



Publishing and consuming Linked Open Data with the LOD2 Statistical Workbench

Valentina Janev, PhD
Senior Researcher

Valentina.Janev@institutepupin.com

Overview

- Linked Data Stack (LOD2, GeoKnow)
- eGovernment Case Study & the Statistical Workbench
- Exploitation of the results (SHARE-PSI)



LOD2 - Creating Knowledge out of Interlinked Data



- Instrument: Large-Scale Integrating Project
- Objective: Intelligent Information Management
- Duration: 09/2010 – 08/2014
- Total Budget: 10,2 M€

- Consortium: **14 Partners from 11 European Countries + 1 Associated Partner from Korea**

UNIVERSITÄT LEIPZIG



NUI Galway
OÉ Gaillimh

OPENLINK
SOFTWARE



Zemanta™
SEMANTIC WEB COMPANY

school • consulting • projects • events • media



Wolters Kluwer
Deutschland

TenForce
Pragmatic Knowledge Management

exalead®
connect the dots

Freie Universität

KAIST

CWI

I²G
INSTYTUT INFORMATYKI
GOSPODARCZEJ

VŠE
SKOLA EKONOMICKA
V PRAZE

iMP
MIHAJLO PUPIN



The Stack

The Linked Data Stack comprises a number of tools for managing the life-cycle of Linked Data. The life-cycle comprises in particular the stages:

- Extraction of RDF from text, XML and SQL
- Querying and Exploration using SPARQL
- Authoring of Linked Data using a Semantic Wiki
- Semi-automatic link discovery between Linked Data sources
- Knowledge-base Enrichment and Repair

<http://stack.linkeddata.org/>
<http://stack.lod2.eu/>



Components in the Stack

Are you interested in using some of the Linked Data stack tools? Check the list of components available in the stack. Most of the tools that are part of the stack have online demos, give them a try!

[View details »](#)

How to Start

The Linked Data Stack is a collection of tools that give support to the Linked Data lifecycle. These tools are created under different research funding programs and third parties, and you can use these tools in your Linked Data creation process.

[View details »](#)

How to Contribute

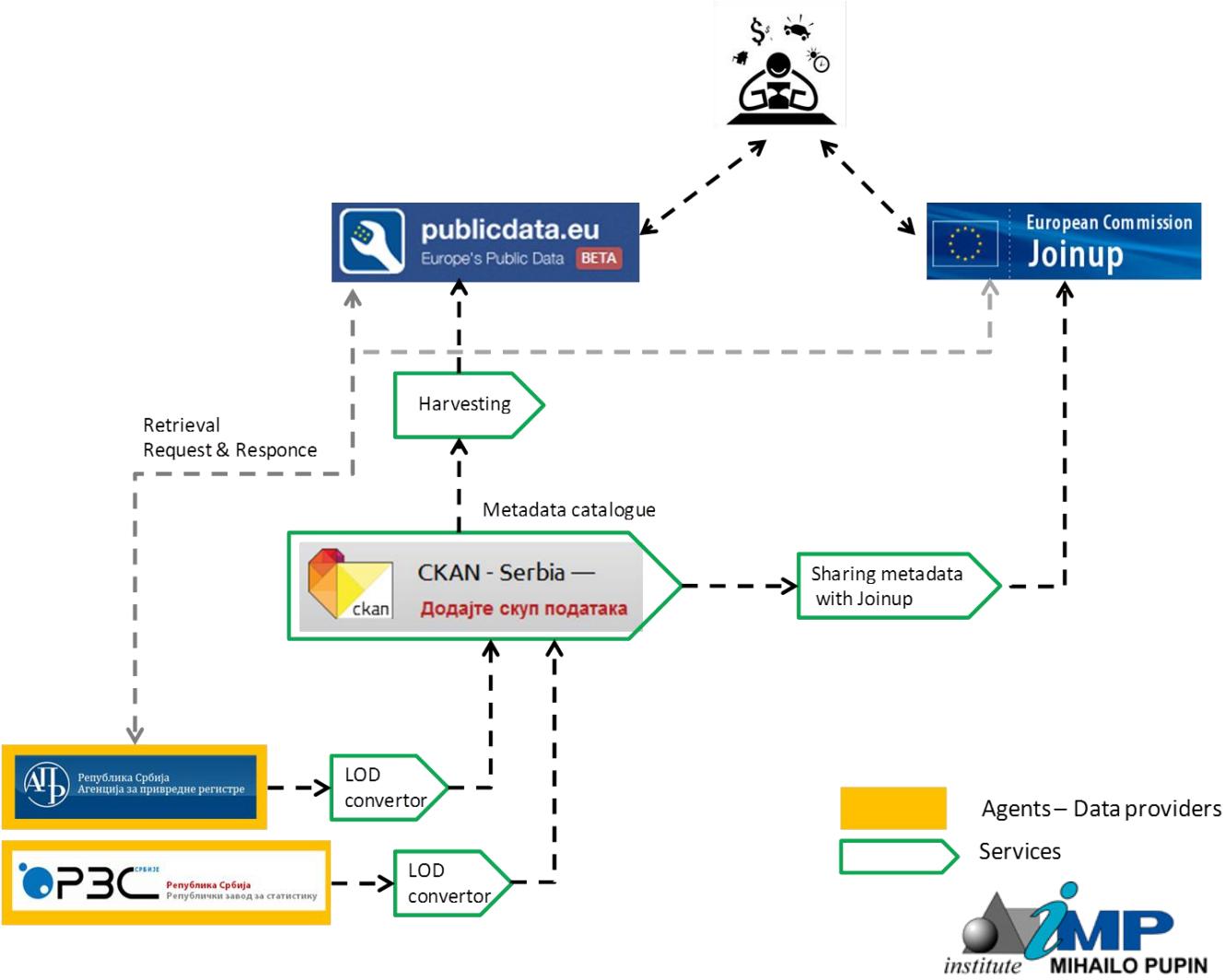
Have you created that missing tool that supports the Linked Data lifecycle?. You can contribute to the Linked Data stack by adding your tool to our repository and reach targeted users from our community.

[View details »](#)

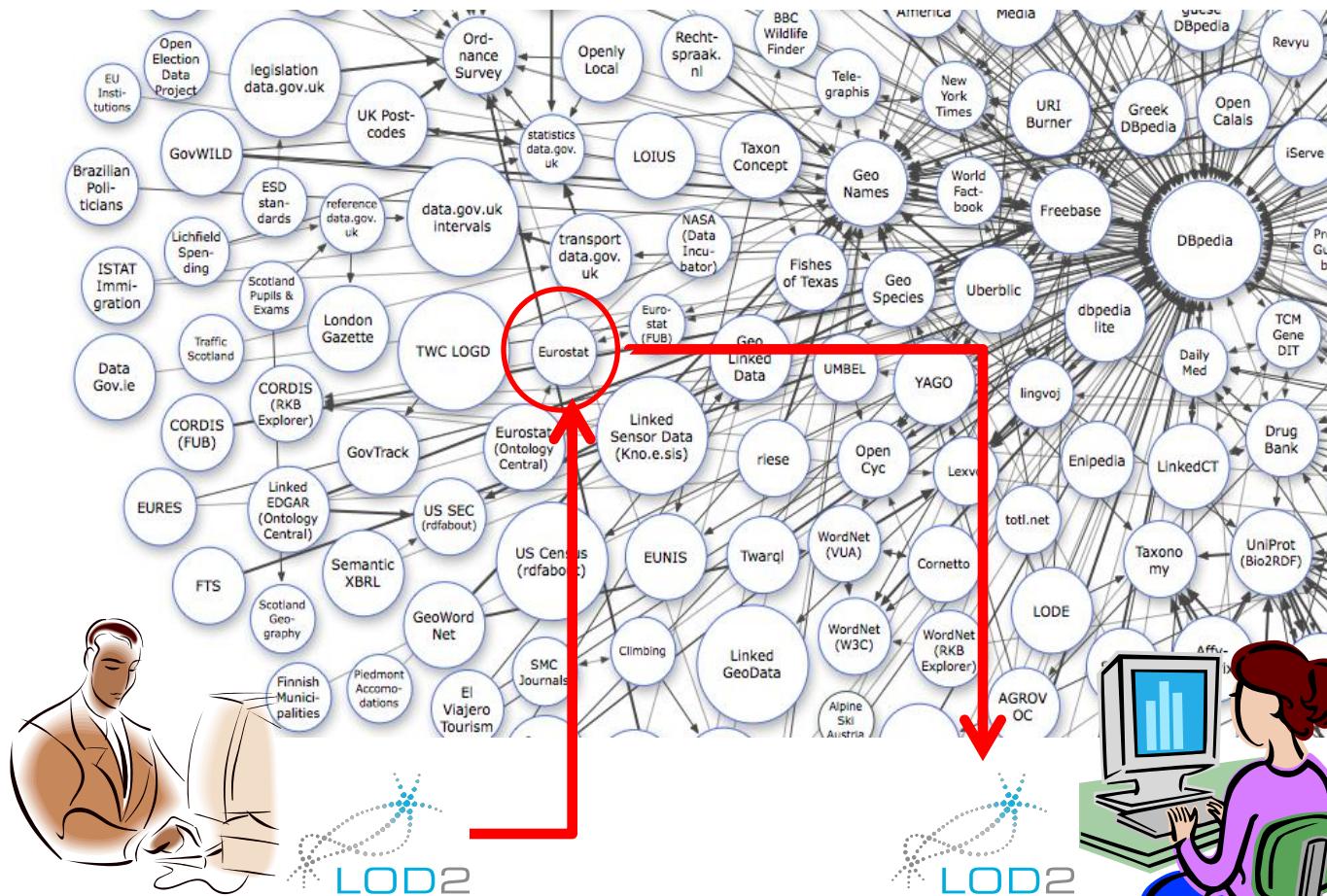
Statistical Office Case Study

Goals

- Contribute to modernization of statistical services
- Define infrastructure for publishing and Integration of public sector information into the LOD cloud
- Provide an integrated collection professional tools for statistical LOD publishing (1) and consuming (2)



LOD2 Statistical Workbench



Statistical Workbench

Statistical Workbench



UNIVERSITÄT LEIPZIG

AKSW

institute
iMP
MIHAJLO PUPIN



Creating Knowledge out of Interlinked Data

Visualization with CubeViz
Geo-Spatial exploration
Publish to CKAN
Publish to Sindice

Manage Graph Find more Data Online Edit & Transform Enrich Datacube Present & Publish Help

Select Default Graph
Create Graph
Import
Export
Validate
Remove Graphs

Manual revision/
authoring

Inter-linking/
Fusing
Edit Graph (OntoWiki)
Edit Code Lists (PoolParty)
Merge datasets
Slice datasets
Transform and Update Graph (SPARQL Update Endpoint)
SparQLed - Assisted Querying
PoolParty Code Lists SPARQL endpoint

Search/
Browsing/
Exploration

Load RDF data from publicdata.eu
Load RDF data from Data Hub
DBpedia
Mondeca SPARQL endpoint Collection

Interlinking dimensions (Silk)

Data enrichment and reconciliation (LODRefine)

Interlinking with Limes

Interlinking with SameAs

- Querying and Exploration using SPARQL
 - Authoring of Linked Data using a Semantic Wiki
- Semi-automatic link discovery between Linked Data sources
- Knowledge-base Enrichment and Repair
- Access tools for each of these stages using the menu

which comprises a
collection of Linked Data.
The process consists of
three main stages:

and SQL

• Querying and Exploration using SPARQL

• Authoring of Linked Data using a Semantic Wiki

• Semi-automatic link discovery between Linked Data

sources

Knowledge-base Enrichment and Repair

Access tools for each of these stages using the menu

The LOD2 Stack is developed by the LOD2 project consortium comprising 15 research groups and companies. The LOD2 project is co-funded by the European Commission within the 7th Framework Programme (GA no. 257934).

You can find further information about the LOD2 Stack at <http://stack.lod2.eu> and the LOD2 project at <http://lod2.eu>.



SORS Dissemination Database

National accounts

[New](#) [Data](#) [Publications](#) [Methodology](#) [Description](#)

SNA2008 / ESA2010 Implementation

As macroeconomic system of data presentation, national accounts offer an insight into overall economy of a country and present the instrument of major importance for analysis, assessments and forecast of economic phenomena. Therefore, the System of National Accounts (SNA) presents a generally adopted international statistical standard. The main aggregates of national accounts and the system of accounts are used for economic policy decision-taking and for international comparisons.

However, new economic environment with the structural changes in the global economy required the new methodology of the system of national accounts. The revision of the SNA was completed in 2008.

The revision of the SNA was completed in 2008. The new system of national accounts (SNA2008) has been implemented in Serbia since 1993, the first year of independence.

The revision of the SNA was completed in 2008. The new system of national accounts (SNA2008) has been implemented in Serbia since 1993, the first year of independence. The revision of the SNA was completed in 2008. The new system of national accounts (SNA2008) has been implemented in Serbia since 1993, the first year of independence.

This document presents the main features of the new system of national accounts (SNA2008). It also describes the main differences between the old system (SNA1993) and the new one.

[SNA2008 Implementation](#)

Data

Annual National Accounts

GDP at current prices

GDP at constant prices

Tourism

SDMX subject	Domain	Properties	Code Lists
1.1	Population and migration	rs:obsIndicator, rs:sex, rs:age, rs:geo, rs:time, qb:measureType sdmx-measure:obsValue	cl:sex, cl:age, cl:geo, cl:time
1.2	Labour statistics on labour force	rs:obsIndicator, rs:age, rs:geo, rs:time sdmx-measure:obsValue	cl:age, cl:geo, cl:time
2.2	Economic accounts	rs:obsIndicator, rs:geo, rs:time, sdmx-attribute:unitMeasure sdmx-measure:obsValue	cl:esa95, cl:geo, cl:time
2.4.5	Tourism	rs:obsIndicator, rs:geo, rs:time, rs:obsTurists, qb:measureType sdmx-measure:obsValue	cl:geo, cl:time, cl:turists

Validation criteria

- [!\[\]\(5c7a31d6d4c42ceb74f6bebcd23d135b_img.jpg\) Summary](#)
- [!\[\]\(103cbb45704d35abee11a17d928d112a_img.jpg\) Provenance information](#)
- [!\[\]\(f2a7d022cb64db9eabeed9791658cced_img.jpg\) IC-1 Unique DataSet](#)
- [!\[\]\(8e8076c207c952aad006095bbb275f6d_img.jpg\) IC-2 Unique DSD](#)
- [!\[\]\(55c24ce002ff25a5f503e93dc461e1cd_img.jpg\) IC-3 DSD includes measure](#)
- [!\[\]\(6188f85c31ba4c153625003e72783792_img.jpg\) IC-4 Dimensions have range](#)
- [!\[\]\(950f2498dbd1808affcb8de0873eebeb_img.jpg\) IC-5 Concept dimensions have code lists](#)
- [!\[\]\(e9bb06b58a413d05269fb2529be12a6f_img.jpg\) IC-6 Only attributes may be optional](#)
- [!\[\]\(c03877dd15b243a95c7fc8753efe237a_img.jpg\) IC-7 Slice Keys must be declared](#)
- [!\[\]\(587f4b559865ad764721ee862f977be3_img.jpg\) IC-8 Slice Keys consistent with DSD](#)
- [!\[\]\(0ce19bf39f39dd27fb4a0ce1ea5f095c_img.jpg\) IC-9 Unique slice structure](#)
- [!\[\]\(4eb929694d5ec364b3439233baaab7ce_img.jpg\) IC-10 Slice dimensions complete](#)
- [!\[\]\(0c73fb0b10bafd8a93bd6fac09b07f3c_img.jpg\) IC-11 All dimensions required](#)
- [!\[\]\(0b19c67214401e22e665feceb1a6c03d_img.jpg\) IC-12 No duplicate observations](#)
- [!\[\]\(633879b8148d5be909c4d8926b1dbad5_img.jpg\) IC-13 Required attributes](#)
- [!\[\]\(7a6147f95f8008ca5eacb81d188e26c2_img.jpg\) IC-14 All measures present](#)
- [!\[\]\(d586b5b88d7806bfdb9a08b0a59d68d3_img.jpg\) IC-15 Measure dimension consistent](#)
- [!\[\]\(a9d46c72cfafda670ab4b7ede64083c8_img.jpg\) IC-16 Single measure](#)
- [!\[\]\(e2fc499d698007d0e7490877a14d7417_img.jpg\) IC-17 All measures present in meas. dim. cu](#)
- [!\[\]\(ac433150f1ee9bb402b2ed0827408c04_img.jpg\) IC-18 Consistent data set links](#)
- [!\[\]\(da79ee53170bc029370e4836be0d645e_img.jpg\) IC-19 Codes from code list](#)

Below is the list of observations that are not linked to exactly one data set. Click on any of them resource in OntoWiki or choose a quick solution

Observations

http://elpo.stat.gov.rs/lod2/RS-DATA/National_accounts/Gross_value_added_by_activities_at_current_prices/d

Details

PROPERTY	OBJECT
http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://purl.org/linked-data/resource#type
http://elpo.stat.gov.rs/lod2/RS-DIC/rs/geo	http://elpo.stat.gov.rs/lod2/RS-DIC/rs/geo
http://elpo.stat.gov.rs/lod2/RS-DIC/rs/time	http://elpo.stat.gov.rs/lod2/RS-DIC/rs/time
http://purl.org/linked-data/sdmx/2009/attribute#unitMeasure	http://elpo.stat.gov.rs/lod2/RS-DATA/National_accounts/Gross_value_added_by_activities_at_current_prices/d
http://purl.org/linked-data/sdmx/2009/measure#obsValue	357785.5

Problem description: The selected observation belongs to 0 data sets. It should belong to exactly one.

Quick Fix

After the fix the selected observation will belong only to the data set selected below or you can manually in OntoWiki

http://elpo.stat.gov.rs/lod2/RS-DATA/National_accounts/Gross_value_added_by_activities_at_current_prices/d

[Quick Fix](#)
[Edit in OntoWiki](#)

Merging RDF Data Cubes



Creating Knowledge out of Interlinked Data

[home](#) http://elpo.stat.gov.rs/lod2/RS-DATA/National_accounts/

[Log in](#)

[Manage Graph](#) [Find more Data Online](#) [Edit & Transform](#) [Enrich Datacube](#) [Present & Publish](#) [Help](#)

Merge datasets

Data cubes from current graph Show data cube labels

Select a (reference) dataset to merge

http://elpo.stat.gov.rs/lod2/RS-DATA/National_accounts/GDP_usage_Exports_of_goods_and_services/data_2011_12_28 (D: 2, M: 1, A: 0)

[Remove](#)

Select a dataset to merge

http://elpo.stat.gov.rs/lod2/RS-DATA/National_accounts/GDP_usage_Imports_of_goods_and_services/data_2011_12_28 (D: 2, M: 1, A: 0)

[Remove](#)

[Add dataset](#)

Select the graph to enter the new datacube in

http://elpo.stat.gov.rs/lod2/RS-DATA/National_accounts/GDP_usage



URI for the new datacube

sage_Exports_Imports

[Merge datasets](#)



Slicing RDF Data Cube



Creating Knowledge out of Interlinked Data

[home](#) http://stat.apr.gov.rs/lod2/RS-DATA/Tourism/RS_Tourism

[Log in](#)

[Manage Graph](#) [Find more Data Online](#) [Edit & Transform](#) [Enrich Datacube](#) [Present & Publish](#) [Help](#)

Choose dataset

Choose dimension

Choose value

[Remove](#)

Choose dimension

Choose value

[Remove](#)

[Add dimension](#)

SliceKey URI

"Total"

"Domestic"

SliceKey Label

Slice URI

Slice Label

[Create slice](#)

[Generate URI\(s\) and label\(s\)](#)



Data Selection

Select a dataset

GDP usage Individual...

CubeViz

THE RDF DATACUBE BROWSER

The selected RDF dataset contains a subdivision called "DataCube datasets" which are used to group different items. The structure of a DataCube dataset corresponds to a specific data structure definition. There may be different DataCube datasets share their corresponding DataStructureDefinition. In the selection dialogue you have to change the selected DataCube dataset. If you do not select one of the subdivision, CubeViz will select one automatically.

1

- GDP usage Individual consumption expenditure
- GDP usage, Import, Export
- Population aged 15 and over, by age and region

4

Select Unit and Measurement

Units

Observed Value

2

Configure the Dimensions

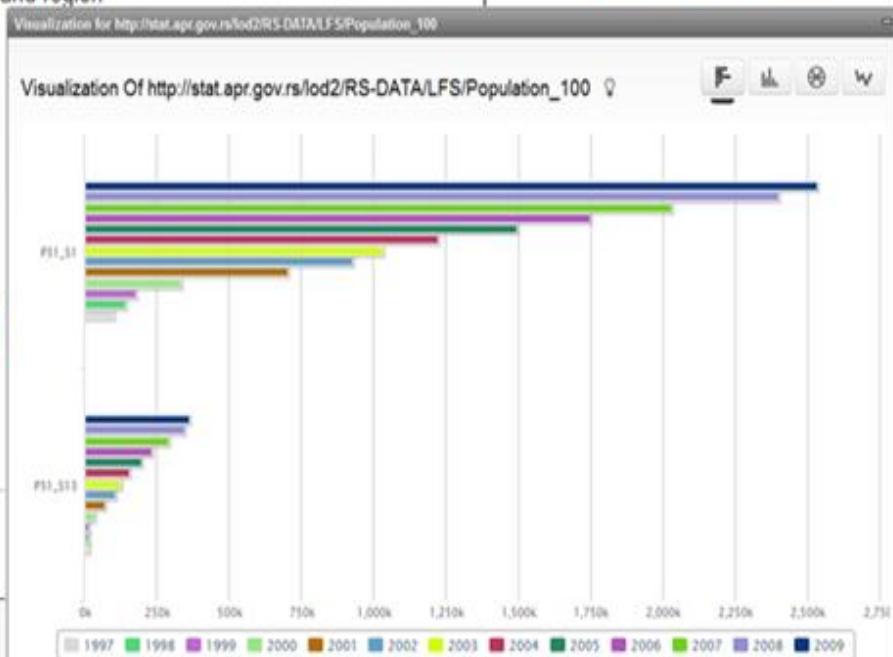
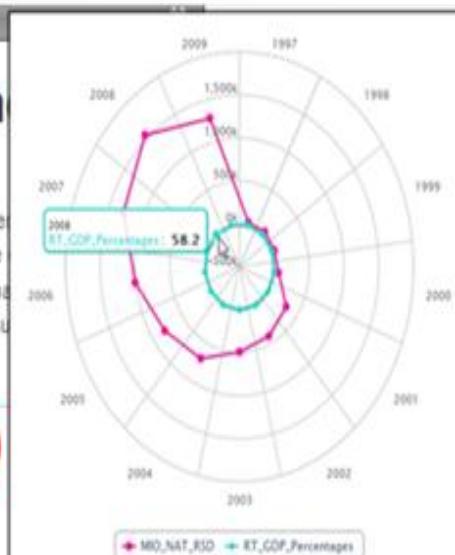
Observed Sector

4 of 4 Selected

Times

13 of 13 Selected

3

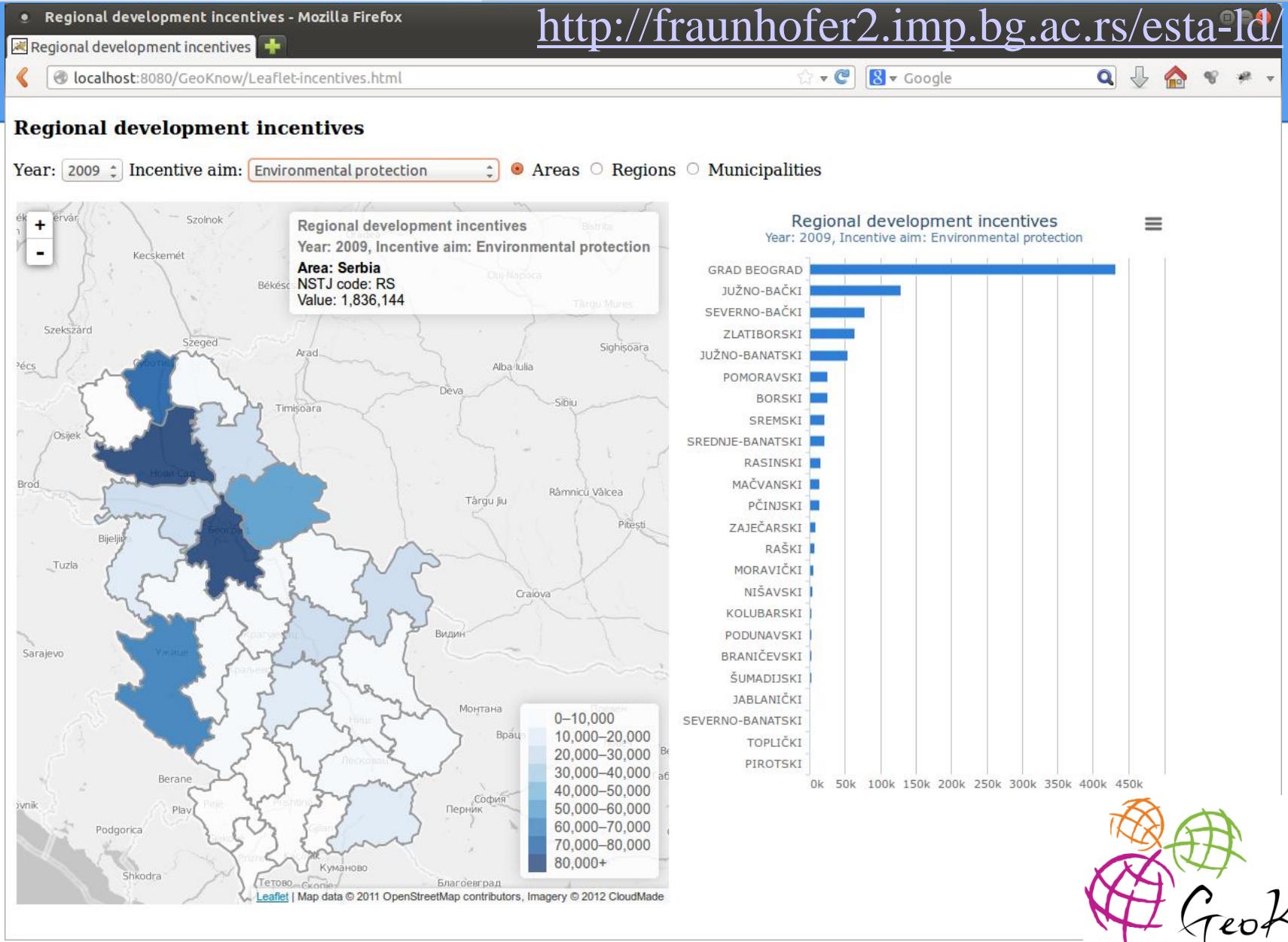


Link

Show Visualization

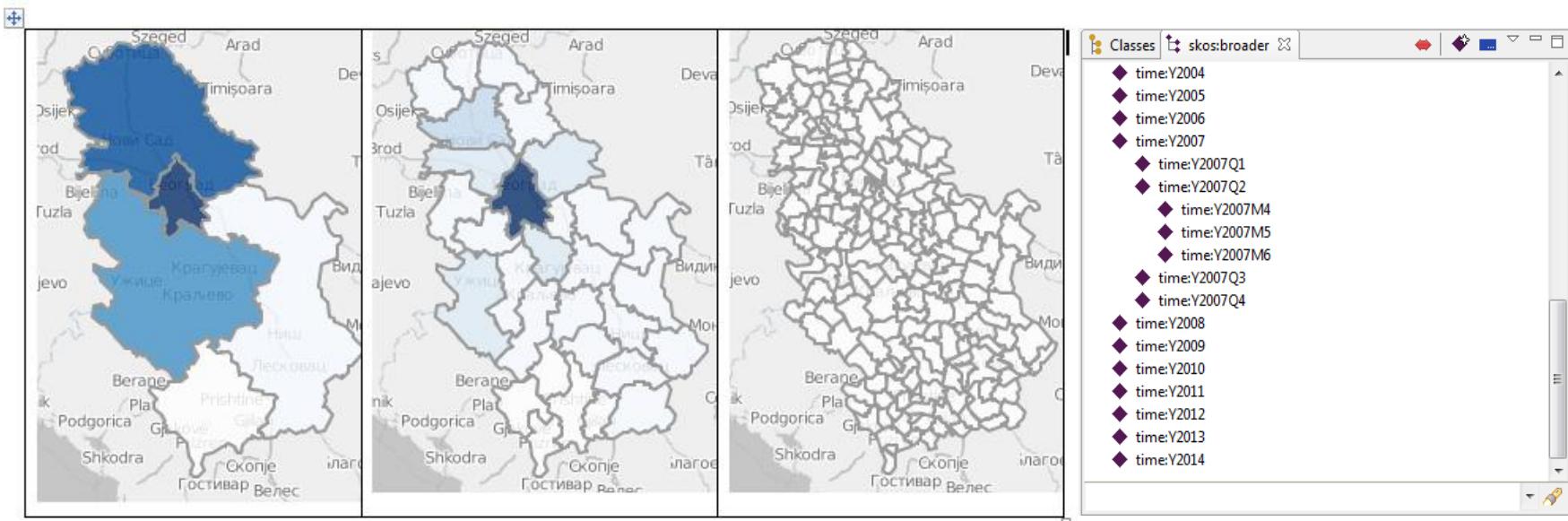
Show Information About Selected Configuration Retrieved Data

Exploratory Spatio-Temporal Analysis



Exploratory Spatio-Temporal Analysis

- ESTA-LD component - to be based on standard W3C Vocabularies (Core Government, RDF Data Cube, Time vocabulary)



CKAN repository information

CKAN repository*

http://rs.ckan.net

Visualization with CubeViz

Geo-Spatial exploration

Publish to CKAN 

Publish to datahub.io

Publish to Sindice

ckan api key*

ff766a71-bbc1-4ebd-bd50-9ce4ecac2921

CKAN package information

name*

rdmi

Extra details?

Select Package



Добродошли на CKAN - Serbia!

Проналажење податка

[Нађи скупове података](#)

СКАН - Serbia садржи **168 скуп(ов)а података** које можете прегледати, учити о њима и преузети.

Дељење података



Add your own datasets to share them with others and to find other people interested in your data.

[Региструјте се »](#)

Сарадња



Find out more about working with open data by exploring these resources:

- [GetTheData.org](#)
- [DataPatterns.org](#)
- [Open Data Handbook](#)

Ко је још овде?

Statistical Office

The Statistical Office of the Republic of Serbia (SORS) is a special professional organization in the system of state administration of the Republic of Serbia that performs the expert...

Statistical Office има 91 скуп(ов)а података.

Ministry of Education and Science

Republic of Serbia

Ministry of Education and Science има 4 скуп(ов)а података.

National Bank of Serbia

The National Bank of Serbia is independent and autonomous in fulfilling its functions stipulated by the NBS Law and other laws, and is accountable for its work to the National Assembly of...

National Bank of Serbia има 41 скуп(ов)а података.

Serbia Investment and Export Promotion Agency

Serbia Investment and Export Promotion Agency (SIEPA) is a government organization dedicated to effectively helping foreign investors and buyers, while raising Serbia's profile in the...

Serbia Investment and Export Promotion Agency има 4 скуп(ов)а података.

Serbian Business Registers Agency

The Serbian Business Registers Agency (SBRA) was established by the Business Registers Agency Law in 2004, while the first electronic register (Register of Companies) started with its...

Serbian Business Registers Agency има 17 скуп(ов)а података.

Serbian Chamber of Commerce

Republic of Serbia

Serbian Chamber of Commerce има 2 скуп(ов)а података.

Benefits for early adopters

- improve the accessibility and transparency of data by extending the existing public services with new features
- standardize the data publishing / consumption process
- ensure interoperability (e.g. integration of data from the Register with data from the Dissemination database of the Statistical Office of the Republic of Serbia, SORS)
- allow advanced analysis and visualization of available indicators (i.e. spatio-temporal data on a geographical map)

More info ...

- <http://stack.linkeddata.org/>, <http://stack.lod2.eu/>
- <http://rs.ckan.net/>
- <http://fraunhofer2.imp.bg.ac.rs/lod2statworkbench>

Thank you for your attention!