

Intelligent and semantic real-time process of the Greek LOD for enhancing citizen awareness in public expenditures

M. Vafopoulos¹, M. Meimaris¹, A. Papantoniou¹, I. Anagnostopoulos²,
I. Xidias¹, G. Alexiou¹, G. Vafeiadis¹ and V. Loumos¹

¹ School of Electrical & Computer Engineering, National Technical University of Athens

² Computer Science and Biomedical Informatics Dpt., University of Central Greece

{vaf, apapant}@medialab.ntua.gr, janag@ucg.gr

Abstract. The provision of publicly available Open Data targets to provide transparency in several public sector decisions and actions. However, this information is served massively and in different forms - mostly due to different bureaucratic procedures- while its diffusion does not occur at regular or at least generally predictable time intervals. Thus, even though the information is available by the involved public sectors and enterprises, citizens are overwhelmed from the size/inconsistency of the information they deal with. In this paper, we present a publicly accessible Web point (publicspending.gr/), which aims to promote transparency and enhance citizen awareness regarding public spending in Greece. The information is based on intelligent and semantic real-time processed open data that are derived from the open API of “Diavgeia” (Clarity) Program (diavgeia.gov.gr/en/) and TAXIS - Taxation Information System (gsis.gr/), while it is served through easily consumed visualization charts and diagrams.

Keywords. Public Spending, Linked Open Data, Transparency, Intelligent Information Processing, Semantics

1 Background

Technologies of the Semantic Web or the Web of Data has significantly expanded and are now applied to bridge previously autonomous business domains and to concatenate independent government activities [1], [2], [3]. A consequence of the Web of Data underlying technologies brought up lately, the term “Linked Open Data” (LOD) or “Big Data”. LOD now form a quite extraordinary cloud, but more than that its datasets involve provenance and governmental data (e.g. openspending.org). In addition, Linked Data enable the creation of better and massive services for use and reuse for many of these data, driving existing infrastructure in its full potential [4]. For government bodies, Linked Data adoption is focused on open, transparent, collaborative and more efficient governance. Related approaches with the one proposed in this paper are “Where does my money go: Showing you where your taxes get spent”

(wheredoesmymoneygo.org), “Open Public Procurement Project” (tender.sme.sk), “Accountability Initiative” (accountabilityindia.in), “Pera Natin ‘to! (It’s Our Money!)” (transparencyreporting.net/), and “Texas Transparency” (texastransparency.org/moneygoes) in U.S.

2 System Description and Data Visualization

Our system is mainly based on data feeds provided by “Diavgeia”, the first Greek Government Open Data API (opendata.diavgeia.gov.gr/). “Diavgeia”, offers the possibility for publicly accessing the full set spending decisions of the Greek Public Sector Organizations in XML files. Metadata include information concerning the names and the VAT registration numbers of payers and payees, the types of government expenditures in a single classification system for public procurement, aimed at standardizing the references used by contracting authorities and entities to describe the subject of procurement contracts (CPVs)¹. Tax Information System (TAXIS, gsis.gr), which is the official web portal where citizens and legal entities submit taxation-related information and documents, is also employed. Through TAXIS, queries are performed in the form of SOAP calls with the VAT registration number of the entity as the reference key. The response contains metadata about the legal entity, including contact details, activity descriptions, registration dates and current operational status. Within the scope of this project, the web service is used for querying legal entities on their first appearance as payment agents, while the response data are “RDFized” and stored as payment agent metadata. The Web point is functional from the beginning of June 2012 under the hosting domain of publicspending.gr. Currently project deliverables run on two “sandbox” domains, one hosting the demo site and one hosting the SPARQL API endpoint of the project (publicspending.medialab.ntua.gr/en/data.php). Through this API one can run SPARQL queries against the dataset and output the results in various formats, such as HTML, spreadsheet, XML, JSON, JavaScript, N-triples, RDF/XML and CSV. A free version of the Highcharts visualization API² was used in order to visualize the resulting JSON data. Diagrams are provided in terms of daily, weekly, monthly and yearly time intervals, as well as since the beginning of the “Diavgeia” project (overall time period). It involves visual information regarding public expenditure according to the respective CPV codes, information about payers, payees, as well as their one-to-one payment connections. Some diagram examples are illustrated in the Appendix (Figure 1). Moreover, Figure 1a presents the total public expenditure on a weekly basis, while Figure 1b depicts the top-10 awardees of public contracts. For both diagrams, the amounts (in MEuros) are calculated from the beginning of “Diavgeia”. Furthermore, Figure 1c illustrates the top-20 expenditure categories according to the respective CPV classification used, for the period between 17/06/2012 and 24/06/2012 (weekly-based information). Finally, Figure 1d highlights through a bubble depiction, the top-10 payees for the payer “Ministry of Finance”.

¹ http://simap.europa.eu/codes-and-nomenclatures/codes-cpv/codes-cpv_en.htm

² <http://www.highcharts.com>

This is an important type of diagrams because it demonstrates the binary relation between the payer and its top-payees. Similarly, in this figure, the amounts are calculated from the inception of “Diavgeia” up to the very last day of submitting this work, thus covering a time period where Greece suffers from the so-called economic “crisis”.

3 Research issues – Future work

Having been inspired from a widely used initiative, the British “Opening up government” project (<http://data.gov.uk/>) and the corresponding “payments” ontology [5], our ontology was developed from scratch. Through the `publicspending.gr` ontology we are initiating cross-governmental interlinking between Greece and UK public spending. However, the expansion of such functions demands to resolve compatibility issues concerning existing “Big Data” datasets. This subject is related to the linkage of our dataset to `openspending.org`, as well as to `opencorporates.com` and the core person, business and location vocabularies³. Parallel to that, we plan to link with the Greek Public Contracts Registry⁴ and the Greek DBpedia (el.wikipedia.org). Apart from the typical graph-based approach described, there will be further improvements/enhancements involving Open GeoSpatial APIs (e.g. `OpenStreetMap`⁵) for comparing public expenditures in respect to demographic data. Future work also involves the connection to the National Typography Service (`et.gr`) and to other similar Open Data portals that are under development in local and global scale. Most importantly, we plan to launch a powerful and elastic dashboard, which can be used by citizens and data journalists to perform comparative queries, to identify irregularities among national and European governmental expenditures (e.g. comparing per unit costs for public spending in health, education, etc.).

4 References

1. A. Passant, P. Laublet, J. G. Breslin, S. Decker, “Enhancing Enterprise 2.0 Ecosystems Using Semantic Web and Linked Data Technologies: The SemSLATES Approach”, *Linking Enterprise Data*, Springer US, 79-102, 2010, doi: 10.1007/978-1-4419-7665-9_5.
2. G. Kobilarov et al., “Media Meets Semantic Web – How the BBC Uses DBpedia and Linked Data to Make Connections”, *Proc. European Semantic Web Conference (ESWC2009)*, Crete, June 2009, doi:10.1007/978-3-642-02121-3_53.
3. D. Siegel, “Pull—the Power of the Semantic Web to Transform your Business”, *Portfolio—Penguin Publishing Group*, New York, 2010, ISBN: 9781591842774.
4. Vafopoulos, M. 2012. The Web economy: goods, users, models and policies. *Foundations and Trends® in Web Science*, Vol. 3: No 1-2, pp 1-136, <http://dx.doi.org/10.1561/1800000015>.
5. <http://data.gov.uk/resources/payments>

³ http://ec.europa.eu/isa/actions/01-trusted-information-exchange/1-1action_en.htm

⁴ dev.opengov.gr/d/agora/?page_id=322

⁵ <http://www.openstreetmap.org/>

Appendix

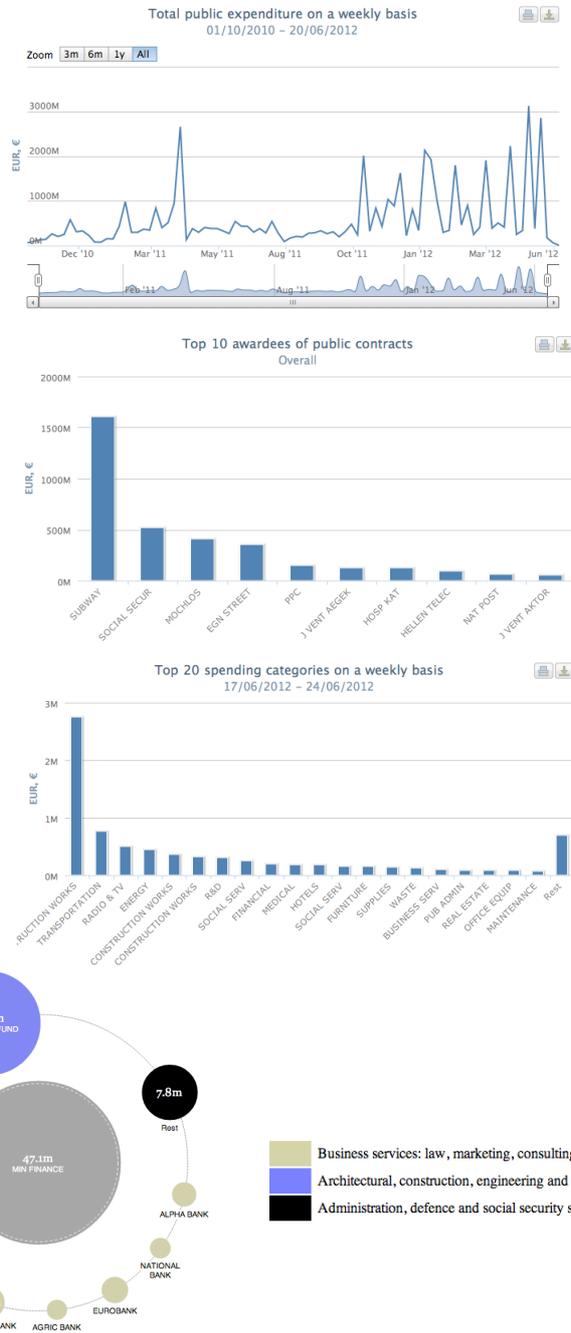


Fig. 1. Indicative visualizations derived from publicspending.gr