Undefined 1 (2013) 1–5
IOS Press

LinkedSpending: OpenSpending becomes Linked Open Data

Konrad Höffner, Michael Martin, Jens Lehmann
University of Leipzig, Institute of Computer Science, AKSW Group
Augustusplatz 10, D-04009 Leipzig, Germany
E-mail: {hoeffner,martin,lehmann}@informatik.uni-leipzig.de

Abstract. There is a high public demand to increase transparency in government spending. Open spending data has the power to reduce corruption by increasing accountability and strengthens democracy because voters can make better informed decisions. An informed and trusting public also strengthens the government itself because it is more likely to commit to large projects. OpenSpending.org is a an open platform that provides public finance data from governments around the world. In this article, we present its RDF conversion LinkedSpending which provides more than 2.4 million planned and carried out financial transactions from nearly 250 datasets from all over the world from 2005 to 2035 as *Linked Open Data*. This data is represented in the RDF Data Cube format and is freely available and openly licensed.

Keywords: government, transparency, finance, budget, openspending, rdf, public expenditure, Open Data

1. Introduction

A W3C design issue [6] motivates making government data available online as Linked Data for three reasons: "1) Increasing citizen awareness of government functions to enable greater accountability; 2) Contributing valuable information about the world; and 3) Enabling the government, the country, and the world to function more efficiently." Increasing the transparency of government spending specifically is in high demand from the public. For instance, in the survey publication [14], "Public access to records is crucial to the functioning government" was rated with a mean of 4.14 (1 = disagree completely, 5 = agree completely). Open spending data can reduce corruption by increasing accountability and strengthening democracy because voters can make better informed decisions. Furthermore, an informed and trusting public also strengthens the government itself because it is more likely to commit to large projects (see [3] for details).

Several States and Unions are bound to financial transparency by law, such as the European Union[1] with its *Financial Transparency System (FTS)*[2] [10]. Public spending services satisfy basic information needs, but in their current form they do not allow queries which go further than simple keyword search or which cannot be answered with data from one system alone. Linked data solves those problems by providing a unified format, a powerful query language and the possibility of integration with services such as CORDIS[3] and linked datasets such as Greece public spending [17].

Our contribution is an RDF transformation of the OpenSpending[4] project which provides government spending financial transactions from all over the world and is thus suitable as a core knowledge base that can be enriched and integrated with other, more focussed

[1] "2. The Commission shall make available, in an appropriate and timely manner, information on recipients, as well as the nature and purpose of the measure financed from the budget[...]" [1]
[2] `http://ec.europa.eu/budget/fts`
[3] The European *Community Research and Development Information Service*, `http://cordis.europa.eu/`.
[4] `http://openspending.org`

datasets. Some advantages of an RDF version over a relational database are (1) querying over multiple datasets at once using interlinks and (2) an additional custom vocabulary for each of the datasets, which are very heterogeneous (see Table 5 for SPARQL queries for common use cases).

The structure of the paper is as follows. Section 2 describes OpenSpending, which is the source of the data, and its statistical data model. Section 3 explains the target RDF data cube vocabulary and the transformation process to it. Section 4 describes, how and where the dataset is published and in which way users can access the data. Section 5 gives an overall view of the data sets, gives details about the licence used and describes the datasets it is interlinked to. The is followed by a description of usage scenarios in Section 6. Section 7 presents related spending datasets as linked open data (LOD). The last section discusses known shortcomings of the datasets and future work. The prefixes used throughout this publication are defined in Table 1.

Table 1
Namespaces

prefix	URL
os	http://openspending.org/
owl	http://www.w3.org/2002/07/owl#
ls	http://linkedspending.aksw.org/resource/
lso	http://linkedspending.aksw.org/ontology/
qb	http://purl.org/linked-data/cube#
sdmxd	http://purl.org/linked-data/sdmx /2009/dimension#
dbpedia	http://dbpedia.org/resource/
dbp	http://dbpedia.org/property/

2. OpenSpending Source Data

OpenSpending is a project which aims to "[...] track every government financial transaction across the world[...]"[5] and, at September 2013, contains more than 15 million government transactions in 321 datasets. It contains this data in structured form stored in database tables and provides searching and filtering as well as visualizations and a JSON REST interface.

The Data Cube Model The domain model of OpenSpending is that of a *data cube* (also *OLAP cube*, *hypercube*) which is a multi-dimensional dataset in which statistical observations are central. This is similar to a star schema in database terms. Each cell corresponds to an observation (an instance of spending or revenue) that contains measurements (e.g. the amount of money spent or received). The context of the measurement is provided by the *dimensions* like the purpose, department and time of a spending item and optionally by *attributes*, which further describe the measured value, e.g., the unit of the measurement.

```
"sub-programme": {
 "label": "Sub-programme",
 "type": "compound",
},
"amount": {
    "datatype": "float",
    "label": "Total",
    "type": "measure",
}
```

Fig. 1. simplified excerpt of an OpenSpending *model*

```
"sub-programme": {
 "label":"Security and safeguarding liberties
     ",
 "html_url":"http://openspending.org/eu-
     budget/sub-programme/security-and-
     safeguarding-liberties",
 "name":"security-and-safeguarding-liberties"
},
"html_url": "http://openspending.org/eu-
    budget/entries/017dfcb58d05671ef9eb5a9f77
    fef39c8b14150c",
"amount": 41.2
```

Fig. 2. simplified excerpt from an OpenSpending *entry*

Figure 1 shows an excerpt from the model of the OpenSpending dataset *eu-budget* with the dimension "sub-programme" and the measure amount. Figure 2 shows the the corresponding part of an *entry* of the dataset, which contains the actual values for the dimension and the measure of the observation.

3. Conversion of OpenSpending to RDF

Because the source data adheres to the data cube model, a conversion of the data to RDF needs an appropriate RDF vocabulary.

The RDF DataCube vocabulary The RDF DataCube vocabulary [2], i.e. an RDF variant of the previously explained data cube model, is an ideal fit for the transformed data.

First and foremost, this vocabulary provides the backbone structure for every LinkedSpending dataset, see Figure 3. Each dataset is represented by an instance of `qb:DataSet` and an associated instance of `qb:DataStructureDefinition` which includes

[5] http://openspending.org/

[6] Simplified version of the structure described in [2].

Table 2
Conversion of OpenSpending to LinkedSpending classes and instances

	Source URL	JSON Path	LinkedSpending class	LS instance scheme
I	os:*name*.json		qb:DataSet	ls:*name*
II	os:*name*/model		qb:DataStructureDefinition	ls:*name*/model
III	os:*name*/model	$.mapping.*	os:{Country,Time}Component Specification or qb:ComponentSpecification	lso:*propertyname*-spec
IV	os:*name*/model	$.mapping.*[?(@.type="compound")]	qb:DimensionProperty	lso:*propertyname*
V		$.mapping.*[?(@.type="date")]	qb:DimensionProperty	
VI		$.mapping.*[?(@.type="measure")]	qb:MeasureProperty	
VII		$.mapping.*[?(@.type="attribute")]	qb:AttributeProperty	
VIII	os:*name*/entries.json	$.results[*].dataset	qb:Observation	ls:observation-*datasetname-hashvalue*

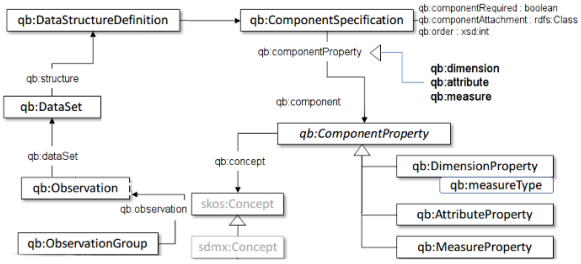

Fig. 3. Used RDF DataCube concepts and their relationships[6]

component specifications. Each component specification is associated to a *component property* which can be either a *dimension*, an *attribute* or a *measure*. RDF Data Cube is supported by tools such as the faceted browser CubeViz as well as other tools in the LOD2 Stack [4].

Transformation All of the OpenSpending datasets describe observations referring to a specific point or period in time and thus undergo only minor changes. New datasets however, are frequently added. Because of this, the huge number of datasets and their size, an automatic, repeatable transformation is required. This is realized by a program[7] which fetches a list of datasets on execution and only transforms the ones who are not transformed yet. Table 2 shows for each class used by LinkedSpending, at which URL the information used to create the instances of those classes is found. In case there are multiple instances described at one URL, a *JSON path*[8] expression is given, that locates the corresponding subnodes. Finally, the table contains the patterns that describe resulting LinkedSpending URLs. For example, the OpenSpending URL `os:berlin_de/model` contains the node `$.mapping.amount` which has a type value of "attribute" and is, thus, transformed to the OpenSpending instance `lso:amount` of the class `qb:AttributeProperty`.

Equivalent component properties (dimensions, attributes and measures) are identified as follows: A configuration file optionally specifies the mapping of dataset and property name to an entity in the LinkedSpending ontology. In most cases such a mapping entry is not specified, however, and the property URI is by default generated based on the property name. Properties with the same name in different datasets not having a mapping entry that states otherwise are assumed to represent the same concept and thus given the same URL.[9]

Error Handling The OpenSpending API lists 321 datasets with 247 of them having a LinkedSpending equivalent. The discrepancy is caused by loss in several stages. In the first step, all datasets are downloaded, in several parts if necessary, which are then joined. 69 of the datasets do not contain any observations, however, and are thus not processed any further. The remaining 252 datasets are then transformed, noting the number of errors such as missing values for a component property. In order to detect faulty datasets and to guarantee high quality data, there is a threshold on the number of errors. If the error count of the transformation of a dataset reaches both at least 30 and 10 % of all observation-property pairs yet processed, the transfor-

[7] written in Java, available as open source at `https://github.com/AKSW/openspending2rdf`

[8] *JSON path* (`http://code.google.com/p/json-path/`) is a a query language for selecting nodes from a JSON documents, similar to XPath for XML

[9] Although that has the possibility of mismatches, such a mismatch has not been spotted yet. Still, evaluating and, if necessary, improving the automatic matching is part of future work.

mation is aborted, which removes 5 additional datasets, leaving 247.

Sustainability The data conversion process of new datasets is started weekly by a chronjob on the server without interrupting the accessibility of the SPARQL endpoint and the services building on it.

4. Publishing

LinkedSpending data is published using OntoWiki [5], which manages several parts of the the Linked Data Lifecycle [4], such as Storage/ Querying and Search/Browsing/Exploration. The interface for human and machine consumption of the data is available at `http://linkedspending.aksw.org`. Depending on the actor and the needs, OntoWiki provides various abilities to gather the published RDF data as described as follows.

The data can be explored by viewing the properties of a resource, its values and by following links to other resources (see Figure 4). Using the SPARQL endpoint[10] provided by the underlaying *Virtuoso Triple Store*[11], actors are able to satisfy complex information needs.

[10] `http://linkedspending.aksw.org/sparql`
[11] `http://virtuoso.openlinksw.com`

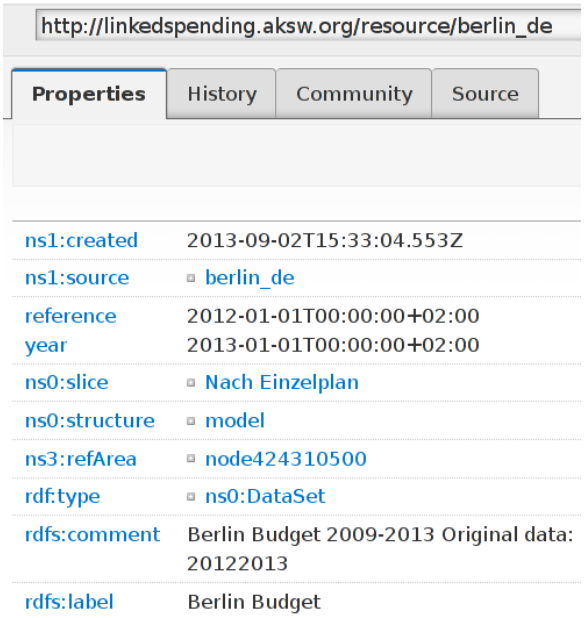

Fig. 4. View of the dataset `berlin_de` in the OntoWiki

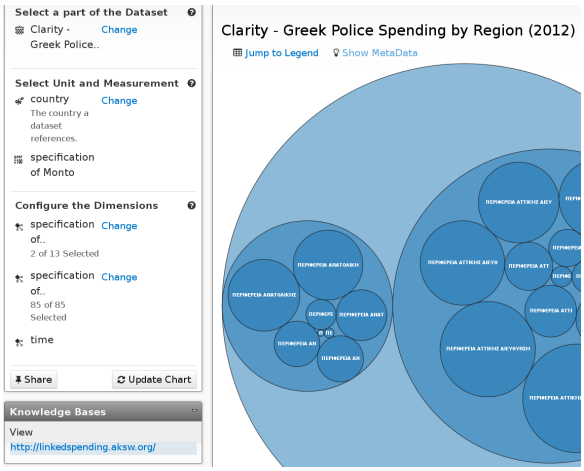

Fig. 5. Facetted browsing in CubeViz by restricting values of dimensions

Faceted search offers a selection of values for certain properties and thus slice and dice of the dataset according to the interests on the fly. Examplarily depicted in Figure 5 is all Greek police spending in a certain region. Visualization supports discovery of underlying patterns and gain of new insights about the data, for example about the relative proportions of a budget (see Figure 6). We set up the RDF DataCube Browser CubeViz [15] as part of the human consumption interface.

Table 3
Technical details of the LinkedSpending dataset

URL	`http://linkedspending.aksw.org`
Version date and number	2013-8-14, 1.0
License	PDDL 1.0[12]
SPARQL endpoint	`http://linkedspending.aksw.org/sparql`
Compressed N-Triples Dump	`http://linkedspending.aksw.org/extensions/page/page/export/`
datahub entry	`http://datahub.io/dataset/linkedspending`

Licensing All published data is openly licensed under the PDDL 1.0. in accordance with the open definition[13].

5. Overview over the Datasets

LinkedSpending consists of 247 datasets (continually growing) with more than 2.4 million observations total. The amount of observations of the individual datasets varies considerably between two (spendings

[12] `http://opendatacommons.org/licenses/pddl/1.0/`
[13] `http://opendefinition.org/`

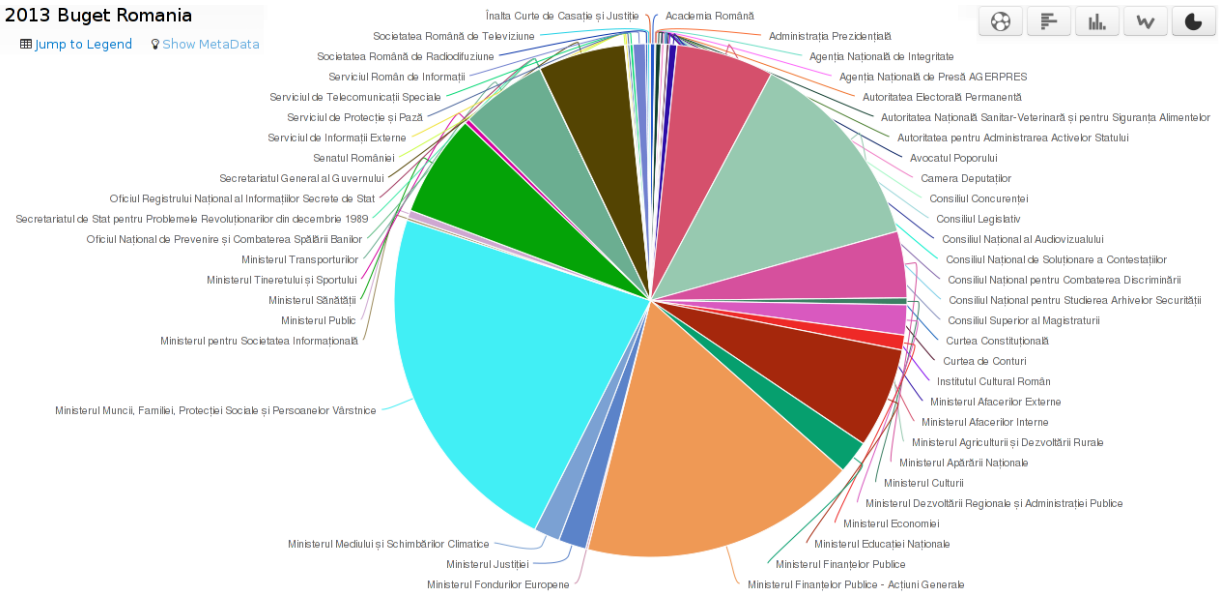

Fig. 6. CubeViz visualization of the romanian budget of 2013

in Prague of about 5000 CZK for an unknown purpose) and 242 209 ("Spending from ministries under the Danish government"). Table 4 details the average and total amount of data in bytes, triples, and observations. Figure 7 shows the distribution of the numbers of measures, attributes and dimensions of the datasets. Measures represent the quantity that an observation describes. All datasets have at least one measure which is the amount of money spent or received and for most of them (217) that is the only one but the maximum is 7 in the `dk-corporate-tax-list` dataset which contains measures for different kinds of profits as well as losses. Attributes give further context to the measurement. The number of attributes is more varied, ranging from 2 to 26, with all datasets having at least a currency and a country, and most of them additionally the time the observations refer to. While the number of dimensions ranges from 0[14] up to 32, almost all of the datasets have between 1 and 6 dimensions, the most common ones being the year and the time the dataset and the observations refers to, respectively.

Technical details about the datasets are described in Table 3.

Use of Established Vocabularies: Apart from standard vocabularies such as RDF, RDFS, OWL and XSD, the datasets are modelled, first and foremost, according to the RDF Data Cube vocabulary (see Section 2). LinkedSpending follows the RDF Data Cube recommendation to make heavy use of the model of the Statistical Data and Metadata eXchange (SDMX) initiative[15] for measures, attributes and dimensions.The deep structure of the datasets is very heterogeneous but there are

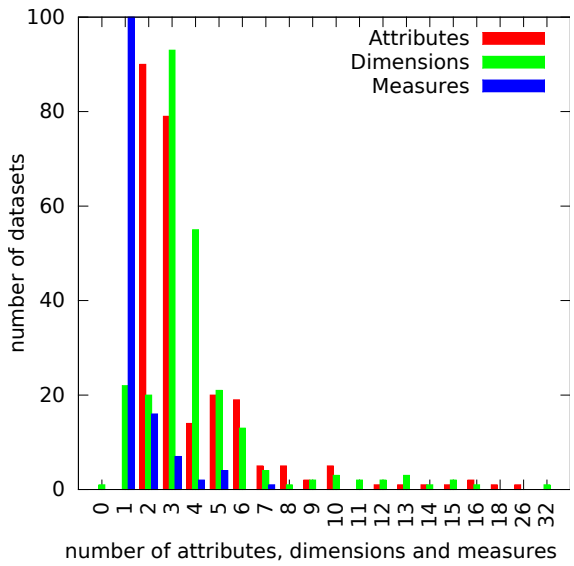

Fig. 7. Numbers of measures, attributes and dimensions of the datasets. 217 datasets have exactly one measure (clipped bar).

[14] there is only one dataset with no dimensions which a test dataset on OpenSpending, as a data cube with no dimensions is not useful

[15] `http://sdmx.org`

Table 4
Amount-of-data (all values rounded to the nearest integer).

	Total	Average
number of datasets	247	
filesize (RDF/N-Triples)	10 012 MB	41 MB
triples	49 619 572	200 889
observations	2 404 313	9734

some properties which are commonly specified and thus modelled with established vocabularies. The year and date, a dataset and an observation refers to, respectively, is expressed by `sdmx-dimension:refPeriod` and XSD. Currencies are taken from DBpedia [12] and countries are represented using the vocabulary of LinkedGeoData [16], which is the main hub for spatial linked data.

Interlinking: There are two distinct possibilities to align entities to another vocabulary: 1) to use the entities of the vocabulary directly and 2) to create an own RDF resource with interlinks, like `owl:sameAs`, to that vocabulary. We generally preferred to use entities of a vocabularies directly, such as the currencies and countries, because this is a more elegant and less convoluted approach as long as the used vocabularies are stable and exactly correspond to the required meaning. While we did not find *sameAs* link targets on observation level, i.e. exactly the same statistical observations described in other datasets, there are many possibilities for interlinks between datasets or dimension values and concepts they refer to. Using the labels of those datasets and dimension values, it is possible, for example, to link values of the dimension "region" of a federal budget, and thus indirectly also the observations which use those values, to the cities in DBpedia or LinkedGeoData whose labels are contained in the label of the region value URI.

Example Queries: Table 5 contains example queries for common use cases. Queries 1–6 are easily possible with a relational database as well. Query 7 however demonstrates the benefit of the interlinked nature of RDF datasets, in this case the interlinking to DBpedia currencies, by querying over two different graphs.[16] Query 8 demonstrates the advantage of having a custom vocabulary[17] for each dataset. An equivalent query using relational databases would thus be more convoluted.

[16]parts of DBpedia and LinkedGeoData describing countries and currencies have been integrated in the SPARQL endpoint. With federated querying however, nearly the whole LOD cloud can be queried.

[17]In this case the "Hauptfunktion" and "Oberfunktion" are unique to the `berlin_de` dataset

6. Usage Scenarios

Apart from the information needs that can already be satisfied using the source data, there are usage scenarios that are easier or only possible with LinkedSpending.

Economic Analysis: LinkedSpending is represented in Linked Open Data which facilitates data integration. Currencies and countries from DBpedia and LinkedGeoData, respectively, are already integrated. This allows queries such as query 7 in Table 5, which asks for datasets with currencies whose inflation rates are greater than 10%.

LinkedSpending can also be used to compute economic indicators across several datasets. A possible indicator about the economic situation of a country is the spending on education per person where the population size can be taken from the LinkedGeoData countries linked from the datasets Such information is often spread across several datasets, e.g. there is a specific dataset for Uganda.[18]. LinkedSpending allows to serve as a hub for the integration of those datasets and their provenance information. More datasets can be integrated with similarity-based interlinking tools such as LIMES [13][19] and Silk [18].

Finding and Comparing Relevant Datasets:

Government spending amounts are often much higher than the sums ordinary people are used to dealing with but even for policy makers it is hard to understand, if a certain amount of money spent is too high or normal. Comparing datasets and finding those which are similar to another one helps separating common values from outliers which should be further investigated. For example, if another country has a similar budget structure but spends way less on healthcare with a similar health level, it should be investigated, whether that discrepancy is caused by inherent differences such as different minimum wages or a different climate or if it is due to preventable factors such as inefficiencies or corruption. While OpenSpending provides several hundreds of datasets which can be searched and it allows browsing and visualization of any single one, it does not provide a comparison function between datasets. Because of the mechanism to identify equivalent properties (see Section 3), SPARQL queries can compare different datasets, e.g. between similar structures in different countries. Query 9 in Table 5 shows a simple

[18]one such dataset is `ugandabudget`, which contains the Uganda Budget and Aid to Uganda, 2003–2006

[19]and its web interface SAIM[9], available at `http://saim.aksw.org`

Table 5
Examplary SPARQL queries for typical use cases.

information need	SPARQL Query
1 list of all datasets	`select * {?d a qb:DataSet}`
2 all measures of the dataset `berlin_de`	`select ?m { ls:berlin_de qb:structure ?s. ?s qb:component ?c. ?c qb:measure ?m.}`
3 all years which have observations in the de-bund dataset from 2020 onwards	`select distinct ?date {?o a qb:Observation. ?o qb:dataSet ls:de-bund. ?o sdmxd:refPeriod ?date. FILTER (xsd:date(?date) >= "2020-1-1"^^xsd:date) }`
4 spendings of more than 100 billion €	`select * {?o lso:amount ?a. ?o dbo:currency dbpedia:Euro. FILTER(xsd:integer (?a)>"1E11"^^xsd:integer) }`
5 datasets with multiple years	`select ?d count(?y) as ?count { ?d a qb:DataSet. ?d lso:refYear ?y. } having (count(?y)>1)`

6 sums of amounts for each reference year of the dataset `berlin_de`	`select ?y (sum(xsd:integer(?amount)) as ?sum)`
	`{?o qb:dataSet ls:berlin_de. ?o lso:refYear ?y. ?o lso:amount ?amount.}`
	`group by ?y`

7 datasets with currencies whose inflation rate is greater than 10 %	`select distinct ?d ?c ?r {?o qb:dataSet ?d. ?o dbo:currency ?c. ?c dbp: inflationRate ?r. filter(?r > 10)}`

8 Berlin city subsectors of research and education that have had their budget reduced from 2012 to 2013	`select ?l (sum(xsd:integer(?amount12)) as ?sum12) (sum(xsd:integer(?amount13)) as ?sum13)`
	`{`
	`?o qb:dataSet ls:berlin_de.`
	`?o lso:Hauptfunktion <http://openspending.org/berlin_de/Hauptfunktion/l>.`
	`?o lso:Oberfunktion ?of. ?of rdfs:label ?l.`
	`{?o lso:refYear "2012"^^xsd:gYear. ?o lso:amount ?amount12.} UNION`
	`{?o lso:refYear "2013"^^xsd:gYear. ?o lso:amount ?amount13.}`
	`}`
	`group by ?l`
	`having (sum(xsd:integer(?amount12)) > sum(xsd:integer(?amount13)))`

9 datasets ordered by their number of properties in common with `2012_tax` (having at least one such common property)	`select distinct ?d count(?c)`
	`{`
	`?d qb:structure ?s. ?s qb:component ?c.`
	`?s qb:component ?c2. ls:2012_tax qb:structure ?s2.`
	`?s2 qb:component ?c. ?s2 qb:component ?c2.`
	`FILTER(?c!=?c2 AND ?d!=ls:2012_tax)`
	`}`
	`order by desc(count(?c))`

query to detect datasets which are most similar to any particular dataset. This is done by calculating the number of common measures, attributes and dimensions.

7. Related Work

The TWC Data-Gov Corpus [7,8] consists of linked government data from the Data-gov project. However, it only contains transactions made in the US and does not overlap with OpenSpending. The publicspending.gr project generates and publishes [17] public spending data from Greece based on the UK payment ontology and without using statistical data cubes. The UK government expenditure dataset COINS[20] is available as Linked Data[21]. *LOD Around-The-Clock (LATC)*[22] is a project, which was funded by the European Union (EU) and converted European open government data into RDF. One of its outcomes is the FTS[23] [10] project, which transforms and publishes financial transparency data of the EU. Furthermore, the Digital Agenda Scoreboard [11] is an EU project which keeps track of the transformation of statistical data to RDF.

[20] `http://data.gov.uk/dataset/coins`
[21] `http://openuplabs.tso.co.uk/sparql/gov-coins`, in a beta version
[22] `http://latc-project.eu`
[23] `http://ec.europa.eu/budget/fts`

8. Conclusions, Shortcomings, Future Work

As shown in Section 3, we converted several hundreds of financial datasets to RDF and, as shown in Section 4, we published them as Linked Open Data in several ways. However, we recognise a few shortcomings and our goal is to enrich the meta data with the help of domain experts and to refine the structure of the individual datasets. Furthermore, we plan to improve the automatic configuration of CubeViz.

Individual Modelling Because the source data is already structured, the transformation of all the datasets without the need of text extraction and in an automatic way was feasible. On a deep level however, there is much unmodelled structure that is unique to each dataset or at most shared between several of them, for instance the categorization of spending into several specific "plans" in German budgets. Because of the amount of datasets, modelling all details, and thus also improving the internal and external connectivity, requires either a large-scale cooperation or a crowd-driven approach, which we did not perform yet.

Languages The source data does not contain language tags and the language of the labels does not always match the country the data refers to. Automatic language detection on single labels did not yield a satisfying success rate and it is not possible to increase the precision of the language detection by combining the estimates about several different labels of an observation because their language is not always identical. Statistical examinations of the relations between labels of different entities and more complex schemes based on those examinations, however, could achieve language detection with a higher success-rate.

Drilldowns Because of the hierarchical organization of the different coded properties "groups" and "functions", the visualizations on `openspending.org` permit "zooming" (drilldown) in and out of the different levels of the data. The RDF Data Cube vocabulary specifies the use of `skos:ConceptSchema` or `qb:HierarchicalCodeList` but neither variant is fully implemented yet and it is not clear, which of those modelling possibilities will win out in the long run and get better tool support.

Acknowledgements: Special thanks goes to the people behind the OpenSpending project, including Friedrich Lindenberg for suggesting the conversion to RDF.

References

[1] Regulation (eu, euratom) no 966/2012, 2012. Article 35: Publication of information on recipients and other information.

[2] The rdf data cube vocabulary. Technical report, W3C, 2013.

[3] J. E. Alt, D. D. Lassen, and D. Skilling. Fiscal transparency, gubernatorial popularity, and the scale of government: Evidence from the states. Technical report, Economic Policy Research Unit (EPRU), University of Copenhagen, 2001.

[4] S. Auer, L. Bühmann, C. Dirschl, O. Erling, M. Hausenblas, R. Isele, J. Lehmann, M. Martin, P. N. Mendes, B. van Nuffelen, C. Stadler, S. Tramp, and H. Williams. Managing the life-cycle of linked data with the lod2 stack. In *Proc. of International Semantic Web Conference (ISWC) 2012*, 2012. 22

[5] S. Auer, S. Dietzold, J. Lehmann, and T. Riechert. OntoWiki: A tool for social, semantic collaboration. In *Proc. of CKC 2007 at WWW2007 Banff, Canada, May 8, 2007*, 2007.

[6] T. Berners-Lee. Putting government data online - design issues, 2009. W3C design issue.

[7] L. Ding, D. DiFranzo, A. Graves, J. Michaelis, X. Li, D. L. McGuinness, and J. Hendler. Data-gov wiki: Towards linking government data. In *AAAI Spring Symposium: Linked Data Meets Artificial Intelligence'10"*, 2010.

[8] L. Ding, D. DiFranzo, A. Graves, J. R. Michaelis, X. Li, D. L. McGuinness, and J. A. Hendler. Twc data-gov corpus: incrementally generating linked government data from data.gov. In *Proc. of the 19th international conference on World wide web*, 2010.

[9] K. Lyko, K. Höffner, , R. Speck, A.-C. Ngonga Ngomo, and J. Lehmann. Saim - one step closer to zero-configuration link discovery. In *Proc. of ESWC Posters & Demos*, 2013.

[10] M. Martin, C. Stadler, P. Frischmuth, and J. Lehmann. Increasing the financial transparency of european commission project funding. *Semantic Web Journal*, Special Call for Linked Dataset descriptions, 2013.

[11] M. Martin, B. van Nuffelen, S. Abruzzini, and S. Auer. The digital agenda scoreboard: A statistical anatomy of europe's way into the information age. Technical report, University of Leipzig, 2012.

[12] M. Morsey, J. Lehmann, S. Auer, C. Stadler, and S. Hellmann. DBpedia and the Live Extraction of Structured Data from Wikipedia. *Program: electronic library and information systems*, 46:27, 2012.

[13] A.-C. Ngonga Ngomo. A time-efficient hybrid approach to link discovery. In *Proc. of OM@ISWC*, 2011.

[14] S. J. Piotrowski and G. G. Van Ryzin. Citizen Attitudes Toward Transparency in Local Government. *The American Review of Public Administration*, 37(3):306–323, sep 2007.

[15] P. E. Salas, F. M. D. Mota, K. Breitman, M. A. Casanova, M. Martin, and S. Auer. Publishing statistical data on the web. *International Journal of Semantic Computing*, 06(04):373–388, 2012.

[16] C. Stadler, J. Lehmann, K. Höffner, and S. Auer. Linkedgeodata: A core for a web of spatial open data. *Semantic Web Journal*, 3(4):333–354, 2012.

[17] M. Vafopoulos, M. Meimaris, I. Anagnostopoulos, A. Papantoniou, I. Xidias, G. Alexiou, G. Vafeiadis, M. Klonaras, and V. Loumos. Public spending as lod: the case of greece. *Semantic Web Journal*, 2013.

[18] J. Volz, C. Bizer, M. Gaedke, and G. Kobilarov. Silk - a link discovery framework for the web of data. In *LDOW*, 2009.