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
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!

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 founding programs and third parties, and you can use these tools in your Linked Data creation process.

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 out community.
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 of tools for statistical LOD publishing (1) and consuming (2)
LOD2 Statistical Workbench
Creating Knowledge out of Interlinked Data

Managing Graph
- Select Default Graph
- Create Graph
- Import
- Export
- Validate
- Remove Graphs

Edit & Transform
- Interlinking/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

Enrich Datacube
- Interlinking dimensions (Silk)
- Data enrichment and reconciliation (LODRefine)
- Interlinking with Limes
- Interlinking with SameAs

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

Help

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.

http://fraunhofer2.imp.bg.ac.rs/lod2statworkbench
National accounts

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 since 1993, the revision frame adjusted countries will be.

The revision frame adjusted countries will be.

Observing the structure contained in the prepared the population and migration.

This document several key steps are documented national accounts.

<table>
<thead>
<tr>
<th>SDMX subject</th>
<th>Domain</th>
<th>Properties</th>
<th>Code Lists</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Population and migration</td>
<td>rs:obsIndicator, rs:sex, rs:age, rs:geo, rs:time, qb:measureType sdmx-measure:obsValue</td>
<td>cl:sex, cl:age, cl:geo, cl:time</td>
</tr>
<tr>
<td>1.2</td>
<td>Labour statistics on labour force</td>
<td>rs:obsIndicator, rs:age, rs:geo, rs:time sdmx-measure:obsValue</td>
<td>cl:age, cl:geo, cl:time</td>
</tr>
<tr>
<td>2.2</td>
<td>Economic accounts</td>
<td>rs:obsIndicator, rs:geo, rs:time, sdmx-attribute:unitMeasure sdmx-measure:obsValue</td>
<td>cl:esa95, cl:geo, cl:time</td>
</tr>
<tr>
<td>2.4.5</td>
<td>Tourism</td>
<td>rs:obsIndicator, rs:geo,rs:time, rs:obsTurists, qb:measureType sdmx-measure:obsValue</td>
<td>cl:geo, cl:time, cl:tourists</td>
</tr>
</tbody>
</table>
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**


**Details**

<table>
<thead>
<tr>
<th>PROPERTY</th>
<th>OBJECT</th>
</tr>
</thead>
<tbody>
<tr>
<td><a href="http://www.w3.org/1999/02/22-rdf-syntax-ns#type">http://www.w3.org/1999/02/22-rdf-syntax-ns#type</a></td>
<td><a href="http://purl.org/linke">http://purl.org/linke</a>...</td>
</tr>
<tr>
<td><a href="http://elope.stat.gov.rs/lo">http://elope.stat.gov.rs/lo</a>...</td>
<td><a href="http://elope.stat.gov.rs/lo">http://elope.stat.gov.rs/lo</a>...</td>
</tr>
<tr>
<td><a href="http://elope.stat.gov.rs/lo">http://elope.stat.gov.rs/lo</a>...</td>
<td><a href="http://elope.stat.gov.rs/lo">http://elope.stat.gov.rs/lo</a>...</td>
</tr>
<tr>
<td><a href="http://purl.org/linked-data/sdmx/2009/attribute#unitMeasure">http://purl.org/linked-data/sdmx/2009/attribute#unitMeasure</a></td>
<td><a href="http://elope.stat.gov.rs/lo">http://elope.stat.gov.rs/lo</a>...</td>
</tr>
</tbody>
</table>

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

**Quick Fix**

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


**Quick Fix**

[Quick Fix] [Edit in OntoWiki]
Merging RDF Data Cubes

Creating Knowledge out of Interlinked Data

Merge datasets

- Data cubes from current graph
- Show data cube labels

Select a (reference) dataset to merge:

Select a dataset to merge:

Select dataset to enter the new datacube in:

URI for the new datacube:
sage_Exports_Imports

Merge datasets
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 the option 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
2. GDP usage Import, Export
3. Population aged 15 and over, by age and region
Exploratory Spatio-Temporal Analysis

Regional development incentives
Year: 2009, Incentive aim: Environmental protection

Area: Serbia
NSTJ code: RS
Value: 1,836,144

Regional development incentives
Year: 2009, Incentive aim: Environmental protection

Grad Beograd
Južno-bački
Severno-bački
Zlatiborski
Južno-banatski
Pomoravski
Borski
Sremski
Srednje-banatski
Rasinski
Macvanski
Pčinjski
Zajecarski
Rački
Moravicki
Nisavski
Kolubarski
Podunavski
Branichevski
Šumadijski
Jablanicki
Severno-banatski
Toplici
Pirotski
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

**ckan api key**
ff766a71-bbc1-4ebd-bd50-9ce4ecac2921

CKAN package information

**name**
rdmi

☐ Extra details?

- Visualization with CubeViz
- Geo-Spatial exploration
- Publish to CKAN
- Publish to datahub.io
- Publish to Sindice
Добродошли на CKAN - Serbia!

Проналажење податка

Нађи скупове података

CKAN - Serbia садржи 168 скуп(ов) података које можете прегледати, уочити о њима и преузети.

Дељење података

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

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 (SDRS) is a special professional organization in the system of state administration of the Republic of Serbia that performs the expert...

**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...

**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...

**Ministry of Education and Science**
The Ministry of Education and Science is a professional ministry that works towards the development of education and science in the Republic of Serbia...

**Serbian Investment and Export Promotion Agency**
Serbia Investment and Export Promotion Agency (S/IEPA) is a government organization dedicated to effectively helping foreign investors and buyers, while raising Serbia’s profile in the...

**Serbian Chamber of Commerce**
The Serbian Chamber of Commerce is an organization that represents the interests of businesses in Serbia.
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!