Linked Open Government Data Research Panorama

Bernardo Todesco, Bruno Blume, Airton Zancanaro, José Leomar Todesco and Fernando Gauthier
Knowledge Engineering Lab, Universidade Federal de Santa Catarina, Florianópolis, Brazil

Keywords: Linked Open Government Data, Transparency, Bibliometry, Semantic Web, Open Government.

Abstract: In order to increase transparency and civic participation, governments around the world sought ways to open their data and allow both to itself as to companies and the civil population a greater share in the maintenance, surveillance and optimization of the services provided. To this end, using a technology called linked data (LD), the data released by governments became easier to be understood and properly used by humans and machines alike, thus creating what today is called linked open government data (LOGD). The purpose of this article is to present the state of art of the research in LOGD through bibliometric research, ultimately presenting a feedback on the matter.

1 INTRODUCTION

The internet has brought to the world substantial changes to social, economic and political structures, presenting an instant communication and data distribution between geographically distant parties. People, companies and, ultimately, governments have become adept to this form of interaction, incorporating such technology as a means to optimize the services provided. In this context, governments have used such technology in its pursuit of enhancing government transparency and civic engagement, enabling the provision of services in a more efficient, effective and, above all, transparent way.

In order to achieve such transparency and enable a more efficient civic engagement, several governments have launched their data online (Davies, 2010; Shadbolt et al., 2012) in the O*pen Data* format. The concept of openness in *Open Data*, although not completely new, carries the idea of providing information free of charge and free of copyright or patents, being the key concept in many other movements such as *Open Source*, *Open Content* and *Open Access*.

Therefore, *Open Data* is the idea of providing data in a manner that allows their free use, re-use and sharing. In other words, without any patents or copyrights (Ferrer-Sapena et al., 2011). Such field has been growing in relevance and research since governments in many countries - such as the United States, the United Kingdom, Australia, New Zealand, Germany, Spain Brazil - have focused their attention on opening governmental data, following *Open Data* guidelines for that purpose.

It was within this context of opening governmental data for use by any interested party that, by using RDF language (Brickley and Guha, 2004; Klyne; Carroll, 2004), the concept of LD was employed (Bizer et al., 2009; Bizer, 2009).

By adding context and standardized formatting to datasets, LD is a method of publishing structured data that enables the creation of relevant information from open data in various sources, subjects and fields, such information being of interest to several different parties, whether governmental, private or civil. (Cerrillo-i-Martinez, 2012; Janssen, 2011). For example, the timetable of a bus route at different points of its path or the monitoring of a project for access to expenses for each step of your process and other relevant information.

Therefore, the so-called LOGD, a specifically governmental form of LD, consolidates itself as the standard format for access and transmission of government data, so as to disseminate and clarify the actions and decisions taken by the government to its citizens (Ding et al., 2012; Hendler et al., 2012).

The present bibliometric research aims to check the state of art of academic research regarding Linked Data and government transparency, in order to favor researchers in the mapping of the profile and characteristics of the publications on the subject.

The purpose of these results is to identify clearly which authors, institutions and countries have

Todesco B., Blume B., Zancanaro A., Todesco J. and Gauthier F..
Linked Open Government Data Research Panorama.
DOI: 10.5220/0004548402780285
In *Proceedings of the International Conference on Knowledge Engineering and Ontology Development* (KEOD-2013), pages 278-285
ISBN: 978-989-8565-81-5

advanced more in the research about LOGD, as well as showing successful methodologies, relevant issues highlighted by authors. This might enable stakeholders to advance in new research or new projects.

The article is divided as the following: Section 2 is a brief explanation about the concept of LOGD, section 3 presents the methodological procedures for the development of this research, while section 4 presents the results obtained and ultimately section 5 presents the closing remarks.

2 LOGD

Before entering in any methodological procedures or results for this article, it would be of great importance to explain the main focus of this work. In other words, to clarify to LD, Open Government or government transparency enthusiasts what LOGD stands for and what it enables. Such clarification also justifies the efforts of performing a research panorama on the field.

Many authors have been writing about LOGD lately, especially after Data.gov and Data.gov.uk came to existence (Ding et al., 2012; Kalampokis et al., 2011; Shadbolt et al., 2012). According to Ding (Ding et al., 2012), these two portals have pioneered the LOGD initiatives worldwide.

The LOGD movement is closely related to the Open Government Data (OGD) initiatives that took place worldwide over the last decade. As a huge amount of datasets was released, governments and civil society faced the difficult task of integrating this material due to the use of different vocabularies, formats and qualities of metadata within what was released (Ding et al., 2012).

As a response to this big challenge, LOGD emerged as "a way of facilitating opening, linking, and reusing OGD. LD offers minimal consensus on data representation (such as using URIs and the Resource Description Framework) and data access (via HTTP), and enables incremental OGD publishing according to Tim Berners-Lee's '5 Stars of Linked Open Data'" (Ding et al., 2012).

Ding explains the process of opening and linking data in three stages. In the first stage, governments play the key role, opening up their data. In the second stage, community must help enhance the quality of the released data. In the last stage, data is reused, in order to build high-value applications from the datasets. Thus, LOGD can reach its full potential only through public participation. Citizens have to get themselves involved in the process, first by pressuring the government to release its data and later by enriching the available data.

3 METHODOLOGICAL PROCEDURES

The present work adopted a combination of different bibliometric techniques and followed some recommendations and descriptions suggested by Macias-Chapula (1998), Vanti (2002) and Francina and Oliveira (2011).

According to Tague-Sutckiffe (apud Macias-Chapula, 1998, p.134), bibliometry is the study of the quantitative aspects of the production, dissemination and use of the information recorded. Moreover, bibliometry develops patterns and mathematical models to measure these processes, using its results for drawing up estimates and supporting decision making.

The present bibliometric study was conducted in two phases: the first phase consisted of searching, filtering and standardizating the articles, while the second phase was the analysis and development of the final work.

The following sub-items describe the steps undertaken.

3.1 Step 1 - Defining the Search of Terms and Database Query

The association of terms related to government and open data aims to identify areas of study and research lines that are indexed in international scientific databases. To perform the search, we used some terms related to government (open government, transparency and e-government) and some terms related to open data (linked data, linked open data and linked government data. These terms were also used in Spanish and Portuguese. Searches were conducted in three international databases: Web of Science (WoS), Scopus and Google Scholar.

3.2 Step 2 - Exporting the Articles and Reading the Abstracts

After making queries in databases, files were generated with the main bibliometric data (title, abstract, keywords, year, author, institution, etc.). Such data were then imported into the bibliographic management software, Mendeley. It enables a more practical and dynamic indexing and use of articles and journals by indexing files in formats such as

PDF or DOC, enabling the creation of automatic references and the categorization of the articles in fields.

While using Mendeley, some criteria were established on the selection of works that were included in the final work. Aside from the removal of articles that were duplicated or without authorship, reading the abstracts permitted the identification of articles that had subjects in accordance with the current research's proposal. This step also included removing papers that weren't available for free download.

3.3 Step 3 - Standardizing Data and Classification of Articles into Macro Themes

By the fact that the research was conducted in three separate databases, the need to standardize the data so that statistical data had greater reliability was noticed. For example, while in Scopus the authorship is identified by surname and first initial letter of the name, in Web of Science, the full name is shown spelled out fully. Moreover, the lack of information such as institutional bond of the authors and lack of keywords forced the categorization to be done using the full text or searches on the Web.

For such standardization, a database was created in Microsoft Access and the fulfillment of information for each article (author, institution, year, country, publication source, etc.) was performed manually.

After the stage of standardization, with more familiarity towards the researchers and the contents, the categorization of the articles according to the macro themes they address was possible.

3.4 Step 4 - Data Analysis and Evelopment of the Final Work

After standardization and categorization, consultations and frequency counts were performed. Several tables, graphs and schematic figures were created from the set of selected papers, thereby allowing several considerations regarding the state of art of the research on LOGD, which was used for the development of the present article.

With all the data gathered, properly standardized and developed, the creation of the current article was made possible.

4 RESULTS

The analysis conducted from the survey of bibliometric data allowed the drawing of conclusions on various aspects of the research regarding LD and government transparency. Some of these conclusions refer to databases, sources of publications, research dates, authors, institutions and key terms that will be presented next.

4.1 Bibliographic Data (Linked Data and Government Transparency)

Almost half of the papers were found in the Scopus database. Out of the 38 articles found there, 21 were selected according to the criteria described above. From the Web of Science database, only one article was included from the eight papers located. The other main source was Google Scholar's database, which resulted in the largest number of selected articles. Thus the total set of works analyzed is 49. Table 1 illustrates the process of selection the papers went through for analysis.

Table 1: Number of publications selected.

Database	Papers found	Papers selected
Scopus	38	21
WOS	8	1
Scholar	27	27
Total	73	49

Most of the works selected are articles indexed in scientific journals, a total of 27. The remaining articles were published in conference proceedings (as Conference Papers) or as book chapters and articles posted on websites such as W3C. The papers came from 40 sources and were prepared by 158 authors, belonging to 65 institutions from 16 different countries.

By summing the articles, 666 references were made, an average of over 13 references per article. Furthermore, 92 keywords were used. The Table 2 summarizes the general bibliometric research.

Table 2: Bibliometric results.

Bibliographic data	Frequency
Papers	49
Publication sources	40
Authors	158
Institutions	65
Countries	16
Keywords	92
References and citations	666

4.2 Publications by Year

The majority of the articles were published after 2009, reaching its peak in 2011 with 16 publications, following 2010 and 2012, with 12 and 9 respectively. This is due to the fact that LD is a very recent concept in the means of information technology. Tim Berners-Lee launched his first considerations on LD in an article from 2006.

Over the last years, the concept of LD also became part of the set of ideas in public administration worldwide, starting with the United Kingdom and the United States of America. Thus, it is understandable that most of the articles were written in the last four years, since the growth in LD research, deals with issues related to governmental LD.

The exceptions to this pattern are seven articles published between 2000 and 2008, which address innovatively on the semantic web, ontologies and other tools now widely used for the preparation of projects linked government data.

4.3 Main Sources of Publications

The articles were originating from 49 different sources. Only three of these had more than one article. The results are shown in Table 3.

Table 3: Main publication sources.

Publication sources	Qty.	Source	JCR
ACM Int'l Conference Proceeding Series	5	Confe-rence	-
IEEE Intelligent Systems	5	Journal	2.154
IEEE Internet Computing	2	Journal	2.000

The source with the highest number of papers are the proceedings of an international conference organized by the ACM Digital Library, a digital publishing company responsible for many journals in several areas - including information technology.

The two other sources with more than one publication obtained from searches are two journals from the Institute of Electrical and Electronic Engineers (IEEE), which have respectable impact factor and recognition in the field of computer science. IEEE is responsible for publishing numerous journals in various fields.

When it comes to the bonds between authors. Figure 1 shows the two main networks identified between the authors in the searches. They were built according to the authorships of the articles. There are two separate networks, named Network 1 and Network 2. In the first of them, many European professors are shown, such as Berners-Lee, Bizer, Heath, Cyganiak, Peristeras, Auer, O'Hara and Shadbolt. Some articles written by those authors are "Linked Data: The Story So Far", "DBpedia: a nucleus for a web of open data", "The emerging web of linked data", among others.

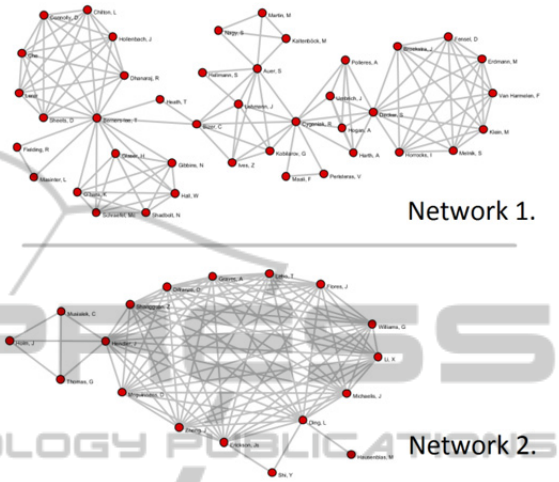

Figure 1: Main networks.

Network 2 is mainly composed by Tetherless World Constellation authors, which have done many articles about LOGD lately. Some highlights from these articles are "TWC LOGD: a portal for linked open government data ecosystems", "Data-gov Wiki: towards linking government data", "Linked Open Government Data" and "US Government Linked Open Data: Semantic.data.gov".

One important aspect that could distinguish both networks is the kind of research developed by its authors. The European network is clearly more conceptual and normative, developing the basis of the LOGD program. Berners-Lee, for example, is the author who has first mentioned the term Linked Data, back in 2006. He also has defined quality standards for the publication of data, with its 5-star classification. Auer is another good example: he developed DBpedia, which has enabled the growth of the web of data since 2009.

On the other hand, TWC researchers have done much more concrete projects, such as portals and catalogs. Among many of those projects, it can be highlighted DIGO, the TWC LOGD Portal and the International Open Government Dataset Catalog.

4.4 Main Authors, Institutions and Countries

In Table 4, shown below, lists the authors with the highest numbers of publications, along with the institutional affiliation, city and country.

Out of 49 selected articles, 31 of them (ie over 60% of articles) were written by a total of nine researchers (out of a universe of 158 in the overall selection), belonging to only four institutions located in two countries.

The research group that stands out most is the Tetherless World Constellation (TWC), which is linked to the Rensselaer Polytechnic Institute. Leading that group are James Hendler and Deborah McGuinness, who also appear in the list of top authors. Among the nine most productive authors identified, six work in TWC.

The remaining authors are Tim Berners-Lee and two German specialists. The academic and professional profile of each of the nine researchers is presented below.

Li Ding is a researcher at the Rensselaer Polytechnic Institute. In the analysis performed, some articles about LOGD were found, aside articles about applications developed specifically for the use of LOGD. In fact, he is a prolific researcher on linked government data. Ding is an author in five articles, four of these as first author, and has 14 connections with other authors about the topics covered.

Tim Berners-Lee, the creator of the World Wide Web, has spread, since the early 2000s, the vision of a Web of data as the evolution of the Web of documents originally conceived by him. LD is an idea developed by Berners-Lee since 2006. In the present study, Berners-Lee is the author of four articles, two of them as first author, and has 16 connections with other authors on the subject addressed.

James Alexander Hendler is a renowned researcher at Rensselaer Institute. He is one of the creators of the Semantic Web. At TWC, Hendler is ahead of the Linking Open Government Data Project, which makes him one of the leading researchers on the use of government data in standard Linked Open Data. In this bibliometric research, Hendler has authored four articles, one of these as the first author. He has connections with 16 other authors on the subject addressed.

Sören Auer is linked to the University of Leipzig, Germany, where he is ahead of some research groups related to the Semantic Web. He also coordinates a project called "LOD2 - Creating Knowledge out of Interlinked Data". In this survey, Auer was the author of three articles, two of these as the first author, and has nine links with other authors on the subject addressed.

Christian Bizer currently works at the University of Mannheim, Germany. He made efforts to create the Linked Open Data community in the W3C, aside from being a DBpedia's co-founder. He is the first author of "Linked Open Data - The Story So Far", most cited work within the articles collected in this research. Bizer authored three articles of the bibliometric selection, two of these as the first author.

The last four authors on the list all work at Rensselaer Polytechnic Institute, more specifically at Tetherless World Constellation: Dominic DiFranzo. Deborah McGuinnes, Alvaro Graves James Michaelis. Each of these names appears in three articles and each author has professional links with at least 12 researchers.

Table 5 identifies the institutions with the biggest number of authors included in the survey. These numbers, however, do not necessarily mean that such bodies have the highest quantities of articles published, since a single article can have multiple authors. Nevertheless, this table can inform us with some accuracy the depth of the workload and research produced by each institution, since a higher number of both authors may represent more researchers engaged in the field as well as more

Table 4: Authors with the most publications.

Author	Publications		Institutional Bond	City	Country
	Total	1st Author			
Ding, L	5	4	Rensselaer Polytechnic Institute	Troy	EUA
Berners-Lee, T	4	2	Massachusetts Institute of Technology	Cambridge	EUA
Hendler, J	4	1	Rensselaer Polytechnic Institute	Troy	EUA
Auer, S	3	2	Universität Leipzig	Leipzig	Alemanha
Bizer, C	3	2	Universität Mannheim	Berlim	Alemanha
Difranzo, D; Graves, A; McGuinness, D; Michaelis, J	3	-	Rensselaer Polytechnic Institute	Troy	EUA

material produced. The next paragraphs present these institutions and its activities.

Table 5: Quantity of authors per institution.

Institution	Authors
Rensselaer Polytechnic Institute	14
Hasso Platner Institute	9
Massachusetts Institute of Technology	8
Polytechnic University of Madrid	8

The Rensselaer Polytechnic Institute is a traditional American institution. The Tetherless World Constellation (TWC) is located there. Among its various research projects, there is one called "Linking Open Government Data", whose focus is to investigate the role of the semantic Web, especially LD, in the production and publication of government data in portals like Data.gov. The researchers leading this project are Jim Hendler and Deborah Mcguinness. The institute has other important researchers such as Li Ding, Peter Fox, Massimo Di Stefano, Xian Li and Dominic DiFranzo.

The Hasso Plattner Institute, located in Germany is a non-profit Information Tecnology System Engineering institute that maintains several lines of research, including LD. It was founded and financiated by the German professor Hasso-Plattner.

The Massachusetts Institute of Technology (MIT) also appears on the list, mainly due to the work of professor Tim Berners-Lee. There are several laboratories and programs related to the field of computer science, including W3C and Decentralized Information Group (DIG), which was created by Berners-Lee and whose mission is to investigate the effects of the use of information on the web.

The Polytechnic University of Madrid appears in fourth place in the list. Most authors of this institution have been involved in the preparation of a paper on the application of spatial data linked to data provided by the Spanish public sector.

4.5 Main Keywords and Topics Addressed by the Articles

As stated earlier, almost a hundred different keywords were identified in the articles selected. With this content, it was possible to create figure 2, a cloud of terms that visually shows the frequency of each term. The bigger the term, the most it has appeared.

The most prominent keyword in the research was “linked data”. Similar terms also were relevant, llike open government data, open government, open data and linked open data. “Semantic Web”, which is closely related to the concept of “linked data”, is also in highlight.

Figure 2: Representation of keywords.

In addressing the different profiles from the selected articles, a thematic classification in macro themes was elaborated (Figure 3), in order to group articles with similar proposals.

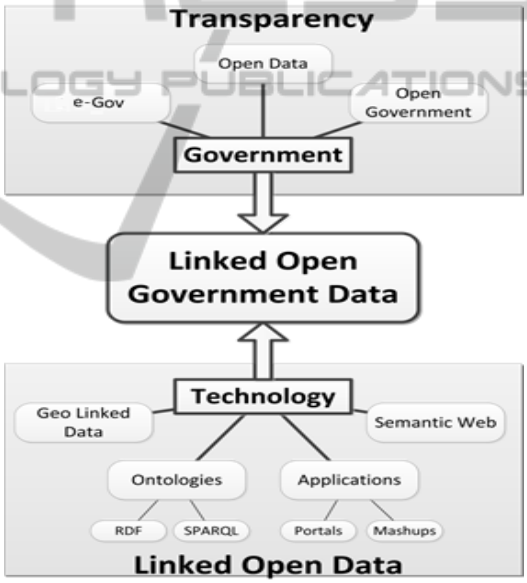

Figure 3: Macro themes and topics covered by the papers.

It was noticed that some articles are mainly concerned with governmental issues related to LD, mostly by addressing the consequences that such concept would bring for government transparency. The main themes identified in this sense were "open government", "open data" (data open term closely connected with the notion of transparency) and e-government.

Other articles, however, addressed technical topics of LD, such as design issues, didactical issues and also the presentation of practical results of applying LD to government data, such as data mashups and portals. Another issue that appeared in

some articles was geolinked data, a LD standard that uses geographical data, which can also be made available by the government.

The convergence of both of these main areas is very well expressed in the central macro theme, which is "linked open government data". Some articles deal specifically with this term - i.e. an article by Ding (2012) and another from Breitman (2012).

4.6 Most Cited References

Table 6 addresses the 15 most cited references by the articles belonging to the conducted bibliometric survey. Most of these 15 publications do not appear in the bibliometric list. However, all the works presented in Table 6 are relevant to the main topic of this article. In fact, most of them present key pieces to the elaboration of the so-called LOGD. Tim Berners-Lee is an author of five of those articles. Tom Heath and Christian Bizer both appear in three of them. The presence of these specialists in this list reveals its importance.

Two main types of publications can be perceived in the group listed above: one from studies that deal specifically with technical aspects of LD and the other from studies regarding the applications of LD in the publication and re-use of government data.

The most referenced work by the selected articles in our bibliometric research deals with LD, written by the biggest name in the bound data, Tim Berners-Lee. In "Linked Data - The Story So Far", the authors are interested in presenting the concepts and principles of LD, which is defined as "a set of best practices for publishing and connecting structured data on the Web" (Bizer et al., 2009). The technological foundations are also presented: the LD is based on Uniform Resource Identifiers (URIs) and HyperText Transfer Protocol (HTTPs).

Moreover, it cites several achievements reached in the first three years since the concept of LD was published. One example of that would be the Linking Open Data project, which has popular background and is supported by the W3C Semantic Web Education and Outreach Group (Bizer et al., 2009).

The second work on the list, "Linked Data - Design Issues," is signed by Berners-Lee and is available online at the W3C Consortium. The four rules for using LD are presented, which form the data default format in the semantic web. Following these rules is not obligatory, but not following them means loss of semantics and data interconnection. Therefore, Berners-Lee highly recommends the observance of those rules.

The article "How to Publish Linked Data on the Web" is the third most cited work. It provides a tutorial on how to publish LD on the Internet and presents the concept of LD, which, according to the author, aims to "enable people to share structured data as easily as they can share documents today"(Bizer, C. Cyganiak, R, Heath, T, 2008). With the purpose of achieving this goal, it presents a series of recipes and methods about publishing information in LD format.

The fourth most cited study, entitled "Putting Government Data Online", provides an overview about linked government data, a default format for government data, whose purpose is to enhance transparency, government efficiency and to disseminate valuable knowledge for the society and the government itself. With this in mind, Berners-Lee presents a series of steps to greater efficiency in opening up data, among these: the preference for offering them as LD, connect these data with other data sources and provide an understandable way of assimilation and its digestion, aside from efforts to improve interoperability – which is the compatibility between different databases.

5 CLOSING REMARKS

The bibliometric method employed in this study enabled the realization of the fundamental aspects of the state of the art research on government data online. In fact, the results revealed the situation of the dissemination of scientific knowledge involving linked government data transparency and government and what progress has already been done in the discussion on this topic.

The results presented in this article point out to the growth of the relevance of such research

Table 6: The 15 most cited references by the articles.

Authors	Year	Title	Publication Source	Cit
Bizer, C; Heath, T; Berners-Lee, T	2009	Linked Data - The Story So Far	Journal; IJSWIS	13
Berners-Lee, T	2006	Linked Data - Design Issues	Web Page; W3C	8
Bizer, C; Cyganiak, R; Heath, T	2008	How to Publish Linked Data on the Web	Web Page	8
Berners-Lee, T	2009	Putting Government Data Online	Web Page; W3C	8

worldwide. The impacts of the discussion on this subject certainly are not restricted to the academic sphere, which can be demonstrated by the evolution of national legislations regarding government data in several countries as well as the settlement of international agreements that begin to set standards on procedures in the public sphere. These movements are in sync with the results of releasing government data experiences on LD format in recent years, especially those made in Britain and in the United States.

Although most studies here raised come from American, British and German institutions the development of research in many other countries, such as Brazil, Chile, Romania, Spain and Albania is noteworthy, because it reveals recognition to linked government data in a global level.

The conclusion reached in this study is that linked government data is a new paradigm that directly affects public sectors, creating new possibilities for interaction between citizens, businesses and government.

However, more extensive bibliometric research can be made, considering the use of other keywords in the databases, such as "Semantic Web", "

The most important consequences from the meeting of this new paradigm with the demand for more efficient public administrations must be studied in detail, possibly with the development of case studies and other convenient methods.

REFERENCES

Bizer, C. (2009). The emerging web of linked data. *Intelligent Systems, IEEE*, 87–92.

Bizer, C., Heath, T., & Berners-Lee, T. (2009). Linked Data - The Story So Far. *International Journal on Semantic Web and Information Systems*, *5*(3), 1–22.

Brickley, D., & Guha, R. V. (2004). RDF Vocabulary Description Language 1.0: RDF Schema.

Cerrillo-i-Martinez, A. (2012). Fundamental interests and open data for re-use. *International Journal of Law and Information Technology*, *20*(3), 203–222.

Davies, T. (2010). Open data, democracy and public sector reform: A look at open government data use from data. gov. uk, 1–47.

Ding, L., Peristeras, V., & Hausenblas, M. (2012). Linked open government data. *IEEE Intelligent Systems*, *27*(3), 11–15.

Ferrer-Sapena, A., Peset, F., & Aleixandre-Benavent, R. (2011). Acceso a Los Datos Públicos y Su Reutilización: Open Data y Open Government. *El Profesional de la Informacion*, *20*(3), 260–269.

Francina, E., & Oliveira, T. De. (2011). Indicadores bibliométricos em ciência da informação : análise dos pesquisadores mais produtivos no tema estudos métricos na base Scopus Bibliometric indicators in information science : analysis of the most productive researchers about metric studies in th. *Perspectivas em Ciência da Informação*, *16*(4), 16–28.

Hendler, J. . J., Holm, J. . J., Musialek, C. . C., Thomas, G. . G., & Hendler J.a Holm, J. . M. C. . T. G. . (2012). US government linked open data: Semantic.data.gov. *IEEE Intelligent Systems*, *27*(3), 25–31.

Janssen, K. (2011). The influence of the PSI directive on open government data: An overview of recent developments. *Government Information Quarterly*, *28*(4), 446–456.

Kalampokis, E., Tambouris, E. . E., & Tarabanis, K. . K. (2011). A classification scheme for open government data: Towards linking decentralised data. *International Journal of Web Engineering and Technology*, *6*(3), 266–285.

Macias-chapula, C. A. (1998). O papel da informetria e da cienciometria e sua perspectiva nacional e. *Ciência da Informação*, *27*(2), 134–140.

Shadbolt, N., O'Hara, K., Berners-lee, T., Gibbins, N., Glaser, H., Hall, W., Schraefel, M. C., et al. (2012). Linked open government data: Lessons from data.gov.uk. *IEEE Intelligent Systems*, *27*(3), 16–24.

Vanti, N. A. P. (2002). Da bibliometria à webometria: uma exploração conceitual dos mecanismos utilizados para medir o registro da informação e a difusão do conhecimento. *Ciência da Informação*, *31*(2), 369–379.