Open Government Data on the Web: A Semantic Approach

Julia Hoxha
Institute of Applied Informatics and Formal Description Methods (AIFB)
Karlsruhe Institute of Technology, Germany
Julia.Hoxha@kit.edu

Armand Brahaj
FIZ Karlsruhe - Leibniz Institute for Information Infrastructure, Germany
Armand.Brahaj@fiz-karlsruhe.de

***Abstract*— Initiatives of making government data open are continuously gaining interest recently. While this presents immense benefits for increasing transparency, the problem is that the data are frequently offered in heterogeneous formats, missing clear semantics that clarify what the data describe. The data are displayed in ways, which are not always clearly understandable to a broad range of user communities that need to make informed decisions.**

We address these problems and propose an overall approach, in which raw statistical data independently gathered from the different government institutions are formally and semantically represented, based on an ontology that we present in this paper. We further introduce the approach deployed in publishing these data in alignment with Linked Data principles, as well as present the methods implemented to query single or combined dataset and visualize the results in understandable ways. The introduced approach enables data integration, leading to vast opportunities for information exchange, analysis on combined datasets, simplicity to create mashups, and exploration of innovative ways to use these data creatively.

***Keywords*— linked open data, open government data, ontology, semantic web, statistical metadata**

I. Introduction

INITIATIVES of making government data open are gaining great interest recently, with more countries constantly embracing the open data paradigm. Open Data is a new form of representation of the traditional public information, which has been the basis of decision making in political, economical or sociological activities for centuries.

The term open has been introduced in the late 2000th to generalize the accessibility of the information that can be public and easily accessible by anyone. The concept of information is also generalized to the concept of data, by referring to raw sets of valuable information which can be structured, analyzed and presented in different forms leading to knowledge representation.

From a technical perspective, we refer as Open Data to any sets of data which can be reused with no restrictions by any form of licensing or patents, data that are well structured and can be easily accessed and reused by institutions, scientist or the web community.

Being driven by benefits of the aforementioned paradigm, we aim to increase openness of data offered by different government institutions, and at the same time, provide means of representing these data in more understandable ways. There are several challenges underlying our work, starting from the way that data is originally provided by the sources. Data are mostly offered in heterogeneous formats, missing clear semantics that clarify what the data describes, and displayed in ways that is not clearly understandable to a broad range user communities that need to make informed decisions. These kinds of representation impede the integration of data coming from different sources, consequently leading to limited opportunities for information exchange, analysis on combined datasets, simplicity to create mashups, and exploration of innovative ways to use these data creatively.

From a technical perspective, our work consists in applying Semantic Web technologies to enable data integration among the different organizations and establish links to interconnect data together on the Web. Since a crucial issue when semantically modeling data is the preservation of their meaning, we present an approach, which deploys Linked Open Data[1] principles and introduces an ontology as foundation for the semantic representation of government statistical data. On the one hand, the ontology offers a formal, explicit description of knowledge enabling in this way a platform for integration, where data can be still interpreted correctly for valid research outcomes. On the other hand, Linked Data provides a technical basis for publishing, sharing and linking data on the web, based on standardized formats and interfaces.

In this paper, we present the ontology that enables formal, semantic representation of government statistical data. We further introduce the approach we deploy in publishing them as linked data, and the methods implemented to query single

[1] http://www.linkeddata.org

or combined dataset and visualize the results. The semantic approach and the ontology are described in details in Section 2, whereas Section 3 gives an overview of Linked Open Data and our work in alignment to those principles. Section 4 presents the technical implementation of our approach focusing on the issues of integration and processing of data from different sources and querying and visualization techniques for representing multiple datasets in a graph. We give an overview of related work in the field of open data and statistical vocabularies in Section 5, concluding in Section 6 with a summary of our work.

II. Semantic Approach

In our work, we are dealing with a large set of sources from which statistical data is gathered. These sources are mainly websites of different government institutions, which offer data online in unstructured or semi-structured formats such as text documents, excel files or XML files. There are very few sources that can provide data structured in Entity-Relationship model.

Our goal, after gathering the raw data independently from the different sources, is to convert them to a representation based on a unified structure and provide semantics for each of the data entries. For this purpose, we have chosen to semantically represent the data based on an ontology that we have built, and then further serialize data as RDF triples following Linked Data principles. In this section, we present the modeled ontology as well as introduce our approach in publishing Linked Open Data.

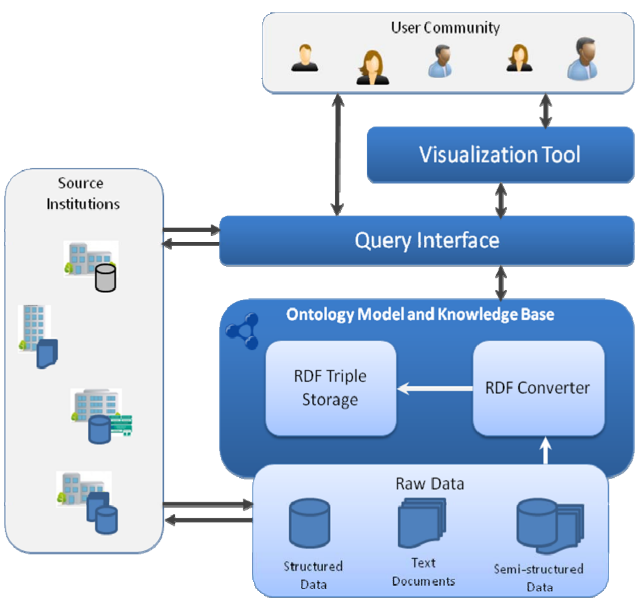

Fig. 1. Overall semantic approach for publishing open government data

In our approach, the knowledge base consists of the Ontology Model and a storage for RDF files. The raw statistical data are converted to RDF[2] using a converter in alignment with Linked Data principles, which we present in Section 3.

Another important component in our approach is the Query Interface, which enables the user community as well as the source institutions that offer these statistical data to access the knowledge base and pose queries upon it. This component consists of an online graphical interface as well as a SPARQL[3] Endpoint, which we both present in Section 4. The results of a query may be displayed as structured Excel and RDF files to the users, but they are additionally visualized in ways that make these data much more tangible and understandable. This approach provides open, structured and linked data to the different institutions, which are then able to use each-others datasets very easily.

A. ODA Ontology

The need to provide a unified representation for the different, heterogeneous data gathered from the various sources, as well as the necessity to provide formal semantics to these data, have greatly motivated us to choose a semantic approach using an ontology, referred as ODA[4] ontology, which we make available online[5]. While modeling the ontology, our goal was to follow linked data principles and reuse as much as possible existing well-known ontologies, but at the same time extending it with other important entities.

An ontology is defined as a "a formal, explicit specification of a shared conceptualization" [1]. It provides a common understanding of the domain knowledge, increasing communication capabilities between people and heterogeneous applications. Using an ontology, we are able to formally describe the semantics of terms and items of statistical datasets and give explicit meaning to the provided information. Besides sharing a common understanding, another direct benefit of a formal semantics is the support for machine-readability, which then enables automated reasoning, information integration and application of intelligent approaches such as semantic search.

The objective of the modeled ontology is to provide a schema (also referred to as vocabulary) for the semantic representation of statistical data, which, on one side, is understandable and simple to be used from different institutions or organizations that offer statistical datasets, and on the other side, is consistent and complete enough to model all the aspects and entries in these datasets.

A dataset is created as a container for the data entries found in a particular statistical file. Each dataset contains information about the date it was created, its creator and publisher, and a title. The dataset contains data entries, which represent the central component modeled in our ontology. A data entry has a fixed value and, among other possible dimensions, is measured for a particular year and a specific country.

In a dataset, the data entries serve to measure particular indicators that monitor for example the socio-economic development of a country, demographic changes, public

[2] Resource Description Framework, http://www.w3.org/RDF/

[3] http://www.w3.org/TR/rdf-sparql-query/
[4] ODA stands for Open Data Albania
[5] http://open.data.al/oda.owl

health, developments in education, and many other aspects. These numerous indicators, also organized among each-other in groups, are additionally assigned particular topics. Each indicator is used to estimate and monitor the progress done in a particular aspect. Besides, a more complete view is provided when looking at the progress of all indicators within a topic.

All these concepts are modeled in the ontology, which is illustrated in Figure 2. It is encoded using the sublanguage *OWL DL* of the Web Ontology Language (OWL)[6], which goes beyond other standards in its power to represent machine interpretable content on the Web. Moreover, OWL DL itself offers high expressiveness.

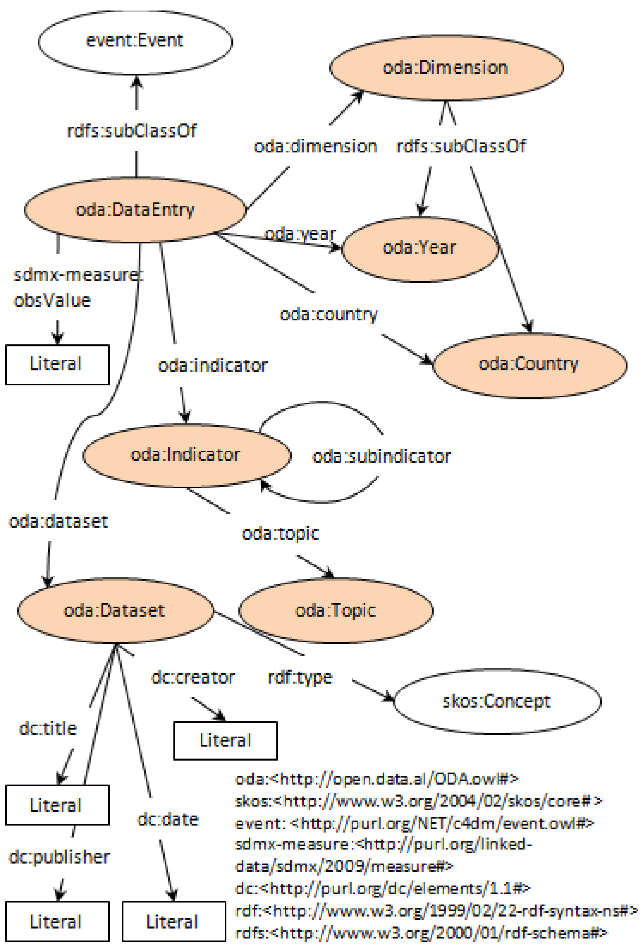

Fig. 2. ODA Ontology

We have modeled the ontology using Protégé[7], a platform that offers support for ontology creation, visualization and manipulation. We explain below the main entities that compose the modeled ODA ontology.

Classes. The concepts and terms of the domain of knowledge are represented in an ontology as classes. Our ontology consists of the following main classes:

Class: oda:Dataset
Dataset - a statistical dataset, represents the container of a set of data
URI: http://open.data.al/oda#Dataset
Superclass: `skos:Concept`

Class: oda:DataEntry
DataEntry- a data entry dataset, represents a single piece of data
URI: http://open.data.al/oda#DataEntry
Superclass: event:Event

Class: oda:Indicator
Indicator - a statistical indicator that is being measured by the entries in the dataset
URI: http://open.data.al/oda#Indicator

Class: oda:Topic
Topic - a theme to which the indicator and dataset belong
URI: http://open.data.al/oda#Topic

Class: oda:Dimension
Dimension - a dimension of a statistical data entry
URI: http://open.data.al/oda#Dimension

Class: oda:Year
Year- time dimension, represents the year to which a statistical data entry belongs
URI: http://open.data.al/oda#Year
Superclass: `oda:Dimension`

Class: oda:Country
Country- spatial dimension, represents the country to which a statistical data entry belongs
URI: http://open.data.al/oda#Country
Superclass: `oda:Dimension`

Properties. The relations among the different individuals of the ontology classes are defined by using the properties construct, distinguishing between two main types: *Object properties* and *Datatype properties*. The *Object properties* link individuals to individuals, whereas the *Datatype properties* link individuals to data values. These are a few of the main Object properties modeled in our ontology:

Property: oda:dataset
dataset- relates a statistical data entry to a dataset
Domain: `oda:DataEntry`
Range: `oda:Dataset`

Property: oda:indicator
indicator- relates a statistical data entry to an indicator
Domain: `oda:DataEntry`
Range: `oda:Indicator`

Property: oda:subindicator
subindicator- relates a parent indicator to other indicators classified as its subcomponents (serves for grouping indicators)
Domain: `oda:Indicator`
Range: `oda:Indicator`

[6] http://www.w3.org/TR/owl-guide/
[7] http://protege.stanford.edu/

Property: oda:dimension
indicator- relates a statistical data entry to an indicator
Domain: `oda:DataEntry`
Range: `oda:Dimension`

We have also defined a set of Datatype properties e.g. `oda:indicatorname` or `oda:topicname` properties, which assign a literal value (in this case string) to the individuals of the class Indicator and Topic, respectively.

Individuals. The classes of an ontology are instantiated, specifying the concrete objects or individuals that belong to a particular class. We have created a large number of topic instances such as *Social Development*, *Economic Development*, *Social Inclusion*, *Public Health*, *Education*, etc. We have related these topics to indicators, such that a topic may have multiple indicators assigned to it. In order to formalize the topics consistently, we have also consulted the way statistical data is being organized into particular themes from Eurostat[8], the statistical office of the European Union.

The indicators are modeled after a thorough analysis of the statistical data offered by the different government institution. They are organized in groups using the `oda:subindicator` relation, forming this way a hierarchy of parent-child indicators.

We have currently created in our ontology instances of more than 100 indicators. Just to give an insight on what do these indicators consist, we mention below a few of them belonging to the topic of *Economic Development*:

- GDP Indicators: GDP nominal per capita, GDP nominal growth rate, GDP per capita in PPS (for EU states)
- Budget Indicators: Expenditures (Total Expenditures, Capital Expenditures, Current Expenditures, Domestic/Foreign Financing Expenditures), Revenue, Deficit (Deficit Growth Rate, Deficit Financing), etc.
- Inflation Indicators: monthly Inflation, yearly Inflation rate.

Aiming to rigorously follow the Linked Data principles, explained in more details in Section 3, one of our goals while modeling the ontology was to reuse existing vocabularies/ontologies. The following entities are used from the ontologies that we have imported:

Class: event:Event
Event – An arbitrary classification of a space/time region, by a cognitive agent. An event may have actively participating agents, passive factors, products, and a location in space/time.
Ontology: The Event Ontology[9]

Class: skos:Concept
Ontology: Simple Knowledge Organization System (SKOS)[10]

Property: sdmx-measure:obsValue
obsValue- The value of a particular variable at a particular period
Ontology: Statistical Data and Metadata Exchange (SDMX) vocabulary[11]

Property: dc:title, dc:publisher, dc:creator, dc:date
Ontology: Dublin Core Metadata Element Set[12]

III. Linking Open Government Data

Making data freely available for everyone on the Web is further empowered when there are well-formed links established among the items of the different data sources. Making data *open* and *linked* is the goal of Linked Open Data paradigm [2], which promotes the ways of providing data that is freely accessible, verifiably, and connected to other data sources.

There are four basic principles expected to be fulfilled when aiming to publish linked data, in order to make these data interconnected [3]:

- Use Uniform Resource Identifiers (URIs) as names to identify things
- Use HTTP URIs so that people can refer to and look up ("dereference") those names
- When someone looks up a URI, provide useful information using standards such as RDF, SPARQL
- Include links to other URIs, so that more things and related information can be discovered on the Web

From a technical perspective [5], the objective is to use common standards and techniques to extend the Web by publishing data as RDF (graph data format for representing information on the Web), creating well-formatted RDF links between the data items, and performing search on the data via standardized languages such as SPARQL query language for RDF.

Recognizing the many benefits of Linked Open Data, we have also followed an approach that is in alignment with these principles, illustrating it with the following example in Figure 3. The figure shows an excerpt of the RDF/XML representation of the information contained in a dataset on the yearly inflation rates of different countries in Europe.

```
<rdf:Description
 rdf:about="http://open.data.al/yearlyinflationrate#dataset">
    <rdf:type rdf:resource="http://open.data.al/oda#Dataset"/>
    <dc:title>Inflation Rate per year and country</dc:title>
    <dc:publisher>International Monetary Fund (IMF)</dc:publisher>
    <dc:creator>Open Data Albania (ODA)</dc:creator>
    <dc:date>2011-03-22T10:56:40+0000</dc:date>
</rdf:Description>

<rdf:Description rdf:nodeID="A0">
  <oda:dataset
    rdf:resource="/yearlyinflationrate#dataset"/>
  <oda:indicator
   rdf:resource="http://open.data.al/oda#YearlyInflationRate"/>
  <oda:country rdf:resource="http://dbpedia.org/resource/Greece"/>
  <oda:year rdf:resource="http://dbpedia.org/resource/2002"/>


  <sdmx-measure:obsValue
   rdf:datatype="http://www.w3.org/2001/XMLSchema#double">
    3.498 </sdmx-measure:obsValue>
```

[8] http://epp.eurostat.ec.europa.eu/
[9] http://motools.sourceforge.net/event/event.html
[10] http://www.w3.org/2004/02/skos/
[11] http://www.sdmx.org/
[12] http://dublincore.org/documents/dces/

```
  <rdf:type rdf:resource="http://open.data.al/oda#DataEntry"/>
</rdf:Description>
```

Fig. 3. Excerpt of a Dataset in RDF/XML Representation

Based on the aforementioned principles, well-formed URI references have been used in our approach to identify resources such as dataset, indicator, country, year, etc. We have invested on a technical infrastructure to make the URIs dereferenceable, making it able to retrieve further an RDF/XML description and various classifications of the identified resource when accessing its URI. RDF links to resources from other data sources are established, enabling visitors and user agents to navigate the Web of data. Examples include direct links to DBpedia[13] for resources such as country and year, or references via the `rdfs:seeAlso` property for the Indicators, in the case when the link exists, e.g. the indicator "*GDP Nominal per Capita*" would have a link to the respective resource[14] in DBpedia.

Reusability of existing vocabularies is also a requirement that tends to facilitate data processing from different client applications. Accordingly, terms from well-known ontologies have been reused, but we have certainly complemented these vocabularies with additional concepts for representing the statistical data. Particularly important to mention in our approach are the concept and taxonomy of indicators as well as topics, which are additionally modeled in the ontology and defined in our own namespace. Common constructs for providing mappings to other ontologies are `rdfs:subClassOf` or `rdfs:subPropertyOf`. As already shown in the presented ontology above, we have modeled a set of terms as subclasses of concepts (same as for properties) belonging to other vocabularies.

The linked data on the web is primarily intended to be processed by machines, but there is an increasing interest from the human visitors to access and navigate through the data. Therefore, we have aimed towards an approach that is convenient for both parts. On the one hand, human readability of the data is facilitated via provision of textual information (possibly in different languages) in forms of comments and labels, using respectively the properties `rdfs:comment` and `rdfs:label`.

On the other hand, our objective is to ease machine interpretability, stating important information explicitly, such as specification of domain, range and cardinality in our ontology. We provide information in such a way that avoids inconsistencies in data representation, but at the same time allows flexibility of the schema to be reused by other parties for data modeling.

IV. Technical Implementation

We aim to provide access to all the citizens, technically inclined or not, to discover structured data, and reuse them. To serve up these data sets, the ODA website accesses a catalog of records and examples of visual representation of datasets generated by visitors or editors of the website.

[13] http://dbpedia.org/
[14] http://dbpedia.org/property/gdpNominalPerCapita

A visualization service is delivered through the site and could include analytics, graphics, charting, and other ways of using the data. The enhanced visualization is built on top of published APIs. in collaboration with third party open source applications (Fig. 4).

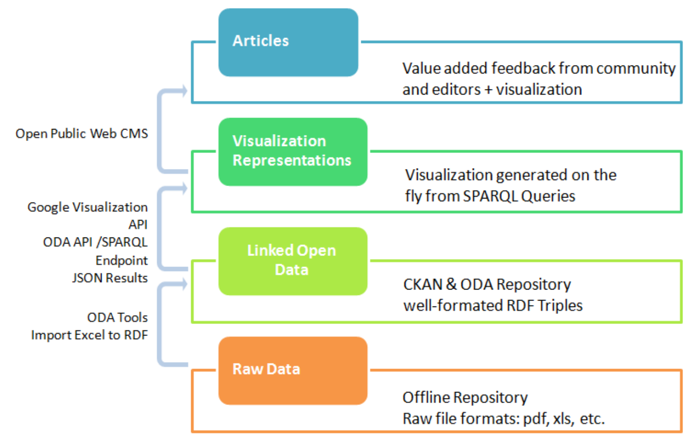

Fig. 4. Data layers and tools implemented

Our work consists of two stages. The first stage is the aggregation of the data and the processing of linked data structures. The second stage consists of providing a set of tools that can make use of these data. As an example of the technology that can be used, a set of visualization representations is generated. The visual representations are published together with a comment providing some additional information on the dataset retrieved. The operations, starting from the aggregation of data up to the publishing of examples, go through the following main steps.

A. *Aggregation of the information*

Data is aggregated from periodical government agency reports. These reports are most of the time provided as PDF files and sometimes in spreadsheet format. All the information is stored in an offline repository for original archiving purposes.

B. *Data Processing*

The information gathered in the aggregation step is semi-automatically processed and transformed into a well-organized Linked Data structure. In order to conform to the designed ontology, a set of spreadsheet templates is initially generated and made available. The information is imported and structured based on these spreadsheets, then it is automatically translated into RDF through a tool referred to as ODA Import to RDF.

The RDF datasets are stored in a CKAN repository, which is made public and can be accessed via the CKAN web interface and CKAN API.

Besides the interface of CKAN, the datasets are also stored in the ODA repository, which is directly connected to an API. The Open Data API is a RESTful, service-oriented platform that allows developers to easily access datasets and create independent services through these calls. REST uses the

HTTP protocol and, as such, requests use the common URL format.

The API provides simple methods that developers can use to tap into the functionality and rich datasets, and gather information, in JSON or XML format, related to different indicators and topics.

All requests on the ODA API should consist of a host name ("http://api.open.data.al "), a path that contains an object and an action (e.g. "/indicator/info" or "/indicator/list"), and a query string that contains the method parameters (e.g. "apiKey", "term", "indicatorId", etc.).

The URL template of a request looks like:

```
http://api.open.data.al/object/action?apiKey=myKey&param1=value1&param2=value2
```

Another way of accessing the information from the RDF datasets is through a SPARQL endpoint. The SPARQL endpoint is protocol service as defined in the SPROT[15] specification. The SPARQL endpoint enables users to query a knowledge base via the SPARQL language. The common result from a SPARQL endpoint is an XML-encoded SPARQL results. In order to access third party Visualization tools, we generate the result sets by default in JSON format.

C. *Graphical Visualization*

One of the values of making data available is that it enables anyone to construct nice visualizations over the data. This is particularly useful for public sector information, since it can provide feedback on how effective particular policies have been.

It is important that the generation of visual graphs should be done on the fly, as soon as result-sets are retrieved by the SPARQL Endpoint. For this reason, a set of open source visual application interfaces were evaluated and we chose to work with Google Visualization API[16].

In order to generate graphs on the fly, the following elements are needed:

- A RDF document: in order to present your data through Google Visualization, a source of RDF document is needed.
- A SPARQL query endpoint: a Sparql query endpoint is needed to process your queries and get the results http://api.talis.com/stores/rdfquery-dev1/services/sparql
- Google Visualization API: http://code.google.com/apis/visualization/

Google Visualization API relies heavily on Data Table, a table arranged in typed columns. An instance of this class is the result of a request to a data source. A Data Table can be rendered in many ways: JSON, HTML, CSV.

In order to feed the Visualization API a similar code should be generated by the application.

```
var data = new google.visualization.DataTable();
data.addColumn('string','Country');
data.addColumn('number','Year');
data.addColumn('number','Value');
data.addRows([
  ["Albania",2010,25],
  ["Albania",2011,25],
  ["Croatia",2010,37],
  ["Croatia",2011,37],
  ["Greece",2010,39],
  ["Greece",2011,42]
  ]);
```

The result provided by the SPARQL Endpoint is modified by a PHP script to generate the Data Table required by the Visualisation API (See Fig 5).

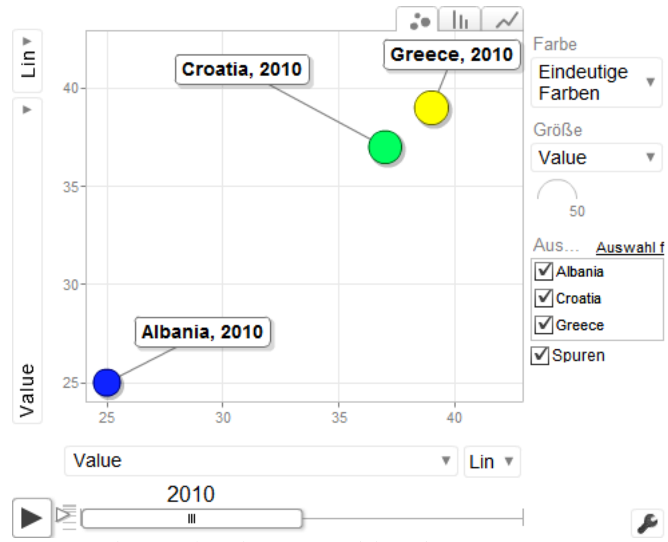

Fig. 5 Motion Chart generated through SPARQL query

D. *Community Contributions*

In order to provide a clear example of how the system will be used, a set of starting articles based on datasets and the respectively generated graphs is created. An article consists of the representation of one or more graphs with a matching description.

The graphic representation of the data assists the audience to draw concrete and logical conclusions on the progress or regress of certain indicators that influence socio-economic activity in a domain that the dataset covers. These articles are examples that the community may follow to contribute with more content, which can be generated by any visitor based on various self-generated graph representations.

V. RELATED WORK

The initiatives of making government data open are gaining a lot of interest recently. While more countries are embracing the Open Government paradigm, among the researchers working with those data there is an increasing awareness in using semantic techniques to represent them. Two of the key leading players that promote the openness of government data combined with the deployment of Semantic Web technologies are the US government [7] and the UK government [6].

In the respective portals[17],[18] manifesting the outcomes of their work, there are numerous data catalogs and respective

[15] SPARQL Protocol for RDF http://www.w3.org/TR/rdf-sparql-protocol/
[16] http://code.google.com/apis/charttools/index.html
[17] Data.gov
[18] Data.gov.uk

metadata published, many of the datasets offered as RDF triples. Data.gov has made a number of its datasets available as Linked Data, altogether summing up to 6.5 billion triples. The goal of these projects is to improve access to the data generated by the government, increase the ability of the people to find desired data, so that they can be further used in various creative ways. In order to access the data, they offer tools to search by keywords, category and agency.

Detailed statistics on the EU and candidate countries are provided by Eurostat[9], which is continuously publishing statistical data in tabular formats, but not in RDF. Unlike other portals, its website provides a clear structure for organizing the datasets by themes and by their respective indicators, but this structure is HTML-based and not semantically described.

Similar to these projects, our approach provides ways to easily access government data. Unlike them, it concentrates on promoting a well-defined, unified schema for the semantic representation of data, in order to increase interoperability and integration among the different providers. A novel aspect of our work consists in providing a semantic representation of the different indicators that can be measured using the datasets, as well as the topics to which they belong. This offers for any specific area a complete picture, making it easier to monitor the progress, make estimations or take a comparative approach.

From the semantic web perspective, there exist a limited number of vocabularies for representing statistical data such as SCOVO [4] and the RDF Data Cube Vocabulary[19]. While SCOVO does not provide semantics about the items that the statistical data describe, the Data Cube vocabulary is more extended and allows modeling of complex data structures and sets, as well as ways to link data with other sources. An approach that uses the Data Cube vocabulary to model statistical data and publish them as Linked Data is the Eurostat wrapper [8]. Another approach uses this vocabulary to semantically represent statistical data from social sciences, focusing on the capabilities of performing statistical calculations and analysis on linked data from heterogeneous and distributed datasets.

While our work encompasses more aspects such as mining of the statistical data, converting and publishing as linked data, data aggregation and integration, as well as search and visualization modules, the underlying semantic approach aims at providing semantics of these data based on an ontology, similar to the Data Cube vocabulary. While this vocabulary presents a core foundation for describing data, our ontology is designed to accommodate extended semantics of government-oriented statistical data. Furthermore, we tend to reuse components of various vocabularies, adhering in this way to linked data principles.

VI. Conclusion

Nowadays, there is a lot of attention to Open Data projects in many countries all over the world. It is a common understanding today that Open Data is simply 'data on the web,' whereas Linked Data is a 'web of data'.

In this paper, we address the problem of data provided in heterogeneous formats, lacking semantics that clarify what the data describes. Our approach is based on an OWL ontology providing semantic description of the government statistical data, which are further published in alignment with Linked Data principles.

The paper also presents an approach for the provision of visualization services deriving from RDF result sets. This is achieved by posing SPARQL queries on the generated RDF triples and passing the resulst to an open-source visualization API.

The ontology created for the ODA and the supporting tools built on top of it describe a method of publishing government data, so that it can be interlinked and become more useful to the community.

Acknowledgments

This work is partly funded by the EU-FP7 Network of Excellence PlanetData (Ref. 257641).

References

[1] Studer, R., Benjamins, R.V., Fensel, D. Knowledge engineering: Principles and methods. Data and Knowledge Engineering, 25, pages 161–197,1998.

[2] Bizer, C., Heath, T., and Berners-Lee, T. Linked Data - The Story So Far. International Journal on Semantic Web and Information Systems (IJSWIS), Vol. 5(3), pages 1-22, 2009.

[3] Berners-Lee, T. Linked Data - Design Issues (2006). Retrieved 2010-10-25 from http://www.w3.org/DesignIssues/LinkedData.html

[4] Hausenblas, M., Halb, W., Raimond, Y., Feigenbaum, L., and Ayers, D. SCOVO: Using Statistics on the Web of Data. Proceedings of the 6th European Semantic Web Conference on the Semantic Web: Research and Applications, pages 708-722, Heraklion, Crete, Greece, 2009.

[5] Bizer, C., Cyganiak, R., and Heath, T. (2007). How to publish Linked Data on the Web. Retrieved 2010-10-25 from http://www4.wiwiss.fu-berlin.de/bizer/pub/LinkedDataTutorial/

[6] Sheridan, J., Tennison, J.. Linking UK government data. Linked Data on the Web (LDOW2010) Workshop at WWW2010, April 2010

[7] Ding, L., DiFranzo, D., Graves, A., Michaelis, J. R., Li, X., McGuinness, D. L., Hendler, J. Data-gov wiki: Towards linking government data. In AAAI Spring Symposium on Linked Data Meets AI, 2010.

[8] Zapilko, B., Harth, A. , Mathiak, B. Enriching and Analysing Statistics with Linked Open Data. In: Eurostat (Hrsg.), NTTS - Conference on New Techniques and Technologies for Statistics, Brüssel, 2011.

[9] Eurostat Wrapper - the experimental Eurostat wrapper to RDF http://estatwrap.ontologycentral.com/

[19] http://publishing-statistical-data.googlecode.com/svn/trunk/specs/src/main/html/cube.html