

## ASPECTS OF PUBLISHING LINKED DATA

**Nicolae Adrian MATEIA**

***Abstract :** Web service users must share the same conceptualization of information (moreover, both humans and machines must be able to use their knowledge in the same way). This implies the existence of a common vocabulary, a common set of constraints and a common conceptual framework offering relationships between classes of information. The common vocabulary ensures that syntactic elements (words, markers etc.) have the same meaning for both the transmitter and the recipient.*

***Keywords:** open data, linked data open government data, triplets, data publication*

**1. Introduction** The concept of Open Data (OD) is rooted in the concept of transparent and participatory governance. Thus, open data should constitute into a bidirectional transfer between the administration and citizens. [1], [2]

Open data, and open government data (OGD), represents an important resource for various entities that collects a large number of data types to meet various objectives. Of these entities, the government administration occupies a significant place within this context, not only due to the data sets quantity and centralization, but also because most government data is public by law, thus being open and publicly available.[5]

Open public data are data produced or held by the public authorities, and are available to citizens free of charge, free to access, reuse and redistribute. They are mainly characterized by the following:

- are processed through automated means,
- are delivered in an open format, over which no entity has exclusive control, and benefit of an open license too, which allows free use of data, including that for commercial purposes, without being limited by intellectual property rights.

Linked Data is a standard way of representing various topic data. Publication of interconnected data gives developers the possibility to link information from different sources, creating thus new and innovative applications.[1]

The basic idea of open government consists in creating a link between politicians, government, industry and citizens aiming the transparency, the democracy, the participation and the collaboration of these entities.

Government data are considered open if made public and respect the following principles [1], [2]:

- Data must be complete
- Data should be primer
- Data must be genuine
- Data must be accessible
- Data should automatically be handled
- The access must be non-discriminatory
- Data is available in a format over which no entity has exclusive control

## 2. Data Publication

RDF (Resource Description Frame) is a framework intended to process metadata, providing interoperability between different applications, which exchange information intelligently, in the sense that the machine understands its semantics. RDF uses XML syntax to syntactically represent the data. One of the framework's purposes is to enable data's semantic specification, based on XML, in an extensible, standardized way, independent of the machine.

The basic model is built using:

- *Resources*: Data described by RDF expressions, which can be web pages, part of a page or an object that is not directly accessible.
- *Properties*: A specific aspect, feature, attribute or relationship, it has a certain semantic but also a set of relationships with other properties.
- *Declarations*: a specific resource together with a property, having a value associated to it.

Additionally, Notation 3 (N3) was introduced as a non XML serialization of RDF models designed to be easier to write manually and in some cases easier to follow. Because it is based on a tabular notation, triplets coded within the document are easier recognizable in comparison with XML serialization.

The publication of open data involves the use of the above principles (standards). Accordingly, interconnected data should allow a higher operability and reuse within the Internet. However, keeping the principles of interconnected data does not involve ditching the existing data management systems and their related applications, but adds superior techniques for connecting to the Web database. There are numerous technical solutions that can be connected to the Web database, but the mechanism behind it is presented in the following figure:

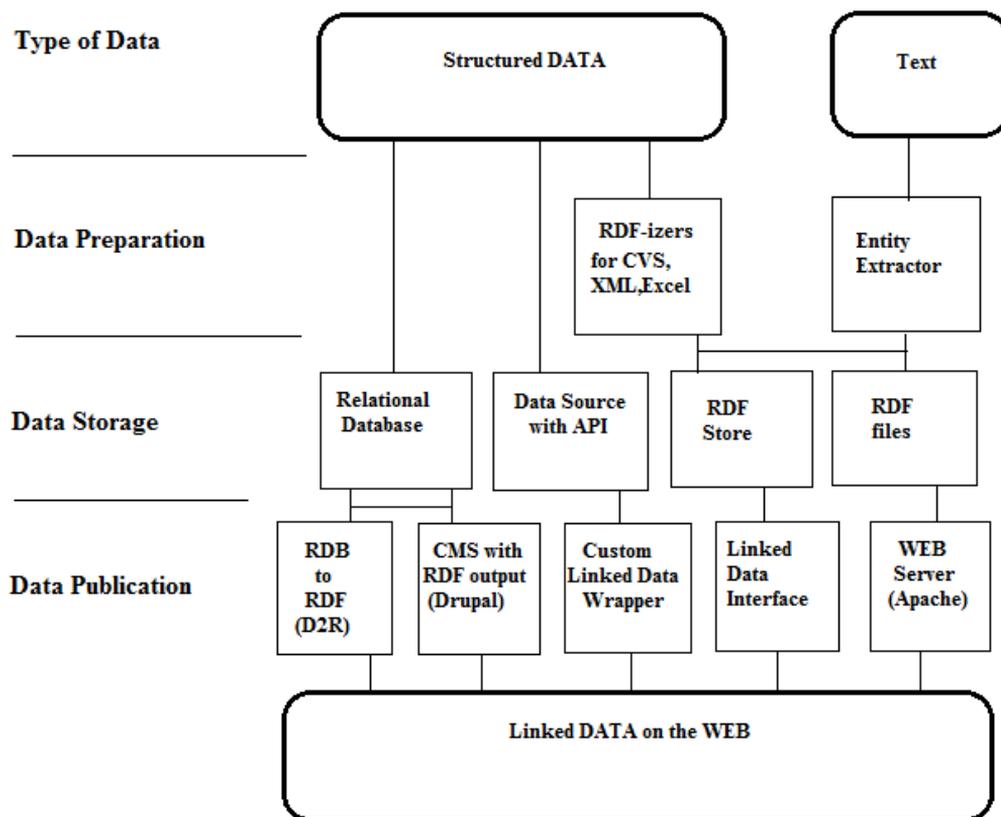


Fig. nr.1. Interconnected data publication and work flux  
 Source: <http://linkeddatatoolkit.com/editions/1.0>

If the data is stored in a relational database, viewing it as interconnected data can be done using a D2R server. It is based on a declarative mapping between database scheme and the imposed conditions of the RDF.

To use it, it is necessary to install the software, then customize the mapping through replacing some terms (well known) in RDF vocabularies, and adding the new ESW Wiki source to the data type list.

Mapping, between a database scheme and a RDFS vocabulary or OWL ontology, is expressed in declarative language as cartography (mapping). The language is designed to provide the necessary expressiveness necessary to create RDF views for a particular target vocabulary given the raw data found online. The language itself is expressed in RDF, and is usually written down as a N3 file. *ClassMaps* defines a set of resources and the way these resources are identified. *PropertyBridges* takes values from the database and attaches itself as the properties of these resources, or link to other resources.

### 3. Conclusions

The interconnected data should allow superior operability and reuse within the Internet environment. However, keeping the principles of interconnected data does not involve ditching the existing data management systems and their related applications, but add superior techniques for connecting the Web database.

Thus, data stored in a relational database, allows their visualization as interconnected data with D2R Server's (or other similar applications) help. This one is based on a declarative mapping between the database schema and the imposed conditions of the RDF.

### Bibliography

- [1.] Bauer F., Kaltenbock M., *-Linked Open Data: The essentials. A Quick Start Guide for Decision Markets*
- [2.] T. Berners-Lee T., J. Hendler, O. Lassila, "The Semantic Web", *Scientific American*, 5, 2001
- [3.] Daum B., U. Merten, *System Architecture with XML*, Elsevier Science, 2003
- [4.] M. Smith *et al.* (eds.), *OWL Web Ontology Language Guide*, W3C Recommendation, 2004: <http://www.w3.org/TR/owl-guide/>
- [5] \* \* \*, <http://opendatahandbook.org/ro/why-open-data/index.html>
- [6.] \* \* \*, *OWL-S Specification*: <http://www.daml.org/services>[7.] \* \* \*, *World-Wide Web Consortium*: <http://www.w3.org/>

Notes on the author:

Conf.univ.dr. Mateia Nicolae Adrian - Universitatea Europeană Drăgan din Lugoj