

# Analysing the Characteristics of Open Government Data Sources in Greece

Charalampos Alexopoulos<sup>1</sup> · Euripidis Loukis<sup>1</sup> · Spiros Mouzakitis<sup>2</sup> · Michalis Petychakis<sup>2</sup> · Yannis Charalabidis<sup>1</sup>

Received: 11 March 2014 / Accepted: 17 September 2015 / Published online: 30 September 2015 © Springer Science+Business Media New York 2015

Abstract Open government data (OGD) initiatives have proliferated over the last years in many countries all over the world as the result of a long-standing movement towards the 'open government' paradigm. These initiatives have been launched and maintained by a variety of government organisations with different strategies and technical capacities and under different social, political and legal conditions. As a result, the OGD sources (defined as various types of portals enabling access to government datasets by the public through the Internet and providing various capabilities/functionalities in this direction) developed through these initiatives demonstrate a great diversity in both content, functionality and technology. However, limited research has been conducted on these OGD sources for understanding better their main characteristics from various perspectives and identifying their strengths and weaknesses. This paper contributes to filling this research gap, by presenting an analysis of the thematic, functional, semantic and technological characteristics of OGD sources in Greece. Sixty OGD sources have been analysed from these perspectives, and statistical analysis of relevant characteristics have been performed. Interesting conclusions have been drawn from this analysis, and based on them, recommendations have been formulated for government policy makers,

Charalampos Alexopoulos alexop@aegean.gr

> Euripidis Loukis eloukis@aegean.gr

Spiros Mouzakitis smouzakitis@epu.ntua.gr

Michalis Petychakis mpetyx@epu.ntua.gr

Yannis Charalabidis yannisx@aegean.gr

- <sup>1</sup> University of the Aegean, Samos, Greece
- <sup>2</sup> National Technical University of Athens, Athens, Greece

in order to enhance OGD provision in Greece and increase the social and economic values that can be generated from them.

**Keywords** Open government data · Public sector information · Linked open data · Semantic web · Transparency

#### Introduction

There has been a long standing movement towards the 'Open Government' paradigm, aiming at increasing interactions and improving relations between government agencies and citizens as well as increasing the trust of the latter towards the former; the main components of this open government paradigm according to the extensively debated and highly influential 'Open Government Directive' of USA (Obama, 2009; Executive Office of the President, 2009; McDermott, 2010; Lathrop & Ruma 2010; Bertot et al., 2014) are transparency, participation of citizens and collaboration with them. The opening of various categories of data possessed by the government (e.g. government spending, economic and business, legal, social, environmental, agricultural and tourism data) to the citizens constitutes a central element of the first component and at the same time a critical precondition for the achievement of the other two. For this reason, Open Government Data (OGD) initiatives have proliferated over the last years worldwide, not only in developed countries but also in developing ones as well (Huijboom & Van Den Broek, 2011; Harrison et al., 2012a; European Commission, 2013). According to the Open Knowledge Foundation, open data are defined as data that are freely accessible online, available without technical restrictions to re-use and provided under open access license that allows the data to be re-used without limitation (OKF 2012). Geiger and v Luecke (2012) define OGD as 'all stored data of the public sector which could be made accessible by government in a public interest without any restrictions for usage and distribution'. A variety of government organisations with different strategies and technical capacities and which are under different social, political and legal conditions are embracing this concept and are opening big numbers of datasets they own to the society, aiming at the achievement of both social and economic benefits, the former being associated with government transparency and accountability as well as citizens' participation, and the latter with development of economic activity in new value added e-services based on the combination of various types of OGD (and possibly private data as well) (Jetzek et al. 2013a and 2013b).

Though the initial motivation for opening government datasets was to promote government transparency and accountability, subsequently there has been an extensive interest in the potential of public sector information (PSI) as an engine for business innovation and job creation, by opening appropriate parts of it as OGD to the society and creating complex ecosystems of public and private actors around them, in order to exploit them and generate social and economic values from them (Harrison et al., 2012b; Chan, 2013). According to the Digital Agenda Assembly (2011), 'the direct annual turnover in the PSI Sector is estimated to be around EUR 28 Billion, with annual growth of around 8 %, and this makes the PSI Sector one of the faster growing sectors of the EU economy', while it is estimated that 'the aggregated macroeconomic footprint of

723

the PSI industry is EUR 140 Billion. This means that the 'spill' or overflow from the PSI industry is even bigger than the sector itself'. A recent study by the McKinsey Global Institute showed that more than 40 countries have developed OGD platforms and over one million datasets have been published worldwide, which is expected to lead to the creation of actual value of about three trillion US dollars in seven industries (education, transportation, consumer products, electricity, oil and gas, health care and consumer finance) (Manyika et al. 2013). For all the above reasons, there is widespread and growing debate on OGD among government practitioners and academics, resulting in a continuously growing research literature (e.g. Allan, 2009; Meijer & Thaens, 2009; Robinson et al., 2009; Lathrop & Ruma, 2010; Parycek & Sachs, 2010; Maier & Huber, 2011; Janssen et al., 2012; Kassen, 2013; Zuiderwijk et al., 2014; Mellouli et al., 2014) as well as in the establishment of opening government datasets as a 'political orthodoxy' in numerous countries worldwide (e.g. in the USA (Obama, 2010), in the UK (Cameron, 2010), in Australia (AGIMO, 2010) and across Europe (European Commission, 2013 and 2014)).

However, despite the big investments that have been made for the development of 'OGD sources', defined as various types of portals enabling access to government datasets by the public through the Internet and providing various capabilities/functionalities in this direction (see "Background Literature Review" for more details) by a variety of government organisations with different strategies and technical capacities and under different social, political and legal conditions worldwide, limited research has been conducted on these OGD sources in order to understand better their main characteristics from various perspectives (a brief review of the limited previous research literature on them is provided in the following 'Previous Related Work') and identify their strengths and weaknesses. This paper contributes to filling this research gap, by presenting an analysis of the characteristics of OGD sources in Greece. These OGD sources constitute a new type of information systems (IS), so according to previous relevant research on IS Success (DeLone & McLean, 1992 and 2003; Urbach & Mueller, 2012), their success relies critically on three main characteristics of them: their 'information quality' (i.e. the quality of the information they provide), their 'system quality' (i.e. their being from a technological perspective) and their 'service quality' (i.e. the support provided to its users, such as training, helpdesk, etc.). Our study focuses on the first two of them. In particular, this first analysis of OGD sources in Greece focuses on two important dimensions of their 'information quality': the thematic width of the datasets they provide and their semantic capabilities; it also focuses on two important dimensions of their 'system quality': the functionality they offer and their technological characteristics. Therefore, the research objective of this study is to analyse the characteristics of OGD sources in Greece from the thematic, functional, semantic and technological perspectives, in order to identify strengths and weaknesses and formulate recommendations for improvements. It should be noted that our study is conducted in a national context different from the ones of the economically and technologically developed countries, having a long tradition in the introduction and exploitation of new technologies and innovations, where most similar previous studies of the main characteristics

of the utilisation of various new ICT have been conducted (see 'The Greek National Context' for more details on this). This is quite interesting, as previous research has revealed the importance of the national context on government information and knowledge sharing (Gharawi and Dawes, 2010; Dawes et al., 2011)

This paper is structured in six sections. In the following second section, a review of the background literature of this study is presented, while in the third section, our research methodology is described. In the fourth section, the results of our analysis are presented, and in the next fifth section, there is a discussion of them. Finally, in the sixth section, the conclusions are summarized and future research directions are proposed.

#### **Background Literature Review**

#### **Open Government Data**

The potential of a significant part of PSI to be re-used beyond the public sector for various commercial and non-commercial social purposes has been recognised more than a decade ago. In 2003, the European Union (EU) adopted the 'Directive on the Reuse of Public Sector Information'<sup>1</sup> (European Commission, 2003), which encourages the member states to make as much of the information they possess available for re-use as possible. It establishes a minimum set of rules as well as practical means for facilitating this re-use, focusing mainly on its economic aspects. An on-line public consultation on this PSI Directive was launched in September 2010, leading to a revision of this Directive<sup>1</sup> (European Commission, 2013), and also signalling its inclusion as one of the key actions of the 'Digital Agenda for Europe'.<sup>2</sup> The new Directive highlights the importance of PSI as a vast, diverse and valuable pool of resources that can benefit the knowledge economy and encourages the proliferation of OGD portals (Zijlstra and Janssen 2013). It includes policies for 'encouraging the wide availability and re-use of PSI for private or commercial purposes, with minimal or no legal, technical or financial constraints, and promoting the circulation of information not only for economic operators but also for the public, which can play an important role in kick-starting the development of new services based on novel ways to combine and make use of such information, stimulate economic growth and promote social engagement'. Also, in July 2014, the European Commission published guidelines in order to help the member states transpose the revised rules and to propose best practices concerning several important aspects for PSI re-use (European Commission, 2014).

As mentioned in the 'Introduction', highly influential for opening PSI to the society has also been the 'Open Government Directive' of USA (Obama, 2009; Executive Office of the President, 2009; McDermott, 2010; Lathrop & Ruma 2010; Bertot et al., 2014), which states that 'to increase accountability, promote informed participation by the public, and create economic opportunity, each agency shall take prompt steps to expand access to information by making it available online in open formats', adding also that 'with respect to information,

<sup>&</sup>lt;sup>1</sup> http://ec.europa.eu/information\_society/policy/psi/rules/eu/index\_en.htm

<sup>&</sup>lt;sup>2</sup> http://ec.europa.eu/information\_society/digital-agenda/index\_en.htm

the presumption shall be in favour of openness (to the extent permitted by law and subject to valid privacy, confidentiality, security, or other restrictions)'. This was followed by the launch of the 'Open Government Partnership'<sup>3</sup> on 2011, when the eight founding countries (Brazil, Indonesia, Mexico, Norway, Philippines, South Africa, UK and USA) endorsed an Open Government declaration, announced their countries' action plans and welcomed the commitment of 38 more countries to join the partnership. A study of the OGD strategies of five countries (Australia, Denmark, Spain, the UK and the USA) conducted by TNO (an independent Dutch public law organisation aiming to enable and support business and government to apply and exploit various types of new knowledge-see https://www.tno.nl/en/about-tno/) concludes that 'in an increasing number of Western countries, "open data" is being placed on the political and administrative agenda', and that 'the focus of strategies is currently on fostering innovation and strengthening democratic participation, whereas some evidence indicates that open data could also contribute to enhancing law enforcement' (Huijboom & Van Den Broek, 2011). Another study concerning 'Open Data in Developing Countries' (Schwegmann, 2012) conducted under the auspices of the European Public Sector Information Platform concludes that 'Open Data seems to be high on the agenda not only in Western countries, but also in developing countries', as civil society organisations and external partners (including the above-mentioned 'Open Government Partnership') of developing country governments are encouraging the use of open data, aiming mainly to increase transparency, accountability and citizen participation. Recently, the new 'Open Data Policy' Memorandum has been issued in the USA (Executive Office of the President, 2013), stating that 'Information is a valuable national resource and a strategic asset to the Federal Government, its partners, and the public', and directing all executive departments and agencies to assign high priority to the management of the information artefacts they generate and maintain as important assets throughout their life cycle, having promotion of openness and interoperability as the main objectives; it also stresses that this will increase operational efficiencies, reduce costs, improve services, support mission needs and increase public access to valuable government information.

Jetzek et al. (2013a and 2013b) argue that there are four types of value that can be generated from the OGD, which differ in the sector generating the value (public or private) and in the kind of generated value (economic or social): (i) transparency-related value (public sector organisations generate social value by offering increased transparency into government actions, which reduces 'information asymmetry' between government officials and citizens and therefore misuse of public power for private benefits and corruption), (ii) efficiencyrelated value (public sector organizations generate economic value through OGD by increasing internal efficiency and effectiveness), (iii) participationrelated value (private sector firms generate social value through participating and collaborating with government), (iv) innovation-related value (private sector firms generate economic value through the creation of new products and services). It is gradually realised that OGD creates big opportunities for both <u>private and public sec</u>tor innovation (see Editorial of Special Issue on

<sup>&</sup>lt;sup>3</sup> http://www.opengovpartnership.org/

Innovation through Open Data of the Journal of Theoretical and Applied Electronic Commerce Research by Zuiderwijk et al. (2014)), being quite an importance for the development of information and knowledge economy. However, a study conducted by Janssen et al. (2012), based on data collected through interviews and workshops, warns that 'a conceptually simplistic view is often adopted with regard to open data, which automatically correlates the publicizing of data with use and benefits'; in particular, they identify five 'myths' that have been gradually developed with respect to OGD: 'The Publicizing of Data will Automatically Yield Benefits', 'All Information Should Be Unrestrictedly Publicized', 'It Is a Matter of Simply Publishing Public Data', 'Every Constituent Can Make Use of Open Data' and 'Open Data Will Result in Open Government'. They finally conclude that the success of the developed OGD infrastructures requires more than the simple provision of access to data: it is necessary to make progress towards the improvement of the quality of government information, the creation and institutionalization of a culture of open government and the provision of the tools and instruments with which to use the data. The realisation of the 'Open Government' paradigm in general seems to be a difficult and complex task, which requires combined efforts of multiple actors not only from the public sector but also from the private sector, as well as a gradual development of 'open government ecosystems' (Harrison et al., 2012b).

### **OGD** Thematic Categories

According to the above-mentioned EU 'Directive on the Re-use of Public Sector Information' (2003) and the relevant report of the OECD Working Party of Information Economy (2006), a wide variety of PSI thematic categories are opened by government agencies and offered as OGD through the Internet, and the most important of them are:

- Economic and Business Information, including financial, public spending, economic activity and statistics, industry and trade data as well as official business registers and public tender databases.
- Geographic Information, including cartographic, cadastral spatial (geographical coordinates), administrative and political boundaries, topographical, public buildings (geo-coordinates) and elevation data.
- Legal Information, including crime/conviction data, laws, rights and duties, legislation and treaties, judicial and patent and trademark information.
- Meteorological and Environmental Information, including oceanographic, hydrographical, environmental (quality), atmospheric and meteorological (weather) data.
- Social Information, including demographic, attitude, health, education and labor data.
- Traffic and Transport Information, including traffic congestion, work on roads, public transport, vehicle registration, transport networks and transport statistics.
- Tourist and Leisure Information, including tourism, entertainment and hotelrelated data.

- Agricultural, Farming, Forestry and Fisheries Information, including cropping/land use, farm income/use of resources, fish farming/harvest and live stock data.
- Natural Resources Information, including biological, ecological, geological and geophysical data, as well as, energy resource/consumption and geodetic networks information.

### **OGD Sources Capabilities/Functionalities**

Research carried out as part of the European project ENGAGE (see www.engageproject.eu, http://www.engagedata.eu/ about/) has revealed that two distinct types of OGD sources/portals have been developed with respect to the capabilities/ functionalities provided to the user (Petychakis et al., 2014):

- 1. OGD direct provision portals: this is the main category of OGD portals, which are 'primary sources' of OGD, publishing original government datasets provided by either one government agency or a small number of similar government agencies (who are the legal owners/licensers of the data); they usually offer a wide range of functionalities supporting the whole lifecycle of OGD, from the creation of datasets to the update and finally to the archiving of them.
- 2. OGD aggregators: this category includes OGD aggregator portals, which are 'secondary sources' of OGD coming from a big number of government agencies, publishing and maintaining lists of other 'primary' OGD catalogues and links to them. They constitute single access points to multiple OGD direct provision portals and make it easier for a user to locate the OGD he/she is interested in. Usually, they include descriptive information about datasets and sources, which is quite useful for the users in order to get a first impression of what is available. Many of them act as highly structured registries of OGD primary sources and datasets, which store structured and machine-processable information, and provide 'index'-like features, such as automated registration and discovery of OGD. Some prominent examples are the widespread CKAN data hub portal,<sup>4</sup> the Open Government Data Initiative<sup>5</sup> (OGDI—a Microsoft initiative/online tool to publish and use a wide variety of public data from government agencies), the European Union Open Data portal<sup>6</sup> and the Freebase.<sup>7</sup>

Another categorisation of the OGD sources with respect to the capabilities/ functionalities offered can be made based on the web paradigm they are based on (the 'traditional' Web 1.0 paradigm or the more recent Web 2.0 paradigm) (Alexopoulos et al., 2013; Alexopoulos et al., 2014):

• The 'traditional' first-generation OGD portals, which have been influenced by the Web 1.0 paradigm, in which there is a clear distinction between content producers

<sup>&</sup>lt;sup>4</sup> CKAN The Data hub, http://thedatahub.org/

<sup>&</sup>lt;sup>5</sup> Open Government Data Initiative, http://ogdisdk.cloudapp.net/DataCatalog/DataSetList

<sup>&</sup>lt;sup>6</sup> European Union Open Data portal, https://open-data.europa.eu/en/data/

<sup>&</sup>lt;sup>7</sup> Freebase http://www.freebase.com

and content users. They are characterised by datasets publishing in non-machineprocessable formats (i.e. PDF), without providing any contextual information or linkage capabilities to other datasets; also, they are limited to offering basic functionalities to data users (consumers) for dataset downloading and to data providers for uploading datasets. They do not support improvements of their published datasets by their users (e.g. through cleaning and further processing) or feedback provision by dataset users to their providers so that the latter can understand better the needs of the former.

• The second generation Web 2.0 OGD portals: The advent of the Web 2.0 paradigm, which facilitates the generation of content of various types by simple and non-expert users, the development of relationships and online communities among them and the extensive interaction, collaboration and sharing of content and information, has led to the emergence of a second generation of OGD portals, which have been influenced by these Web 2.0 principles. They provide (in addition to the above-mentioned basic functionalities of the traditional first generation OGD portals) functionalities for commenting and rating datasets, forming groups around common interests, visualising and processing datasets, improving or adapting them to specialised needs and then publishing them again, uploading new datasets and enabling OGD users to become data 'pro-sumers' (both consuming and producing datasets). Their main objective is to support and facilitate extensive communication between OGD users (citizens, journalists, businesses, scientists, etc.) and providers (government agencies) as well as collaborative value generation from OGD.

#### Semantic Technologies and OGD

The use of Semantic Web technologies (such as 'Uniform Resource Identifiers' (URI), the 'Resource Description Framework' (RDF), vocabularies and ontologies) in OGD enables a more effective browsing and discovery of datasets through distributed SPARQL queries as well as linking and combining OGD from multiple sources across the Web, which can increase significantly the usefulness of the OGD and the value generated from them (e.g. it allows discovering new correlations and gaining deeper insights or developing new advanced value-added e-services by combining different datasets from multiple OGD sources); also, the value of any kind of data (including OGD) increases each time it is being re-used and linked to another resource, and this can be facilitated and triggered by providing informative and explanatory data about each available dataset (i.e. metadata). These ideas gave rise to the development of the Linked Open Government Data (LOGD) concept (an application/elaboration of the more general 'Linked Data' concept for the OGD) (Wood, 2011; Geiger and v Luecke 2012; Bauer and Kaltenböck, 2012; Zuiderwijk et al., 2012a; Zuiderwijk et al., 2012b). Geiger and v Luecke (2012) define LOGD as 'all stored data of the public sector connected by the World Wide Web which could be made accessible in a public interest without any restrictions for usage and distribution' and argue that 'the cross linking of Open Data via the Internet and the World Wide Web as 'Linked Open Data' (LOD) offers the possibility of using data across domains or organisational borders for statistics, analysis, maps and

publications', which can lead to the generation of more insight, knowledge and innovation from OGD. The W3C Government Linked Data Working Group<sup>8</sup> is developing standards and best practices in order to assist government agencies to open and publish their data as effective and usable Linked Data using Semantic Web technologies (for more details on methodologies for this see Wood (2011) and Bauer and Kaltenböck (2012)).

#### **Previous Related Work**

However, the degree of realisation of the above concepts and methods in 'real life' in various contexts has been investigated to a very limited extent. Limited research has been conducted on the existing OGD sources that have been developed by a variety of government organisations in order to understand better their main characteristics from various perspectives and identify strengths and weaknesses. Zuiderwijk et al. (2013) investigate and compare the functionality offered by three OGD portals: the 'Open Data Hub' of the European Union, the 'Junar' commercial open data platform and the OGD infrastructure developed in the European research project ENGAGE (www. engage-project.eu and www.engagedata.eu). They conclude that all the examined OGD portals have basic functionalities for uploading and managing data as well as some additional specific features for supporting the specific goals and targets groups of each of them (e.g. two of these OGD portals offer functionalities for the analysis and visualisation of OGD that users retrieve from them). Petychakis et al. (2014) analyse the OGD sources developed in the European Union (in its 27 member states) from a functional, semantic and technical perspective, in terms of their thematic content, licencing, multilinguality, data acquisition, data discovery, data provision and data formats. They conclude that most of the datasets of the European OGD sources are published without a clearly defined or open licence and that the thematic categories with the highest relative frequency are 'Law Enforcement, Courts and Prisons', 'Budget, Revenues and Expenditures' and 'Business, Economics and Trade'; about half of these OGD portals in their user interface support the native language of the corresponding country, while the other half are multilingual (they support one or more foreign languages as well). With respect to data acquisition, most of the published datasets originate from the internal back office IS of government agencies that support daily operations, while the most frequently provided means for datasets search and discovery by prospective users are groupings in categories and free text search. These OGD datasets are provided to interested users mainly through download of various types/formats of files, the most frequent of them being non-machine-processable ones (PDF and HTML, 38 and 28 % respectively), while there is limited dataset provision in machine-processable formats (mainly Excel XLS/XLSX,8 %). Our study makes a contribution to this limited body of research and knowledge on the characteristics of existing OGD sources by presenting an analysis of OGD sources in Greece (see 'The Greek National Context') using a wider range of

<sup>&</sup>lt;sup>8</sup> http://www.w3.org/2011/gld/charter.html

perspectives and indicators (see following 'Research Method' section) than that of the previous studies.

### The Greek National Context

As mentioned in the 'Introduction', this study has been conducted in a national context that differs from the ones of previous studies on the utilisation of various new ICT. Greece has a lower level of economic and technologic development than the developed Western countries have and is now experiencing a severe economic crisis and an increasing citizens' distrust in government. According to the World Bank (http://data.worldbank.org/indicator/ NY.GDP. PCAP.CD) the gross domestic product (GDP) per capita (a primary indicator of a country's economic development) of Greece in 2013 was US\$21,956 (undergoing a strong decrease in the last 4 years due to a severe economic crisis: in 2010, at the beginning of the Greek economic crisis, its GDP per capita was US\$26,861), much lower than those of the developed western countries (e.g. in USA, it was as high as US\$53,042, in UK it was US\$41,787 and in Germany US\$46,269). Also, the e-Government Development Index (an indicator of technological development of government) value for Greece, according to the 'United Nations E-Government Survey 2014' (United Nations, 2014), was at the level of 0.7118, lower than those for the developed western countries (e.g. for USA it was 0.8748, for UK it was 0.8695 and for Germany 0.7864). Furthermore, Greece experiences a growing citizens' distrust in government (which has grown further recently due to the existing economic crisis); according to the 'Transparency International' (http://www.transparency. org/country#GRC) in Greece, 'The public sector suffers from substantial integrity gaps in both law and practice' since 'some public officials have acted without transparency or effective oversight for decades', so 'citizens' distrust in public service has proliferated'. Therefore, the development of OGD sources in Greece can be quite beneficial in terms of addressing the above problems and contributing to government transparency and accountability improvement, so it is interesting and useful to analyse the current situation with respect to OGD sources from various perspectives.

## **Research Method**

The research method we adopted for conducting a multi-perspective analysis of the Greek OGD Sources landscape consists of four distinct stages: (i) Identification of OGD Sources, (ii) Formulation of Analysis Framework (= definition of the analysis perspectives and the particular indicators of each of them), (iii) Data Collection and (iv) Data Processing. These four research stages are described in the following paragraphs. Our study has been based on quantitative techniques, as, according to the relevant literature (Maylor and Blackmon, 2005; Ragin and Amoroso, 2011), they are the recommended approach for collecting data from a large quantity of units (in our case the whole population of OGD sources in Greece—see following description of stage 1) concerning a limited number of features of them (in our case the selected thematic,

functional, semantic and technological characteristics of these OGD sources—see following description of stage 2), and then condensing/summarising these data into a few numbers (e.g. relative frequencies).

#### **Stage 1: Identification of OGD Sources**

The first stage of our research method refers to the identification of the OGD Sources in Greece. For this purpose, we adopted a top-down approach based on desk-based research, starting from Greek Ministries websites, in which we searched for published OGD datasets, and then following links to other government organizations (e.g. supervised by the particular Ministry or Regional Administrations, Municipalities, etc.), in which we performed similar searches for finding OGD datasets and links to other government organizations, etc., in an effort to identify websites and portals containing and providing OGD. As a result of this stage, we identified 60 OGD sources, which are shown in the Appendix (for each of them, we can see the corresponding government agency, a description, the thematic categories of the OGD it provides and finally its URL).

#### **Stage 2: Formulation of Analysis Framework**

The second stage of our research method is the construction of the analysis framework. As mentioned in the 'Introduction', OGD sources constitute a new type of IS, so for analysing them we can use as our theoretical foundation the models that have been developed in previous research on IS Success (DeLone & McLean, 1992 and 2003; Urbach & Mueller, 2012). According to them, the success of an IS relies critically on three main characteristics of it: its 'information quality' (i.e. the quality of the information it provides to its users), its 'system quality' (i.e. its quality viewed as a technological system) and its 'service quality' (i.e. the support provided to its users, such as training, helpdesk, etc.); these affect 'users satisfaction' and the 'actual use' of the IS and finally its 'individual impact' and 'organisational impact'. The above model has been used extensively for studying various types of IS (with appropriate elaboration and adaptation to the characteristics and objectives of each), including government IS providing various electronic services to citizens (e.g. Wang & Liao, 2008; Rana et al., 2013; Chena et al., 2015). Our analysis of Greek OGD sources has focused on the first two of the above critical characteristics. the information quality and the system quality, elaborating and adapting them to the particular characteristics and objectives of the OGD sources. In particular, we examine two important dimensions of the Greek OGD sources' 'information quality': (i) the thematic range of the datasets they provide (since as mentioned in the previous 'Background Literature Review' section, there is a wide variety of PSI thematic categories opened by government agencies and offered as OGD through the Internet, and this is an important quality and value dimension of OGD: more OGD thematic categories means more potential usefulness and value for more user groups, e.g. citizens, journalists, businesses, scientists, etc.); (ii) their semantic capabilities (since as mentioned in the previous section the use of Semantic Web technologies in OGD can significantly increase their

usefulness and value). We also examine two important dimensions of the Greek OGD sources' 'system quality': (iii) the functionality they offer (i.e. to what extent are offered on one hand the basic and on the other hand the advanced features mentioned in the previous section, which is quite important for the users); (iv) their technological characteristics (as they are critical for the performance of the OGD sources, for their adaptability to new needs and for the efficient provision of high quality functionality to the users). Therefore, our analysis framework consists of the above four analysis perspectives (thematic, functional, semantic and technological), which are described in more detail in the following paragraphs.

- 1. Thematic Analysis Perspective: It includes analysis of the thematic categories of the datasets provided by the OGD sources. It has been conducted using the nine main thematic categories of OGD identified by the EU 'Directive on the Re-use of Public Sector Information' (2003) and the relevant report of the OECD Working Party of Information Economy (2006), which have been mentioned in 'OGD Thematic Categories' section. Therefore, it has been based on nine binary indicators, each of them assessing whether or not an OGD source provides datasets of one of these thematic categories, and then the calculation of their frequencies.
- Functional Analysis Perspective: It includes analysis of the functionalities 2. provided by the OGD sources. It has been conducted using five indicators: three of them concern fundamental functionalities of the OGD sources and have been previously used in the above-mentioned analysis of the European OGD sources by Petychakis et al. (2014): they concern dataset discovery, dataset provision and also multi-linguality-related capabilities. The other two indicators correspond to two important features of the new and more advanced generation of OGD sources mentioned in 'OGD Sources Capabilities/Functionalities' section: visualisation and user feedback-related capabilities. Therefore, the corresponding five indicators of the functional analysis of the OGD sources (it should be noted that their possible values have been defined based on the study of Petychakis et al. (2014) as well as the analysis of a representative subset of the identified Greek OGD sources, from which the main possible values have been determined, and then these have been enriched with new values we found during the process of analysing the identified OGD sources (in stage 3)):
  - Datasets Discovery: it concerns the tools provided for discovering the datasets the user is interested in; its main possible values (not mutually exclusive) were simple document list, free text search, browsing through categories, browsing through filters, browsing through interactive map and SPARQL search.
  - Data Provision: it concerns the ways of data provision to the users; its main possible values (not mutually exclusive) were download file, online view of dataset and on map.
  - Language: it concerns the language(s) the user interface is available in; its main possible values are (not mutually exclusive) Greek, English, French and German.

- Visualisations: it concerns the dataset visualisation capabilities provided; one possible value is 'not existing', while other main possible values (not mutually exclusive) are visualisations in charts and visualisations in maps.
- Feedback: it concerns the existing tools allowing feedback from OGD users to the providers; its two main possible values were 'not existing' and 'existing' (the latter has been further refined during the process of analysing the identified OGD sources in stage 3).
- 3. Semantic Analysis Perspective: It includes analysis of the use of Semantic Web technologies for the representation and structure of OGD (see 'Semantic Technologies and OGD'). In particular, it includes initially an overall assessment of the OGD sources using the well-established five-star Berners-Lee's rating system for open data (Bauer and Kaltenböck, 2012) and then an analysis of the metadata (which are critical for linking OGD and obtaining higher levels of value from them as mentioned in the previous section) and of licence information. Therefore, the corresponding four indicators of the semantic analysis of the OGD sources are as follows:
  - Sources rating according to the five-star Berners-Lee's Rating Scheme for open data<sup>9</sup>:

\*Make your stuff available on the web (whatever format) \*\*Make it available as structured data (e.g. excel instead of ima

\*\*Make it available as structured data (e.g. excel instead of image scan of a table)

\*\*\*Using non-proprietary format (e.g. csv instead of excel)

\*\*\*\*Use URLs to identify things, so that people can point at your stuff \*\*\*\*\*Link your data to other people's data to provide context

• Sources metadata rating according to the five-star Maturity Scheme of Metadata Management (ISA, 2011).

\*Metadata Ignorance \*\*Scattered or Closed Metadata \*\*\*Open Metadata for Humans \*\*\*\*Open Reusable Metadata

- \*\*\*\*\*Linked Open Metadata
- RDF compliance: it concerns the use of technologies that support RDF, including technical products of open data initiatives publishing structured data in a way that it can be interlinked, which, as mentioned in the previous 'Background Literature Review' section, is quite important for enabling more effective browsing and discovery of datasets and for linking and combining OGD from multiple sources (e.g. see Wood (2011); Bauer and Kaltenböck (2012)); it is a binary indicator.
- Data License: it concerns license information related to the use of the published datasets; this is one of the most important characteristic of OGD sources, since it defines the allowed ways of OGD use and

<sup>&</sup>lt;sup>9</sup> http://lab.linkeddata.deri.ie/2010/star-scheme-by-example/

exploitation for generating various types of social and economic value and reduces all relevant legal uncertainties and risks (e.g. see Wood (2011); Bauer and Kaltenböck (2012)); its two main possible values were 'not existing' and 'existing' (the latter has been further refined during the process of analysing the identified OGD sources).

- 4. Technological Analysis Perspective: It includes analysis of the technologies and products that have been used for the development of the OGD source at the main technological layers: web server, content management system (CMS) or platform, user interface, data format and API. The five indicators of the technological analysis are (their possible values have been formulated during the process of analysing the identified OGD sources):
  - Web Server: it concerns the web server that the OGD site is hosted on
  - CMS/Platform: it concerns the CMS or Platform that the OGD source has been based on.
  - User Interface (UI): it concerns the technologies used for the presentation layer of the OGD source.
  - Data Format: available data representation formats of the published information
  - API: available application programming interfaces (API) and web service interfaces.

### **Stage 3: Data Collection**

This third stage includes the collection of all the necessary data for the analysis defined in stage 2 from the OGD sources identified in stage 1. For the first three analysis perspectives (thematic, functional and semantic), a desk-based research approach was used, which included manual examination of all OGD sources and assessment of all indicators. In order to ensure the validity of the results, two different persons (being also among the co-authors of this paper) made these assessments independently and then compared their results; in cases of differences (there were differences in less than 5 % of the results), they were discussed by the whole authors' group, which made the final decisions. For the fourth analysis perspective (technological), we used the software tool Wapallyser (https://wappalyzer.com/) in order to collect from all examined Greek OGD sources all the required information for assessing its five indicators. The collection of all these data from the OGD sources as well as the access to all websites mentioned in the previous 'Background Literature Review' section and in the References section has taken place in January 2014.

### Stage 4: Data Processing

In the final stage, statistical analysis of the data collected in stage 3 was conducted, which included calculation of various descriptive statistics, such as frequencies and relative frequencies of all values for each of the above-mentioned indicators, and

construction of various charts, using the Excel software. These statistics were discussed

extensively among the whole authors' group, and conclusions were drawn from them, as well as recommendations for enhancing OGD provision in Greece.

## Results

### Thematic Analysis Perspective

In Fig. 1, we can see for each of the nine main thematic categories of OGD identified by the EU 'Directive on the Re-use of Public Sector Information' (2003) and the relevant report of the OECD Working Party of Information Economy (2006) its frequency, i.e. how many of the identified 60 OGD sources in Greece provide such data. Due to the fact that an OGD source can provide OGD from more than one of these thematic categories, the sum of their frequencies exceeds the number of the OGD sources (60). We remark that the thematic category with the highest frequency (36 (60 %)-having a significant difference from the second one) is the economic and financial one, concerning mainly public spending data for various government agencies as well as data about economic activity and firms. This is strongly associated with two important characteristics of the Greek national context mentioned in the 'Introduction': the growing citizens' distrust in government (so many government agencies respond by publishing data on their spending) and the existing severe economic crisis (which necessitates an increase in economic activity, so it is useful to provide data on existing economic activity/firms, which allow a better understanding of it, and support a better design and planning of its increase). It should be noted that this finding differs from the corresponding one of the analysis of the European Union member states' OGD sources conducted by Petychakis et al. (2014), which found that the thematic category with the highest relative frequency is 'Law Enforcement, Courts and Prisons' (probably reflecting the increasing criminality and security concerns in many EU countries). We also remark that there are also four thematic categories (social, natural resources, legal and geographic information) with much lower frequencies (found in 6 to 11 OGD sources), while the remaining four thematic categories (traffic/transport, meteorological/environmental, agricultural/farming/ forestry/fisheries and tourism/leisure) have quite low frequencies, despite their importance (e.g. the importance of agriculture and tourism for the Greek economy). Therefore, it is concluded that the thematic range of OGD sources is rather narrow, focusing mainly on the provision of economic/financial data. Our analysis also indicates that 2 out of the 60 identified OGD sources are OGD aggregators: the ones of the Initiative of the Ministry of Environment and Climate Change<sup>10</sup> and the geodata.gov.gr, both concerning geospatial data; all the others are OGD direct provision portals.

<sup>&</sup>lt;sup>10</sup> http://www.inspire.okxe.gr/index.php?option=com\_content&view=article&id=30&Itemid=43



Fig. 1 Thematic analysis of OGD sources in Greece (frequencies of the nine main thematic categories)

#### **Functional Analysis Perspective**

In Figs. 2, 3, 4, 5 and 6, we can see the results of the analysis of the Greek OGD sources from a functional perspective. We can see from Fig. 2 that the main tool for dataset discovery is the simple document list, which is provided by all the examined OGD sources, followed by free text search, provided by a much smaller number of them (16 (27 %)); the extensive use of dataset categorisation in the EU countries OGD sources according to the above-mentioned study by Petychakis et al. (2014) does not appear in the Greek OGD sources. In general, it is concluded that there is a lack of advanced datasets' discovery tools. With respect to data provision, Fig. 3 indicates that nearly all examined OGD sources (59) offer the capability of downloading the datasets as files and only few of them to view data on maps or in web forms. With respect to multilinguality, Fig. 4 indicates that all OGD sources have designed their user interfaces (UI) in Greek, while half of them offer also English UIs. The range of visualisation capabilities offered is significantly limited, as we can see in Fig. 5: 54 (90 %) of the examined OGD sources do not offer such capabilities, and only 6 (10 %) offer some visualisation capabilities, such as 'View Geospatial data on map' 'View data on Charts', 'Predefined charts and statistics', 'Dynamically create charts' and 'Modifying the type and the scale of the graphs'. In the latter group of OGD sources, more than one visualisation method can be provided. Although in the last years visualisation engines are widely used and have become more comprehensive, flexible and light-weighted, there is limited use of them by the Greek OGD sources for the graphical representation



**Fig. 3** Data provision (frequencies of main tools)



of geospatial data on a map or on charts providing statistical analysis. Lastly, from Fig. 6, we can see that there is limited interest in receiving feedback from the users in a systematic manner, e.g. concerning their needs for new datasets or improvements in the provided datasets. About half of these OGD sources (28 (47 %)) provide no user feedback mechanism, while the other half (32 (53 %)) provide simple 'Contact us forms' for comments and suggestions from users. The purpose of these feedback forms is to receive comments mainly on technical problems issues, rather than on the actual datasets. Only one OGD source provides such advanced feedback capabilities for rating datasets.

### Semantic Analysis

In Figs. 7, 8, 9 and 10, we can see the results of the analysis of the Greek OGD sources from a semantic perspective. From Fig. 7, we can see that most of the examined OGD sources (41 of them (68 %)) are at the lowest level (one star) of the five-star Berners-Lee's rating scheme (see 'Research Method'), while a smaller number (9 (15 %)) are at the second level (two stars) and only a few at higher levels. This indicates a lower level of maturity of the Greek OGD sources from an open data viewpoint. Focusing now on metadata, from Fig. 8, we can see that the majority of examined OGD sources (46 of them (77 %)) are at the initial level of 'metadata ignorance'; a much smaller number have reached higher maturity levels (7 (12 %)) are at the fourth level of 'open re-usable metadata', while only 2 (3 %) are at the highest level of 'linked open metadata'). By examining the minority of the OGD sources that provide metadata for their datasets, we





found that most of them have custom metadata schemas, and only 2 (3 %) adopt an existing metadata standard (both for describing geospatial information, based on the INSPIRE Directive). Also, from Fig. 9, we can see that the majority of the examined OGD sources (46 (77 %)) do not support RDF and SPARQL, which, as mentioned in the previous section, are quite important technologies for enabling more effective browsing and discovery of datasets and for linking and combining OGD from multiple sources. Lastly, Fig. 10 indicates that more than half of them (37 (62 %)) do not provide license information, while those that have no common policy on this (their licenses for use and reuse of data vary significantly, with the 'free personal use provided suitable acknowledgement of the source and copyright owner is given' being the dominant licence model).

### **Technological Analysis**

Finally, in Figs. 11, 12, 13, 14 and 15, we can see the results of the analysis of the Greek OGD sources from a technological perspective. It indicates (Fig. 11) that there is a strong preference for using custom platforms and content management systems (CMS) for the development of these OGD sources (used in 22 of them (37 %)), which



**Fig. 7** Open data level (frequencies of levels)



have been developed internally by the corresponding government agencies, followed by the Joomla (used in 19 sources (32 %)) and the Dreamweaver (used in 5 sources (8 %)) CMS, both mature and widely used products. This is associated with the specific characteristics of the Greek national context: due to the existing severe economic crisis in Greece, government agencies have to reduce drastically their budgets, so they tend to resort to internal developments of these OGD sources (by their own IS personnel); this results in savings of financial resources on one hand, but also in low exploitation of existing more advanced and mature technologies on the other hand. Most of these OGD sources (39 (65 %)) are hosted on the Apache web server, followed by the Internet Information Server (14 (23 %)) (Fig. 12), which are both mature and widely used ones. With respect to the format of the published datasets, the majority of OGD sources publish their data in non-machine-processable formats, such as PDF and HTML (Fig. 13), and only a small number of them publish datasets in machine-processable formats, such as Excel/XLS (11 sources (18 %)), XML (9 sources (15 %)) and CSV (4 sources (7 %)). This constitutes a significant weakness, as it reduces considerably the usefulness of the published datasets and the value that can be generated from them; however, it should be noted that the same weakness has been identified, as mentioned in 'Background Literature Review' section, by the study of the OGD sources of the European Union member states conducted by Petychakis et al. (2014). The most frequently used technologies for the development of the user interfaces in the examined





OGD sources are PHP (used in 38 of them (63 %)) and jQuery (used in 35 (58 %)), also mature and widely used (Fig. 14). Lastly, only a few of these OGD sources (5 (8 %)) offer APIs for data and metadata interactions (Google Font or Restful APIs), though this would significantly enhance the usefulness and value of the published datasets (e.g. allowing the development of value added e-services based on them).

### Discussion

From the analysis presented in the previous sections, it is concluded that Greece has made some substantial first steps towards opening a number of interesting and useful datasets possessed by government agencies of various levels (e.g. Ministries, Regional Administrations, Municipalities, etc.). However, our study indicates that there are some important improvements that should be made, in order to enhance OGD provision in Greece and increase the social and economic value that can be generated from them. Some recommendations in this direction are provided in this section.

More government datasets should be opened, from a wider range of thematic categories (since the thematic range of the existing OGD sources is rather narrow, focused on the provision of primarily economic/financial datasets and secondarily social, natural resources and legal datasets). It will definitely be important to open more economic/financial datasets concerning government spending (leading to higher government transparency and accountability, which will be quite important due to the growing citizens' distrust in government and the existing economic crisis that necessitates the rationalisation of government expenses) as well as on economic activity and firms (the existing economic recession makes it more necessary than ever to understand better economic activity, identify sectors and regions in which he had more severe economic activity, and based on the conclusions design policies and measures for expansion). However, at the same time, more emphasis should be placed on opening datasets of some highly important thematic categories that have been neglected, such as agriculture and tourism-related datasets (as both of them are quite important economic sectors for the Greek economy, with a strong potential for further expansion and development) and environment-related datasets (which is quite important for the development of our tourism industry, and also for citizens' quality of life).

At the same time, the functionality of the existing OGD sources should be enhanced, providing more advanced tools mainly for data discovery (so that potential users can find more easily and quickly the datasets they are interested in), data visualisation (for instance on maps and charts, so that potential users can easily and quickly get a first understanding about the datasets and decide whether it is worth continuing with a more detailed analysis of them) and user feedback (so that OGD users can provide feedback to their providers, about the quality of the datasets they have used, existing weaknesses and necessary improvements and needs for additional datasets—as the collaboration between OGD users and providers has been recognised as critical for the generation of value from them (e.g. Zuiderwijk et al., 2012c; Harrison et al., 2012b)). Also, it Fig. 9 RDF compliance (frequencies)



is quite important to place more emphasis on the multi-linguality of Greek OGD sources (with at least one language beyond the Greek), so that they can be used and exploited not only by Greek actors (e.g. analysts, scientists, e-services development firms, journalists, etc.) but also by foreign ones as well (as Greece is a member of the European Union, the exploitation of these OGD sources beyond Greece across this wider geographic region can be highly beneficial).

Furthermore, since there is a low maturity of Greek OGD sources from a linked open data viewpoint (as concluded in 'Semantic Analysis' section, it is important to make a progress in this direction and expand the use of linked open data technologies (moving towards more stars in the above-mentioned Berners-Lee's rating scheme). More emphasis should be placed on the use of structured and machine-processable file formats in publishing datasets and metadata (adopting existing metadata standards) as well as on the support of RDF and SPARQL, which will enable more effective browsing and discovery of datasets as well linking and combining OGD from multiple sources, leading to a big increase of their usefulness and value for various groups. Another recommendation is to develop rich API in the OGD sources, which will enable the automated use of the capabilities as part of various user developed programs, which will promote more advanced exploitation of datasets, as well as



30

40

0

10

20





the development of value added e-services based on them. Finally, it is necessary to place more emphasis on the provision of license information in OGD sources and the development of a common policy on this across the Greek public sector (having the fewest possible limitations and restrictions on datasets use), aligned with the recent relevant documents and guidelines of EU (European Commission 2013 and 2014); this is going to clarify the allowed ways of OGD use and exploitation for generating various types of social and economic values and therefore reduce all relevant legal uncertainties and risks, leading finally to a wider and more intensive and innovative use of the published datasets. In general, it is necessary the Greek to make progress towards the 'second generation OGD portal' paradigm (outlined in 'Background Literature Review' section).

The ability to create a socially and commercially added value from OGD (e.g. new insights on important social problems and useful applications, products and services) will depend not just on the government's willingness to open data (which exists in Greece) but also on how this opening and provision will be executed in order to provide higher capacity to businesses, journalists, civil society and individuals to effectively use and re-use these data. However, it should be recognised that a possible barrier to this proposed improvement of OGD provision in Greece will be the reduced budgets of government agencies, due to the existing severe economic crisis; so government agencies might decide to use their limited ICT budgets to other kinds of ICT projects perceived as being of higher visibility and priority, such as the enhancement of their websites, or e-transaction services to the citizens and firms. Government agencies traditionally consider their websites and e-transaction services to be of higher priority than are technical infrastructures that open up their data to the society. However, it is widely agreed that opening government data to citizens





Fig. 13 Data formats (frequencies)



and firms can lead to significant social and economic benefits, since 'the better use of your data will be pointed by someone else'. Already there are some encouraging results of these first initiatives of the Greek government agencies to open their data: a good example is the 'public spending' initiative, which is a free, open and objective web application aiming to provide data and visualisations of them along with research results concerning the Greek public expenditure (offering advanced capabilities for opening, downloading, interlinking and visualising data about expenditures of Greek government

**Fig. 14** User interfaces (frequencies)







agencies), based on the ideas of the UK's 'Where does my money go?', and contributing to increased government transparency.

### Conclusions

It is gradually realised that the extensive data collected by government agencies as part of their operations and for the support of public policy making can be of critical importance for the development of the emerging information and knowledge-based economy as well as for progressing towards the 'open government' paradigm (characterised by transparency, participation of citizens and collaboration with them). For this reason, many government agencies of various levels (central government, regional government, municipalities, etc.) all over the world have undertaken big open data initiatives and made big investments for the development and operation of various types of OGD sources/portals enabling access to numerous government datasets through the Internet. However, limited research has been conducted on these OGD sources in order to understand better their main characteristics, identify strengths and weaknesses and formulate recommendations for their improvement. Our paper aims to contribute to filling this research gap. It presents an analysis of the thematic, functional, semantic and technological characteristics of OGD sources that have been developed in Greece, in a national context characterised by lower levels of economic and technological development compared to that in the highly developed countries, and also experiencing a severe economic crisis and a growing citizens' distrust in government. Sixty OGD sources have been identified and analysed from the above four perspectives, leading to interesting conclusions and recommendations for enhancing OGD provision in Greece.

Our analysis indicates that in Greece, substantial first steps have been made in this area, despite the limited financial resources available to the government agencies in the last 4 years due to the existing economic crisis. However, both the thematic content of the OGD sources (focused mainly on economic/financial datasets) and their technology/functionality (there is limited exploitation of the most advanced technologies in this area and provision of the most advanced relevant functionality) has been influenced strongly by the above-mentioned special characteristics of the Greek national context (as explained in more detail in 'Results' section). In particular, the thematic range of OGD sources is rather narrow, focused mainly on the provision of economic/financial data, while there is limited provision of other thematic categories of data which are significant for the Greek economy and society, such as agricultural, tourism and environmental data. With respect to the functionality provided, there is a lack of advanced dataset discovery, visualisation and user feedback capabilities as well as limited multi-linguality of the OGD sources (only half of them offer a second language beyond Greek, mainly English). Also, there is limited use of linked open data technologies, despite the important capabilities they provide for more effective browsing and discovery of datasets and for linking and combining OGD from multiple sources, which can lead to new and deeper insights and to innovative and valuable e-services. There is limited provision of metadata as well as support of RDF and SPAROL, while most of the examined OGD sources do not publish the datasets in machine-processable file formats and do not provide license information (creating important uncertainties for the users) and APIs (limiting the capabilities for development of value added eservices from the published datasets).

Our study has interesting implications for both research and practice. With respect to research, it contributes to the limited existing empirical research on the characteristics of the OGD sources that have been developed so far, studying a special national context that is different from the ones of the highly developed countries (where most studies of the characteristics of the exploitation of new technologies are conducted); also, for this purpose, we developed a multi-perspective research framework based on previous IS success research, which can be used (possibly with extensions and adaptations) for future research on OGD sources. With respect to practice, our conclusions and recommendations can be useful for organising and enhancing OGD provision in Greece (with respect to content selection for opening, and OGD portals functionality development), and also in other similar countries, and also at the individual government agency level as well. Further research is required concerning the characteristics of existing and under-development OGD sources in various national contexts (in countries of various levels of economic and technological developments), from more perspectives, using more indicators; also, it is important to proceed beyond the manual examination and assessment of the OGD sources, and collect additionally assessment information from their users and providers using both quantitative and qualitative techniques. Furthermore, it is necessary to conduct research not only for understanding the characteristics of the OGD sources from various perspectives but also for understanding better their determinants (i.e. the actors and factors shaping them).

Acknowledgments This paper is related with the ENGAGE FP7 Infrastructure Project (An Infrastructure for Open, Linked Governmental Data Provision towards Research Communities and Citizens). The main goal of the ENGAGE project is the deployment and use of an advanced service infrastructure, incorporating distributed and diverse public sector information resources as well as data curation, semantic annotation and visualisation tools, capable of supporting scientific collaboration and governance-related research from multi-disciplinary scientific communities, while also empowering the deployment of open government data towards citizens. More information about this project can be found at www.engage-project.eu and www.engagedata.eu.

No.	Government agency	Description	Thematic categories	URL
-	Transparency Program	Laws/regulations/ministry decisions and public bodies spending	Economic and Business Information; Legal Information	http://diavgeia.gov.gr http://greek-lod.math.auth.gr/diavgeia/
7	General Secretariat of Information Systems	Taxation statistics for Greece	Economic and Business Information	http://www.gsis.gr
ŝ	Geodata project	A project funded by Greek government in order to host geospatial public cross- sector information	Geographic Information	http://geodata.gov.gr
4	Hellenic Statistical Authority	The official source of statistical information for almost every activity of Greek economy	Social Information; Economic and Business Information	http://www.statistics.gr
5	Ministry of Finance	Country's economic data	Economic and Business Information; Legal Information	http://www.minfin.gr
9	Ministry of Citizen Protection—Greek Fire Department	Provide information on fires that have been recorded in Greece	Legal Information; Geographic Information	http://www.fireservice.gr http://greek-lod.math.auth.gr/fire-brigade/
2	Ministry of Labour and Social Insurance	Sharing legal information on employment rights and decisions by the ministry	Legal Information	http://www.ypakp.gr
×	Ministry of Education, Lifelong Learning and Religion	Provide information of the location of all schools in Greece	Geographic Information	http://www.minedu.gov.gr/
6	Ministry of Citizen Protection-Greek Police Department	Sharing information on traffic accidents and criminality	Legal Information; Geographic Information	http://www.astynomia.gr http://greek-lod.math.auth.gr/police/
10	Ministry of Interior, Decentralisation and e-Government	Census about public sector employees	Social Information; Economic and Business Information	http://www.ydmed.gov.gr/ http://greek-lod.math.auth.gr/kalikratis/

Appendix

No.	Government agency	Description	Thematic categories	URL
11	Ministry of Infrastructures andTransportations	Detail invoices concerning public works	Economic and Business Information; Legal Information; Traffic and Transport Information	http://www.yme.gr/
12	Ministry of Health	Provide aggregated information about surgeries and tests conducted in hospitals	Social Information	http://yyka.gov.gr
13	Hellenic Ministry of Rural Development and Food	Subsidy provision per prefecture and financial affairs	Economic and Business Information	http://www.minagric.gr
14	Ministry of Environmental Protection and Climate Change	Information about protected areas, such as Mountain Protection Zones, National Parks and Forests Borderline, Controlled Hunting Regions, etc.	Meteorological and Environmental Information; Geographic Information; Natural Resources Information	http://www.ypeka.gr
15	Observatory for Digital Greece	Key point of reference for accurate and up-to-date information on Information Society indicators	Economic and Business Information; Social Information	http://www.observatory.gr/
16	National Observatory for SMEs	Various types of data on SMEs	Economic and Business Information; Social Information	http://pforumgr.commex.gr
17	General Secretariat for Sports	Census of and subsidy provision for athletic unions	Social Information	http://www.gss.gov.gr/
18	Greek LOD Cloud	First attempt for provision of open-linked data (not only governmental) (Police, Fire-Brigade, Diavgeia (Transparency programme), Kalikratis)	Social Information; Geographic Information; Economic and Business Information; Legal Information	http://greek-lod.math.auth.gr/
19	Athens Urban Transport Organisation	Public transport routes and stops	Traffic and Transport Information; Economic and Business Information	http://www.oasa.gr
20	Archaeological Cadastre	Quantitative and geospatial data on archaeological sites	Geographic Information	http://archaeocadastre.culture.gr/el/data
21	Ministry of Aegean	Information on various topics concerning Aegean islands	Natural Resources Information; Agricultural, Farming, Forestry and Fisheries	http://www.ypai.gr/

Table 1 (continued)

Tabl	e 1 (continued)			
No.	Government agency	Description	Thematic categories	URL
			Information; Tourist and Leisure Information	
22	Ministry of Development	Information about public works	Economic and Business Information	http://www.ypoian.gr/
23	General Secretariat for Development	Information about public works	Economic and Business Information	http://www.gsrt.gr/
24	Ministry of National Defence	Information about the Greek army	Economic and Business Information	http://www.mod.gr/
25	Ministry of Citizens' Protection	Information about Greek Police, Fire Service, Coast Police etc.	Legal Information	http://www.minocp.gov.gr/
26	Ministry of Justice, Transparency and Human Rights	Information about citizens' rights, laws, transparency, justice	Legal Information	http://www.ministryofjustice.gr
27	Ministry of Education and Religious Affairs	Information about educational system, public schools, public universities, educational laws	Legal Information	http://www.ypepth.gr/
28	Ministry of Maritime Affairs, islands and fisheries	Information about Greek islands, fisheries and sea affairs	Agricultural, Farming, Forestry and Fisheries Information	http://www.yen.gr
29	Ministry of foreign Affairs	Information on foreign individuals, information about other countries	Legal Information	http://www.mfa.gr/
30	Ministry of interior	Information about national situations	Economic and Business Information	http://www.ypes.gr/
31	Ministry of Macedonia and Thrace	Information about various topics concerning North Greece	Economic and Business Information	http://www.mathra.gr/
32	Ministry of culture and Tourism	Information about Greek culture and tourism	Economic and Business Information; Traffic and Transport Information; Tourist and Leisure Information	http://www.culture.gr
33	Ministry of the Press and Media	Information about Media	Economic and Business Information	http://www.minpress.gr/
34	Athens' Prefecture	Information about various topics concerning Athens	Economic and Business Information	http://nom-athinas.gov.gr
35	West Macedonia's Region	Information about various topics concerning West Macedonia	Economic and Business Information	http://www.grevena.gov.gr/

No.	Government agency	Description	Thematic categories	URL
36	East Macedonia's and Thrace's Region	Information about various topics concerning in East Macedonia and Thrace	Economic and Business Information	http://www.pamth.gov.gr/
37	South Aegean's Region	Information about various topics concerning South Aegean	Economic and Business Information	http://www.notioaigaio.gr/per/
38	Epirus' Region	Information about various topics concerning Epirus	Economic and Business Information	http://www.epirus.gov.gr
39	Thessaly's Region	Information about various topics concerning Thessaly	Economic and Business Information	http://www.thessalia.gov.gr/
6	Ionian Islands' Region	Information about various topics concerning Ionian Islands	Economic and Business Information	http://www.pin.gov.gr/
41	West Greece Region	Information about various topics concerning Western Greece	Economic and Business Information	http://www.pde.gov.gr/gr/index.php
4	Attica's Region	Information about various topics concerning Attica	Economic and Business Information	http://www.attiki.gov.gr/
43	Peloponnese Region	Information about various topics concerning Peloponnese	Economic and Business Information	http://ppel.gov.gr/
4	North Aegean's Region	Information about various topics concerning North Aegean	Economic and Business Information	http://www.northaegean.gr/
45	Crete's Region	Information about various topics concerning Crete	Economic and Business Information	http://www.crete-region.gr/
46	Central Macedonia's Region	Information about various topics concerning Central Macedonia	Economic and Business Information	http://www.pkm.gov.gr/
47	Municipality of Keratsini and Drapetsona	Information about various topics concerning Keratsini and Drapetsona	Economic and Business Information	http://www.keratsini-drapetsona.gr/
48	Municipality of Athens	Information about various topics concerning Athens	Economic and Business Information	http://www.cityofathens.gr/
<del>1</del> 9	Municipality of Thessaloniki	Information about various topics concerning Thessaloniki	Economic and Business Information	http://www.thessaloniki.gr

Table 1 (continued)

Tab	le 1 (continued)			
No.	Government agency	Description	Thematic categories	URL
50	Municipality of Chania	Information about various topics concerning Chania of Crete	Economic and Business Information	http://www.chania.gr/
51	Municipality of Rethymno	Information about various topics concerning Rethymno of Crete	Economic and Business Information	http://www.rethymno.gr/
52	Municipality of Heraklion	Information about various topics concerning Heraklion of Crete	Economic and Business Information	http://www.heraklion.gr/
53	Municipality of Larissa	Information about various topics concerning Larissa	Economic and Business Information	http://www.larissa-dimos.gr/
54	Municipality of Messini	Information about various topics concerning Messini	Economic and Business Information	http://www.messini.gr
55	Municipality of Kozani	Information about various topics concerning Kozani	Economic and Business Information	http://www.kozanh.gr
56	Municipality of Samos	Information about various topics concerning Samos	Economic and Business Information	http://www.vathi.gr/
57	Municipality of Chios	Information about various topics concerning Chios	Economic and Business Information	http://www.chioscity.gr/
58	Municipality of Mytilene	Information about various topics concerning Mytilene	Economic and Business Information	http://www.mytilene.gr/
59	Municipality of Halkidiki	Information about various topics concerning Halkidiki	Economic and Business Information	http://www.halkidiki.gov.gr
60	Municipality of Drama	Information about various topics concerning Drama	Economic and Business Information	http://www.dimos-dramas.gr

#### References

- AGIMO (2010). Declaration of Open Government, retrieved February 25, 2011 from http://www.finance.gov. au/e-government/strategy-and-governance/gov2/declaration-of-open-government.html
- Alexopoulos, C., Loukis, E., Charalabidis, Y., & Zuiderwijk, A. (2013). An evaluation framework for traditional and advanced open public data e-infrastructures. Como, Italy: Paper presented at the 13th European Conference on eGovernment – ECEG.
- Alexopoulos, C., Zuiderwijk, A., Charalabidis, Y., Loukis, E., Janssen, M. (2014). Designing a second generation of open data platforms: integrating open data and social media. In Electronic Government (pp. 230–241). Dublin, Ireland: Springer Berlin Heidelberg.
- Allan, R. (2009). The power of government information. In J. Gøtze & C. B. Pedersen (Eds.), State of the eUnion: government 2.0 and onwards. AuthorHouse.
- Bauer, F., & Kaltenböck, M. (2012). *Linked open data: the essentials—a quick start guide for decision makers*. Vienna, Austria: Monochrom Editions.
- Bertot, J. C., Gorham, U., Jaeger, P. T., Sarin, L. C., & Choi, H. (2014). Big data, open government and egovernment: issues, policies and recommendations. *Inf Polity*, 19(1–2), 5–16.
- Cameron, D. (2010). Letter to Government departments on opening up data, retrieved from https://www.gov. uk/government/news/letter-to-government-departments-on-opening-up-data
- Chan, C. (2013). From Open Data to Open Innovation Strategies: Creating e-Services Using Open Government Data, 46th Hawaii International Conference on System Sciences.
- Chena, J. V., Jubiladoa, R. J. M., Capistrano, E. P. S., & Yen, D. C. (2015). Factors affecting online tax filing—an application of the IS Success Model and trust theory. *Comput Hum Behav*, 43, 251–262.
- Dawes, S., Gharawi, M., & Burke, B. (2011). Knowledge and information sharing in transnational knowledge networks: a contextual perspective. In *Proceedings of the 44th Hawaii International Conference on System Sciences*.
- DeLone, D. H., & McLean, E. R. (1992). Information systems success: the quest for the dependent variable. Inf Syst Res, 3(1), 60–95.
- DeLone, D. H., & McLean, E. R. (2003). The DeLone and McLean model of information systems success: a ten-year update. J Manag Inf Syst, 19(4), 9–30.
- Digital Agenda Assembly (2011). Report from the workshop "Open Data and re-use of public sector information", Brussels 16–17 June 2011 http://ec.europa.eu/information\_society/events/cf/daa11/itemdisplay.cfm?id=5963.
- European Commission (2003). Directive 2003/98/EC of the European Parliament and of the Council of 17 November 2003 on the Re-use of Public Sector Information, available at: http://eur-lex.europa.eu/ LexUriServ/LexUriServ.do?uri=OJ:L:2003:345:0090:0096:EN:PDF.
- European Commission (2013). Directive 2013/37/EU of the European Parliament and of the Council of 26 June 2013 amending Directive 2003/98/EC on the re-use of public sector information, available at: http:// eur-lex.europa.eu/LexUriServ/LexUriServ.do?
- European Commission (2014). Commission Notice Guidelines on recommended standard licenses, datasets and charging for the reuse of documents of 24 July 2014, 2014/C 240/01, available at: https://ec.europa. eu/digital-agenda/en/news/commission-notice-guidelines-recommended-standard-licences-datasets-andcharging-re-use
- Executive Office of the President (2009). Memorandum for the Heads of Executive Departments and Agencies: Open Government Directive, Washington D. C., Memorandum No. M-10-06. Retrieved from http://www.whitehouse.gov/open/documents/open-government-directive.
- Executive Office of the President (2013). Open Data Policy Managing Information as an Asset. Washington D. C., Memorandum No. M-13-13. Retrieved from http://www.whitehouse.gov/sites/default/files/omb/ memoranda/2013/m-13-13.pdf
- Geiger, C. P., & v Luecke, J. (2012). Open government and (linked) (open) (government) (data). J e-Democracy Open Gov, 4(2), 265–278.
- Gharawi, M., & Dawes, S. (2010). Conceptualizing knowledge and information sharing in transnational knowledge networks. In Proceedings of the 4th International Conference on Theory and Practice of Electronic Governance (ICEGOV '10).
- Harrison, T. M., Guerrero, S., Burke, G. B., Cook, M., Cresswell, A., Helbig, N., Hrdinova, J., & Pardo, T. (2012a). Open government and e-government: democratic challenges from a public value perspective. *Inf Polity*, 17(2), 83–97.
- Harrison, T. M., Pardo, T. A., & Cook, M. (2012b). Creating open government ecosystems: a research and development agenda. *Future Internet*, 4, 900–928.

- Huijboom, N., & Van Den Broek, T. (2011). Open data: an international comparison of strategies. Eur J Epractice, 12, 1–13.
- Janssen, M., Charalabidis, Y., & Zuiderwijk, A. (2012). Benefits, adoption barriers, and myths of open data and open government. *Inf Syst Manag*, 29, 258–268.
- Jetzek, T., Avital, M., & Bjørn-Andersen, N. (2013a). The generative mechanisms of open government data. In Proceedings of ECIS (p. 156). Atlanta, GA: Association for Information Systems. AIS Electronic Library (AISeL), 2013.
- Jetzek, T., Avital, M., & Bjørn-Andersen, N. (2013b). Generating value from open government data. In proceedings of the 34th International Conference on Information Systems. ICIS 2013. Atlanta, GA: Association for Information Systems. AIS Electronic Library (AISeL), 2013.
- Kassen, M. (2013). A promising phenomenon of open data: a case study of the chicago open data Project. Gov Inf O, 30, 508–5.
- Lathrop D., & Ruma L. (2010). Open government: collaboration, transparency, and participation in practice, 1st ed. Sebastopol, CA: O'Reilly Media, Inc.
- Maier, U., & Huber, S. (2011). Open: the changing relation between citizens, public administration, and political authority. *JeDEM*, 3(2), 182–191.
- Manyika J, Chui M, Groves P, Farrell D, Van Kuiken S, and Almasi Doshi E (2013) Open data: Unlocking innovation and performance with liquid information. McKinsey Global Institute, McKinsey Center for Government and McKinsey Business Technology Office. Available at: http://www.mckinsey.com/insights/ business technology/open data unlocking innovation and performance with liquid information.

Maylor, H., & Blackmon, K. (2005). *Researching business and management*. New York: Palgrave-Macmillan. McDermott, P. (2010). Building open government. *Gov Inf Q*, 27(4), 401–413.

- Meijer, A., & Thaens, M. (2009). Public information strategies: making government information available to citizens. Inf Polity, 14(1-2), 31–45.
- Mellouli, S., Luna-Reyes, L. F., & Zhang, J. (2014). Smart government, citizen participation and open data. Inf Polity, 19(1–2), 1–4.
- Obama, B. H. (2009). Transparency and open government. Memorandum for the Heads of Executive Departments and Agencies. Available at: http://www.whitehouse.gov/the-pressoffice/transparencyandopen-government.
- Obama, B. (2010). Memo from President Obama on Transparency and Open Government. In D. Lathrop & L. Ruma (eds.), Open Government. http://www.whitehouse.gov/open/documents/open-governmentdirective. Accessed 18 April 2010.
- Open Knowledge Foundation (OKF) (2012) "Open Data Handbook version 1.0.0" (revised November 14, 2012). Available at http://opendatahandbook.org/pdf/OpenDataHandbook.pdf
- Parycek P, & Sachs M (2010) Open government information flow in web 2.0. European Journal of ePractice, 9(1), 1–70.
- Petychakis, M., Vasileiou, O., Georgis, C., Mouzakitis, S., & Psarras, J. (2014). A state-of-the-art analysis of the current public data landscape from a functional, semantic and technical perspective. J Theor Appl Electron Commer Res, 9(2), 34–47.
- ISA Programme (2011). Interoperability Solutions for European Public Administrations, "Towards Open Government Metadata". https://joinup.ec.europa.eu/sites/default/files/towards\_open\_government\_metadata\_0.pdf.
- Ragin, C., & Amoroso, L. (2011). Constructing social research: the unity and diversity of method (2nd ed.). California: Pine Forge Press – Sage Publications.
- Rana, N., Dwivedi, Y., & Williams, M. (2013). Evaluating the validity of IS success models for e-government research: an empirical test and Integrated Model. *Int J Electron Gov Res*, 9(3), 1–22.
- Robinson, D. G., Yu, H., Zeller, W. P., & Felten, E. W. (2009). Government data and the invisible hand (Fall 2009). Yale J Law Technol, 11, 160. Available at SSRN: http://ssrn.com/abstract=1138083.
- Schwegmann, C. (2012). Open data in developing countries. In European Public Sector Information Platform - Topic Report 2013/02.
- United Nations (2014) United Nations E-Government Survey 2014 E-Government for the Future we Want, Department of Economic and Social Affairs of the United Nations, accessed from http://unpan3.un.org/egovkb/
- Urbach, N., & Mueller, B. (2012). The updated DeLone and McLean Model of information systems success. In Y. Dwivedi, M. Wade, & S. Schneberger (Eds.), *Information systems theory—explaining and predicting our digital society*. New York: Springer.
- Wang, Y. S., & Liao, Y. W. (2008). Assessing eGovernment systems success: a validation of the DeLone and McLean model of information systems success. *Gov Inf Q*, 25, 717–733.
- Wood, D. (Ed.). (2011). Linking Government Data. New York: Springer.
- Zijlstra T, and K Janssen (2013) The new PSI directive as good as it seems? Open Knowledge Foundation Blog. [Online]. Available: http://blog.okfn.org/2013/04/19/the-new-psi-directive-as-good-as-it-seems/

- Zuiderwijk, A., Janssen, M., van den Braak, S., & Charalabidis, Y. (2012a). Linking open data: challenges and solutions. In Proceedings of the 13th Annual International Conference on Digital Government Research (dg.o '12) (pp. 304–305). New York, USA: ACM.
- Zuiderwijk, A., Jeffery, K., & Janssen, M. (2012b). The potential of metadata for linked open data and its value for users and publishers. J e-Democracy Open Gov, 4(2), 222–244.
- Zuiderwijk, A., Janssen, M., Choenni, S., Meijer, R., & Alibaks, R. S. (2012c). Socio-technical impediments of open data. *Electron J e-Government*, 10(2), 156–172.
- Zuiderwijk, A., Janssen, M., & Parnia, A. (2013). The complementarity of open data infrastructures: an analysis of functionalities. In *Proceedings of the 14th Annual International Conference on Digital Government Research (dg.o '13)* (pp. 166–171). Quebec City, Canada: ACM.
- Zuiderwijk, A., Helbig, N., Gil-García, R., & Janssen, M. (2014). Special Issue on Innovation through Open Data - A Review of the State-of-the-Art and an Emerging Research Agenda: Guest Editors' Introduction. Special Issue on Innovation through Open Data - A Review of the State-of-the-Art and an Emerging Research Agenda: Guest Editors' Introduction. J Theor Appl Electron Commer Res, 9(2), 1–13.