Mining and Linking Open Economic Data from Governmental Communities

Michalis Vafopoulos[1(✉)], Stylianos Rallis[2],
Ioannis Anagnostopoulos[3], Vassilios Peristeras[4], Dimitrios Negkas[1],
Ilias Skaros[1], and Aggelos Tzani[1]

[1] Software and Knowledge Engineering Laboratory, IIT,
NCSR-"Demokritos", Athens, Greece
vaf@aegean.gr, dimneg@gmail.com,
skaros.ilias@gmail.com, tzaniaggelos@hotmail.com
[2] School of Engineering, Department of Production and Management
Engineering, Democritus University of Thrace, Xanthi, Greece
strallis@gmail.com
[3] School of Sciences, Department of Computer Science and Biomedical
Informatics, University of Thessaly, Lamia, Greece
janag@dib.uth.gr
[4] School of Science and Technology, International Hellenic University,
Thessaloniki, Greece
v.peristeras@ihu.edu.gr

Abstract. In this paper, we propose a model conceptualization (Linked Open Economy - LOE) capable of exploiting the massive amount and variety of open economic data that are gradually becoming available by governments and open source communities. The main aim is to unleash the power of open data and open source systems and create a common ground to serve as a catalyst in providing more efficient answers in important economic activities.

Keywords: Open data · Open systems · Government · Economy

1 Introduction

In order to ensure transparency and provide valuable means of innovation, many governments demand from their public authorities to open their data (e.g. data.gov.uk), as well as to support their provision and maintenance with crowd-source independent repositories and open-source systems (like CKAN[1] etc.). Towards this end, there are also many organized initiatives and projects. To name a few, the IBP[2] promotes public access to accountable budget systems, thus allowing citizens to investigate whether central governments manage properly procedures in respect to public finances. Similarly, the Open Budgets[3], the Big Data Europe[4] and the Your Data Stories[5] H2020

---

[1] ckan.org.
[2] internationalbudget.org/.
[3] OpenBudgets.eu.
[4] BigDataEurope.eu.
[5] YourDataStories.eu.

projects provide a scalable platform for public administrations to publish and preserve open economic data under open-source support (e.g. GitHub). A comprehensive review of other similar initiatives can be found in [1].

2 Our Proposed Model

There are several standards, controlled vocabularies and ontological schemes for achieving semantic linking between open data. In the economy domain, the most important are: (i) the GoodRelations ontology [2], (ii) the Public Contracts ontology [3], (iii) the Organization ontology, and (iv) the vCard ontology. All above ontologies and schemas are cross-linked with several domain-independent controlled-vocabularies such as FOAF, DC Terms, SKOS [4], and DBpedia.

Based on the above (data and interoperability standards), we propose the Linked Open Economy (LOE) model (see Fig. 1), which is an economy-driven rather than a data- or statistics-driven approach. The model has been designed to better balance the trade-off of, being as generic as to be scalable to future open data categories, and as specific as to be compatible with existing initiatives. Furthermore, its top-level conceptualization interlinks the main publicly available data for the economy, as well as it models the flows incorporated in public budgeting, procurement and subsidies/aid. Detailed information about LOE modeling, as well as its main classes and properties can be found at https://github.com/LinkedEcon/LinkedEconomyOntology-ELOD.
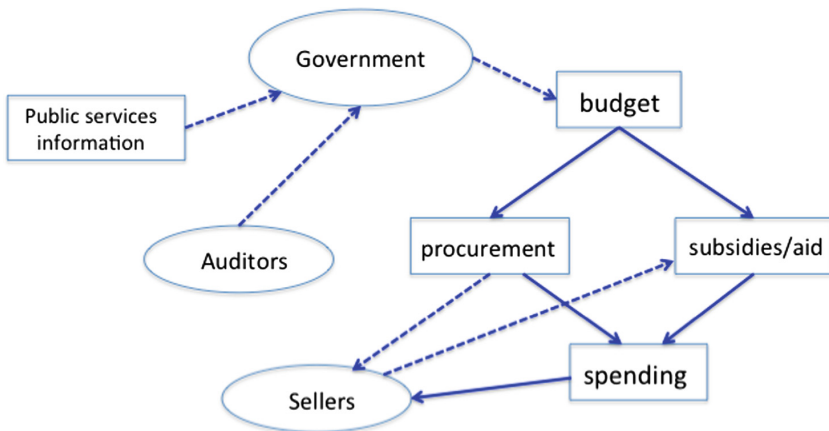

Fig. 1. The Linked Open Economy (LOE) model.

3 Application Case Study: Greek Public Procurement

The Central Electronic Registry of Public Procurement (KHMDHS) (eprocurement. gov.gr) is the official procurement portal of the Greek government which contains information on all public contracts for goods, services, and public works worth more than €1,000. It was established to reinforce transparency and effective management of

public procurement. According to the Linked Economy Ontology (hereafter under the *elod* prefix), the basic model consist of:

elod:DecisionDomain

In this dataset the "tenders" concept is available. In order to model this part of the LOE flow, we use the *pc:Contract* class. The main parts are the following:

- the contracting authority which issues a contract,
- the CPV code which characterizes the products or services,
- the specified price of the contract including the currency specification, and
- the tender to whom a contract is awarded.

elod:SpendingDomain

The main components are:

- the authority which pays for the products or services,
- the sponsor who is paid for the offered products or services,
- the CPV code which characterizes the products or services, and
- the expense amount including the currency specification.

Organization Classes

The essential elements of the organization are given by the vat Id, the name, the vat Id type, and the country registered at.

elod:PriceDomain

In the same sense as above, we use the *gr:Offering* class as well as the following:

- *gr:BusinessEntity:* This class represents the organizations. It is one of the four classes used for their modeling.
- *gr:SomeItems:* This class contains the information of a service or product.
- *elod:SpendingItem:* This class represents decisions that involve spending actions.
- *gr:UnitPriceSpecification:* This class represents a conceptual entity that specifies the price asked for a service or product. It specifies the price, the currency and whether the price includes VAT or not.

In order to combine and compare products which belong to a *gr:Offering* from different instances of *elod:PriceDomain* class we use the class *elod:CPV* in the second level of our modeling. For example the instances of *gr:Offering* class in KHMDHS Data set are linked to an *elod:CPV* instance, the same connection exists in eprices Data set as well. Thus, we can extract information based on a common *elod:CPV* instance from different *elod:PriceDomain* resources.

The basic part of *elod:CPV* instance is:

- The unique code of cpv and the property which specifies it is *elod:cpvCode*.
- The Greek and English name of this category by using the properties *elod: cpvGreekSubject* and *elod:cpvEnglisSubject*.

4   Conclusions

LOE can be used to (i) enable information exchange between open source systems, (ii) integrate open data from heterogeneous sources and (iii) publish semantic and linked data related to economic activities. It can also be connected to market processes as analytically described in [1] where the generic LOE model is introduced. At the end of this short paper, we depict a SPARQL query for an open API endpoint (http://143. 233.226.49:8890/sparql). This query asks for retail tomato prices from the official Greek Price Observatory and compares them with wholesale prices provided by the Central Market of Thessaloniki/ KATH (in three different qualities) for a specific day. Retail stores with the highest and the lowest prices are also provided. Finally, the query returns the price paid for a specific public contract for tomatoes, thus combining information from governmental communities and market itself.

Beyond government and public authorities, we do believe that LOE can provide high valuable insights to other domain stakeholders (e.g. citizens, developers, researchers, journalists, companies etc.), who can use it as an establish and compact common ground to get and compile valuable economic information.

Endpoint: http://143.233.226.49:8890/sparql
PREFIX pc: <http://purl.org/procurement/public-contracts#>
PREFIX elod: <http://linkedeconomy.org/ontology#>
PREFIX gr: <http://purl.org/goodrelations/v1#>
PREFIX vcard: <http://www.w3.org/2006/vcard/ns#>
SELECT DISTINCT ?signatureDate ?buyer (str(?buyerLegalName) as ?buyerName)
(str(?sellerLegalName) as ?sellerName)
?cpvEnglishSubject ?amountOfThisGood
(xsd:decimal((1+?vatPercentage/100)?amount) AS ?amountInclVat)*
?minPriceWholesail ?maxPriceWholesail
FROM <http://linkedeconomy.org/EprocurementProper>
FROM <http://linkedeconomy.org/Organizations>
FROM <http://linkedeconomy.org/KATH>
FROM <http://publicspending.net/DiavgeiaI/CPV>
WHERE {
<http://linkedeconomy.org/resource/Contract/15PAY002574748> elod:buyer ?buyer ;
pc:item ?offering ;
elod:signatureDate ?signatureDate .
?offering gr:includesObject ?tqn ;
gr:hasPriceSpecification ?ups ;
elod:seller ?seller .
?tqn gr:typeOfGood ?someItems ;
gr:amountOfThisGood ?amountOfThisGood .
?ups elod:hasVat ?vat ;
gr:hasCurrencyValue ?amount ;
gr:valueAddedTaxIncluded ?taxIncluded .

```
?vat elod:vatPercentage ?vatPercentage .
?someItems elod:productCategory <http://linkedeconomy.org/resource/CPV/03221240-0> ;
      gr:description ?description .
OPTIONAL {
?someItems elod:producedAt ?producedAt } .
<http://linkedeconomy.org/resource/CPV/03221240-0> elod:cpvEnglishSubject
?cpvEnglishSubject .
OPTIONAL {
?buyer gr:legalName ?buyerLegalName .
}.
OPTIONAL {
?buyer vcard:hasAddress ?address .
?address vcard:postal-code "55337"^^xsd:string .
}
OPTIONAL {
?seller gr:legalName ?sellerLegalName .
}
#KATH
?offerWholesail gr:includesObject ?tqnWholesail; gr:hasPriceSpecification ?upsWholesail ;
elod:year <http://linkedeconomy.org/resource/Year/2015> .
?upsWholesail gr:hasMinCurrencyValue ?minPriceWholesail ; gr:hasMaxCurrencyValue
?maxPriceWholesail;
gr:validThrough "2015-02-11T00:00:00"^^xsd:dateTime ;
rdfs:label "Category II"^^xsd:string. ?tqnWholesail gr:typeOfGood ?someItemsWholesail.
?someItemsWholesail elod:hasCpv <http://linkedeconomy.org/resource/CPV/03221240-0> ;
elod:productEnglishName ?nameProductWholesail .
}
```

References

1. Vafopoulos, M., Vafeiadis, G., Razis, G., Anagnostopoulos, I., Negkas, D., Galanos, L.: Linked Open Economy: Take Full Advantage of Economic Data (2016). https://doi.org/10.2139/ssrn.2732218
2. Hepp, M.: GoodRelations: an ontology for describing products and services offers on the web. In: Gangemi, A., Euzenat, J. (eds.) EKAW 2008. LNCS (LNAI), vol. 5268, pp. 329–346. Springer, Heidelberg (2008). https://doi.org/10.1007/978-3-540-87696-0_29
3. Nečaský, M., Klímek, J., Mynarz, J., Knap, T., Svátek, V., Stárka, J.: Linked data support for filing public contracts. Comput. Ind. **65**(5), 862–877 (2014)
4. Miles, A., Matthews, B., Wilson, M., Brickley, D.: SKOS core: simple knowledge organisation for the web. In: International Conference on Dublin Core and Metadata Applications, pp. 3–10 (2005)