A taxonomy of open government data research areas and topics

Yannis Charalabidis¹, Charalampos Alexopoulos¹, Euripidis Loukis¹

¹University of the Aegean
{yannisx, alexop, eloukis}@aegean.gr
30 Voulgaroktonou Str., GR-11472, Athens, Greece

Abstract

The opening of government data, in order to have both social and economic value generated from them, has attracted the attention and interest of both researchers and practitioners from various disciplines, such as information systems, management sciences, political and social sciences and law. Despite the rapid growth of this multidisciplinary research domain, which has led to the emergence and continuous evolution of technologies and management approaches for open government data (OGD), a detailed analysis of the specific areas and topics of this research is still missing. In this paper, a detailed taxonomy of research areas and corresponding research topics of the OGD domain is presented: it includes 4 main research areas (ODG management & policies, infrastructures, interoperability and usage & value), which are further analysed into 35 research topics. An important advantage of this taxonomy, beyond its high level of detail, is that it has been developed through extraction and combination of relevant knowledge from three different kinds of sources: important relevant government policy documents, research literature and experts. For each of these 35 research topics we identified, its research literature has been summarized and main research objectives and directions have been highlighted. Based on the above taxonomy, an extension of the existing in the literature OGD lifecycle has been proposed; also, under-researched topics that require further research have been identified. Our OGD research taxonomy extends and elaborates previous research taxonomies for the ‘ICT-enabled Governance’ and ‘Policy Making 2.0’ domains,
which have been developed in the European projects CROSSROAD and CROSSOVER.

**Keywords:** open government data, research taxonomy, open government data management, open government data technology, open government data interoperability open government data value.

1. Introduction

Open Government Data (OGD) has been attracting a growing attention and interest of both researchers and practitioners from various disciplines, such as information systems, management sciences, political and social sciences and law, due to its widely recognised potential to generate public value through driving innovation and economic growth, and also scientific research, and by promoting transparency and substantial evidence-based political dialogue (Stevens 1984; Conradie et al. 2012; Janssen K. 2011). The concept of open data is strongly associated with innovative capacity and transformative power (Davies et al. 2013). It is increasingly recognized that proactively opening public data can create considerable benefits for several stakeholders, such as firms and individuals interested in the development of value added e-services or mobile applications, by combining various types of OGD, and possibly other private data, or scientists, journalists and active citizens who want to understand better various public problems and policies through advanced data processing and production of analytics (Zuiderwijk et al. 2014; Janssen K. 2011). However, it should be noted that at the same time there are a few articles discussing unintended consequences and negative side effects of opening data (Blakemore et al. 2006; Zuiderwijk et al. 2014b). At the same time, as mentioned in the ‘Call for Papers’ of this Special Issue ‘Yet organizations are struggling to generate value from big and open data’.

Furthermore, OGD, as a new organizational invention gradually diffusing in government is under a continuous renegotiation over its meanings and practices, and therefore a gradual formulation of its ‘organizing vision’ (using the term proposed by Swanson et al. 1997). According to Tammisto et al. (2012) the first level of renegotiation in the context of OGD took place initially in relevant policy discussions, public and professional press, and consultancy. The second level of renegotiation is taking place when organizations gradually understand how to benefit from open
data and drive the development of social and economic value from it. This renegotiation and evolution of this new domain can be greatly assisted by establishing a common code of understanding concerning the main areas and topics of research on OGD. The development of a detailed taxonomy of current research areas and topics in the domain of OGD will address the communication gap in this new domain, and facilitate better interaction among researchers, and also with interested practitioners. Also, it can provide a solid base for future research in this domain, and thus contribute to reaching higher levels of maturity in the practices of opening and exploiting government data, and in the generation of social and economic value from them; in general, it can contribute to the development of a body of knowledge in this area, which will enable improving and optimizing the technology, and also the design, operations and performance of the units of government agencies responsible for opening data. Such a taxonomy is of critical importance for the development of a ‘science base’ (see Charalabidis et al. (2011)) in the OGD domain.

Furthermore, it can be useful for ICT firms (assisting them in developing better OGD technological infrastructures), government agencies (for improving their OGD practices) and firms developing innovative value added e-services or mobile applications based on OGD. However, despite the rapid growth of this multidisciplinary research domain, which has led to the emergence and continuous evolution of technologies and management approaches for open government data (OGD), a detailed analysis of the specific areas and topics of this research is still missing (see section 2 for more details).

This paper contributes to filling the above research gaps. In particular, it makes the following contributions:

i) It develops a detailed taxonomy of research areas and corresponding research topics of the OGD domain has been developed, including 4 main research areas, which are further analysed into 35 research topics.

ii) This taxonomy has been developed through extraction and combination of relevant knowledge from three different kinds of sources: important relevant government policy documents, research literature and experts.

iii) For each of these 35 research topics we identified, its research literature has been summarized and the main research objectives and directions have been highlighted; also, under-researched topics that require further research have been identified.
Based on the above taxonomy an extension of the existing in the literature OGD lifecycle has been proposed (including important additional stages).

Our OGD research taxonomy extends and elaborates previous research taxonomies for the ‘ICT-enabled Governance’ and ‘Policy Making 2.0’ domains, which have been developed in the European projects CROSSROAD and CROSSEOVER.

Finally, directions have been formulated for future multi-disciplinary research based on OGD aiming to address current societal challenges. The research presented in this paper has been conducted within the FP7 ENGAGE project (‘An Infrastructure for Open, Linked Governmental Data Provision towards Research Communities and Citizens’ – see http://www.engage-project.eu/ and http://www.engagedata.eu/about/).

This paper is structured as follows. Section 2 describes the methodology we followed for developing the taxonomy. In section 3 the main findings of literature review we have conducted for this purpose are presented and discussed. Then section 4 presents the taxonomy, including descriptions of the identified main research areas, and the particular research sub-areas/topics for each of them. Finally, a discussion of findings is provided in section 5, while section 6 concludes the paper.

2. Methodology

This study is focused on two main research questions, which constitute a first step towards the creation of a ‘descriptive theory’ of the OGD domain that will enable the development of a science base of it: (a) what are the main research areas and topics of the OGD domain, and (b) how they can be categorized? Gregor (2002) proposes five types of theories that need to be developed in the information systems domain; the first and more fundamental of them, which is necessary for the development of the other four more advanced ones, is the ‘descriptive theories’, which ‘describe or classify specific dimensions or characteristics of individuals, groups, situations, or events’.

There are two categories of descriptive theories: naming theories and classification theories (Stevens, 1984). A naming theory is a description of the main dimensions or characteristics of some phenomenon. A classification theory is more elaborate in that it also includes interrelations between such dimensions or characteristics of given phenomena.

This paper contributes to the development of description theory for the OGD domain, both a naming and classification theory, which are of critical importance for the development of more
advanced types of theories in this domain (e.g. concerning relationships between various dimensions or characteristics of them), and in general for the development of its scientific base. In particular, we developed an OGD research areas taxonomy (OGDRAT), based on relevant government policy documents, previous relevant research literature and also experts’ knowledge. For this purpose we followed the bottom-up approach to taxonomy development proposed by Ramos et al. (2003) and Sujatha et al. (2011), which includes the four stages shown in Figure 1 (our research has focused on the first three of them). The scope of this study focuses on the first three steps of this approach, as it is illustrated with the dashed box in Figure 1.

**Figure 1**: Taxonomy development approach

In particular, the methodology we followed for the development of the OGDRAT was based on content analysis (Krippendorff, 2013) of different kinds of documents (government policy documents, previous relevant research literature and minutes of experts’ workshops). It consisted of the following eight steps (shown also in Figure 2):

1. Initially we identified and analysed important relevant government policy documents concerning OGD, which define the main terms, issues and perspectives, and also the main problems and challenges posed in this domain. The most important of them were: (a) European Commission Directives and Communications (European_Commission, 2010; 2011a; 2011b; 2012; 2013a; 2013b), (b) US Government documents (Executive Office of
2. Then we identified and analysed previous research papers that propose categorisations of research areas and perspectives of the OGD research domain. Additionally, we identified and analysed previous research literature concerning barriers to OGD publishing and exploitation, and also uptake of OGD and value generation from them. A brief review of this literature is presented in the following section 3 (while we refrain from presenting a review of the relevant government policy documents on OGD identified and analysed in step 1, in order to keep the paper in acceptable level). The outcome of this step was another set of OGD related terms (having some overlapping with the ones of the set produced in the previous step), which were used as well for the construction of the first version of OGDRAT in step 3.

3. After realising the above first two steps, the main research topics in the OGD domain were defined, and then were grouped in higher level research areas; this was a first version of the OGDRAT.

4. A thorough literature search was then conducted, based on the E-Government Reference Library (EGRL - faculty.washington.edu/jscholl/egrl/), which is a widely recognized and frequently updated electronic library of peer-reviewed papers in the electronic government/governance domain, using as keywords the terms of the above first version of the OGDRAT. In particular, the EGRL was searched by paper title and abstract for each of these terms, and the most relevant papers were retained and read in detail. This led to the identification of additional research topics in the OGD domain, which were used for the construction of a second version of the OGDRAT.

5. The realisation of the fourth step resulted in the second version of the OGDRAT.

6. A workshop was organised for the discussion, evaluation and validation of the above second version of the OGRAT, aiming at the assessment of its main research topics, and the possible proposition of new ones, and also at the assessment of their grouping, and the possible proposition of changes. In this workshop participated 20 OGD experts from 11 different EU
countries (NL, UK, DE, GR, BE, IT, AU, RO, ES, BG, LV), from different kinds of organizations (public administrations, universities and firms) and different educational levels (Professors, PhD and MSc holders), in order to validate and further elaborate second version of the OGRAT. All of the participants were selected based on their experience in the area of OGD and they are characterised as very experienced in the OGD domain, having been or currently being involved in OGD related projects (national or European).

7. Based on feedback collected from this workshop (which included the proposition of new research topics, such as the topics 2.7 (‘citizen-generated open data’) and 2.8 (‘sensor-generated open data’) described in section 4, and also of changes in their grouping in research areas), the final version of the OGRAT was produced, which is presented in the section 4.

8. Finally we proceeded to further processing and exploitation of it, and the results are presented in section 5.
3. Literature Review

In the step 2 of our methodology (described in the previous section 2) we have identified four previous research papers that propose categorisations of OGD research in areas and themes (Davies et al. 2013; Zuiderwijk et al. 2014a; Lindman et al. 2014; Harrison et al. 2012), which were reviewed as they include elements that can be useful for the development of the OGDRAT.

The study of Davies et al. (2013, p.11) argues that “over its short history as a field of action a number of distinct fronts of research into open data have developed, responding to different practice, policy and knowledge needs. These can be usefully classified into three broad groups: 1) open data readiness assessments, 2) open data implementation studies and 3) impact studies”. Readiness studies aim to assess whether the conditions in public administrations are appropriate for the effective development of open data initiatives. Implementation studies aim to assess whether the conditions for open data itself actually exists in terms of open data availability, extent of publishing government agencies and importance of published datasets. Finally, impact studies aim to assess to what extent open data initiatives have led to change and public value.

The second study by Zuiderwijk et al. (2014a, p.2) identifies seven different perspectives of OGD research, namely, (a) political, (b) social, (c) economical, (d) institutional, (e) operational, (f) legal and (g) technical and argues that “combining perspectives may be more effective in dealing with the issues related to open data and stimulating innovation”. Furthermore, it also identifies a number of OGD research directions, and categorises them under three major topics: i)
open data theory and development, ii) open data policies, use, and innovation, and ii) open data infrastructures and technologies.

Another study conducted by Lindman et al. (2014 p.4) focuses on the research challenges concerning Open Data Services, and categorises the relevant issues based on the work systems framework (Alter 2010). It argues that "there are two basic approaches for organizing the research issues according to the challenges that emerge when data is made available to the public, and further provided as services. These are: 1) an analysis of the life-cycle of the data and 2) an analysis of the levels of inquiry at which the open data phenomenon is studied". The proposed categories for the organisation of open data services research are: 1) Technologies, 2) Information, 3) Processes and Activities, 4) Products and Services, 5) Participants, 6) Customers and 7) Environment; each of them includes several research questions.

Finally, the study of Harrison et al. (2012, p.23) examines the Open Government ‘ecosystem’, concluding that OGD emerges as an essential dimension of the open government concept, arguing that "the importance of developing the social and material infrastructures for creating, managing, and sharing data in the short term, along with the governance structures through which innovative architectures, infrastructures, and standards will be negotiated for the future". Then they define the main themes of the research required in order to realise this vision, along with the workflow of defining data of interest, prioritizing data collection, conducting data collection, publishing the data, and then using them and generating value.

Furthermore, there is another research stream dealing with the barriers to OGD publishing and exploitation (Conradie et al. 2012; Janssen K. 2011a; Janssen M. et al. 2012; McDermott 2010; Barry et al. 2014). We reviewed this research stream, as the main findings of it (e.g. identified barriers) might correspond to important research topics (e.g. concerning new ways of overcoming these barriers), so they can be useful for the development of OGDRAT. Finally, for the same reason we also reviewed another research stream dealing with the uptake and use of OGD, and their exploitation for innovation and value generation (Bason 2010; Borins 2001; Hartley 2005; Kundra 2012; Mohr 1969; Windrum et al. 2008; Yang et al. 2013). The main conclusions of this stream of research indicate that the uptake and use of the OGD, and also the generation of innovation and value in general from them, are not straightforward, being complex, and requiring the collaboration of several actors.
From the above literature review it has been concluded that although there are some previous studies that propose categorisations of OGD research of OGD in areas and themes, these are at a too high level, and lack the detail required for directing future research, for facilitating a better interaction among researchers, and also with interested practitioners, and in general for providing the development of a ‘science base’ in this domain. Our research, as mentioned in the Introduction, contributes to filling this gap.

4. The OGD Research Areas Taxonomy

The final version of the OGDRAT we developed (= the outcome of step 7 of our methodology described in section 2) consists of four major research areas (in its first level): OGD Management and Policies, OGD Infrastructures, OGD Interoperability and OGD Usage and Value (shown in Figure 3), which include 35 research topics (in the second level). These 35 identified research topics were initially divided into two categories: the technological and non-technological ones; the latter correspond to the abovementioned OGD Usage and Value research area.

By examining the former we distinguished two clear sub-groups of research topics, concerning the interoperability and the management of the OGD respectively, which lead to the definition of the OGD Interoperability and the OGD Management and Policies areas; the remaining technological factors concerned the OGD infrastructures, so they were grouped in a separate research area. This grouping of the identified research topics into the above four research areas has been confirmed by the experts who participated in the workshop mentioned in the ‘Methodology’ section 2; however, for some research topics were proposed changes of the research area they were associated with. The OGDRAT has been constructed in an online tool “mind42.com” and is available for commenting1.

1http://mind42.com/public/f2a7c2f6-63ec-475f-a848-7ed5abe6c5a4
4.1. OGD Management and Policies

The first research area of the OGDRAT has been named “Open Government Data Management and Policies”. Data and information Management is an important research topic in the broader information systems domain, from which concepts, theories and frameworks can be borrowed and elaborated for further analysis and investigation of OGD management challenges.

Policy issues are closely related to the data management, in a broader definition, since policy decisions create the context of OGD management, so it affects data management procedures. Data management is a challenge both for OGD providers (public organizations) and for OGD users (e.g. scientists, analysts, journalists, active citizens). Therefore this research area includes several research topics corresponding to important OGD management challenges (such as methods for OGD anonymisation, cleansing, visualization, linking, publishing, mining, and also quality assessment). It is worth mentioning that within the workshop there were comments on whether we should put some of the research topics, such as OGD linking and mining in the category of infrastructures, since they are supported and provided by the developed infrastructures.

Finally, it was agreed that the OGD management capabilities, due to their importance for the use and the generation of value from OGD, should be viewed as a separate research area. In Figure 4 we can see the research topics of the ‘OGD Management and Policies’ research area, while in Table 1 these OGD research topics are described in more detail, supported by some representative relevant literature from the EGRL.
Table 1: Description of the research topics of OGD Management & Policies research area

<table>
<thead>
<tr>
<th>Research Topic</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1 Policy &amp; Legal Issues for OGD</td>
<td>This research topic concerns the investigation of different policies, strategies and principles for opening data, as well as specific measures and instruments in this direction (Blakemore &amp; Craglia, 2006; European Commission, 2013a and 2013b; Zuiderwijk &amp; Janssen, 2014c). Formulating an OGD policy is a complex multidisciplinary problem, and as such it is associated with many of the following research topics.</td>
</tr>
<tr>
<td>1.2 OGD Anonymisation Methods</td>
<td>The current practice in data publishing relies mainly on policies and guidelines as to what types of data can be published and on agreements concerning the use of published data. A major precondition for opening data of government agencies is not to disclose sensitive private data of citizens and firms. Therefore this research area focuses on methods for the anonymisation of opened data. Privacy-preserving data publishing (PPDP) provides methods and tools for publishing useful information while preserving data privacy (Benjamin et al., 2010).</td>
</tr>
<tr>
<td>1.3 OGD Cleaning Methods</td>
<td>This research topic deals with data cleaning methods for OGD, which aim to correct errors in quantitative attributes of datasets, or even other types of attributes (Hellerstein, 2008). Data cleaning is a process used to determine inaccurate, incomplete or unreasonable data, and then improve their quality through correcting of detected errors and omissions. Generally data cleaning reduces errors and improves the data quality (Natarajan et al. 2010).</td>
</tr>
<tr>
<td>--------------------------</td>
<td>-------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>1.4 OGD Quality Assessment Frameworks</td>
<td>This research topic deals with data quality, a major issue in information management in general, highly important for OGD in particular. Data quality problems occur anywhere in information systems, and they are solved by data cleaning (see previous research topic). After applying data cleaning, the quality of the data can be assessed in a number of ways, based on the internal consistency of the data and comparison of the corrected intensities with the corrected standard deviations (Chapman et al. 2005).</td>
</tr>
<tr>
<td>1.5 OGD Visualisation methods and tools</td>
<td>Visualization methods and tools is an important research topic, aiming to provide simple mechanisms for understanding and communicating large amounts of data. There is a need for exploratory mechanisms to navigate the data and metadata in these visualizations. It is therefore highly important to develop features and tools for facilitating the creation of visualizations by users on OGD (Graves et al. 2013).</td>
</tr>
<tr>
<td>1.6 OGD Linking</td>
<td>The principles, frameworks, techniques and tools for OGD linking are the subjects of this research area (Kalampokis et al. 2013; Bojārs et al. 2008). The term linked data refers to data published on the Web so that they are machine-readable, their meaning is explicitly defined, can be linked to (and from) other external datasets (Bizer et al., 2009). The advancements on this research topic concentrate on how we can structure our data so that we can find, link and process them more easily. Knowledge management</td>
</tr>
</tbody>
</table>
representation systems have been created and continue evolving in order to link different kinds of data.

1.7 OGD Publishing
The OGD Publishing research deals with and investigates all the issues of the publishing workflow and its involved actors (Bizer et al. 2009; Dawes, 2010; Helbig et al. 2012). It also examines the interconnection between the OGD publishing processes and their context (main actors and their interests and goals), and also their effects on OGD use and outcomes, and on their dynamics.

1.8 OGD Mining
The OGD mining research aims to exploit and elaborate the algorithms and methods developed in the area of data mining, in order to extract useful patterns and knowledge from OGD. Data mining uses a broad family of computationally intensive methods which include decision trees, neural networks, rule induction, machine learning and graphic visualization (Bakirli et al. 2012; Mostafa et al. 2013).

1.9 OGD Rating and Feedback
This research focuses on policies and mechanisms for closing the feedback loop between OGD users and providers, through establishing communication channels between them. Another important objective of this research is to enable OGD providers to manage efficiently comments and requests from OGD users. Thus, tools for supporting the rating of OGD and their infrastructures, providing feedback to the corresponding public organizations are more than essential. The use of OGD users–providers collaboration techniques for the above purposes are also investigated in this research area, e.g. through Web 2.0 oriented mechanisms (Alexopoulos et al. 2014; Charalabidis et al., 2014).

4.2. OGD Infrastructures
The second research area of OGDRAT has been named “Open Government Data Infrastructures”. It includes research topics concerning various important technological aspects of the ICT
infrastructures developed by government agencies in order to make OGD accessible to different groups of actors, such as their architectures, APIs provision and personalisation capabilities; another important research topic is OGD storage and long – term preservation, and also the use of cloud services in this domain. Furthermore, though the main source of OGD is the information systems of government agencies, two more sources are gradually emerging, sensors and citizens, so researching them and their exploitation is an important research challenge. In Figure 5 we can see the research topics of the ‘OGD Infrastructure’ research area, while in Table 2 these OGD research topics are described in more detail, supported also with representative literature from the EGRL.

![OGD Management and Policies](image)

**Figure 5**: Research topics for the OGD Infrastructures research area

**Table 2**: Description of the research topics of the OGD Infrastructures research area

<table>
<thead>
<tr>
<th>Research Topic</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1 OGD Portals Architecture</td>
<td>This research aims at defining the architectures of OGD portals, with respect to their scope and provided data and functionalities (Alexopoulos et al. 2014; Charalabidis et al., 2014; Helbig et al. 2012). Various types and generations of architectures are proposed and discussed from various perspective. Additionally, some research is conducted concerning the development of architectures of ICT Infrastructures that allow for and support application development utilising OGD.</td>
</tr>
<tr>
<td>2.2 Open Web Services / APIs</td>
<td>This research aims at facilitating and providing well-designed standards for Application Programming Interfaces (APIs) in OGD platforms, in order to ensure the exploitation and re-usability of published data. It is of high importance to use APIs for machine-to-machine operations for OGD. Unfortunately many of the OGD are not machine readable or the data are provided in a proprietary format (Braunschweig et al., 2012). Open Web services in this domain should conform to a set of conventions that define how a client searches for and interacts with a service (Paolucci et al. 2002; Kleijnen et al. 2003).</td>
</tr>
<tr>
<td>2.3 OGD User Profiling and Service personalisation</td>
<td>This research focuses on user profiling, which can offer big opportunities to make OGD related services more personalised, to infer and predict citizens’ behaviour, and to even influence their behaviour (Pietersen et al. 2005). Like the private sector, the public sector makes more and more use of user profiling in order to personalise the electronic services that are being offered to citizens (Mostafa and El-Masry, 2013).</td>
</tr>
<tr>
<td>2.4 OGD Long-term Preservation</td>
<td>This research topic can be found in every ICT related research domain, dealing with the ways and methods for the long-term preservation of data, which is particularly important for OGD (Agrawal et al. 2000).</td>
</tr>
<tr>
<td>2.5 OGD Storage</td>
<td>This research topic concerns the optimization of OGD storage, combining knowledge from various domains, such as databases and algorithms.</td>
</tr>
<tr>
<td>2.6 Cloud computing for OGD</td>
<td>The use of private and public cloud computing technologies and services (Lewis 2013) for hosting and providing OGD is an important research challenge, taking into account the increasing adoption of cloud in the public sector (Joshi 2012). The Linked Open Data Cloud creation supporting the vision of the Web of Data is also a research challenge classified under this research topic (Sorrentino et al. 2013; Jain et al. 2010).</td>
</tr>
</tbody>
</table>
This research aims to investigate the emerging and continuously growing volunteered user-generated content, which is often used to replace existing commercial or authoritative datasets, for example, Wikipedia\(^2\) as an open encyclopaedia, or OpenStreetMap\(^3\) as an open topographic dataset of the world (Richter et al. 2011). Open data generated by citizens, e.g. through e-participation platforms and social media, and their use for ‘crowdsourcing’ purposes, are an emerging research topic of this research area (Heipke 2010).

This emerging research topic involves tools, methods and techniques for OGD generation through sensors, which will be made freely available to the public. Big data is becoming of critical importance for science and commercial applications development (e.g. Elgendy and Elragal, 2014), so exploiting the knowledge developed in this domain and elaborating it for the OGD can be quite useful. This research topic also includes the development of methods of processing such data, calculation of analytics, and finally exploitation of them (for scientific and business purposes).

### 4.3. OGD Interoperability

Interoperability is a highly important feature of all types of information systems, and this gave rise to the development of a well-established research domain, which attracts considerable research interest, motivated by the increasing need of data exchange among organizations (both of the private and the public sector) (Jardim-Goncalves et al. 2013). Interoperability has many aspects, mainly technical, semantic and organisational. It becomes increasingly important in government, since “The divergent interpretations of data, the lack of common metadata and the absence of universal reference data hinder governments from seamless data exchange, information systems integration and the delivery of cross-border public services” (Shukair et al. 2013, p.10). So our third research area deals with the interoperability issue in the specific domain of OGD. It includes

---

\(^2\) [http://en.wikipedia.org/wiki/Main_Page](http://en.wikipedia.org/wiki/Main_Page)

\(^3\) [http://www.openstreetmap.org/](http://www.openstreetmap.org/)
research topics concerning OGD metadata, semantic annotation, ontologies and controlled vocabularies and codelists, and also on OGD platforms technical interoperability, services interoperability standards and organizational interoperability. In Figure 6 we can see the research topics of the ‘OGD Interoperability’ research area, which are described in more detail, and also supported with relevant literature from the EGRL, in Table 3.

**Figure 6:** Research topics for the OGD Interoperability research area

**Table 3:** Description of the research topics of the ‘OGD Interoperability’ research area

<table>
<thead>
<tr>
<th>Research Topic</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1 Metadata for OGD</td>
<td>This research topic includes various OGD metadata related research sub-topics: data models, schemata, taxonomies, codelists and ontology-based extended metadata sets for OGD, and also other types e-Government Resources. The term semantic interoperability asset is widely used to refer to these types of resources (Charalabidis et al. 2009; Zuiderwijk et al. 2012; Robertson et al. 2001).</td>
</tr>
<tr>
<td>3.2 Multi-linguality</td>
<td>Multilinguality is a research topic that has been attracting a growing interest by supranational institutions, such as the European Union. It includes research associated with using, extending, combining and developing semantic assets towards the support of multi-linguality in the domain of OGD (Houssos et al. 2012).</td>
</tr>
<tr>
<td>3.3 Service Interoperability Standards</td>
<td>This research topic is concerns mainly the identification, composition and execution of various applications (designed and implemented independently) offered as services. This research investigates standards that can be used for seamless interconnection among OGD related services, in order to serve different OGD uses and user scopes (Jardim-Goncalves et al. 2013). It includes the development of information systems and registries consisting of workflow models and process descriptions in an integrated knowledge base (Sourouni et al. 2008).</td>
</tr>
<tr>
<td>3.4 Semantic Annotation</td>
<td>This research focuses on methods and tools for the semantic annotation of OGD generated by public organisations and sensors, as well as the semantic annotation of User-Generated Content (UGC) (Dong-Po Deng 2013). Semantic annotation techniques capture not only the semantics, but also the pragmatics of the resources, such as who, when, where, how and why the resources are used (Kiryakov et al. 2004; Warner et al. 2009; Dill et al. 2003). The major objective of this research is the development of algorithms and tools for semantic integration (Bergamaschi et al. 1999), and also for automated extraction of metadata (self-extracted metadata).</td>
</tr>
<tr>
<td>3.5 OGD Ontologies</td>
<td>This research topic includes investigation of the proper release of OGD and the use of ontologies behind these sources (Parundekar et al. 2010). Ontologies for the description and use of OGD as well as, the sense of ontology alignment are under investigation in this research (Jain et al. 2010; Alexander et al. 2009). The Linked Open Data (LOD) paradigm is the major outcome of this research area.</td>
</tr>
<tr>
<td>3.6 Platform technical Interoperability</td>
<td>This research examines various technical issues involved in linking OGD systems and services, such as open interfaces, interconnection services, data integration, middleware, data presentation and exchange, accessibility and security services (Sarantis et al. 2008; Jardim-Goncalves et al. 2013).</td>
</tr>
<tr>
<td>3.7 Organisational Interoperability</td>
<td>The main objective of this research is the investigation of the processes by which different organisations, such as different government agencies,</td>
</tr>
</tbody>
</table>
3.8 Controlled Vocabularies and Codelists Preservation

This research includes investigation regarding preservation, indexing, and retrieval of semantic assets, such as vocabularies and codelists (Kiryakov et al., 2004).

4.4. OGD Usage and Value

The fourth research area of OGDRAT is directed towards the measurement and deeper understanding of the use of OGD, as well as the impact and value generation from them. It includes research topics concerning on one hand OGD needs, readiness, use, skills management and reputation management, and on the other hand OGD related value and impact, innovation, entrepreneurship and contribution to accountability/transparency. In Figure 7 we can see the research topics of this ‘OGD Usage and Value’ research area, while an elaboration of them and EGRL literature support are provided in Table 4.

Figure 7: Research topics for the OGD Usage and Value research area

<table>
<thead>
<tr>
<th>OGD Management and Policies</th>
<th>OGD Infrastructures</th>
<th>OGD Interoperability</th>
<th>OGD Usage and Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Policy and Legal Issues for OGD</td>
<td>• OGD Portals Architecture</td>
<td>• Metadata for OGD</td>
<td>4.1 Skills management for OGD</td>
</tr>
<tr>
<td>• OGD Anonymisation Methods</td>
<td>• Open Web Services / APIs</td>
<td>• Multilinguality Issues</td>
<td>4.2 Reputation Management</td>
</tr>
<tr>
<td>• OGD Cleansing Methods</td>
<td>• OGD User Profiling and Service personalisation</td>
<td>• Services Interoperability Standards</td>
<td>4.3 OGD Use</td>
</tr>
<tr>
<td>• OGD Quality Assessment Frameworks</td>
<td>• OGD Long Term Preservation</td>
<td>• Semantic Annotation</td>
<td>4.4 OGD-based Entrepreneurship</td>
</tr>
<tr>
<td>• OGD Visualisation methods and tools</td>
<td>• OGD Storage</td>
<td>• OGD Ontologies</td>
<td>4.5 OGD Value &amp; Impact Assessment</td>
</tr>
<tr>
<td>• OGD Linking</td>
<td>• Cloud Computing for OGD</td>
<td>• Platform &amp; Technical Interoperability</td>
<td>4.6 OGD Needs Analysis</td>
</tr>
<tr>
<td>• OGD Publishing</td>
<td>• Citizen - generated Open Data</td>
<td>• Organisational Interoperability</td>
<td>4.7 OGD-based Accountability</td>
</tr>
<tr>
<td>• OGD Mining</td>
<td>• Sensor - generated Open Data</td>
<td>• Controlled Vocabularies / Codelists Preservation</td>
<td>4.8 OGD Readiness Assessment</td>
</tr>
<tr>
<td>• OGD Rating and Feedback</td>
<td></td>
<td></td>
<td>4.9 OGD Portals Evaluation Frameworks</td>
</tr>
</tbody>
</table>

4.10 OGD Innovation |

Table 4: Description of the research topics of the ‘OGD Usage and Value’ research area

<table>
<thead>
<tr>
<th>Research Topics</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td></td>
</tr>
</tbody>
</table>
### 4.1 Skills Management for OGD
This research aims to identify and understand better the necessary skills required for OGD analysis and processing (by OGD users’ side), and also for OGD publishing and management (by OGD providers’ side). They are usually defined in terms of skills frameworks (also termed as competency frameworks or skills matrices); each of them consists of a list of skills, and a grading system, with a definition of what it means to be at particular level for a given skill.

### 4.2 Reputation Management
This research includes investigation of the use of reputation systems in OGD value chain. It examines various algorithms and methods for the reputation management of various OGD stakeholders (Bani et al. 2013; Hansson et al. 2013).

### 4.3 OGD Use
It includes studies that describe and analyse examples, ways and paradigms of OGD use for various purposes, not only by citizens (e.g. scientists, journalists, active citizens, firms active in the development of value-added e-services and mobile applications), but also by government as well (e.g. for policy making: Kalampokis et al. (2011) combined social data and ODG for participatory decision-making in government).

### 4.4 OGD-based Entrepreneurship
This research topic concerns mainly business models for exploiting the potential value of OGD and initiating OGD value chains (Ferro et al. 2012; 2013).

### 4.5 OGD Value and Impact Assessment
The current OGD research on this topic focuses on analysing OGD initiatives that have led to the generation of some kind of public value (Davies et al. 2013; Jetzek et al. 2012; 2013; Charalabidis et al., 2014), analysing the positive – and sometimes the negative as well - aspects of OGD use and impacts.

### 4.6 OGD Needs Analysis
This research includes studies of OGD users’ needs, with respect to both government datasets, and also functionalities of OGD infrastructures, aiming to lead to further developments of OGD strategies of public organizations, and also functionalities of ODG
infrastructures/portals. For instance this research led to the identification of needs for collaboration workflows and feedback mechanisms (Alexopoulos et al. 2014), and also needs for better metadata and semantic annotation mechanisms (Zuiderwijk et al. 2012).

### 4.7 OGD-based Accountability

This research investigates the use of OGD as part of anti-corruption programmes, in order to increase public sector accountability and credibility. Many government organizations publish a variety of datasets on the web, in order to promote transparency, accountability, and satisfy relevant legal obligations (Bohm et al. 2012; Alon 2011).

### 4.8 OGD Readiness Assessment

The main objective of this research is to develop frameworks and methods for assessing from various viewpoints (both ‘internal’ and ‘external’ ones) the degree of readiness of a national, regional or municipal government - or even individual agencies - to implement OGD initiatives (Davies et al. 2013; World Bank, 2013).

### 4.9 OGD Portals Evaluation Frameworks

This research aims at the creation of roadmaps, guidelines and benchmarking frameworks for the evaluation of OGD portals and infrastructures from various viewpoints (Charalabidis et al. 2014; Alexopoulos et al. 2013; Kalabokis et al. 2011).

### 4.10 OGD Innovation

The main objective of this OGD research is to identify and analyse innovations driven by OGD, both in the private sector (e.g. e-services innovations), and in the public sector (Zuiderwijk et al. 2014a). According to this literature, OGD innovation concerns mainly three domains: (a) research, (b) business and (c) transparency (Jetzek et al. 2012; 2013). While US literature and practice focuses mainly on (b), EU tends to focus on (a), but both of them are equally interested the potential of promoting (c) through OGD.)
5. Discussion

In this section are presented the outcomes of the further processing and exploitation of the OGDRAT conducted finally as part of the step 8 of our research methodology (see section 2): analysis of EGRL publications for each of the identified research topics (section 5.1); development of an extended OGD lifecycle (5.2); exploitation of OGDRAT for OGD Science Base Creation (5.3); association of OGDRAT with the ICT-enabled Governance research taxonomy developed in the CROSSROAD and the CROSSOVER projects, and also use of the former in order to extend the latter (5.4); and formulation of direction for multi-disciplinary research on important societal challenges using OGD (5.5).

5.1 EGRL publications for research topics

For all the OGD research areas identified and presented in the previous section (of the final version of the OGDRAT produced in step 7 our methodology (section 2)) we searched for relevant publications in the EGRL. In Figure 8 we can see the number of publications found for each topic (the topics are sorted in descending order of publications’ number); for the few publications that concern more than one of these topics we proceeded to their classification in the one judged as dominant (after discussion and consensus reaching among the authors). We remark that there are significant differences among these research topics as to the number of relevant publications: for some of them we have found more publications, e.g. for research topics concerning OGD use, portals evaluation frameworks, publishing, policy and legal issues, while for some others we found significantly less or even no publications, e.g. for research topics concerning sensor-generated OGD, OGD storage, long-term preservation, reputation management and skills management (for these five research topics there is no relevant literature in the EGRL, and were proposed in the workshop (step 6 of our OGDRAT development methodology (section 2)) by the experts who participated as major issues of OGD). Also, from Figure 8 we can see that there are many under-researched topics with very small numbers of relevant publications. Therefore further research is required on these research topics with very small numbers or even no publications, since they constitute interesting emerging topics, which can be significant for the achievement of higher maturity in OGD practices and value generation from them.
Figure 8: Ranking of OGD research topics based on EGRL relevant literature
5.2 An Extended OGD Life Cycle

Taking into account the OGD research topics identified during the development of our taxonomy, and also the discussions we had in the workshop with the experts who participated in it, which revealed a wide range of tasks to be performed during the life cycle of OGD, we proceeded to the development of an extended OGD life cycle, based on the combination of the Linked OGD Life Cycle\(^4\) (Open Data Support Working Group) and the Curation Life Cycle\(^5\). It is shown in Table 5: it consists of nine stages (create, pre-process, curate, store/obtain, publish, retrieve/acquire, process, use, and collaborate with users), and for each of them associated tools and methods that can be used in the particular stage are shown.

Table 5: Extended OGD Life Cycle and Associated Tools and Methods

<table>
<thead>
<tr>
<th>Life Cycle Stage</th>
<th>Tools</th>
<th>Methods</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Create</strong></td>
<td>• Sensors, RFID, IS, Human, Connection with already gathered OGD</td>
<td>• Automated data creation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Manual data entry</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Linking with OGD Portals</td>
</tr>
<tr>
<td><strong>Pre-process</strong></td>
<td>• Detailed Metadata Standards, Evaluation Metrics and Models, Maturity Matrices</td>
<td>• Conceptualisation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Structuring</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Evaluation</td>
</tr>
<tr>
<td><strong>Curate</strong></td>
<td>• Web Tools for LOD (Open Refine External Tool), Individual / Native Tools</td>
<td>• Metadata Refinement</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Change Data Format</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Data Cleansing</td>
</tr>
<tr>
<td><strong>Store / Obtain</strong></td>
<td>• Repository and Data Centre, Cloud infrastructures</td>
<td>• Versioning</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Data Linking</td>
</tr>
<tr>
<td><strong>Publish</strong></td>
<td>• Data publication mechanisms, Data publication sites</td>
<td>• Licensing</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Intellectual Property Rights</td>
</tr>
<tr>
<td><strong>Retrieve / Acquire</strong></td>
<td>• Advanced search techniques (i.e. multilingual)</td>
<td>• Open Access</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• 3-layer Metadata Schema</td>
</tr>
</tbody>
</table>

\(^4\)https://joinup.ec.europa.eu/sites/default/files/D2.1.1%20Training%20Module%202.1%20The%20Linked%20Open Government%20Data%20Lifecycle_v0.11_EN.pdf

\(^5\)http://www.dcc.ac.uk/resources/curation-lifecycle-model#sthash.FnrCA3Kf.dpuf
<table>
<thead>
<tr>
<th>Process</th>
<th>Use</th>
<th>Collaborate</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Download capabilities</td>
<td>• Internal Visualisation tools</td>
<td>• Web 2.0 Capabilities and Tools</td>
</tr>
<tr>
<td>• External Data Processing tools</td>
<td>• External Visualisation tools</td>
<td>• Collaboration Workflows</td>
</tr>
<tr>
<td>• Web Tools for LOD (Open Refine External Tool)</td>
<td>• Statistical Packages</td>
<td>• Feedback mechanisms and tools</td>
</tr>
<tr>
<td></td>
<td>• Linking with external artefacts</td>
<td>• Exchange notes/emails/ideas</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Data Quality Rating</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Create Groups of common interests</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Needs &amp; Requests on OGD</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Data enrichment</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Create Linked Open Data</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Different Datasets combination</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Statistical Analysis</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Map Visualisation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Chart Visualisation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Plot Visualisation</td>
</tr>
</tbody>
</table>

5.3 Contribution to OGD Science Base Creation

As mentioned in section 2 the research presented in this paper contributes to the development of ‘description theory’ for the OGD domain, so it constitutes the first step towards the creation of a Science Base for it. According to Charalabidis et al. (2011) the science base of a domain should include the main concepts, methods, tools and standards of the domain, and also supportive relevant experiments, surveys and case studies that have been conducted and produced a body of knowledge in the domain, and also various types of ‘proofs of concept’, aiming all to assist practitioners in this domain to solve particular problems and generate value. Our OGDRAT contributes in the above directions, as i) it identifies the main concepts, methods and tools in OGD, and ii) provides directions for future research in this domain, aiming to increase maturity of these methods and tools, so that finally OGD stakeholders (government, scientific communities, journalists, active citizens, and e-/m-services development firms) can be systematically assisted in their relevant activities, leading to higher value generation from OGD.
5.4 Extension of ICT-enabled Governance Taxonomy

The OGDRAT is associated with and extends/elaborates the ICT-enabled Governance research taxonomy developed in the CROSSROAD\(^6\) and the Crossover\(^7\) European projects. In particular, the CROSSROAD project has developed a research areas taxonomy for the ICT-enabled Governance domain, which consists of five main research themes, seventeen research areas and more than eighty research sub-areas (Lampathaki et al., 2010). One of the research themes of this taxonomy is “Open Government Information & Intelligence for Transparency”, which includes three research areas concerning “Open and Transparent Information Management”, “Linked Data” and “Visual Analytics”. The OGDRAT extends and elaborates this research theme, as the main research areas and topics of the former can replace the research areas and sub-areas of the latter, providing a higher level of detail and adding recently emerged research topics.

Also, the CROSSOVER project developed a taxonomy of research challenges in a related but narrower domain, concerning the next generation of public policy making in the Web 2.0 social media context (policy making 2.0) (CROSSOVER Project Deliverable 2.2.2, 2013), which categorises these research challenges under two research themes: (a) Data-powered Collaborative Governance and (b) Policy Modelling, in order to develop a roadmap on policy making 2.0. The OGDRAT extends and elaborates the “Link Open Government Data” research challenge of the “Data-powered Collaborative Governance” theme.

5.5 Multi-disciplinary Research on Societal Challenges Based on OGD

In the workshops it was emphasized by the participating experts that the most important and socially beneficial research OGD research can be conducted by using them as a basis of multi-disciplinary research on important societal problems and challenges that modern societies face. These data can be used by multi-disciplinary scientific teams, e.g. including members from various ‘neighbouring scientific domains’, such as economic, political, social, management and behavioural sciences (and using theoretical foundations from these sciences) in order to perform various sophisticated analyses from various disciplinary perspectives and gain useful synthetic


\(^7\) http://www.crossover-project.eu/ResearchRoadmap.aspx
insights into serious problems and challenges of modern societies; these can be quite important for the design of effective solutions and public policies for addressing them. Some directions for such multi-disciplinary research were mentioned, and are summarized in the following Table 6.

**Table 6: Directions of Multi-disciplinary Research on Societal Challenges Based on OGD**

<table>
<thead>
<tr>
<th>Societal Challenge</th>
<th>ICT-enabled Governance Research Topic</th>
<th>OGD Research Topic</th>
<th>Neighbouring Scientific Domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language divide and lack of cross-communities communication</td>
<td>Language and Cultural Interoperability</td>
<td>Metadata for OGD</td>
<td>Information Intelligence</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Multilinguality</td>
<td>Computer Science</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Controlled Vocabularies and CodelistsPreservation</td>
<td>(Translation tools)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Preservation</td>
<td>Behavioural sciences</td>
</tr>
<tr>
<td>Anticipating unexpected crises</td>
<td>Social – Economic Simulation Models</td>
<td>Semantic Annotation</td>
<td>Social and economic sciences</td>
</tr>
<tr>
<td></td>
<td>Policy Modelling</td>
<td>Organisational Interoperability</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Process Optimization for OGD</td>
<td>Sensor-generated open data</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(Accurate provision)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enhanced collective cognitive intelligence (human / ICT-enabled) for better Governance</td>
<td>Modelling and Simulation</td>
<td>OGD Mining</td>
<td>Economics</td>
</tr>
<tr>
<td></td>
<td>Policy Analysis</td>
<td>Citizen-generated open data</td>
<td>Mathematics</td>
</tr>
<tr>
<td></td>
<td>Identity Management</td>
<td>Visualization</td>
<td>Sociology</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Computer Science</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
5 Conclusions

As mentioned in the Introduction, the OGD research domain is still in its early stages, so it is important to develop a taxonomy of its main research areas and topics. This paper makes the following contributions in this direction:

- It develops a detailed taxonomy of research areas and corresponding research topics of the OGD domain (which was missing from OGD previous literature, despite its importance for the progress of this domain towards higher levels of maturity), through extraction and combination of relevant knowledge from three different kinds of sources: important relevant government policy documents, research literature and experts. For each of these OGD research topics relevant literature from the EGRL has been identified and analysed, which enables a better understanding of them and their main research objectives and directions.

- Based on the above taxonomy initially an extension of the existing OGD lifecycle has been developed, which includes important additional stages that can contribute to higher maturity in the OGD related practices of both providers and users of them, and finally more social and economic value generation from them. Furthermore, our OGD research taxonomy has been connected with two previous research taxonomies for the ‘ICT-enabled Governance’ and ‘Policy Making 2.0’ domains respectively, which have been developed in the European projects CROSSROAD and CROSSOVER, providing extensions and elaborations of them for the OGD domain. Finally, directions have been formulated for future multi-disciplinary research based on OGD for addressing important challenges that modern societies face.

The findings of our study reveal the interesting thematic ‘richness’ of the OGD research domain, which currently includes a wide range of research topics, both technological and non-technological ones, concerning both the opening and publishing of government datasets, and also their usage (by various actors, such as e-service or mobile apps developers, scientists, analysts, journalists, active citizens, etc.), exploitation and value generation from them. This reflects the inherent complexity of opening of government data to the society and the economy, and then creating value from them, which the OGD research aims to address. In particular, we identified a
multitude of technological research topics in the OGD research domain, with most of them concerning the exploitation of existing or emerging technologies, on one hand in the opened datasets (e.g. anonymisation, cleansing, mining, metadata, linking and semantically enriching technologies), and on the other hand in the OGD infrastructures (e.g. web services, storage, cloud computing, interoperability technologies), in order to enrich their usefulness. Furthermore, we identified a multitude of non-technological OGD research topics, which concern mainly OGD needs, use, impact, value and entrepreneurship.

However, our study has revealed significant differences among the above identified OGD research topics as to the ‘quantity’ of the research conducted on them. For some of these topics there are limited or even no publications at all (e.g. for research topics sensor-generated OGD, OGD storage, long-term preservation, reputation management and skills management); so further research is required on these under-researched topics.

Our research taxonomy has interesting implications for research and practice. With respect to research it provides directions and structure for future research in the OGD domain, and also facilitates communication and interaction among researchers (through the ‘common language’ it introduces), and also with interested practitioners. Also, it contributes to the development of a ‘description theory’ of the OGD domain, which can be useful for the development of other more advanced types of theories (as mentioned in section 2). Finally, it identifies important under-researched topics, on which further research is required. With respect to practice our OGDRAT is useful to government agencies, as it proposes to them possible dimensions of their OGD strategies, practices and infrastructures, on which they should focus their attention, in order to improve the value generated from them. Also, this detailed taxonomy can contribute to the development of new knowledge in this domain, which will enable improving and optimizing the technology, and also the design, operations and performance of the units of government agencies responsible for opening data. Finally, OGDRAT is useful to ICT firms developing OGD technological infrastructures, as it provides them directions for improving their products and services.

A limitation of our study is that for practical reasons we organized only one workshop (though we had participants from 11 EU countries, and from different types of organizations, such as public administrations, universities and firms). So it is necessary to organize more workshops in order to further validate the OGDRAT, and probably have proposals for additional research topics, with participants from all major stakeholder groups, such as such as e-service or mobile apps.
developers, scientists, analysts, journalists, active citizens and public servants. In this direction the proposed taxonomy is available on the Web and can be accessed by the following link: (http://mind42.com/public/f2a7c2f6-63ec-475f-a848-7ed5abe6c5ad), so that we can collect ratings, comments and ideas from the OGD community for further elaboration and update. Another limitation is that we have identified and analysed relevant literature for all the research topics of OGDRAT only from the EGRL; so it would be good to exploit other research libraries as well. Also more research is needed in order to map the multiple OGD research projects which are currently in progress (e.g. supported by European Commission or USA research programs) to the first level research areas and the second level topics of the OGDRAT, and possibly based on them elaborate and update the taxonomy.

6 References


Chapman, A.D., (2005), “Principles and Methods of data cleaning - Primary Species and Species - Occurrence Data, Version 1.0”, Report for the Global Biodiversity Information Facility, Copenhagen


Janssen, K., (2011 a), "Legal interoperability – barriers to the harmonization of licences" presented at the ICRI - Share PSI workshop, Brussels 2011


