



**Power to the people!  
Introducing open linked data services to the National  
Legal Database of the Library of the National Congress of  
Chile**

**Christian Sifaqui**  
Biblioteca del Congreso Nacional de Chile  
Valparaíso, Chile

**Session:**

**214 — *Libraries for the law and for lawmakers* — Law Libraries with  
Library and Research Services for Parliaments**

**Abstract:**

National legal databases aim to democratize the use of the law by providing easy, free and up-to date access to legislation. In 2008 the Chilean Library of Congress released the system LeyChile, a legislative database which contains the whole text of legal norms, their versions (available from 1998), as well as their amendments (modifications, regulations and concordances). The first objective of this system was to bring a solution to the “juridical certainty” for the National Congress of Chile and also citizens in general, especially with regard to in-force texts. In other words to ensure that both Parliament and citizens have access to the latest and most up-to date version of a law.

LeyChile offers a search box that allows for free text searching, and also web services that provide the norms in XML and other complementary services and applications such as widgets/gadgets to facilitate the consumption and use of the legal information stored in the database. Upon its launch, LeyChile was taken on enthusiastically by the Parliamentary Community, legal professionals and citizens alike. It has become the authoritative resource when it comes to citing a law, being used by a diverse universe of users such as official bodies or tweeters.

In 2011 the Library introduced open-linked data as a new service. To this end we defined an ontology of legal norms stored in LeyChile. This ontology was written using both RDF Schema and OWL, making the application of inferences to RDF graphs possible. Our open-linked data portal also provides full access to the RDF data through an SPARQL endpoint, allowing so the creation of new services and visualizations.

# 1. Introduction

The Chilean Library of Congress (BCN) subscribes the concept of Open Government because it considers it is a useful philosophy for working at empowering citizens and providing them with access and licenses to use the data generated by public entities, in a manner that they can use, store, redistribute and integrate to other data sources. This opening of information is justified in providing means to favor citizen participation, strengthening democracy and being a driver for innovation by allowing the creation of new initiatives and business.

Today the concept of Open Government is related to the concepts of open data and linked data<sup>1</sup>. We understand that, while the concept of “Open Data” is oriented to data that must be available to all, in a free and unrestricted manner, “Linked Data” is a way of publishing data that facilitates the relationship between different data sources.

In 2010, the World Wide Web Foundation<sup>2</sup> through the Center of Technology, CTIC, carried out a study on the viability and potential of an Open Government Data (OGD) in Chile, Ghana y Turkey. The report [1] indicates that Chile shows favorable conditions for liberating public information for OGD, recommending its development, given the knowledge, people, technology and disposition to do so.

As the Report indicates, BCN believes that state organizations in Chile are prepared to fully embark in OGD. In this line, the Library has realized the following actions.

- Free access Portals / Sites: All of the information that BCN provides in the portals ([www.bcn.cl](http://www.bcn.cl), [www.leychile.cl](http://www.leychile.cl), etc.) is complete, reliable, and reusable, can be distributed, allows for interoperability and is based upon the concept of a common cultural heritage.
- Markup: BCN places marks and semantic attributes on digital objects to obtain precise results, related between themselves. This process of “semantification of the information” allows search engines to access with precision, the information available in our portals. .
- Widgets/Gadgets: BCN provides in its portals [www.bcn.cl](http://www.bcn.cl) y [www.leychile.cl](http://www.leychile.cl) a set of computer applications in web platform as a means to provide simple, automatic, distributed access.
- Semantic Web: As a starting point in the use of technologies that support this concept, BCN released the site <http://datos.bcn.cl>, that offers datasets with public ontologies, to facilitate the computerized automatic analysis, and the data is being published in the linked data model (RDF). Some of them can be accessed with the SPARQL query language. Today, <http://datos.bcn.cl> includes data from the LeyChile portal and will be gradually including other datasets, such as Parliamentary Biographical Profiles. The centralized publication of <http://datos.bcn.cl> is also going to include de markup of contents with RDFa in several of the portals managed by BCN, which will facilitate the distributed publication of data and its latter integration with <http://datos.bcn.cl>.

This article describes de design of the <http://datos.bcn.cl> project and the services available to the citizens from the legal database [www.leychile.cl](http://www.leychile.cl).

---

<sup>1</sup> Linked Data - Connect Distributed Data across the Web [www.linkeddata.org](http://www.linkeddata.org)

<sup>2</sup> World Wide Web Foundation. <http://www.webfoundation.org/>

## 2. LeyChile

Legislation is defined as the set of positive norms (acts, laws, decrees, resolutions, etc), which make up the national legal code.

Legislation is information that comes from the public sector, is generated by the State Organs and is financed by public resources. It is of public interest, operating on general use subjects that affect the lives of citizens. This information is of high public value because it provides an experience that citizens consider valuable. Having said this, legislation is public information and therefore must be in the public domain.

The concept of “concretion mechanism of the principle of security” can be derived in a natural way, supported on the expectations of citizens that the in-force legislation must be complied. To meet this expectation, two points of view must be considered: the concept of “legal security” (founded and guaranteed certitude that the norm will be respected), and the concept of “legal certainty” (perceptibility and certitude of the content of the legal norm). To provide for this, countries provide the means of publicizing the law, for both points of view, with the Official Law Bulletin or Official Gazette.

In the Chilean case, three articles of the Civil Code refer to this publicity mechanism, and the other mentioned outlooks:

- Art. 7°. The publication of the law is in the Official Gazette, and from that date, it is understood that it is known by all, and mandatory.
- Art. 8°. Nobody can argue ignoring the law after it has coming into force.
- Art. 706° [...] an error in matters of the law is presumption of bad faith, and does not admit proof of the contrary.

The articles mentioned lead to what is known as “legal fiction of knowledge”, since in Chile the access to the Official Gazette is paid for, and the modifying laws are published instead of the current complete texts, so the access to previous norms to have the complete text is difficult.

BCN began in the fifties a card based compilation of references of links between norms and classified them by subject as a way to solve the “legal certainty” for the Chilean National Congress as well as for the citizens.

In the eighties the card system was replaced by an automated STAIRS based system, to automate the cards and its annotations. In midst nineties, the system was replaced by a client-server BASIS PLUS architecture, which could reconstruct online the whole texts of the norms.



Fig. 1. Homepage of www.leychile.cl

The LeyChile web system was released in 2008 (see figure 1), which holds all the legal norms in full text and their versions (available since 1998), as well as the links (modifications, regulations, combined texts and concordances). This system offers a search box as well as web services with XML formatted text. It also provides additional services such as widgets and gadgets for an easier consumption and use of legal information stored in this database.

Today, the system receives an average of 14.000 visits per day, with peaks of 20.000 visits. This is quite high, considering the type of contents and that there are 7.3 million internet users in Chile [2].

## Details of the implementation

The basic concept considered was that the norm is not a static object. On the contrary, during its life cycle, from publication in the Gazette until eventual repeal or recombination, it suffers modifications. As shown in figure 2, a norm can be modified by another norm (modifying norm), which generates a new version. The XML model of the norm allows the modification to parts (heading, structure, promulgations, annexes, see figure 3), which are affected by the modifying norm. Thus, LeyChile dynamically builds the complete text of a version of the norm (current version, intermediate versions or original version).

Another detail is that in some cases the Executive Branch generates an official version of the norm, called the combined norm, with the objective of systematizing, coordinating and organizing the content of a norm that has had an important amount of modifications. LeyChile models this type of association (link), between the combined text and the originating norm.

It is worth noticing that the normative document model is in its conclusive stage, (promulgated), without considering the legislative iterations.

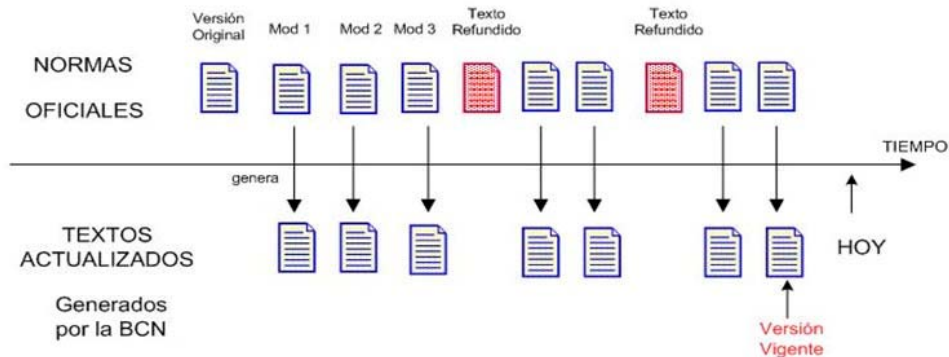


Fig. 2. Model, life cycle of a norm

Based upon these details, the XML schema of LeyChile has three layers or mark levels.

- **Text:** versions, hyperlinks, references, notes.
- **Structure:** Hierarchical organization of the parts of a norm (see figure 3).
- **Metadata:** additional information of the formal document, such as identification of the norm, subjects, free retrieval terms, etc.

This strategy permits different actors within the organization to “enrich” the legal text, correct and increase these marks as necessary. LeyChile has a production module that manages the workflow of specialized analysts that complete the markup of the norm.

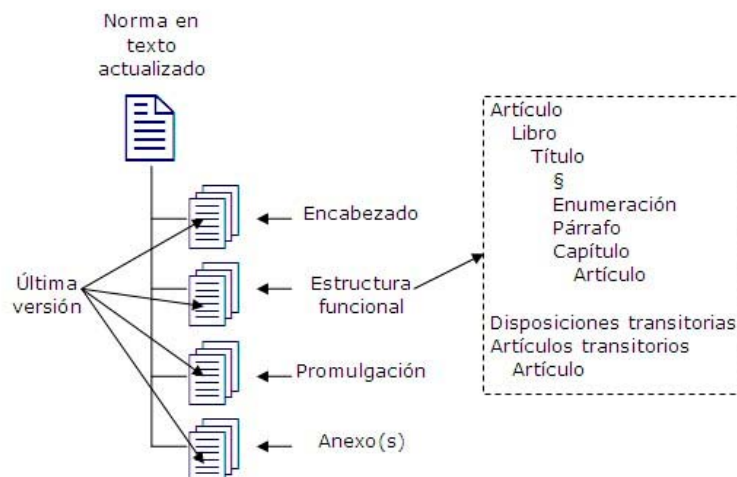


Fig. 3. Model of the structure of a legal norm

One of the greatest challenges for LeyChile is to offer system interoperability with external and internal systems. This is done with web services. The web services use the message languages in XML. Some services use the REST mechanism and others the SOAP protocol. By using this integration standard, the interaction between a specific technology within each application of BCN and the technologies in each

institution can be made independent. Figure 4 shows the SOAP services definition. For example, getNorma permits any system in the web to consume the exchange<sup>3</sup> XML of a particular norm.

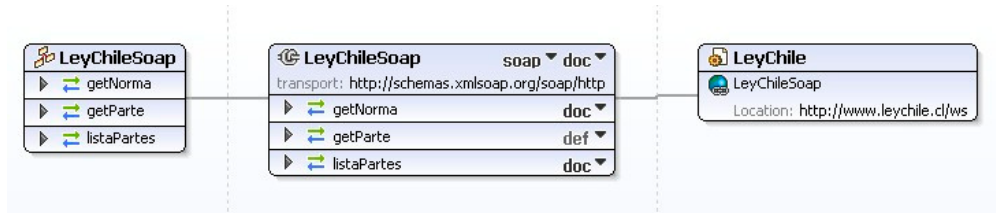


Fig. 4. SOAP, available in <http://www.leychile.cl/ws/LeyChile.wsdl>

Most services are invoked by an HTTP call, providing an XML as answer. For example, last published laws<sup>4</sup>, bills sent by Congress to the Executive or the Constitutional Court, norm<sup>5</sup> metadata, amongst others.

In this way, the user of LeyChile is not limited to using search boxes only, to obtain answers to more complex requests.

### 3. datos.bcn.cl

It has the objective of providing citizens with access to data sources such as Linked Open Data. It is operational since May 2011 and is the first BCN initiative related to the publication of linked data (see figure 5).



Fig. 5. Homepage of <http://datos.bcn.cl>

Offering this new layer of services, all the system will be a reference model in terms of entities and rules, published as ontologies, including the possibility of executing complex information searches with

<sup>3</sup> Accepted schema as the standard model of the chilean norm for interchange, by the Administradora de Esquemas y Metadatos (inscripción 354) <http://www.aem.gob.cl/index.html>

<sup>4</sup> <http://www.leychile.cl/Consulta/obtxml?opt=3&cantidad=5>

<sup>5</sup> <http://www.leychile.cl/Consulta/obtxml?opt=3&cantidad=5>

information of legal norms through an endpoint SPARQL, allowing access to results in several output formats such as RDF/XML, JSON, HTML+RDFa or N3.

In a first stage, the process of contextualization was developed for a very punctual domain, that of norms, to offer a new layer of services for LeyChile. Each layer of services will provide access of the type of open linked data, to the metadata of each norm in [www.leychile.cl](http://www.leychile.cl)

In the contextualization process, a document was prepared that describes the three principal elements of the context: the data to be provided; the way of providing them and who will consume them. According with the previous questions, and in a very synthesized way, the data to be provided is the metadata of the norms and their relationships, without considering the internal structure of a norm. The way of providing the data is by means of a RDF graph on HTTP and an Endpoint SPARQL. Lastly, those who consume the data are visualization applications of the Library and applications oriented to searching laws that can be implemented by public administrations and by the community. An ontology was defined later (see figure6), and a namespace for the ontology of norms, in the particular national reality context.

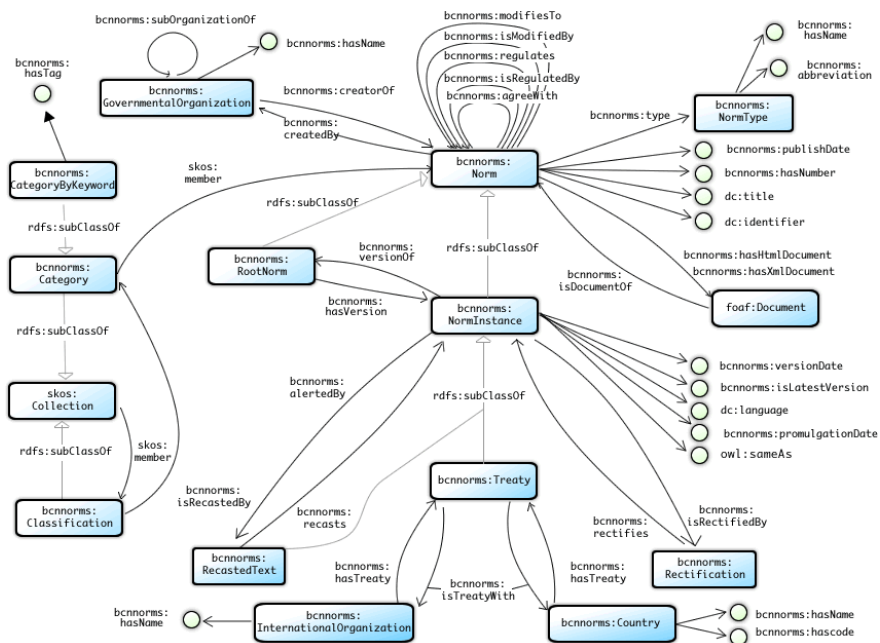


Fig. 6. Ontology of norms

An ontology structure extensible to other domains such as congress, education, health and others, has been considered. This ontology has been written using RDF schema and OWL, permitting RDF graph inference application. Another important feature of this ontology is that it has been built by using previous ontologies and datasets such as SKOS, Dublin Core, FOAF, Geonames, Organization and DBPedia. In the two last mentioned, it was possible to link data from the graph of legal norms to sets of external data, specifically international treaties and countries. This task was not trivial because it required an intense manual intervention. Finally, the ontology was stored in the RDF store, to allow inferences such as those already published using text files in RDF/XML and syntax N3, whereas the documentation was published in English and Spanish.



Once the ontology was structured, the RDF output graph was modeled. In practice, an URI schema with all the possible URI patterns for valid searches was defined. Figure 7 shows an example of a URI pattern which considered the use of standard IFLA FRBR as a URI of legal norms.

`cl/{type}/{organization}/{publish_date}/{number}/data.{ext}`

Fig. 7. URI pattern

In general terms, the graph follows a hierarchical scheme in each of the resources available for searches. On the other side, some searches were modeled (for example, obtaining legal norms of a specific date). Thus, for each URI pattern an RDF output using syntax N3 was defined. Finally, the output formats for the resources were defined. For this project, they were RDF/XML, JSON, Ntriples, N3 y HTML+RDFa.

The process of transformations and data loading was then built. An update service in Java using the API Kettle was built for loading, updating and transformation. Thus, with the ETL designer, different transformations were implemented, that generate RDF triples in N3 syntax for the initial loading and the updating of the triples (which usually will only add new triples). In these conditions, the updating service executes transformations and loads the triples in the RDF store.

In the next stage, the RDF output graph on HTTP was implemented, according to the model designed for this purpose. The WESO DESH tool, which is a front end for linked data was used for the implementation. This will be shortly liberated as open software. Finally, this implementation of linked data was certified by linked data validators such as Vapour of the Fundación CTIC and RDF/XML of W3C.

The project has a documentation web portal with new documentation permanently included in Spanish and English about using the Linked Data infrastructure.

It is worth noting that this project was entirely developed with open source software, so a similar endeavor could be replicated without paying for software licenses.

On feature of this project is that given the nature of the data, it is normal to encounter discrepancies in the styles of writing or typing errors, so there was special care in the design of the model for manually editing data in the future. For example, the instance GovernmentalOrganization can be mentioned, because the name variations for the same resource, different classes were defined.

A complete description of the architecture defined for the development of the project is described in [3].

At the end of section 2, it was mentioned that [www.leychile.cl](http://www.leychile.cl) offers different mechanisms for searching, that allow users to access resources in different ways, but there is a limitation as to the complexity of searches.

It is here that <http://datos.bcn.cl> shows its potential by allowing access to execute more complex searches. Let us suppose that you want to retrieve all of the norms generated by a municipality between 1995 and 2000, which have been modified after 2005. This search cannot be done with [www.leychile.cl](http://www.leychile.cl) in spite of the multiple parameters available in its advanced search.

But, they can be solved with the following SPARQL search, written directly in <http://datos.bcn.cl>:

```
PREFIX dc: <http://purl.org/dc/elements/1.1/>
PREFIX n: <http://datos.bcn.cl/ontologies/bcn-norms#>
```



```

SELECT distinct str(?normTitle) as ?Titulo str(?creatorName) as ?Municipio ?pubDate as ?Fecha_Publicacion
?pubDateOther as ?Fecha_modificacion
WHERE {
  ?norm n:createdBy ?creator .
  ?creator n:hasName ?creatorName .
  ?norm dc:title ?normTitle .
  ?norm n:publishDate ?pubDate .
  ?norm n:isModifiedBy ?otherNorm .
  ?otherNorm n:publishDate ?pubDateOther .
  FILTER (regex(?creatorName,"MUNICIPALIDAD","i"))
  FILTER (?pubDate > xsd:date("1995") && ?pubDate < xsd:date("2000") && ?pubDateOther > xsd:date("2005"))
}
ORDER BY (?pubDate)

```

In this way, the ordinances that meet the criteria are retrieved.

## 4. Visualizations

<http://datos.bcn.cl> represents a good technological advance and offers unrestricted access to the stored datasets. But as seen at the end of the previous section, generating searches in SPARQL languages is as arid as when HTML pages were written in the nineties.

In [4] it says “Open data does not create its own demand. The Government must commit data developers”. Technology and datasets, and the development made <http://datos.bcn.cl>, is a necessary condition to develop new applications or new services.

One way for attracting developers to published datasets is also attracting interest in the development of showing visualizations of these datasets and the possibility of linking information available with other data sources. Two visualizations were developed for this, described as follows:

### Law links graph

This visualization (see figure 8) shows relationships that BCN has defined in the metadata of the main norm.

The objective is to show, in a simple manner, the relations between norms. To obtain this, it is illustrated as a directed graph, representing with a circle each norm, and the edges in different colors, the relationship between norms. The central node represents the chosen norm.

The user inputs the number of the law in the search box. Then, the algorithm sends a SPARQL query to our endpoint and the results obtained are shown graphically. In this way you can rapidly see the norms that modify the main norm, its regulations and related norms. By pressing on each node the metadata of the norm is reached, within which there is a link to [www.leychile.cl](http://www.leychile.cl) which will show the complete text of the chosen norm.

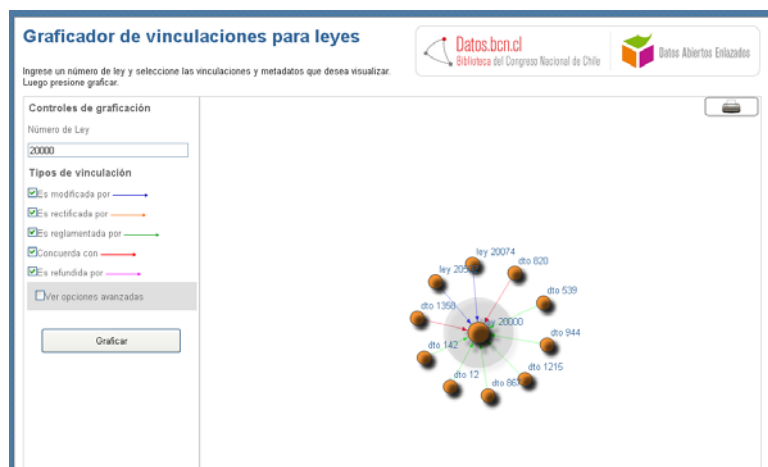


Fig. 8. Visualization of all of the linked norms for law 20.000

## County ordinance map

Another visualization developed (see figure 9) allows the retrieval of county ordinances by county, navigating the map of Chile obtained from Google maps. Each blue dot in the map indicates the center of the county. When a county is selected, a window appears showing the ordinances available in the database. The red area shows the comparative amount of municipal ordinances of the county in our database. If the size of the red circle is large, it indicates a larger amount of available ordinances.

For this visualization, we had to do a previous metadata analysis. Each ordinance has the name of the county as metadata, so a list of counties had to be prepared. Then, the counties had to be georeferenced. A County is defined by a polygon and the georeference consisted in finding the latitude and longitude of the center of the polygon which represents the county, and then these coordinates were given to the map in Google maps. .

When the user chooses a county, a SPARQL search returns the list of ordinances for that county. Then the titles are visualized in a window and the user then is sent to [www.leychile.cl](http://www.leychile.cl) to see the complete text.



Fig. 9. Visualization of ordinances by county

At present, lines that indicate the limits of the counties (polygons) are in development.

## 5. Conclusions

BCN has been introducing semantic web standards gradually and recurrently developing practical applications. This strategy has permitted BCN to travel through the different states that lead to the implementation of a semantic architecture:

- Universal identification (URI) and a set of universal characters (Unicode)
  - BCN web platform
- Representation, exchange and metadata formats
  - XML: [www.leychile.cl](http://www.leychile.cl)
- Linked data, endpoint
  - RDF, SPARQL: <http://datos.bcn.cl>
- Semantic modeling
  - OWL: Ontology for LeyChile

These projects have provided experience and grounds for BCN to move successfully in the path towards the Semantic web.

Other areas for development are mentioned in [5], such as the RDFa offer in our Transparency portal, and the intensive use of semantic web technologies for a new project, of a new system for History of the Law and Parliamentary Work. As to interoperability, a RIF application is being studied to relate different legal data bases in other State organisms.

As indicated at the beginning of this article, BCN has in its ADN to “semantize” its holdings, and the technologies used in this perspective have allowed a development to offer new products and services for Congress, citizens, and all of the country. We believe this experience can be replicated in other organizations of the State, and thus achieve a real Open Government. We also believe in our role of creating an environment that provides to every type of user more flexibility for using a wide range of software tools to manipulate and add value to our institutional information resources..

## 6. Bibliography

- 1.- IGLESIAS, C. (Ed.) Open Government Data. Feasibility Study in Chile. 2011, [http://public.webfoundation.org/2011/05/OGD\\_Chile.pdf](http://public.webfoundation.org/2011/05/OGD_Chile.pdf)
- 2.- FOSK, A. Latin america's internet population grows 15 percent in past year to 112 million people. [en línea] Marzo 2011, [http://www.comscore.com/Press\\_Events/Press\\_Releases/2011/3/Latin\\_America\\_s\\_Internet\\_Population\\_Grows\\_15\\_Percent\\_in\\_Past\\_Year\\_to\\_112\\_Million\\_People](http://www.comscore.com/Press_Events/Press_Releases/2011/3/Latin_America_s_Internet_Population_Grows_15_Percent_in_Past_Year_to_112_Million_People)
- 3.- CIFUENTES F., SIFQUI C. y LABRA J. 2011. Towards an architecture and adoption process for Linked Data technologies in Open Government contexts – A case study for the Library of Congress of Chile. En: Proceedings of the 7th International Conference on Semantic Systems, 2011. pp 79-86.
- 4.- ROBINSON, D. y YU, H. Apertura de datos: primeras lecciones para el diseño de políticas. En: CONCHA G. y NASER A. (Eds.) El desafío hacia el gobierno abierto en la hora de la igualdad. Santiago, Cepal, 2012, pp 87-99. <http://www.eclac.cl/ddpe/publicaciones/xml/9/46119/W465.pdf>

5.- SIFQUI C., OTTO E., ALMAZAN F. y HERNANDEZ D.. 2011. En camino hacia la Web semántica: experiencias de la Biblioteca del Congreso Nacional de Chile. Bits de Ciencia, 6: 35-43. 2011 [http://dcc.comopapel.com/revista\\_bits\\_de\\_ciencia/6/#/page/34-35](http://dcc.comopapel.com/revista_bits_de_ciencia/6/#/page/34-35)