Available online at www.sciencedirect.com

ScienceDirect

Procedia - Social and Behavioral Sciences 147 (2014) 167 – 174

Procedia
Social and Behavioral Sciences

ICININFO

Enriching the e-GIF Ontology for an Improved Application of Linking Data Technologies to Greek Open Government Data

Pavlina Fragkou, Eleni Galiotou*, Michalis Matsakas

Dept. of Informatics, Technological Educational Institute of Athens, Ag. Spyridonos, Egaleo Athens GR-122 10, Greece

Abstract

We present the current state of an ongoing project which involves the application of linked data technologies on a portion of data found in the ERMIS Greek portal for Public Administration. A starting point for linked open data technologies is data organization i.e. vocabulary definition that covers all classes and properties via an RDF schema or an ontology. More specifically, the focus of this paper is the enrichment of the e-GIF ontology, modeling the concepts and relations that are used to organize the information appearing in the ERMIS Greek portal for Public Administration with equivalent information required for the purposes of license provision according to the Directive 123/2006/EC. The enrichment process also involves the examination of the information appearing in the sites of other European countries. As it is already known, ontology alignment is a prerequisite step for finding schema-level links between Linked Open Data (LOD) datasets. The aim here is to identify those sites that contain the most valuable information and to include their concepts and relations in our ontology. This type addition will be further used for the interconnection of information between different sites via linked data.

Selection and peer-review under responsibility of the 3rd International Conference on Integrated Information.
Keywords: Ontologies; e-GIF; Open government data; linked data; Greek one stop shop governmental portal; Directive 123/2006/EC

1. Introduction

Open Government Data (Berners-Lee, T. (2006) provide useful information to citizens and enterprises for their transaction with the public sector. Although this public sector information is available to anyone for analyze and reuse, it is organized and published in such a chaotic way that its exploitation in its current state remains a difficult task. This situation will be facilitated only if public sector data are transformed to Linked Open Data (LOD) in order to satisfy the minimum requirements for data linkage and reuse. The term "*linked data*" refers to "*data published on the web in such a way that it is machine-readable, its meaning is explicitly defined, it is linked to other external datasets and can in turn be linked to from external datasets*" (Bizer, C., Heath, T. & Berners-Lee, T.

Selection and peer-review under responsibility of the 3rd International Conference on Integrated Information.
doi:10.1016/j.sbspro.2014.07.141

2009). According to Tim Berners-Lee (Berners-Lee, T. 2006; Linked Data Paradigm) the four "*Linked Data Principles*" are:

1. Use URIs as names for things
2. Use HTTP URIs so that people can look up those names.
3. When someone looks up a URI, provide useful information, using the standards (RDF, SPARQL)
4. Include links to other URIs, so that they can discover more things.

In Galiotou, E., & Fragkou, P. (2013) we have proposed a case study on the creation of LOD concerning data hosted in ERMIS Greek Government portal for public administration. In particular, we focused on data regarding service provision according to Directive 123/2006/EC, which aims at simplifying the procedure of practicing a profession by a European citizen in another member state. In order to implement this directive, each member state is required to provide a dedicated portal named Point of Single Contact (Greek PSC), which contains information regarding the required supporting documents for each service activity and the related information in two languages (that of the member state and English). Each service activity is categorized into one or more categories, which follow the European reference framework for the production and dissemination of statistics related to economic activities- NACE (Nace codes). The terms that were chosen in order to describe each service activity, were based on the e-GIF ontology which is part of the Greek e-Government Interoperability framework, a survey containing a list of rules for the provision of e-Government services to public bodies, businesses and citizens in a unified manner. In our approach (Galiotou, E., & Fragkou, P. 2013) to transforming data of the ERMIS portal into Linked Open Data, we have come across a number of issues such as:

- Web page retrieval and storage
- Eventual enrichment of the e-GIF ontology in order to cover all types of concepts and relations necessary
- Transformation of web pages into RDF using the (enriched) eGIF ontology
- Choice of appropriate tools to efficiently perform all the aforementioned in order to create Linked Open Data

A detailed examination of the eGIF ontology (used as a basis for organizing information in ERMIS) has led us to the conclusion that an enrichment is necessary in order cover all necessary cases of semantic interconnections. In this way, the enriched ontology will serve as a guideline in order to create links via RDF annotations, such as: links between a service activity in Greek and its translation in English, links involving the Public Sector responsible for the legal framework of the service activity, the Public Sector responsible for reception and expedition of the authorization of the service activity, the service activity type, each type of supporting document as well as the NACE codes. A first step towards such an enrichment is proposed in this paper, which is organized as follows:

In the beginning of section 2, we provide a brief description of the Greek e-GIF ontology. Then, an attempt to propose the appropriate update is made. In section 3 we perform a comparison between the updated e-GIF ontology used in our point of single contact (PSC) and corresponding ones used by the PSCs of other European countries. The purpose of this comparison is to locate common information between sites. Finally, in section 4, we draw conclusions and point to future work.

2. The Greek e-Gif ontology

The eGovernment Knowledge Interoperability Ontology (eGKI or Greek E-Gif) is a two-layer ontology whose aim is to capture and link all knowledge elements that are essential to describe services provided to citizens or businesses either manually or electronically (Sourouni, A-M., Lampathaki, F., Mouzakitis, S., Charalabidis, Y., & Askounis, D., 2008). In its original status, the Greek E-Gif Ontology contained 37 classes, 131 data type properties, 83 bidirectional object properties (reflecting the relations between the classes). The OWL formalization was used (Web Ontology Language) due to its wide acceptance in representing ontologies on the web. The open source ontology editor Protégé was used for its creation. For inconsistency checking, the trial version of the Description Logic Reasoner RacerPro was used (RacerPro Reasoner). Each layer of the ontology corresponds to a