

# A Data-Driven Model for Linking Open Economic Information

M. Vafopoulos<sup>1</sup>, A. Koukourikos<sup>1(✉)</sup>, G. Vafeiadis<sup>1</sup>, D. Negkas<sup>2</sup>,  
I. Skaros<sup>2</sup>, and A. Tzani<sup>2</sup>

<sup>1</sup> Institute of Informatics and Telecommunications, National Centre for Scientific Research “Demokritos”, Agia Paraskevi, Greece

vafopoulos@gmail.com, kukurik@iit.demokritos.gr,  
vafeiadis.giorgos@gmail.com

<sup>2</sup> University of Piraeus, Piraeus, Greece

dimneg@gmail.com, skaros.ilias@gmail.com,  
tzaniaggelos@gmail.com

**Abstract.** While public finance data are becoming openly available as part of the broader promotion of fiscal transparency, there is little effort towards maximizing their potential value by interlinking them under a concrete framework and establishing the means to extract interesting insights. The Linked Open Economy model (LOE) aims to act as a top-level conceptualization that connects economic flows with open economic data and as an adaptable and extensible underlying model for modelling different scenarios. The paper presents the LOE model, emphasizing its theoretical foundations. Furthermore, it presents the usage of the model in realistic settings, showcasing its extensibility and its ability to address interesting questions.

**Keywords:** Open data · Linked Data · Semantic web · Economy · Public procurement · Prices · Circular financial flow model

## 1 Introduction

Eight years after the call from [1] for public investments in promoting open data-driven “computational social science”, their ascertainment that “...in the leading disciplinary journals in economics, sociology, and political science, there would be minimal evidence of an emerging computational social science engaged in quantitative modelling...” seems to be still valid.

Despite the plethora of open and big data initiatives that publish an impressive number of datasets, there is not yet much evidence of related scientific research. Maybe because published datasets are of low quality and relevance to research goals or established methodologies in science have not yet adapted to the so called “data revolution” [2]. In any case, it is interesting to try to bridge the gap between the mass available data and research efforts.

Particularly in the field of Economics, only lately some scholars have been starting to discuss the need of incorporating new data in empirical research (see for instance [3, 4]. The recently coined terms “Open data business models” and “Open data economy”

analyse the economic opportunities emerging by the increasing open data provision [5] and the economic benefits of open data [6], respectively. Research in modern disciplines such as Information systems, Business Intelligence and Finance have shown faster reflexes in addressing big data issues, mainly focusing on analytics (see for instance [7–9]).

Today, the most active efforts in open economic data are coming from Open Government Data initiatives that promote innovation, transparency and accountability (e.g. Open Government Partnership [10] and Open Ownership) and mainly refer to the fields of public budgeting, procurement, spending, Official Development Assistance, subsidies and corporate information.

Tygel et al. [11] analyse open budget data initiatives and conclude that special attention should be paid to user feedback, semantics standards and linking possibilities. There are also some research efforts on semantic modelling of budgets that are mainly focused on the available data from specific countries and regions (see for instance [12–17]). In the field of data standards, the Fiscal Data Package is developed as a simple, open technical specification for government budget and spending data.

Alvarez-Rodríguez et al. [18] review the efforts of implementing semantic technologies in the field of e-procurement. Indicatively, the list of projects includes LOTED2, Public Contracts Ontology, Methods On Linked Data for E-procurement Applying Semantics (MOLDEAS) and PPROC ontology.

Recently, the Open Contracting Partnership has developed the Open Contracting Data Standard (OCDS) [19] that sets out key documents and data that should be published at each stage of a contracting process. The Standard is backed up by a documented specification describing data fields and structures that publishers should use to increase the accessibility, usability and interoperability of their disclosures.

In the field of public spending, [OpenSpending.org](http://OpenSpending.org) [15] is an open platform for government expenditure tracking created by the Open Knowledge Foundation. It offers an easy system to upload, explore and share public finance data. The PublicSpending.net project [20–22] cleans, analyses and converts to LOD public spending data from seven governments, both local and national, with total value almost 1.5 trillion euros.

With regards to corporate information, [Opencorporates.com](http://Opencorporates.com) [23] is an effort of aggregating company information from different countries and jurisdictions and releasing it as open data. The opencorporates team works on creating Linked Data representations out of their databases, by mapping company metadata to certified ontologies such as the Core Business Vocabulary and linking them to other data hubs, such as [DBpedia.org](http://DBpedia.org) and Geonames.

[Openownership.org](http://Openownership.org) has been recently established to create a global beneficial ownership register.

The aforementioned initiatives are paving the way for open data in diverse aspects of economic activity, but in practice act as e-catalogues that are fragmented into topic, place and time since they do not share common standards and methodologies. In cases where open data exist, the basic obstacle is the fact that there are not even common practices for representing the main actors (e.g. payers and payees) and the type of payments. Therefore, it is impossible to interlink the available data in meaningful ways and support services and decision-making. Surely, the cost of data discovery and collection has been substantially decreased, but getting valuable insights in public

finance still demands high expertise and timely efforts; and this is a serious danger that may undermine the further development of LOD in general.

Hence, it is time to guide our focus in developing more comprehensive approaches to interconnect the stylized facts of open economic data. The potential usages of such conceptualization ranges from crowd-sourced monitoring and risk assessment of public finance to real-time integration in Business Intelligence systems for more efficient resource allocation.

For instance, subsidies to public and private organizations and the provision of aid to third countries and international organizations can be considered at the same conceptual level of public procurement because they are both money transfers (payments) through a predefined process (e.g. open call). The main difference between public procurement and subsidies or aid provisioning is that in the first case the buyer (or payer) receives direct compensation (e.g. product or service) for its payments. In the case of subsidies and aid provisioning the direct benefits stay within the beneficiary while the society or specific groups of it enjoy the indirect benefits (e.g. social inclusion, economic development).

In this context, the proposed Linked Open Economy (LOE) model aims to bridge theory-driven approaches that offer generality and scalability with the more readily applicable data-driven approaches that cover specific, realistic modelling needs.

More specifically, the LOE model addresses economic open data orchestration by providing a series of coherent and scalable conceptualizations that are based on economic theory and business practice, and extending them to reflect real-world practices and requirements.

## 2 The Linked Open Economy Model

### 2.1 Model Principles

LOE builds on top of the four-sector Circular Flow of Income (CFM) conceptualization used in basic macroeconomics (Fig. 1), where major exchanges are modelled as flows of money, goods and services between economic agents falling under four distinct sectors: (1) Households, (2) Firms, (3) Government and (4) Rest of the world. These agents cooperate over their activity in specific markets of (i) Goods and Services, (ii) Factors or Resources (iii) Financial constructs and participate in flows of economic activity, either financial (direct or indirect exchange of money) or real (exchange of goods, services or factors).

Focusing on the part of CFM that involves governmental actions, we specialize the income flow process as follows: governments form and publish budgets, partially targeting projects and works that are assigned through calls for tenders. The transfer of funds, specified via signed contracts, is realized after the completion of the projects.

The following subsection elaborates on the conceptualization of agents, markets and flows in the LOE model, discussing on the correspondences with CFM and arguing on the assumptions adopted for building the model, based on the realistic restrictions posed by the nature, scope and range of the openly available economic data (marked with red bubbles in Fig. 1).

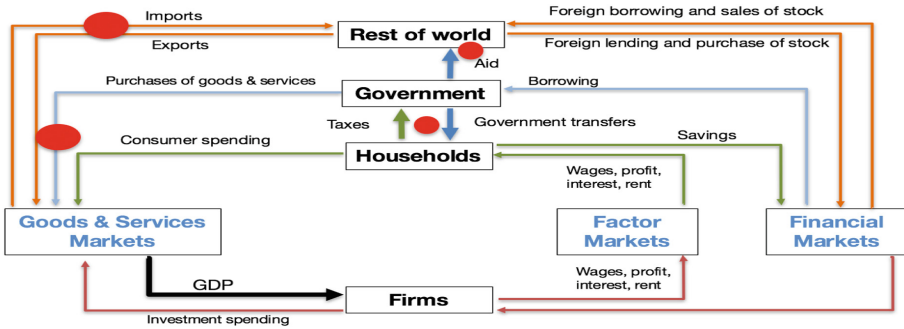


Fig. 1. Extended circular flow model

## 2.2 Adapting CFM to the Open Economic Data Environment

Taking into account the intricacies of the open government data, the LOE model proposes the following conceptualization for the various constructs defined in CFM.

### Economic Agents

Economic agents activated in the CFM flow are represented as instantiations of a generic Agent class. Households are modelled as Persons, as the available open economic data do not include information at the level of households, but rather at the level of a single person as an economic agent. On the other hand, firms are modelled as organizations and business entities. In particular, available data involve a Business firm as a seller of services or products to the Government. Furthermore, a Business firm or a Government can benefit from Subsidies provided by a Government which, if provided to foreign countries, are considered to be International Aid. In this context, a Government acts mainly as a buyer of services and products from Business firms or - in relatively few cases - from other Governments or Governmental units. Note that, in the broader picture, a Business firm routinely acts as a buyer of goods but B2B transaction data are not publicly available.

Another limitation stemming from the scope of publicly available data is the fact that open international trade data describe solely the cumulative value of bilateral trades, with no further information on the specific Business firms that participate in import and export activities. Consequently, LOE groups the Business firms that export and import goods and services in a country as Group National Agents with the given country code, in order to capture the Rest of the World, Imports and Exports constructs in the CFM.

### Flows of Economic Activity

In LOE, direct or indirect money flows are represented as Amounts exchanged between economic agents. Additionally, real flows (i.e. goods, services or factors) are modelled either as foreign Trade Activities or Public Procurement Activities. The latter activities are modelled using the following constructs, reflecting the budgetary and procurement process followed by Greek central and local government, while being straightforwardly extensible to cover other public administrations.

Expense items can be distinguished as Budget Items, Committed Items, Expense Approval Items, or Spending Items. Additionally, Revenue Items can be either Budget Items, Revenue Recognized Items, or Collected Items.

- A Budget Item models a part of a Budget, i.e. an allocation of funds for specific purposes based on an accounting system and according to an annual time plan.
- A Committed Item represents a firm, written obligation from a public organization (buyer) to provide a specified amount of funds, related to a specific Budget Item, under particular terms and conditions and for specific purposes. The recipient (beneficiary or seller) may be defined at this or at a later stage. Commitments can be issued before or after the procurement process. In the case of public budgeting in Greece, commitments are issued before the procurement process.
- The Expense Approval Item construct is introduced to model cases where further approvals are required to proceed with payments. The approvals are bureaucratic administrative decisions issued after the delivery of a contract and before a payment.
- A Spending Item represents the final stage of an Expense Budget Item. Part or the total of related expense approval items proceed to payment from buyer to seller.
- A Revenue Recognized Item represents a revenue that has been recognized by the public organization and is qualified for collection (e.g. a fine or a tax).
- A Collected Item represents the final stage of a Revenue Budget Item, i.e. part or the total of revenue budget items are collected by the public organization from a third party (e.g. taxpayers and central government).

## Markets

The concept of markets is included in the LOE model primarily through open data for prices for specific products (e.g. basic goods, agricultural products and fuel), since there is no other source of publicly available data.

Taking into account the aforementioned remarks and assumptions the LOE model proposes a specialization of the CFM tailored to the specificities of the targeted domain, as depicted in Fig. 2.

The following section presents the formal specification of the LOE model, adapting the generic CFM approach to the actual data being made available.

## 2.3 Formal LOE Model Specification

Given the aforementioned definitions and assumptions, the LOE model formally implements the resulting conceptualization as an ontology using the Web Ontology Language (OWL) W3C specification. LOE extensively uses established ontologies and is fully conforming with the Core Vocabularies developed by the European Union.

### Usage of Existing Ontologies

The following table summarizes the external ontologies used in LOE for defining the concepts and roles foreseen by the model (Table 1).

FOAF is used to describe agents responsible for specific actions as defined in LOE. Specializations relevant to LOE are foaf:Person and foaf:Organization. In the same fashion, the GoodRelations ontology is used to describe the Business Entities involved in a commercial activity, the type of their services, and the financial details of the

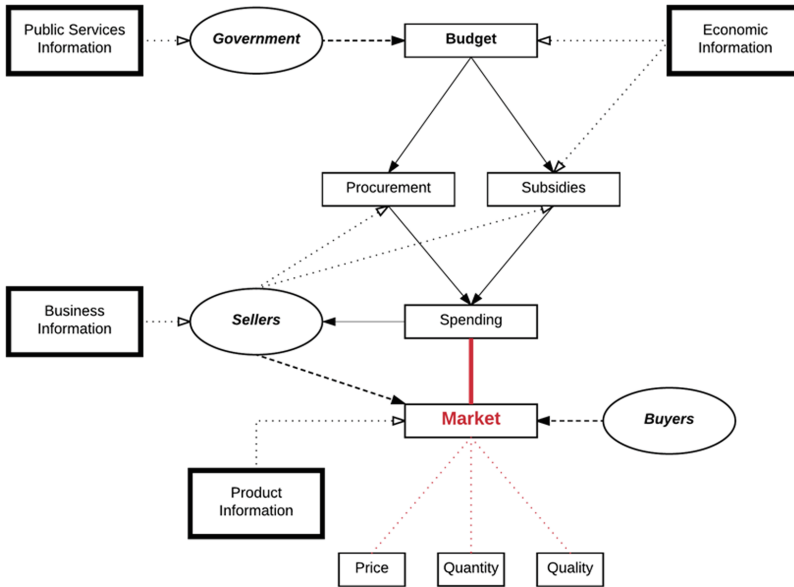


Fig. 2. CFM model specialization for public spending connected to the market

Table 1. Namespaces used at the general LOE model

Ontology	Namespace	Prefix
Friend of a friend	<a href="http://xmlns.com/foaf/0.1/">http://xmlns.com/foaf/0.1/</a>	foaf
Good relations	<a href="http://purl.org/goodrelations/v1#">http://purl.org/goodrelations/v1#</a>	gr
Organizations	<a href="http://www.w3.org/ns/org#">http://www.w3.org/ns/org#</a>	org
Registered organization	<a href="http://www.w3.org/ns/regorg#">http://www.w3.org/ns/regorg#</a>	rov
Public contracts	<a href="http://purl.org/procurement/public-contracts#">http://purl.org/procurement/public-contracts#</a>	pc
Linked economy	<a href="http://linkedeconomy.org/ontology#">http://linkedeconomy.org/ontology#</a>	elod

contract or of the payment. The Organization ontology is used to define the organizations along with their organizational units. Public Contract ontology is used to define the following types of information:

- public contracts during all stages of their existence
- procedures specifying how the details of a contract is published and how a supplier is selected
- main object of the contract (e.g. works, supplies or services)
- contract’s price, depending on its stage (before or after the offer)
- award criteria that define the conditions under which the best offer will be selected and awarded, along with their weights and
- main and supplementary products or services purchased by the contract (as determined by their CPV codes).

The conceptualization of organizational entities is carried out through the usage of the Organization ontology and the Registered Organization vocabulary.

### Conformance with Core Vocabularies

In order for a model to conform to the e-Government Core Vocabularies as developed by the European Union, it needs to publish a mapping to the conceptual model of the Core Vocabularies, as a self-conformance statement. The statement must cover two basic requirements:

- Each data element in the model is required to have an identifier, a label and a definition;
- The provided mapping should include the following information: Core Vocabulary identifier, Core Vocabulary version, mapping relation, Identifier, Identifier of the data element that is mapped, and a comment on the mapping.

The LOE specification includes a self-conformance statement in the form of a mapping spreadsheet, publicly available via the LOE model distribution.

## 3 The LOE Model in Practice

While LOE is based on a concretely defined economic abstraction as the CFM, its true aim is to bridge this theoretical foundation with the current and future reality of the open financial data ecosystem and, on an even broader context, with the Web 3.0 practices. As a testbed for assessing the model's adaptability and extensibility, it has been used as the backbone of the YourDataStories (YDS) project platform ([platform.yourdatastories.eu](http://platform.yourdatastories.eu)), a framework aiming to serve stakeholders of different communities under different use cases. The following sections summarize the use cases defined in the context of YourDataStories, the respective data to be processed, and the extensions posed by the specific requirements. Detailed access information is summarized in Sect. 4.

### 3.1 YourDataStories Use Cases and Objectives

YDS operates over three distinct but connected use cases, all handling economic data of different scope and granularity and targeting different operations where the usage of open data can produce significant added value.

**YourDataStories Pilot #1: Follow Public Money.** Often, information about public projects resides in not connected systems owned by different ministries and public agencies. In the case of Greece, this information can be mainly found in the form of open data in two systems: (1) the NSRF portal ([anaptyxi.gov.gr](http://anaptyxi.gov.gr)) and (2) the Greek Transparency Portal (Diavgeia, [diavgeia.gov.gr](http://diavgeia.gov.gr)).

In particular, the NSRF portal presents project-based information such as the title, budget, completion rate, related subprojects and the involved public and private organizations of a specific project. Contrarily, the Diavgeia dataset has been designed to be organization-based in the sense that every public organization has to publish all its administrative decisions. Thus, data are organized in decision types fitted to the Greek public organizations. Despite the fact that any administrative decision related to NSRF projects must be uploaded and distributed through the Diavgeia website, practically, this cannot be validated in project base because there is not a unique identifier (e.g. project code) to interconnect NSRF data to Diavgeia decisions.

The first YDS pilot is focused on bridging this gap between NSRF and Diavgeia data by identifying -through text mining- related administrative decisions for each NSRF project.

As a next step, geographical information for each project was manually added through a customized web interface for important public projects (e.g. highways and tube stations) in order to build mobile services for receiving comments and evaluations by the users.

**YourDataStories Pilot #2: Official Development Assistance.** The Dutch Ministry of Foreign Affairs supports many development projects across a broad range of countries, but has announced that there is a 50% cut in Dutch Official Development Assistance (ODA) starting from 2015. The Ministry also decided that at the same time it would be keen to support new projects that provide a ROI for the Dutch economy and work with particularly innovative methods. Accordingly, Dutch as well as local development NGOs are protesting. They argue that crucial humanitarian projects will come to a premature end, thus invalidating experiences and efforts built up over long time periods. This debate has attracted the attention of a team of Dutch journalists. Aim of the use case is to increase the transparency of public spending through data-driven journalism.

Initially, in pilot 2, is selected a set of 5–8 sample countries that receive development assistance from the Netherlands. Next, we retrieve detailed information on all projects run in these countries between 2011 and 2016. In step 3, we analyse information on trade relations between the sample countries and the Netherlands between 2011 and 2016. Consequently, it is retrieved detailed information on all projects run in the sample countries by Germany, France, and Denmark and on trade relations during the same period. In step 6, we made a flexibly experiment with simple charts to analyse the information above and share some expressive charts on social media. We also generated a set of research questions based on the “five Ws”, with the help of social media users and conducted a journalistic investigation.

**YourDataStories Pilot #3: Cross-Europe Financial Comparability.** The third pilot focus on the comparability of financial data across EU member-states, specifically looking at Ireland and Greece. Financial data from the Greek ‘Follow Public Money’ pilot will be compared with budget and spending data from the Irish national and local government, with a particular focus on construction and road infrastructure projects. Issues reported by the public via FixMyStreet.ie will also be incorporated into the data stream. Additionally, cross EU and international comparisons in public procurement are made by modelling and analysing Tenders Electronic Daily (TED) and Australian contract data.

This intelligent use of big and publicly available economic data aims to stimulate smart services for the following target users: (a) journalists to search for new stories and additional sources, (b) civil society to act more effectively in transparency and accountability issues, (c) auditors to better evaluate effectiveness and corruption in public bodies, (d) web and mobile developers to get access in cleansed and structured economic data and (e) suppliers to search for business opportunities in public procurement.

### 3.2 Pilot-Driven Model Extensions

The open nature of LOE allowed its direct extension with additional concepts and constructs, in order to reflect the needs of the aforementioned use cases. The following



figure summarizes the model supporting the YDS platform, showcasing the involved entities and their relationships in the broader YDS context (Fig. 3).

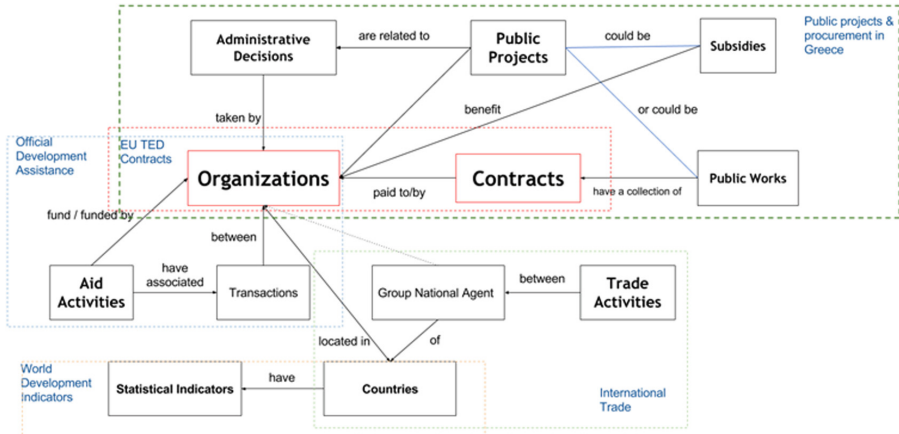


Fig. 3. YourDataStories project extension of the LOE Model

### 3.3 Datasets

The relevant datasets for each pilot are retrieved via dedicated, source-tailored harvesters (<https://github.com/YourDataStories/harvesters>) that consume the APIs and repositories of each source. The retrieved data are semi-automatically checked for correctness and consistency and are transformed into the respective RDF representations that are subsequently stored in the YourDataStories knowledge base. The following Tables 2, 3, 4 and 5 summarize the dataset size retrieved and incorporated in the YDS ecosystem.

Table 2. Diavgeia harvester results (November 2010 to November 2015)

Description	Count
Buyers	18,706
Sellers	406,955
Financial decisions	6,269,062
Produced RDF triples	166,435,634

Table 3. NSRF harvester results

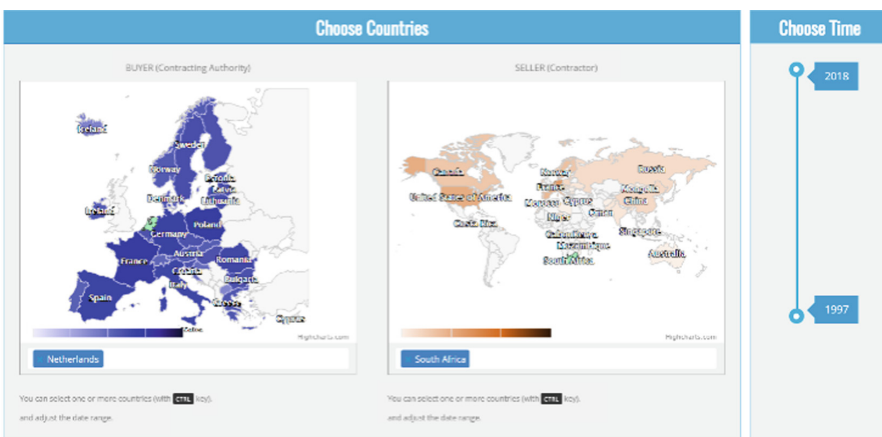
Description	Count
Buyers	3,766
Sellers	17,484
Beneficiaries	126,773
Public works	12,411

**Table 4.** Triple count for pilot 2 data sources (graphs)

Graph	Triple count
<a href="http://yourdatastories.eu/WDI">http://yourdatastories.eu/WDI</a>	3,816,489
<a href="http://yourdatastories.eu/OpenCorporates">http://yourdatastories.eu/OpenCorporates</a>	10,103
<a href="http://yourdatastories.eu/WorldFactbook">http://yourdatastories.eu/WorldFactbook</a>	8,606
<a href="http://yourdatastories.eu/taxonomies">http://yourdatastories.eu/taxonomies</a>	180,695
<a href="http://yourdatastories.eu/trade/ZW">http://yourdatastories.eu/trade/ZW</a>	2,112,174
<a href="http://yourdatastories.eu/trade/NL">http://yourdatastories.eu/trade/NL</a>	17,816,427
<a href="http://yourdatastories.eu/ODA/ZW">http://yourdatastories.eu/ODA/ZW</a>	139,700
<a href="http://yourdatastories.eu/ODA/NL">http://yourdatastories.eu/ODA/NL</a>	668,919
<a href="http://yourdatastories.eu/countries">http://yourdatastories.eu/countries</a>	8657
<a href="http://yourdatastories.eu/regions">http://yourdatastories.eu/regions</a>	236

**Table 5.** Highest amount of EU Tender amount per country

Country	Notices	Award notices	Buyers	Sellers
FR	431,770	1,199,603	218,123	520,690
PL	176,607	997,541	26,394	265,935
UK	114,299	290,423	57,337	179,406
DE	213,520	290,956	63,286	157,393
IT	91,478	180,133	47,073	94,146
ES	93,322	198,991	37,164	91,059
BG	18,854	68,399	10,899	42,790
RO	38,418	160,304	17,213	41,814
EU	1,640,996	4,287,115	642,359	2,037,866



**Fig. 4.** YourDataStories platform interface

```

PREFIX elod: <http://linkedeconomy.org/ontology#>
PREFIX dcterms: <http://purl.org/dc/terms/>
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
PREFIX gr: <http://purl.org/goodrelations/v1#>
PREFIX foaf: < http://xmlns.com/foaf/0.1/>
PREFIX pc: < http://purl.org/procurement/public-contracts#>

select distinct ?title ?buyerName (xsd:decimal(?amount) as
?amount) ?graph
where {
  {
    graph ?graph {
      ?decision elod:hasExpenditureLine ?expLine ;
      dcterms:subject ?title ; dcterms:publisher ?unit .
      ?buyer org:hasUnit ?unit ; skos:prefLabel ?buyerName .
      ?expLine elod:seller
      <http://linkedeconomy.org/resource/Organization/999995379> ;
      elod:amount ?ups . ?ups gr:hasCurrencyValue ?amount .
    }
  }
  union {
    graph ?graph {
      ?project elod:hasRelatedContract ?subproject ; elod:buyer
      ?buyer ; dcterms:title ?title .
      ?subproject elod:price ?budgetUps ; dcterms:title ?titleSub-
      project ; elod:seller
      <http://linkedeconomy.org/resource/Organization/999995379>.
      ?buyer foaf:name ?buyerName .
      ?budgetUps gr:hasCurrencyValue ?amount .
      filter langMatches(lang(?title),'el')
    }
  }
  union {
    graph ?graph {
      ?awardNotice elod:documentType ""Contract award no-
      tice""^^<http://www.w3.org/2001/XMLSchema#string> ;
      elod:buyer ?buyer ;
      elod:seller
      <http://linkedeconomy.org/resource/Organization/999995379> ;
      elod:contractId ?contractId .
      OPTIONAL{?buyer skos:prefLabel ?buyerName .}
      OPTIONAL {
        ?awardNotice dcterms:title ?title
      } .
      OPTIONAL {
        ?awardNotice pc:agreedPrice ?ups .
        ?ups gr:hasCurrencyValue ?amount
      } .
    }
  }
}

```

**Fig. 5.** Combining seller information from different data collections

### 3.4 YourDataStories Platform and Queries

The overall LOE model, as well as, the partitions referring to the different pilot-specific domains, are openly accessible via the YDS GitHub repository ([github.com/YourDataStories/ontology](https://github.com/YourDataStories/ontology)). The data pertaining to YDS pilots is exposed via the YDS platform ([platform.yourdatastories.eu](https://platform.yourdatastories.eu)), from where the relevant stakeholders can access the underlying repositories via a SPARQL endpoint. The platform allows the immediate observation of various metrics related to the scenarios of the pilots, while the endpoint allows the direct querying of the YDS datasets (Fig. 4).

The YDS ontology distribution provides indicative queries, aiming to showcase the completeness and the added value of the model and its population. Some exemplary interesting queries are depicted in the following figures. Figure 5 depicts the query for combining seller information from different collections (graphs in the YDS repository).

Similarly, Fig. 6 presents the query for retrieving the aggregate amounts per country made available for a given CPV code.

## 4 Accessing and Using the LOE Model and YDS

Complete information on the LOE model and its applications can be found through the YourDataStories project website<sup>1</sup> and the project's GitHub repository dedicated to the development and maintenance of the ontology assets used throughout the YDS applications and use cases<sup>2</sup>.

YDS also offers a range of dashboards for the covered datasets, where the interested users can obtain customized views for the data, essentially via a visual query editor. Direct access to the SPARQL endpoint is provided at: <http://143.233.226.60:8890/>, where the end user can execute the exemplary queries provided through this paper or the YDS GitHub repository, or formulate and run custom queries.

## 5 Conclusions

Publicly available open data are growing rapidly in quantity, but their quality can be further improved to unveil their strong potential. One particular value aspect of open data is related to their ability to address unanswered questions and provide more effective solutions in crucial policy and management issues. The LOE model is proposed as a high-level conceptualization that incorporates major economic open data, by including them in specifications that adheres to the generic CFM model. LOE, as evidenced by its support for different use cases in the context of the YDS project, is designed as the foundation for a compact but extensible common ground for journalists, professionals and public authorities to import, consume and customise open economic data.

<sup>1</sup> <https://yourdatastories.eu/>.

<sup>2</sup> <https://github.com/YourDataStories/ontology>.

```

PREFIX elod: <http://linkedeconomy.org/ontology#>
PREFIX skos: <http://www.w3.org/2004/02/skos/core#>
PREFIX gr: <http://purl.org/goodrelations/v1#>
PREFIX pc: < http://purl.org/procurement/public-contracts#>

SELECT ?countryName (sum(xsd:decimal(?agreedPriceValue)) as
?sumOfAmounts)

FROM <http://yourdatastories.eu/TEDUpdate>
FROM <http://yourdatastories.eu/TEDGreece>
FROM <http://yourdatastories.eu/TEDIreland>
FROM <http://yourdatastories.eu/taxonomies>
FROM <http://yourdatastories.eu/countries>
WHERE {
?awardNotice elod:documentType "'Contract award no-
tice'"^^<http://www.w3.org/2001/XMLSchema#string> ;
elod:buyer ?buyer ;
pc:mainObject <http://linkedeconomy.org/resource/CPV/45233100-
0> ;
pc:agreedPrice ?ups .
?buyer elod:countryIsoCode ?countryIsoCode .
?ups gr:hasCurrencyValue ?agreedPriceValue .
?countryIsoCode skos:prefLabel ?countryName
}
GROUP BY ?countryName

```

**Fig. 6.** Cross-country CPV code budgets

## 6 Future Work

As stated in Sect. 2, the central ambition of the LOE model is to incorporate under a theory-wise sound conceptualization real-world data of different forms and functions. Consequently, and as financial and economic data are increasingly becoming openly available, our efforts will focus on analysing further open datasets and incorporating them in the LOE knowledge base, extending or modifying accordingly the overall model.

Another major step towards the expansion of the model and its ambition to reflect the realistic nature of public spending, is the introduction of social data in the model. Web 2.0 increasingly becomes a major source of information for facilitating the transition to the Web 3.0 vision to which LOE adheres. The initial step towards the linking of open economic data with social media will revolve around the first YDS pilot, and focus on the discovery and processing of user activity relevant to the public works incorporated in the YDS repository. To this end, the LOE model will be extended to accommodate (a) social media content and (b) its links to specific public works and their properties. The analysis of the extracted content will be modelled quantitatively (as a progress/evaluation score) and qualitatively (via the extraction of important features for the given project and the assignment of comments to the different features). A strong basis of this LOE extension is the SIOC ontology, a W3C specification and ongoing project trying to model the activities and properties of online communities, further extending the linking of the LOE schema with external conceptualizations.

**Acknowledgements.** Research and outcomes reported in the present manuscript have received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement No. 645886 (YourDataStories – YDS).

## References

1. Lazer, D., Pentland, A.(S.), Adamic, L., Aral, S., Barabasis, A.L., Brewer, D., Christakis, N., Contractor, N., Fowler, J., Gutmann, M., Jebara, T., King, G., Macy, M., Roy, D., Van Alstyne, M.: Life in the network: the coming age of computational social science. *Science* (N.Y. NY) **323**(5915), 721–723 (2009)
2. Kitchin, R.: *The Data Revolution: Big Data, Open Data, Data Infrastructures and Their Consequences*. Sage Publications, Thousand Oaks (2014)
3. Wolfers, J.: Big data needs economics and vice versa. Blanchard-fest Celebration. MIT, 4 June 2016
4. Smith, N.: *Economics Struggles to Cope With Reality* (2016)
5. Zeleti, F.A., Adegboyega, O., Curry, E.: Exploring the economic value of open government data. *Gov. Inf. Q.* **33**(3), 535–551 (2016)
6. World Bank: *Open Data for Economic Growth*. World Bank, Washington, D.C. (2014)
7. Hsinchun, C., Chiang, R.H., Storey, V.C.: Business intelligence and analytics: from big data to big impact. *MIS Q.* **36**(4), 1165–1188 (2012)
8. Agarwal, R., Dhar, V.: Editorial—big data, data science, and analytics: the opportunity and challenge for IS research. *Inf. Syst.* **25**(3), 443–448 (2014)
9. Fang, B., Zhang, P.: Big data in finance. In: Yu, S., Guo, S. (eds.) *Big Data Concepts, Theories, and Applications*, pp. 391–412. Springer, Cham (2016). doi:[10.1007/978-3-319-27763-9\\_11](https://doi.org/10.1007/978-3-319-27763-9_11)
10. Manolea, B., Cretu, V.: Topic Report No. 2013/10: The influence of the Open Government Partnership (OGP) on the Open Data discussions. Misc, European Public Sector Information Platform (2013)
11. Tygel, A.F., Attard, J., Orlandi, F., Campos, M.L.M., Auer, S.: How much? Is not enough: an analysis of open budget initiatives. In: *Proceedings of the 9th International Conference on Theory and Practice of Electronic Governance* (2016)
12. Bolgov, R., Chugunov, A., Filatova, O.: Open budgets in Eurasian countries: Cloacae or Klondike? In: *Conference for E-Democracy and Open Government* (2015)
13. Brusa, G., Caliusco, M., Chiotti, O.: A process for building a domain ontology: an experience in developing a government budgetary ontology. In: *Proceedings of the Second Australasian workshop on Advances in Ontologies* (2006)
14. de Oliveira Araújo, L.S., Santos, M.T., Silva, D.A.: The Brazilian federal budget ontology: a semantic web case of public open data. In: *Proceedings of the 7th International Conference on Management of computational and collective intelligence in Digital EcoSystems* (2015)
15. Höffner, K., Martin, M., Lehmann, J.: LinkedSpending: OpenSpending becomes Linked Open Data. *Semant. Web* **7**(1), 95–104 (2015)
16. de Santana, M., da Silva Craveiro, G.: Challenges and requirements for the standardisation of open budgetary data in the Brazilian public administration. In: *GI-Jahrestagung* (2013)
17. Vafopoulos, M., Meimaris, M.: Weaving the economic Linked Open Data. In: *7th International Workshop on Semantic and Social Media Adaptation and Personalization (SMAP)* (2012)
18. Alvarez-Rodríguez, J.M., Labra-Gayo, J.E., de Pablos, P.O.: New trends on e-Procurement applying semantic technologies: current status and future challenges. *Comput. Ind.* **65**(5), 800–820 (2014)

19. Open Contracting Partnership: Open Contracting Data Standard (OCDS), version 1.0 (2015)
20. Vafopoulos, M., Meimaris, M., Papantoniou, A., Alexiou, G., Avraam, I., Xidias, I., Loumos, V.: Publicspending.gr : Greek public spending in a nutshell. In: Samos Summit on Open Data and Interoperability (2012)
21. Vafopoulos, M., Rodríguez, J., Meimaris, M., Xidias, I., Klonaras, M., Vafeiadis, G.: Insights in global public spending. In: 13th International Conference on Knowledge Management and Knowledge Technologies and 9th International Conference on Semantic Systems, Messe Congress Graz, Austria, 2013
22. Vafopoulos, M., Meimaris, M., Anagnostopoulos, I., Papantoniou, A., Xidias, G.A.I., Loumos, V.: Public spending as LOD: the case of Greece. *Semant. Web J.*, 155–164 (2014)
23. Taggart, Ch., Peltola, V.: OpenCorporates: why we're crying out for this database of companies. *The Guardian* (2011). <http://bit.ly/2ilSwdD>. Accessed 10 Oct 2017