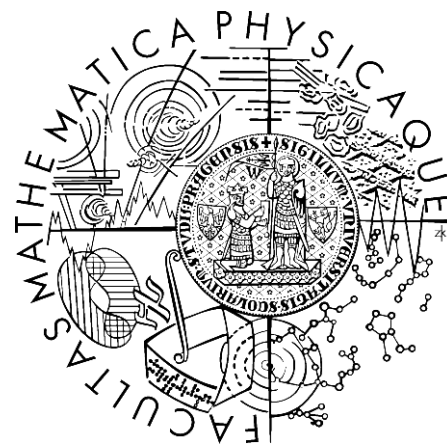
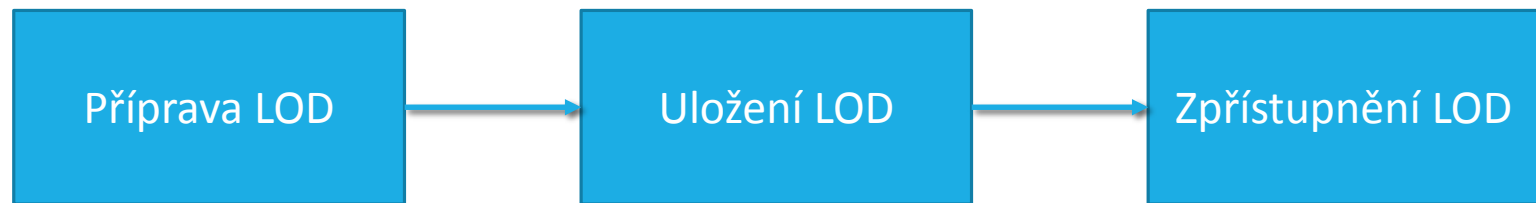


Standards and tools for LOD publications

Jakub Klímeck
Mathematico-physical faculty
University of Charles in Prague

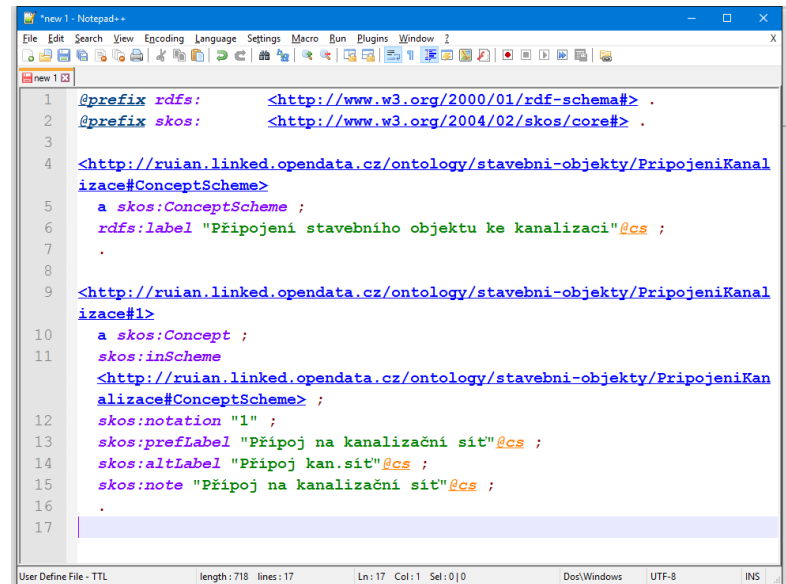


Typický proces publikace LOD



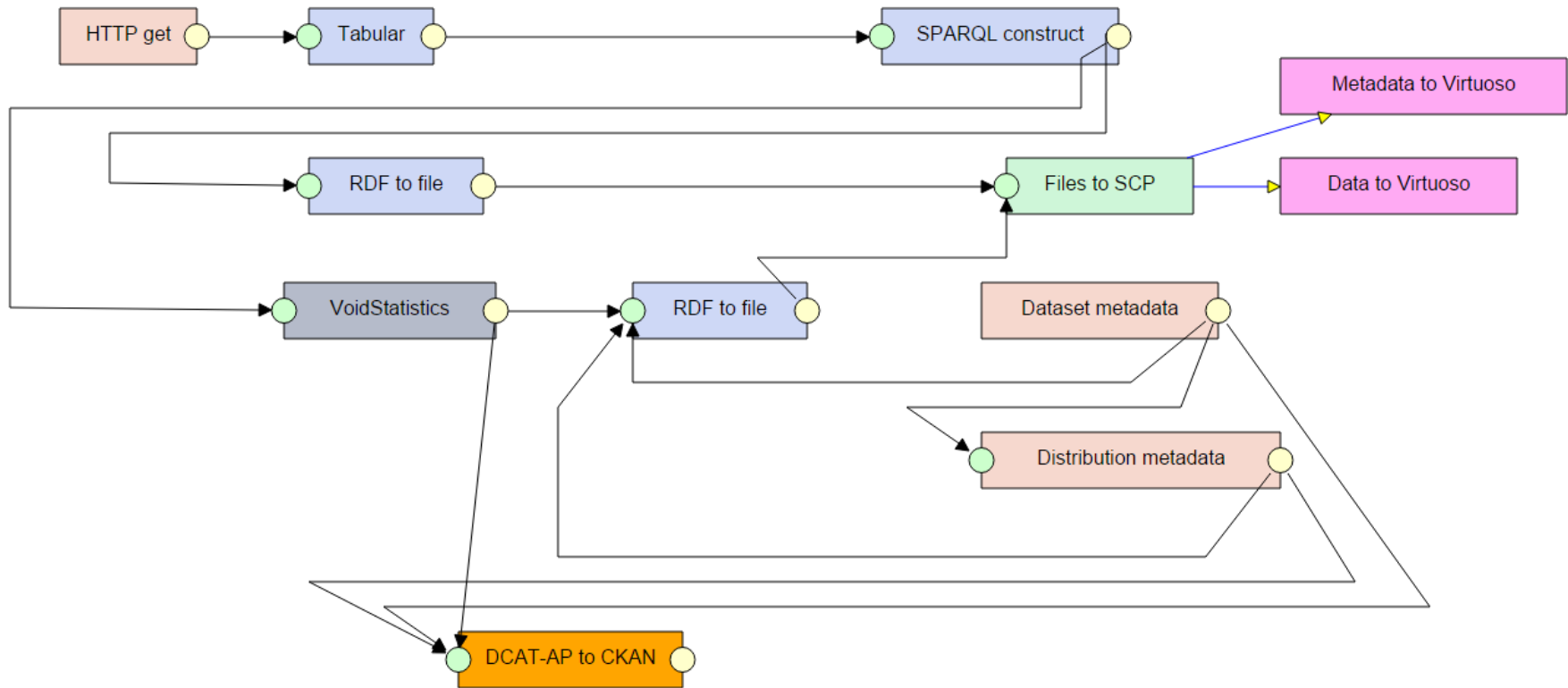
Příprava LOD – Nástroje

- Notepad++
 - [Open source](#)
 - Pro psaní jednoduchých číselníků
 - Pro psaní slovníků
- LinkedPipes ETL
 - [Open source](#)
- LinDa
 - [Open source](#)
- Swirrl PublishMyData
 - [Open source](#) (dále nerozvíjený), [Komerční cloudové řešení](#)



```
1 @prefix rdfs:      <http://www.w3.org/2000/01/rdf-schema#> .
2 @prefix skos:     <http://www.w3.org/2004/02/skos/core#> .
3
4 <http://ruian.linked.opendata.cz/ontology/stavebni-objekty/PripojeniKanalizace#ConceptScheme>
5   a skos:ConceptScheme ;
6   rdfs:label "Připojení stavebního objektu ke kanalizaci"@cs ;
7   .
8
9 <http://ruian.linked.opendata.cz/ontology/stavebni-objekty/PripojeniKanalizace#1>
10  a skos:Concept ;
11  skos:inScheme
12    <http://ruian.linked.opendata.cz/ontology/stavebni-objekty/PripojeniKanalizace#ConceptScheme> ;
13  skos:notation "1" ;
14  skos:prefLabel "Připoj na kanalizační síť"@cs ;
15  skos:altLabel "Připoj kan.síť"@cs ;
16  skos:note "Připoj na kanalizační síť"@cs ;
17  .
```

Příprava LOD – LinkedPipes ETL

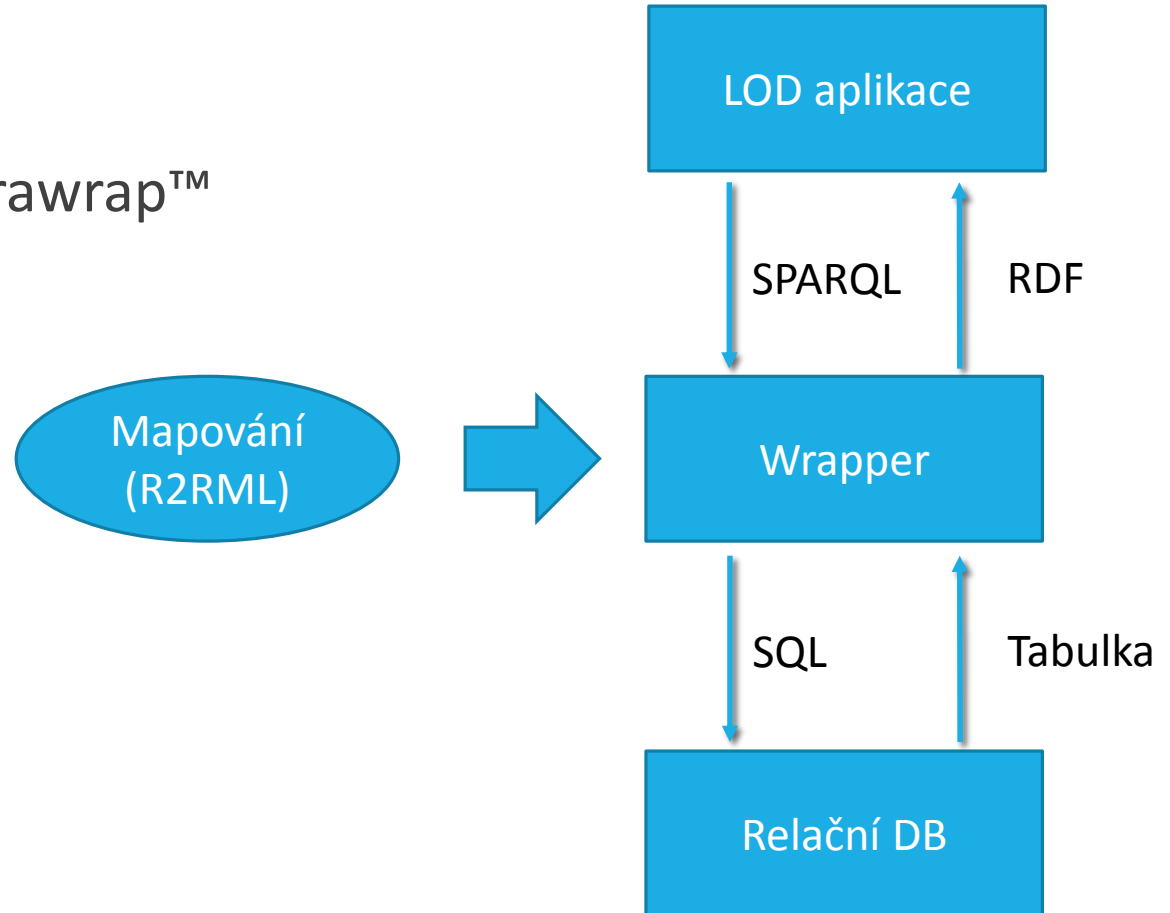


Uložení LOD – RDF databáze

- Openlink Virtuoso
 - [Open source](#), [Komerční](#)
- Blazegraph
 - [Open source](#), [Komerční](#)
- Apache Jena Fuseki
 - [Open source](#)
- Eclipse RDF4J (OpenRDF Sesame)
 - [Open source](#), [Komerční podpora](#)
- Ontotext GraphDB (OWLIM)
 - [Free verze](#), [Komerční](#)

Wrappery relačních databází

- D2RQ
 - [Open source](#)
- Capsenta Ultrawrap™
 - [Komerční](#)



Publikace LOD – SPARQL endpoint

<http://linked.opendata.cz/sparql>

Virtuoso SPARQL Query Editor [About](#) | [Namespace Prefixes](#) | [Inference rules](#)

Default Data Set Name (Graph IRI)

Query Text

```
select * where {?s ?p ?o} LIMIT 100
```

(Security restrictions of this server do not allow you to retrieve remote RDF data, see [details](#).)

Results Format:

Execution timeout: milliseconds (values less than 1000 are ignored)

Options: Strict checking of void variables

(The result can only be sent back to browser, not saved on the server, see [details](#))

Copyright © 2016 [OpenLink Software](#)
Virtuoso version 07.20.3214 on Linux (x86_64-unknown-linux-gnu), Single Server Edition

Publikace LOD – RDF dump

<https://data.cssz.cz/dump/nove-priznane-duchody-dle-veku.trig>

@prefix dcterms: <<http://purl.org/dc/terms/>> .
@prefix qb: <<http://purl.org/linked-data/cube#>> .
@prefix rdf: <<http://www.w3.org/1999/02/22-rdf-syntax-ns#>> .
@prefix rdfs: <<http://www.w3.org/2000/01/rdf-schema#>> .
@prefix xml: <<http://www.w3.org/XML/1998/namespace>> .
@prefix xsd: <<http://www.w3.org/2001/XMLSchema#>> .

<<https://data.cssz.cz/ontology/dimension/druh-duchodu>> a qb:DimensionProperty,
rdf:Property .

<<https://data.cssz.cz/ontology/dimension/pohlavi>> a qb:DimensionProperty,
rdf:Property .

<<https://data.cssz.cz/resource/dataset/nove-priznane-duchody-dle-veku>> a qb:DataSet ;
dcterms:description "Počet nově přiznaných důchodů v České republice podle roku, druhu penze, věkové kategorie a pohlaví" @cs,
"Numbers of newly granted pensions in the Czech Republic by year, age, kind of pension and sex" @en ;
dcterms:publisher <<http://www.cssz.cz>> ;
dcterms:subject <http://dbpedia.org/resource/Czech_republic>,
<<https://data.cssz.cz/ontology/sdmx/subject/1.6>> ;
dcterms:title "Počet nově přiznaných důchodů v České republice dle věkové kategorie" @cs,
"Numbers of newly granted pensions in the Czech Republic according to age" @en ;
qb:structure <<https://data.cssz.cz/resource/data-structure-definition/nove-priznane-duchody-dle-veku>> .

Publikace LOD – Dereferencovatelné URI

<http://ruian.linked.opendata.cz/resource/staty/1>

HTML (prohlížeč)



About: Česká republika [Goto](#) [Sponge](#) [NotDistinct](#) [Permalink](#)
An Entity of Type : [ruian:Stat](#), within Data Space : [ruian.linked.opendata.cz:8890](#) associated with source [doi](#)
Type:

Attributes	Values
rdf:type	Území států
skos:notation	1
Definiční bod	http://ruian.linked.opendata.cz/resource/staty/1/point/DST.1
ID transakce v RUIAN	0(xsd:double)
NUTS kód	http://ec.europa.eu/eurostat/ramon/rdfdata/nuts2008/CZ
Začátek platnosti	2011-07-01 00:00:00(xsd:dateTime)
s:name	Česká republika
ogcgm:ld	ST.1
is Nadřazený stát of	Praha Středočeský Jihočeský Západočeský Severočeský »more»
is dct:spatial of	RUIAN - registry of land identification, addresses and properties of the Czech Republic

RDF (aplikace)

```
@prefix gml: <http://www.opengis.net/ont/gml#> .  
@prefix ns1: <http://ruian.linked.opendata.cz/ontology/> .  
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .  
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .  
@prefix schema: <http://schema.org/> .  
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .  
@prefix xml: <http://www.w3.org/XML/1998/namespace> .  
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
```

```
<http://ruian.linked.opendata.cz/resource/staty/1> a ns1:Stat ;  
  ns1:definicniBod <http://ruian.linked.opendata.cz/resource/staty/1/point/DST.1> ;  
  ns1:idTransakce 0.0 ;  
  ns1:nuts <http://ec.europa.eu/eurostat/ramon/rdfdata/nuts2008/CZ> ;  
  ns1:platiOd "2011-07-01T00:00:00+02:00"^^xsd:dateTime ;  
  schema:name "Česká republika" ;  
  gml:id "ST.1" ;  
  skos:notation "1" .
```

Nástroje pro katalogizaci

- CKAN
 - [Open source](#), [Komerční podpora](#)
 - Nejpopulárnější
 - Python, Java, PostgreSQL, Apache Solr
 - CKAN API pro čtení i zápis
- NuCivic DKAN
 - [Open source](#), [Komerční](#)
 - PHP+MySQL (Drupal)
 - Emuluje CKAN API pro čtení, ale ne dobře
 - Vlastní Drupal API pro zápis
- Vlastní řešení emulující CKAN API
 - Soubory odpovídající CKAN API
- NKOD
 - Záznamy pouze v NKOD, odkazy na soubory na webu
 - Záznamy pouze v NKOD, malá data v NKOD



Katalog pro otevřená data

The screenshot shows the Opendata.cz website interface. At the top, there is a blue navigation bar with the 'OPENDATA CZ' logo on the left and a search bar on the right. Below the navigation bar, the breadcrumb trail reads: 'Organizace / Opendata.cz / Léčivé přípravky obchodované ...'. The main content area is divided into a left sidebar and a main panel. The sidebar contains information about the organization 'Opendata.cz', including a description and social media links. The main panel displays the dataset title 'Léčivé přípravky obchodované v ČR' and a list of data sources with 'Prozkoumat' buttons. A filter bar at the bottom of the main panel shows tags for 'lék', 'léčivý přípravek', and 'zdraví'.

OPENDATA CZ Datové sady Organizace Skupiny O nás Vyhledat

Organizace / Opendata.cz / Léčivé přípravky obchodované ...

Léčivé přípravky obchodované v ČR

Následovníci
0
Sledovat

Organizace

OPENDATA CZ

Opendata.cz
We are a group of enthusiasts interested in Linked Data and Open Data. We come from the Charles University in Prague, the Czech Technical University in Prague and the University...
[načíst další](#)

Sdílet
Google+
Twitter

Datová sada Skupiny Události Související Spravovat

Léčivé přípravky obchodované v ČR

Léčivé přípravky obchodované v ČR podle Státního ústavu pro kontrolu léčiv

Data a zdroje

- SPARQL Endpoint Prozkoumat
- Example resource in Turtle Prozkoumat
- Example resource Prozkoumat
- Example resource in Turtle Prozkoumat
- Example resource Prozkoumat
- Léčivé přípravky obchodované v ČR Prozkoumat
Léčivé přípravky obchodované v ČR podle Státního ústavu pro kontrolu léčiv

lék léčivý přípravek zdraví

Doplňující informace

Pole	Hodnota
------	---------

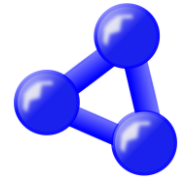
Podpora pro RDF v nástrojích pro katalogizaci

- [DCAT-AP v1.1](#)
 - RDF slovník vyvinutý Evropskou komisí
 - Pro všechny datové portály v Evropě
- CKAN
 - DCAT-AP v1.1 rozšíření v iniciální verzi slíbeno na konec února
- DKAN
 - Na [rozšíření DCAT-AP v1.1 se pracuje](#) ...
- Vlastní řešení
 - RDF soubory odpovídající slovníku

Webové standardy: W3C Recommendation



- Jak webová data zapisovat
 - [RDF 1.1](#)
 - ([Turtle](#), [TriG](#), [N-Triples](#), [N-Quads](#), [RDF/XML](#), [JSON-LD](#), [RDFa](#))
 - Aktuální 2014, první 2004
- Jak se dotazovat a jak měnit
 - SPARQL 1.1 - [Query Language](#), [Update](#), [Service Description](#)
 - Aktuální 2013, první 2008
- Jak webová data získat
 - (2012) Z relační databáze
 - Automaticky: [Direct Mapping](#), Mapování: [R2RML](#)
 - (2015) Z CSV: [Generating RDF from Tabular Data on the Web](#)
 - (2014) Z JSON: [JSON-LD](#)
 - Z XML: (obecně) [XSLT 2.0](#)
 - (2015) [XSLT 3.0](#) je W3C Candidate Recommendation



Webové standardy: Modelování a struktura



- W3C Recommendation
 - (2014) Organizace a jejich struktury - [The Organization Ontology](#)
 - (2014) Statistická data - [The RDF Data Cube Vocabulary](#)
 - (2014) Datové katalogy - [Data Catalog Vocabulary \(DCAT\)](#)
 - (2009) Hierarchie, číselníky - [SKOS Simple Knowledge Organization System Reference](#)
- W3C Group Note
 - (2014) Kontaktní údaje - [vCard Ontology - for describing People and Organizations](#)
 - (2013) Registrované organizace - [Registered Organization Vocabulary](#)
 - (2013) Schémata, modely, slovníky - [Asset Description Metadata Schema \(ADMS\)](#)

Číselník dle SKOS

@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .

@prefix skos: <http://www.w3.org/2004/02/skos/core#> .

<http://ruian.linked.opendata.cz/ontology/stavebni-objekty/PripojeniKanalizace#ConceptScheme>

a skos:ConceptScheme ;

rdfs:label "Připojení stavebního objektu ke kanalizaci"@cs ;

.

<http://ruian.linked.opendata.cz/ontology/stavebni-objekty/PripojeniKanalizace#1>

a skos:Concept ;

skos:inScheme <http://ruian.linked.opendata.cz/ontology/stavebni-objekty/PripojeniKanalizace#ConceptScheme> ;

skos:notation "1" ;

skos:prefLabel "Přípoj na kanalizační síť"@cs ;

skos:altLabel "Přípoj kan.síť"@cs ;

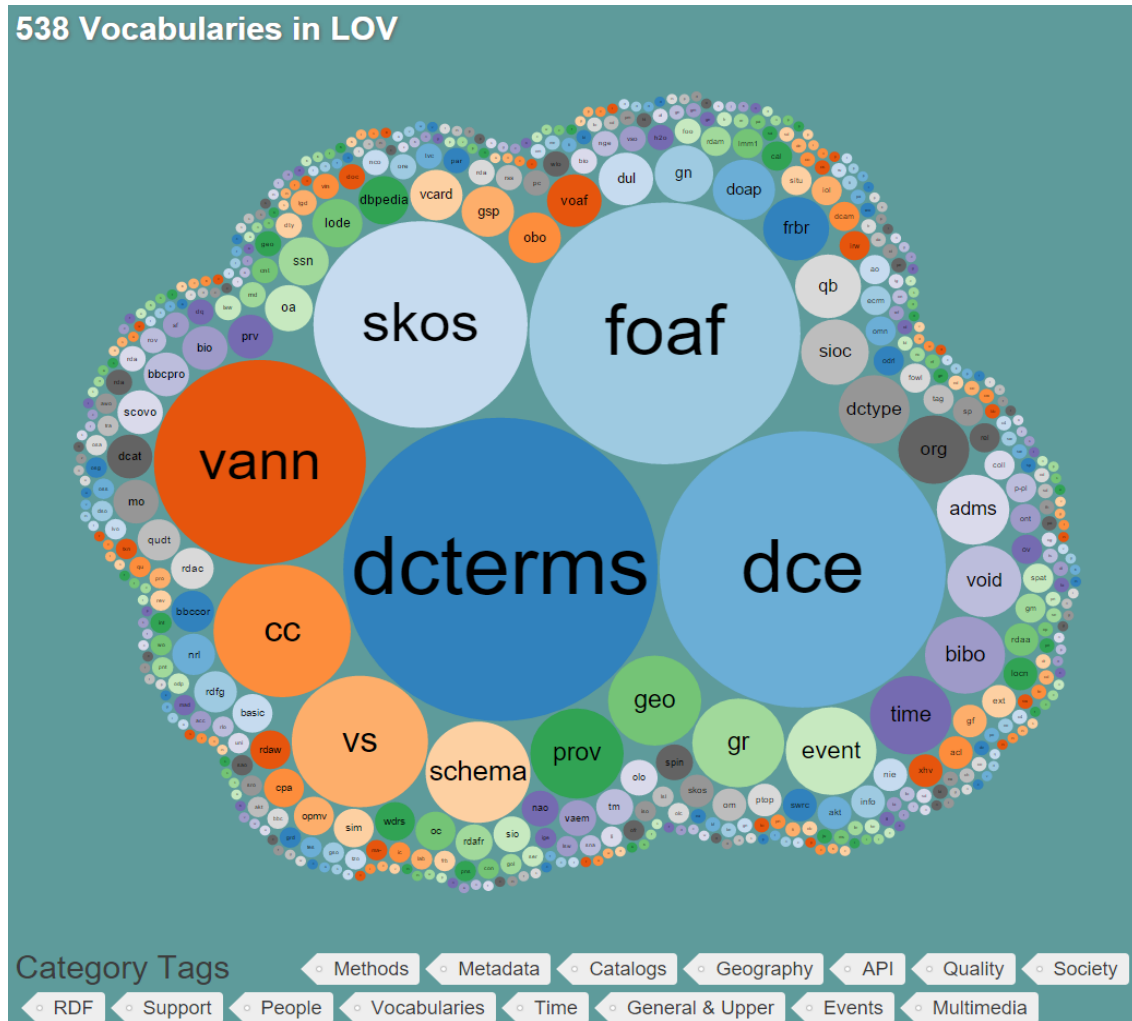
skos:note "Přípoj na kanalizační síť"@cs ;

.

Jiné, známé slovníky

- Schema.org
 - Široký záběr
- Dublin Core
 - Metadata dokumentů
- GoodRelations
 - e-Commerce - zboží, ceny, firmy, pobočky, nabídky, poptávky

Linked Open Vocabularies



Evropské slovníky: EU Publications Office

- Core vocabularies pro eGovernment
 - Business, Location, Person, Public Service, Criterion & Evidence, Public Organisation
- DCAT-AP v1.1
 - Profil DCAT pro Evropské datové portály
- Metadata Registry
 - Named Authority Lists
 - Nejrůznější číselníky, v RDF používají SKOS
 - EuroVoc
 - Tezaurus, pojmy ve všech jazycích EU, opět používá SKOS

Standards and tools for LOD publications

Jakub Klímek
Mathematico-physical faculty
University of Charles in Prague

