

Challenges and requirements for standardisation of open budgetary data in the Brazilian public administration

Marcelo Tavares de Santana, Gisele da Silva Craveiro

School of Arts, Sciences and Humanities
University of São Paulo
Rua Arlindo Béttio, 1000
03828-000 – São Paulo – SP – Brazil
marcelo.tavares@usp.br
giselesc@usp.br

Abstract: Open Government Data put some new elements to the discussion of budget transparency. The reality in Brazil, which has a legal framework that enforces all public entities to web publish detailed budgetary data in real time, shows that effective transparency will be reached only if a standardization progress is met. This paper discusses the challenges for the standardization of open budgetary data in Brazil, and derives a set of requirements that should be addressed by a standardization proposal in this context. Furthermore, we evaluate how some existing standards are capable to meeting these requirements and propose lines for future work in the standardization of open budgetary data in Brazil.

1 Introduction

The Web is a cheap and flexible medium for publishing, editing and providing interaction among people, businesses and government. Its use may potentialize the development of democracy through public oversight, which can be understood as citizen participation in public management, surveillance, monitoring and control of public administration activities. Also it can be an important mechanism to Prevent corruption and strength to citizenship [Oe02].

Many public sector information is published on the web, and budgetary information being has received attention since 2000 years, the many web portals, like financial transparency portals [So04]. Brazil have a good Recognized Legal framework about fiscal transparency, as in 2009 the Supplementary Law No. 131 requires that all brazilian public entities release, in real time, detailed information on budget execution on the web [Br09]. Brazilian Law No. 12,527 also enforces that many public sector information, expenditures, financial transfers and procurements must be published in the primary, authentic and updated way [Br11a].

A recent study surveyed have budgetary execution data quality in 88 transparency web portals in Brazil [CSA12] and, although in some aspects a good evaluation is done, one of the main findings leads to the lack of a suitable data primary level, met only by 17.16% of all analyzed datasets.

Also, another result shows 30% of all datasets were machine processable, but it is important to note that it is relatively weak, since in most cases the data were published in semi-structured files like CSV format and without metadata, neither common vocabulary or structure.

The Brazilian reality in budget data dissemination unfortunately is still an opaque transparency, as a federation of the size of Brazil, with a federal district, 26 states and 5,570 municipalities, must necessarily face a standardization effort for anyone interested in following the public budget can get the complete scenario of investment in public policies. This requires the integration of federal, state and municipal contexts through an appropriate level of information available for each public organization and also for interoperability between them.

It is important to add that budget transparency, disclosure open data and recognition of the need for standards are concerns in a global context, as can be seen in guidelines of the Open Government Partnership (OGP) [Og11] and national policies of open data, such as the United States [Us13]. The United States and Brazil were the first chairs of OGP and this adds an important element in the national political scenario.

We agree with [WS96] that poor data quality can have substantial social and economic impacts and require that the conceptualization of this quality should emerge from the perspective of data consumers. This article aims to make a contribution in this direction by presenting some requirements to a good open budget data publishing related to the specific case of Brazilian reality.

In this paper, in Section 2 we present the methodology undertaken for gathering requirements for budget data publishing. Section 3 contains the requirements raised presented chronologically. Section 4 lists the requirements with layers in a information system's model. Section 5 presents initiatives related with data publishing. Section 6 discusses the main challenges encountered in budget data publication with open formats in Brazil.

2 Methodology

This study was an exploratory research of requirements from the Brazilian legal framework in relation to budget data publication on the web. Later on some papers that present a broad survey of the of Data Quality and Information Quality were studied, [MA09], [GH07] and [Zh12], but at the current stage of the work there is still no full correspondence with all information presented by articles in that area.

We decided at that time to prioritize open data consumers viewpoint and we started from open data eight principles [Og07] and documents from organizations that work with publication and consumption of open data [Ok11] [MC13]. Important elements in this vision of consumer of data were also found in the model of the five stars [Be09] and some work in the area of linked open data [Ok12] [ST10] [GG09].

In these works were collected contributions to requirements of open data and then researched references of this initial literature, in search of more requirements. Requirements were organized in a list in chronological order, eliminating similar requirements.

The requirements raised were related to a development model of information system from classical literature, which is used to identify existing standards related to the requirements.

3 Requirements raised

This section presents elements got from the preliminary survey conducted in this work and represent the challenges and requirements to be implemented in a platform for publishing data. They will be presented chronologically and, whenever possible, relevant information will be related to the brazilian reality.

In 2007, [Og07] published open data eight principles that represent a starting point in gathering requirements:

Data Must Be Complete: "All public data are made available.[...] Public data are data that are not subject to valid privacy, security or privilege limitations, the Governed by other statutes." In the context of brazilian law [Br09], means publishing information of revenues and expenses of budget execution, not including previous steps as budget planning.

Data Must Be Primary: "Data are published as collected at the source, with the finest possible level of granularity, not in aggregate or modified forms." In Brazil these data should be detailed [Br09], which in portuguese is not just raw data.

Data Must Be Timely: "Data are made available as quickly as necessary to preserve the value of the data." In Brazil, it's considered one business day [Br10].

Data Must Be Accessible: "Data are available to the widest range of users for the widest range of purposes."

Data Must Be Machine processable: "Data are reasonably structured to allow automated processing of it."

Access Must Be Non-Discriminatory: "Data are available to anyone, with no requirement of registration".

Data Formats Must Be Non-Proprietary: "Data are available in a format over which no entity has exclusive control".

Data Must Be License-free: "Data are not subject to any copyright, patent, trademark or trade secret regulation. Reasonable privacy, security and privilege restrictions may be allowed the governed by other statutes."

[Be09] in 2009 developed a model of linked open data into five levels (called stars) to classify progressively ease and power usage of data by data consumers. This model is a benchmark qualification of open data in Linked Data approach and it complements the eight principles in data access evaluating. The first three correspond to License-free, Machine processable and Non-Proprietary Formats, but three issues are added by this model:

Available on the web: the principle differs Date Must Be Accessible because define that the way of access is the Web.

Identify things: "Use open standards from W3C to **Identify Things**, so that people can point at your stuff." Create a unique identifier is a means to allow other datasets can be linked to the same resource.

Link data: "Link your data to other people's data to provide context."

Also in 2009, [Ha09] developed a model for Web data provenance, which includes two dimensions - the creation of data and access to data - in order to "support the assessment of data qualities such as accuracy, reliability and timeliness". This model is distinguished provenance model applied to the database to include information data access [Ha09], as type of data access service, if it is a server, a web service or a search service. So we added the publication of **Provenance Information** as a requirement.

Sun Light Foundation in 2010 added two more principles to the previous eight [SF10]:

Permanence: "The capability of finding information over time is referred to the remains.[...] Often times, information is updated, changed or removed without any indication that an alteration has been made [...] information made available online should remain online, with appropriate version-tracking and archiving over time."

Usage Costs: "One of the greatest barriers to access to publicly-available information ostensibly is the cost imposed on the public for access - even when the cost is *de minimus*."

Also in 2010, [ST10] reports the UK experience in publishing Linked Data and states that data formats as RDF/XML and Turtle are "impenetrable without special parsers, making it hard to use [the data] by non-experts". Due this difficulty, and to broad civic

participation, the UK government decided to provide data in CSV and XML formats. So we include **Providing Alternatives** of access for society as requirements, for better defines what should be done for the user in relation to Data Must Be Accessible.

[Ok12] in 2012 converges with the idea that data should be Available on the Web and License-free, and adds one more question:

Make data discoverable: "post on the web and perhaps organize a central catalog to list your open datasets". This catalog will have the function of a reference search for data, facilitating the discovery by citizens.

The big volume and complexity of governmental structure can make finding information a hard task, even impossible for citizens [Br11b]. Controlled vocabularies are a way to assist citizens in search for information. "Changes in the semantics of terms in Linked Data vocabularies can have a dramatic influence on the interpretation of remote datasets using them" [KA12]. Due this, it is important to point out which **Vocabulary Versioning** is being used, in the dataset that is being published. We consider it more one requirement.

[ZJJ12] claim that linking data is not enough, metadata is required to make effective use of data, metadata also facilitates use to create new hypotheses to improve the ability to find patterns. Therefore, we add the requirement **Use metadata**. Metadata can also be divided into collection and variable-level [Po09], discovery (flat), contextual and detailed [ZJJ12], supporting information such as license, update, provenance, permanence, together dataset to be published.

4 Mapping of requirements in layers of an Information System

In this section we associate the requirements raised previously with layers of its implementation in an information system. It is important to note that standardization efforts will transcend these limits, because many of the requirements depend on policy decisions reflected in legal frameworks, best practice recommendations among others.

Figure 1 shows a model of information system into four layers, it will be used in discussing where the requirement should to be implemented. The top-down design has been chosen by placing the focus of developing on end user, as shown in Figure 1 as client platforms that use the access channels. The presentation layer is accessed by client, where is defined formats and protocols for data that the client accesses through the channels by its platform. The application logic must deliver the content in the format specified in the presentation layer from resource management layer, which defines the data sources and the organization necessary to application logic.

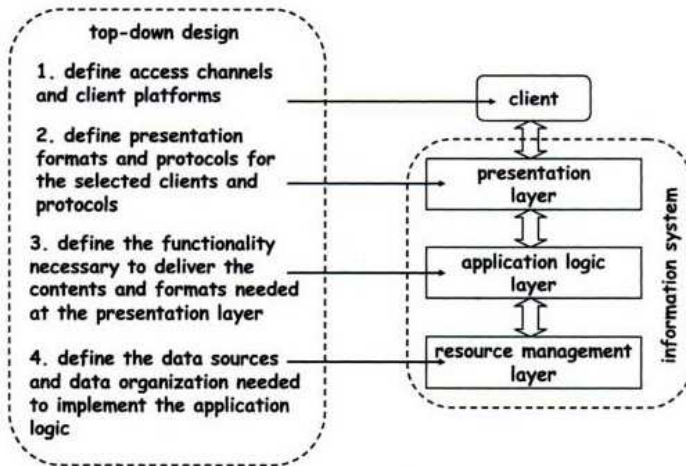


Figure 1: Information System model[AI04]

Table 1 lists the requirements that have been identified in the literature and where the requirements should be implemented in viewpoint of information systems in Figure 1.

The completeness and primarity of data should be presented to data user, but for this to be possible it is necessary that the data are complete and primary in the resource, because the application logic should not create aggregated information. Therefore, the data must be complete and primary in resource management and presentation layers. In the same way, the data structure that exists in resource management should be represented in the presentation layer to allow the client easily reorganizing data to (re) use in its analysis tools. This reuse can be improved by the use of metadata, which can only be presented to the client if they exist in the resource management.

Besides the completeness and primarity, issues like creation of data, access, data versioning, vocabularies versioning, update and duration of resources, including as Provenance Information [Ha09] and Permanence [SF10], should be implemented in resource management and properly presented to client as metadata.

The frequency that data is updated and the variety of formats that the same resource is presented depends on the application logic, considering that the resource exists in the resource management layer. It is important to note that some information produced in application logic should be represented in dataset because they can be copied and passed from hand to hand, from system to system, for example, if the update date are not represented on dataset, this information is lost in the redistribution of information. Therefore, some metadata are generated in the application logic and should be represented in the dataset.

Table 1 - requirements identified in literature

Requirement	Layer os IS
Data Must Be Complete	resource management, presentation
Data Must Be Primary	resource management, presentation
Data Must Be Timely	application logic, presentation
Data Must Be Accessible	application logic, presentation
Data Must Be Machine processable	resource management, presentation
Access Must Be Non-Discriminatory	application logic
Data Formats Must Be Non-Proprietary	presentation
Data Must Be License-free	presentation
Available on the Web	presentation
Identify things	presentation
Linked data	application logic, presentation
Provenance Information	resource management, presentation
Permanence	resource management, presentation
Usage costs	*
Providing alternatives	application logic, presentation
Make data discoverable	*
Vocabulary versioning	resource management, presentation
Use metadata	resource management, presentation

* this requirement is beyond the model of information system

In application logic layer is defined how content will be delivered in the presentation layer, so in this layer must be guaranteed non-discriminatory access to the data. The resource management and application logic can be implemented with proprietary technology, but the dataset must be in non-proprietary formats and have free license, be on the Web and be identified in the presentation layer.

The cost of data access should be the minimum possible, this requirement depends on the platform and client access channels, but provides guidance on how the information

should be presented and where (Web) in the information system. The requirement Make Data Discoverable is a service that should be implemented as data catalog, in other systems of information that does not contain the dataset, due it this requirement it is not directly associated with a standard, it's not related with a layer in Table 1.

Presented the requirements of each layer in an information system based on this perspective, the following are standardization initiatives in other contexts and also we suggests some technologies that can be used in implementation of these requirements for a data publishing standard of budget.

5 Some initiatives

Some organizations publish recommendations or standards without which it would be impossible to create a network with computers interoperable with each other. Standards are mandatory in communication systems.

The most important organization that publish recommendations of standards for Web is the World Wide Web Consortium (W3C¹). It recommends standards ranging from HTML used by end users in Web browsers, XML to create other patterns of marking data, XML Schema to define schemas for XML-based standards, and others correlated standards, also used in the business . The W3C also aggregates work groups related with government and linked data, as eGovernment Interest Group and Government Linked Data Working Group.

A standard based on W3C recommendations for publication of reports is the eXtensible Business Reporting Language (XBRL), which can be used for financial reporting through the Internet [Ri06]. [Si04], on managing information electronically by the Central Bank of Brazil, said XBRL makes the exchange of information more efficient, facilitates analysis and decision making.

Furthermore, XBRL uses XML Schema to represent taxonomies [Xb12], is machine processable and uses standardized vocabularies such as the three letter naming of currency ISO 4217. XBRL was also adopted by the National Treasury in the project SICONFI [Bra13], a system of accounting and tax information. Data published in XBRL can be converted into RDF (another W3C recommendation) and Web ontology, and published as Linked Data [GG09]." XBRL interoperable with Linked Data allows ecosystem evolution that will contribute to wider use and consumption of XBRL" [OCH12].

Another initiative, the project Popolo² is developing a specification that is based on other nine standards focusing on legislative data, including legislators, organizations and

¹ <http://www.w3.org>

² <http://popoloproject.com>

membership with the "goal is to make it easier for developers to create civic government transparency, monitoring and engagement websites, by developing reusable open source components that implement the specifications." This project also aims to create a permanent unique identifier for each instance of each class: person, organization, address, etc.

The International Aid Transparency Initiative (IATI) also created a standard based on XML (and XML Schema)³, used to share detailed data for their own projects. This standard supports transactions containing detailed information on financial flows, and also a geolocation element. In IATI there is also the concern to allow a direct translation from its standard for JSON and CSV formats, allowing multiple forms of access.

Even important, some standard data formats can be used as open data, but not as linked data from the perspective of the five star [Be09], others are not structured and can not be used as the default format. CSV, JSON and XML standards are semi-structured and can not directly be used as standards for linked data. Using XML Schemas, such as the standards of IATI, XML data become structured, the same happens with JSON data associated to JSON-Schema. However, CSV format hasn't a schema option, therefore, it's not a option to become structured.

Another issue is the ability to link data. XML data can be linked with other data using XPointer and XLink, both W3C specifications. However, JSON does not has capability to make links, it is necessary to use JSON-LD⁴. How reuse XML, XML Schema, XLink and XPointer in a specification, XBRL⁵ is a standard that can be used as Linked Data with pointed by [GG09].

6 Challenges encountered

Brazil is a federation with continental size and thousands entities including executive, legislative and judicial branches at federal, state and municipal level, that have theirs autonomy guaranteed in Brazilian Constitution Law [Br88]. All data that is generated by these entities, the Law No. 131 [Br09] only requires the publication of data on budget execution.[CSA12] did an extensive survey of published data of budget execution, on three levels of government, as discussed below.

We note that the publication of data in Brazil compared to the Linked Data model can not be classified as one star, because despite the datasets of budget execution were published on the Web, 0% of these had free license [CSA12]. However, the Access to Information Act [Br11a] ensures that all government information is public domain and considering this as free license, even if the datasets are not explicitly under a license we

³ <http://iatistandard.org>

⁴ <http://json-ld.org>

⁵ <http://xbrl.org>

can assess whether the data collected could be two stars. If we consider that structured data is only when there is an explicit scheme on them, the datasets are not two stars, because 49.27% of the datasets of budget in Brazil were published in CSV or XLS, the rest in PDF [CSA12], therefore, less than half of these sets can be considered machine-processable. Accepting these data as structured, only 30.07% of the datasets can be considered three-star, or used a non-proprietary format, but CSV format does not support links. Thus, it is not possible in this scenario have Linked Data in Brazil.

Brazilian law requires federal agencies to follow a classification of budget implementation for revenue and expenses. These datasets were 59.46% complete and 17.16% had all elements of this classification [CSA12], as disaggregated data. Furthermore, only 4.4% of the datasets of budget execution were updated.

Complete and primary data mandatorily must exist in the resource management layer, public managers must plan and execute the budget with the classification of revenues and expenses set out in law. A first challenge in Brazil is charge public managers to guarantee that the application logic layer solutions implemented extract all data from resources and represent complete and primary data in presentation layer. Another challenge is to convince managers to publish data into standards that are structured and can be linked. The lack of a standard also become hard the delivery of data in other formats, including different levels of aggregation in studied datasets [CSA12] that increase data usage cost.

In UK example, it is necessary that there are standards to identify data[ST10]. We not found any recommendation regarding identification of budget data in Brazil. There is a Brazilian catalog of data⁶ but without any guarantee that the locations of these data can not be changed, the failure on 'Identify things' weakens the requirement 'make it discoverable' on the catalog.

Of publishing formats found only XLS have some property fields on the file, but it is a poor choice to be placed metadata of provenance, collection, creation, access, update, versioning, etc. Therefore, we consider that the formats used in Brazil do not support the metadata requirements presented.

Datasets surveyed [CSA12], 86.73% were accessible, 98.82% were got in a non-discriminatory way and all were found on the Web, so the requirements associated with these issues are addressed in the Brazilian case.

⁶ <http://dados.gov.br>

7 Conclusions

This paper presents some elements on web publishing of budget execution data in Brazil and discusses the need for standardization that is made from the perspective of data consumer.

Were derived and discussed some requirements of literature about open government data and linked data and were subsequently mapped onto the layers of an information system for their subsequent implementation.

Set of 18 requirements raised in this paper, three are met in the Brazilian case: 'data must be accessible', 'access must be non-discriminatory' and 'available on the Web'.

The requirement 'make data discoverable' is met, but so weak because there are currently no regulations that place the need for 'identify things'. The remaining 14 are not met: 'data must be complete', 'data must be primary', 'data must be timely', 'data must be machine processable', 'data formats must be non-proprietary', 'data must be license-free', 'identify things', 'linked data', 'provenance information', 'permanence', 'usage costs', 'providing alternatives', 'vocabulary versioning', and 'use metadata'

On balance, are listed the major challenges presented in brazilian case for a publication budget and standardized data that provides real transparency of public policy investments.

This work does not present an exhaustive survey of the literature and its sequel will further review the field of data quality and information quality. As the guideline adopted emphasizes the vision of data consumer, other studies like Open Budget Index⁷, which performs the scoring of budget transparency in world should, also be explored.

References

- [Al04] Alonso, G.: Web Services. Springer-Verlag Berlin Heidelberg, 2004
- [Be09] Berners-Lee, T.: Linked Data. 2009. Available: <<http://www.w3.org/DesignIssues/LinkedData.html>>. Accessed: 10 mar 2012.
- [Br13] Brasil. Tesouro Nacional. Grupo Técnico de Sistematização de Informações Contábeis e Fiscais: Projeto SICONFI: Taxonomia XBRL. 2013. Available: <https://www.tesouro.fazenda.gov.br/images/arquivos/Responsabilidade_Fiscal/Contabilidade_Publica/arquivos/Apresentacao_GTSIS-Taxonomia.pdf>. Accessed: 11 mar 2013.
- [Br11a] Brasil: Lei nº 12.527. 2011. Available: <http://www.planalto.gov.br/ccivil_03/_Ato2011-2014/2011/Lei/L12527.htm>. Accessed: 20 feb 2012.
- [Br09] Brasil: Lei Complementar nº 131. 2009. Available: <http://www.planalto.gov.br/ccivil_03/Leis/LCP/Lcp131.htm>. Accessed: 20 feb 2012.

⁷ <http://internationalbudget.org/what-we-do/open-budget-survey/rankings-key-findings/rankings/>

- [Br10] Brasil: Decreto nº 7.185. 2010. Available: <http://www.planalto.gov.br/ccivil_03/_Ato2007-2010/2010/Decreto/D7185.htm>
- [Br11b] Brasil, Ministério do Planejamento, Orçamento e Gestão, Secretaria de Logística e Tecnologia da Informação: Vocabulário Controlado do Governo Eletrônico. 2011. Available: <www.governoeletronico.gov.br/biblioteca/arquivos/vocabulario-controlado-do-governo-eletronico-vcge/download>
- [Br88] Brasil: Constituição da República Federativa do Brasil de 1988. Available: <www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm>
- [CSA12] Craveiro, G. S., Santana, M. T. e Albuquerque, J. P.: Assessing Open Government Budgetary Data in Brazil. In The Seventh International Conference on Digital Society 2013.
- [GH07] Ge, M., Helfert, M.: A Review of Information Quality Research - Develop a Research Agenda. In MIT Conference on Information Quality - IQ, pp. 76-91. 2007.
- [GG09] Garcia, R., Gil, R.: Publishing XBRL as Linked Open Data. In Linked Data on the Web. 2009. Available: <http://events.linkedata.org/ldow2009/papers/ldow2009_paper6.pdf>
- [Ho09] Hartig, O.: Provenance Information in the Web of Data. In Linked Data on the Web. 2009. Available: <http://events.linkedata.org/ldow2009/papers/ldow2009_paper18.pdf>
- [Ma09] Madnick, S. E., Wang, R. Y., Lee, Y. E., Hongwei, Z.: Overview and Framework for Data and Information Quality Research, 2009, ACM Journal of Data and Information Quality. v.1, n.1.
- [Kä12] Käfer, T., Umbrich, J., Hogan, A., Polleres, A.: Towards a Dynamic Linked Data Observatory, 2012. In 5th Linked Data on the Web Workshop at WWW.
- [Mc13] McKinney, J.: Legislative Open Government Data Standards. 2013. Available: <<https://www.newschallenge.org/open/open-government/submission/legislative-open-government-data-standards/>>. Accessed: 25 mar 2013.
- [Oe02] OECD: OECD Best Practices for Budget Transparency. OECD publications, France. 2002. Available: <www.oecd.org/gov/budgetingandpublicexpenditures/Best%20Practices%20Budget%20Transparency%20-%20complete%20with%20cover%20page.pdf>
- [OCH12] O'Riain, S., Curry, E. e Harth, A.: XBRL and open data for global financial ecosystems: A linked data approach, International Journal of Accounting Information Systems, Volume 13, Issue 2, June 2012, Pages 141-162, ISSN 1467-0895, 10.1016/j.accinf.2012.02.002.
- [Og11] Open Government Partnership: Open Government Declaration. 2011. Available: <<http://www.opengovpartnership.org/open-government-declaration>>
- [Og07] Open Government Working Group: 8 Principles of Open Government. 2007. Available: <<http://www.opengovdata.org/home/8principles>>
- [Ok11] Open Knowledge Foundation: Beyond Access: The Right to (Re)Use Public Information (p. 8-9). 2011.
- [Ok12] Open Knowledge Foundation: Open Data Handbook Documentation – Release 1.0.0, 2012. Available: <<http://opendatahandbook.org/pdf/OpenDataHandbook.pdf>>
- [Po09] Powell, L.: Building a strong foundation with Dublin Core. In Proceeding DCMI '09 Proceedings of the 2009 International Conference on Dublin Core and Metadata Applications, p. 87-95.
- [Ri06] Riccio, E.: Introdução ao XBRL – nova linguagem para a divulgação de informações empresariais pela internet. 2006. Available: <www.scielo.br/pdf/ci/v35n3/v35n3a16.pdf>. Accessed: 20 mar 2013
- [Sf10] Sunlight Foundation: 10 principles of Open Government Data. 2010. Available: <<http://sunlightfoundation.com/policy/documents/ten-open-data-principles>>

- [Si04] Silva, P. C.: A gestão eletrônica da informação financeira no Banco Central do Brasil apoiada por XBRL e regras de inferência. 2004. Available: <www4.bcb.gov.br/pre/inscricaoContaB/trabalhos/A%20Gest%C3%A3o%20Eletr%C3%B4nica%20da%20Informa%C3%A7%C3%A3o%20Financeira%20no%20Banco%20Central%20do%20Brasil%20Apoiada%20por%20XBRL%20e%20Regras%20de%20Infer%C3%Aancia.pdf>. Accessed: 25 mar 2013.
- [So04] Solana, m.: Transparency Portals - Delivering Public Financial Information to Citizens in Latin America, 2004, Thinking out Loud, Vol. 5 Spring, p. 71-81.
- [ST10] Sheridan, J. e Tennison, J.: Linking UK Government Data. In Linked Data on the Web 2010, Raleigh, North California.
- [Us13] United States of America: Open Data Policy-Managing Information as an Asset. 2013. Available: <<http://www.whitehouse.gov/sites/default/files/omb/memoranda/2013/m-13-13.pdf>>
- [WS96] Wang, R. Y., Strong, D. M.: Beyond accuracy what data quality means to data consumers, 1996, Journal of Management Information Systems, v. 12, n.4, p. 5-34.
- [Xb12] XBRL International: Extensible Business Reporting Language (XBRL) 2.1. 2012. Available: <<http://www.xbrl.org/specification/xbrl-recommendation-2003-12-31+corrected-errata-2012-01-25.htm>>. Accessed: 08 jan 2012.
- [ZJJ12] Zuiderwijk, A., Jeffery, K. e Janssen, M.: The necessity of metadata for linked open data and its contribution to policy analyses. 2nd Conference for E-Democracy and Open Government (CeDEM 2012). Danube-University Krems, Austria, pp 281-294.
- [Zh12] Zhu, H., Madnick, S. E., Lee, Y. W., Wang, R. Y.: Data and Information Quality Research: Its Evolution and Future, 2012, In MIT Information Quality Program Proceedings. Available: <http://mitiq.mit.edu/Documents/Publications/Papers/2012/Madnick_2012_Data%20and%20Information%20Quality.pdf>