Exploiting Linked Statistical Data in Public Administration: The Case of the Greek Ministry of Administrative Reconstruction

Emergent Research Forum (ERF) Paper

Eleni Chaniotaki
Athens University of Economics and Business, Greece
echaniotaki@aueb.gr

Evangelos Kalampokis
University of Macedonia and CERTH/ITI, Greece
ekal@uom.edu.gr

Efthimios Tambouris
University of Macedonia and CERTH/ITI, Greece
tambouris@uom.gr

Konstantinos Tarabanis
University of Macedonia and CERTH/ITI, Greece
kat@uom.gr

Antonios Stasis
Ministry of Administrative Reconstruction, Greece
a.stasis@ydmed.gov.gr

Abstract

Opening Government Data is a political priority in many countries. A large part of open data is numerical and concerns statistics thus can be easily processed and visualized. At the technical level, statistical data can be represented using linked open technologies. The resulting Linked Open Statistical Data (LOSD) enables developing data-driven services that require automatic data processing. However, in many cases public agencies confront organizational, structural and cultural challenges when adopting novel technologies. The aim of this paper is to investigate the challenges faced by a ministry that is traditionally organized when exploiting LOSD. The preliminary results suggest that major challenges include the absence of a favorable culture and skills mainly related to data governance and software development. Therefore, essential ingredients seem to be missing to create a fertile ground for the use and exploitation of LOSD technology.

Keywords

Linked Open Statistical Data, Open Government Data, Public Administration, Traditional Bureaucratic Model, Data-Driven Services

Introduction

Opening up government data for others to reuse is a political priority in many countries around the globe. *Open Government Data (OGD)* promise, among others, modernizing Public Administration and improving decision-making processes (Burdon, 2009). However, the potential of OGD has been unrealized to a large extend. This is explained by a number of barriers that hamper the implementation of sophisticated solutions in Public Administration. The relevant literature suggests that OGD barriers can be categorized at the institutional level, the task complexity of handling data, the use of open data and participation in the open-data process, legislation, information quality, as well as at the technical level (Janssen et al., 2012).

Recent research suggests that a promising path to overcome current OGD barriers is to focus on numerical data and, more specifically, statistics (Kalampokis et al., 2016a). Open statistics constitute a large part of OGD according to policy documents. In addition, their added value is related to the fact that they are numerical and highly structured hence they can be easily processed and visualized. From a technological perspective, linked open data technologies can be employed to represent statistical data. Using *Linked Open Statistical Data (LOSD)* enables the integration of data residing in disparate data stores by exploiting semantic technologies, such as the W3C RDF Data Cube Vocabulary[1]. As a result, LOSD enable realizing data-driven scenarios that require automatic data processing (Kalampokis et al., 2016b).

Although LOSD has been successfully exploited to support research problems in a number of areas (Koho et al. 2014; McCusker et al. 2013), it has not yet found much application in Public Administration. Public Organizations are characterized by many deficiencies that hamper their effective functioning (Avgerou 2007). In addition, a number of them fall into the category of *organized anarchies* (Zhu and Kindarto 2016), where the decision making process is described by the *garbage can* model, according to which organizations operate under severe ambiguity as they are characterized by problematic preferences, unclear technology and fluid participation (Cohen et al. 1972).

The aim of this paper is to investigate the challenges faced by a traditionally-organized ministry when exploiting LOSD. To achieve this aim, we developed and deployed LOSD services at the Greek Ministry of Administrative Reconstruction (MAREG) regarding Government Vehicles (GV) management and monitoring. More specifically, we initially studied the existing setting in the ministry in terms of Information Systems, available datasets, as well as cultural and organizational issues and we, thereafter, implemented and deployed LOSD services. This provided us with some interesting findings on the challenges encountered when attempting to implement a sophisticated technology to an organization operating in a traditional and rather outdated mode.

This research is part of a 36-month, in-depth qualitative field study of MAREG's involvement in the Greek pilot of OpenGovIntelligence project. This emergent research forum paper was written immediately after the completion of the first year of the project and presents preliminary results.

This paper is structured as follows: The opening section provides the basic theoretical background of Linked Open Statistical Data (LOSD) and it is followed by a section presenting the research methodology. The following section gives insight to the preliminary findings of this research including context description, LOSD solution description and obstacles for its implementation. The paper closes with a short section with conclusions from the current research.

Methodology

The Research field site

The current research was undertaken in the context of the Greek Ministry of Administrative Reconstruction (MAREG), which runs the Greek pilot of the OpenGovIntelligence project. Access to the site was obtained through the first and last authors of the current paper.

Data collection methods

Qualitative data about the organization were mainly collected through participant observation, which was conducted during project implementation and was facilitated by diary keeping of important facts and events related to GV marking that time period. Given the fact that the first author also belonged to MAREG's staff, introspection or auto-observation was also undertaken as a method, whereby the researcher's own experiences and attitude changes while sharing the field become part of the data (Adler and Adler 1998).

---

[1] https://www.w3.org/TR/2012/WD-vocab-data-cube-20120405/

During this observation phase, the researcher had the opportunity to interact with informants in their milieu, study first-hand the every-day behavior of subjects in particular situations and question them about their feelings and interpretations (Taylor and Bogdan 1984). The research identity adopted was the 'participant-as-observer', where researchers participate in activities in the research setting without revealing their intention to conduct research (Burgess 1984). As the primary means of data collection was participant observation, the researcher was engaged in note-keeping, a process that included descriptions of people, events and conversations, as well as the observer's actions, feelings and hunches or working hypotheses (Taylor and Bogdan 1984). For data triangulation purposes (Denzin 1978), additional sources of data were used, such as official documents on GV depicting the standard operating processes, as well as discussions with GV officials. The purpose of the selected methodology was to gain a thorough view of all aspects related to GV management and monitoring, namely the culture shared, the technology used, the data possessed, as well as the methods and processes followed by officials to conduct their job-related tasks. The ultimate goal of this exploration was to attempt to identify the burdens placed by the context and the current situation in general to effective internal decision making in the domain of GV.

Preliminary Findings

MAREG context and challenges

The Ministry of Administrative Reconstruction (MAREG) became an independent Public Agency for the first time in 2011 as a result of the First Economic Adjustment Program for Greece, having as its main mission to consolidate the Greek public sector through various Administrative Reconstruction projects. Since then, it has changed its status four times, alternating between a Ministry and a General Secretariat, as a result of the high government turnover rate, the political insstability predominant in Greece during the past years and the tendency of different political parties governing to annul actions taken by previous governments and even utterly change the entire map of Greek Public Administration.

At an organizational level, MAREG follows Weber's traditional bureaucratic model, which was established on the principles of fixed and official jurisdictional areas, office hierarchy, document-centric office management, thorough and expert training, full working capacity of officials and strict and exhausting rules for office management (Gerth and Mills 1970). This model is now characterized as outdated and receives criticism as being rigid, narrowly focusing on structure, hierarchy and process rather than being outcome-oriented (Hughes 2012).

At a cultural level, MAREG displays most of the attributes of Greek Public Organizations: direct political control of public administration with top managers appointed politically, *ad hoc* recruitment and placement procedures non-consistent with general human resource management principles and techniques, inadequate staff training, non-defined and formal processes, as well as formalistic rather than effectiveness-oriented functioning (Avgerou 2007).

GV department context and challenges

MAREG's mandate includes the management and monitoring of Government Vehicles (GV) used by all Public Agencies with the exception of ambulances, police patrol cars and fire trucks. However, GV processes are not supported by any Information System and, thus, are performed manually. Counter to this, the GV department follows faithfully the traditional bureaucratic model, based upon devoutly preserved written documents, and *ad-hoc*, rather than scientific decision making. As such, GV officers confine themselves in administering written requests on various related issues from all Public Agencies, which, due to their limited access to information, are usually approved. Poor management of GV caused by the absence of structured and well-defined data is manifested through a number of inefficiencies, such as lack of information regarding important measures (e.g. count of GVs, number of vehicles per region, etc.), inability to match offer and demand, reduced control of operational costs, limitations in policy making and limited transparency towards the public.

GV data and challenges

The GV Registry is merely maintained in a spreadsheet. The spreadsheet contains a large number of entries (more than 30,000) thus transforming its data to linked open data could be only done automatically. Automatic transformation of a spreadsheet to linked open data is easy provided the entries are structured and homogeneous. Unfortunately, the GV spreadsheet had predefined columns which however did not contain predefined values, although this would be possible in most cases, e.g. the terms "Unleaded Catalytic" and "Catalytic Unleaded" were both used. Thus, an interesting finding was that MAREG had not acknowledged before the need to properly collect and maintain GV data. This can be attributed to MAREG officials' limited appreciation of the potential and value of data along with limited knowledge of proper data governance. This may be due to the fact that the majority of MAREG's employees share certain common attributes, such as unilateral educational background in social sciences, lack of job experience prior to joining the public service, rigid and traditional way of thinking and attachment to the letter of law.

To overcome this problem, the research team investigated alternative data sources. Consequently, it was found that the Ministry of Transport maintained basic descriptive data about all vehicles in operation in the Greek territory, including GVs. This, in itself, was an interesting finding as these data are not directly related to that ministry's policy objectives.

The LOSD solution and challenges

The implementation of the LOSD solution included two main steps: (a) the creation of the LOSD store with GV data, and (b) the development and deployment of data-driven services to support decision-making related to GV management and monitoring.

Linked Statistical Data Store

This step included the cleansing of available raw data. We used OpenRefine to homogenize the data and thus improve the quality. The cleansed data were then transformed to RDF following the W3C RDF Data Cube Vocabulary using Grafter[2]. The transformation process relied on best practices for publishing LOSD including re-use of existing dimensions, code lists, etc. The creation of code lists for Greek Public Agencies including connections between agencies (e.g. agency X supervises agency Y) was also an essential step. Finally, the transformed data were imported to an RDF data store in order to be made publicly available.

The final result of this step was an RDF data cube that provided information about the number of GV per registration year, public authority, fuel type, vehicle type, brand, and geographical area. These data were modelled as LOSD and they can be easily combined with other data. The challenges faced at this stage were related to the expected technological obstacles from using the underlying technology. All technical work has been performed by a team of experts including Ministry's employees and researchers from CERTH, another project partner.

Linked Statistical Data-driven Services

The traditional development of LOSD services includes directly accessing LOSD stores, and therefore requires excellent skills in semantic Web technologies, such as RDF and SPARQL. These skills however are absent in many public agencies, including MAREG. To overcome this obstacle, the OpenGovIntelligence project implemented an Application Programming Interface (API) that hides the complexity of accessing LOSD. Using this API, LOSD services can be implemented using mainstream Web technologies (e.g. JavaScript and HTML).

Therefore, MAREG LOSD services were implemented without major technical problems and without the need for sophisticated technical skills. These services enable advanced explorations of the underlying

---

[2] http://grafter.org/

data, such as typical OLAP operations (e.g. roll-up, drill-down) and dynamic visualizations (e.g. bar chart, pie chart, area chart) that change according to user's selections.

The services developed provide valuable insights and can support relevant decision-making processes. For example, they can empower users to answer questions, such as "How old is the vehicle fleet?", "Which fuel type is mostly used?", "Can the maintenance cost be reduced by renewing the vehicle fleet?" etc. At the time this paper was written, LOSD services had been only used by project members. In the next project phases, the services will be used by domain experts and decision makers.

An interesting finding here was associated with the skills of the officials to whom the LOSD Solutions would eventually be delivered. It became evident to the research team that the employees of the Government Vehicles department did not possess the needed technical skills necessary for future use of the project tools, which constitutes a risk for the sustainability and maintenance of the tools developed.

Conclusions

The current study attempts to shed light to the challenges encountered during the deployment of an advanced technology (Linked Open Statistical Data) at a traditional Greek Public Organization. Preliminary findings indicate the existence of major challenges, such as the absence of a favorable culture and the absence of skills mainly related to data governance and software development, which would create a fertile ground for the use and exploitation of the technology. Given that this is an ongoing research, further contacts will be made between MAREG employees and the research team and more services will be co-created. Future work includes investigating whether the obstacles and challenges encountered during the first year of the project will be surmounted and whether the values and benefits of Linked Open Statistical Data as perceived by users will finally bend their technological resistance and enable cultural change in the organization.

Acknowledgements

REFERENCES

Adler, P.A. and Adler, P. 1998. "Observational techniques", in N. Denzin and Y. Lincoln (eds), *Collecting and Interpreting Qualitative Materials*, London: Sage.

Avgerou, C., and McGrath, K. 2007. "Power, rationality, and the art of living through socio-technical change". *MIS Quarterly*, pp. 295-315.

Burdon, M. 2009. "Commercializing public sector information privacy and security concerns", Technology and Society Magazine, (28:1) pp.34–40, IEEE.

Burgess, R. 1984. In the Field: An Introduction to Field Research, London: George Allen & Unwin.

Cohen, M. D., March, J. G., and Olsen, J. P. 1972. "A garbage can model of organizational choice". *Administrative science quarterly*, pp.1-25.

Denzin, N.K. 1978. *The Research Act*, second edition, New York: McGraw-Hill.

Gerth, H. H. and Mills, C. Wright (eds) 1970. *From Max Weber: Essays in Sociology*. London: Routledge & Kegan Paul.

Hughes, O. E. 2012. *Public management and administration*. Palgrave Macmillan.

Janssen, M., Charalabidis, Y., and Zuiderwijk, A. 2012. "Benefits, Adoption Barriers and Myths of Open Data and Open Government" *Information Systems Management* (29:4), pp. 258-268.

Kalampokis, E. Tambouris, E., Karamanou, A., and Tarabanis, K. 2016a. "Open Statistics: The Rise of a new Era for Open Data?" EGOV2016, LNCS 9820, pp.31-43, Springer.

Kalampokis, E. Tambouris, E., and Tarabanis, K. 2016b. "Linked Open Cube Analytics Systems: Potential and Challenges" *IEEE Intelligent Systems*, (31:5), pp.89-92.

Koho, M., Hyvonen, E., and Lehikoinen, A. 2014. "Ornithology Based on Linking Bird Observations with Weather Data," Proceedings of the 4th Workshop on Semantic Publishing CEUR Vol. 1155

McCusker, J. P. et al., 2013. "Towards Next Generation Health Data Exploration: A Data Cube-Based Investigation into Population Statistics for Tobacco," 46th Hawaii International Conference on System Sciences, Wailea, HI, USA, 2013, pp. 2725-2732.

Taylor, S.J. and Bogdan, R. 1984. Introduction to Qualitative Research Methods: The Search for Meanings, second edition, New York: Wiley.

Zhu, Y. Q., & Kindarto, A. 2016. A garbage can model of government IT project failures in developing countries: The effects of leadership, decision structure and team competence. *Government Information Quarterly*.