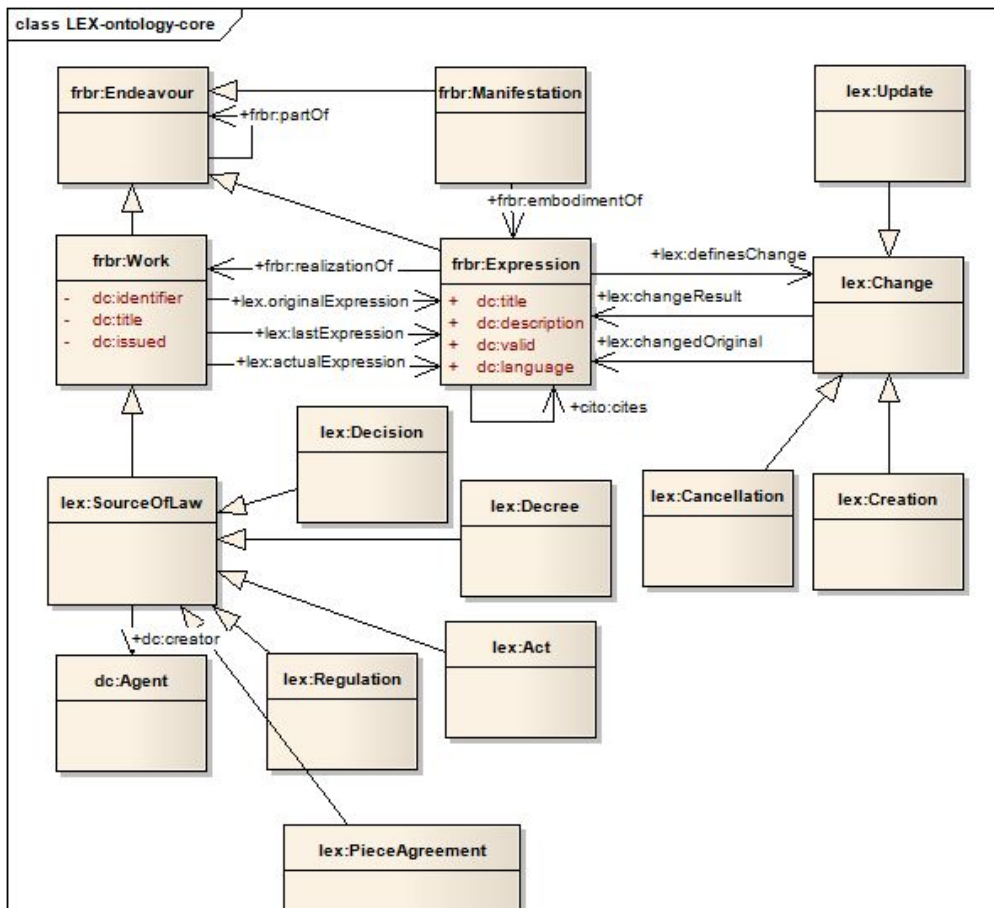


- Will be published on 1 Jan 2019
- SPARQL endpoint
- 25 datasets
 - JSON-LD distributions
- Documentation
 - official standard for each dataset
 - auto-generated from RDFS/OWL vocabularies, JSON @context and JSON Schema
 - <https://data.gov.cz/otevřené-formální-normy/registr-práv-a-povinností/orgány-veřejné-moci>

Auto-generating documentation of LOD datasets



Collection of Law



- Structured and machine readable representation of Collection of Laws
- Each legal document and its part with dereferencable persistent IRI
- 10/2018 - 12/2020



Thank You!

Martin Nečaský

martinnec@gmail.com - necasky@ksi.mff.cuni.cz