3
Related Work

The related work is divided into two main approaches: the OLAP data approaches and the Linked Data with RESTful approaches.

3.1 OLAP data approaches

(Wood 2011) discusses both the costs and benefits of using linked data techniques with government data. Chapter 6 (StdTrip: Promoting the Reuse of Standard Vocabularies in Open Government Data) discusses the mapping from database concepts to a vocabulary for generating the RDF representation and presents the StdTrip framework (P. E. Salas 2011), a tool that supports the conceptual modeling stages by reusing the W3C RDF vocabularies. Chapter 7 (Official Statistics and the practice of data fidelity) written by Richard Cyganiak, Michael Hausenblas, and Eoin McCuirc presents the field of official statistics and discusses the modeling of statistical data in RDF including its integration with other kinds of government data, tools for data conversion and publishing of statistics as Linked Data.

(Kämpgen and Harth 2011) presented a mapping from statistical Linked Data, which conforms to the RDF Data Cube vocabulary to a common Multidimensional Model used in data warehouses. (Kampgen et al. 2012) define common OLAP operations on data cubes modeled in RDF and show how a nested set of OLAP operations lead to a SPARQL query which generates all required facts from the data cube.

Considering that the RDF Data Cube Vocabulary (QB) is not sufficient for modeling and querying OLAP data cubes, (Etcheverry and Vaisman 2012a) address these issues in the *Open Cubes Vocabulary (OC)*, an extension of QB (at the expense of not being compatible with previous applications based on QB).

Although the Open Cubes Vocabulary adheres to the classic Multidimensional models for OLAP and allows for implementing OLAP

operators as SPARQL queries, it does not provide a mechanism for reusing data already published using QB. To deal with this issue, (Etcheverry and Vaisman 2012b) proposed a new vocabulary, denoted QB4OLAP, which extends QB to fully support OLAP models and operators. They created algorithms to transform cubes based on QB into equivalent QB4OLAP cubes (and vice versa). Figure 19 shows the classes and properties in light gray background (with prefix `qb4o`), which were added to the QB vocabulary in order to support the OLAP models.

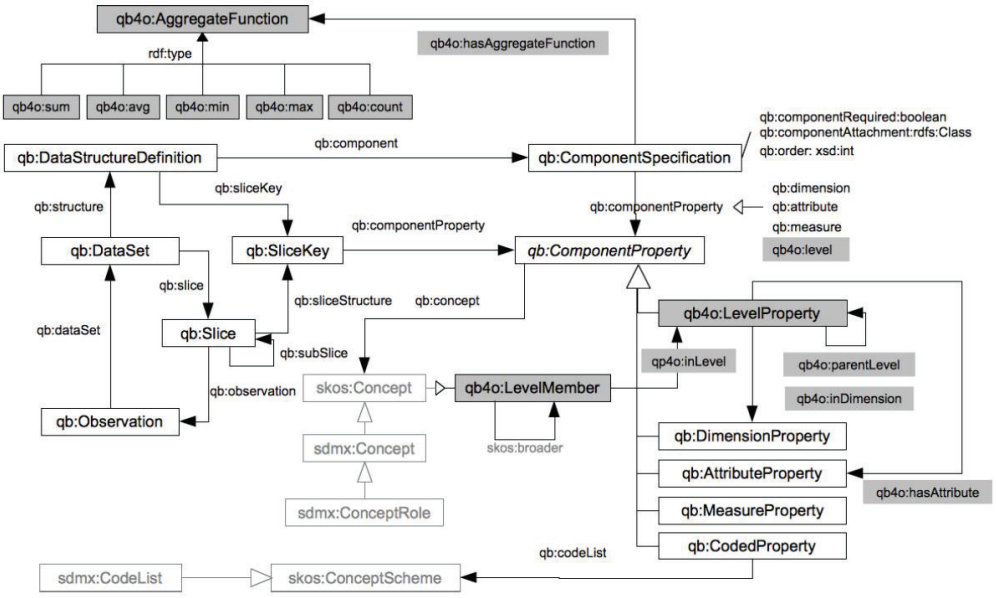

Figure 19: The QB4OLAP Vocabulary (Etcheverry and Vaisman 2012b)

In 2011, Brazil joined to the Open Government Partnership, a multinational initiative to promote worldwide adoption of OGD committed to public transparency and action on securing open publication of official data. The Brazilian Open Government Data Portal [22] is a tool made available by the government in order facilitate access to public information (K. Breitman et al. 2012). The Brazilian access information law that regulates the provision of open government data in the Web ("Lei de Acesso à Informação"[23]) represents a big leap towards transparency of government official data.

[22] http://dados.gov.br/
[23] http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/l12527.htm

In the same direction, (Zancanaro et al. 2013) present a process for identifying sources, ontology generation, mapping and publishing statistical linked data using the RDF Data Cube Vocabulary.

OLAP2DataCube (P.E.R. Salas et al. 2012) is an Ontowiki plug-in for statistical data publishing for extracting and publishing statistical data on the Web. The architecture described in this dissertation is based on the experience obtained during the development of OLAP2DataCube and aimed at complementing it.

The *OLAP2DataCube Catalog On Demand* (Ruback et al. 2013) focuses on mitigating the problems created by redundantly maintaining both the relational data cubes and their triplifications and is presented in Chapter 4. Comparing with OLAP2DataCube, data integration is performed in an evolutionary way and data cubes are triplified on demand, so that the problems created by data redundancy are bypassed. The LDC Mediator is the component of this architecture that mediates access to the underlying statistical relational databases and exposes data cubes to the applications through a RESTful API. Table 1 presents a comparison of some of the aspects covered in these previous approaches.

Table 3: Comparing the related OLAP approaches

Related Work	OLAP operators	RDB Mapping	R2RML mapping	RESTful methods	QB compatibility
Kampgen et al 2012 (Open Cubes Vocabulary)	✓	×	×	×	×
Kampgen et al 2012 (QB4OLAP Vocabulary)	✓	×	×	×	✓
Percy et al 2012 (Olap2DataCube)	×	✓	×	×	✓
Ruback et al 2013 (Olap2DataCube Catalog On Demand – with the **LDC Mediator**)	×	✓	✓	✓	✓

3.2 RESTful Linked Data approaches

The *Linked Data Platform Working Group* is a W3C group announced in August 2012 that has as mission to produce a W3C Recommendation for HTTP-

based (RESTful) application integration patterns using read/write Linked Data[24]. The group's goal is to define a writable Linked Data API, which shares some characteristics with Linked Data (such as the use of HTTP and URIs) but still relies on a flexible RDF-based data model that allows for multiple representations. They created the *Linked Data Platform (LDP) Specification*[25], a document (intended to become a W3C Recommendation) that describes a set of the best practices and simple approaches for a read-write Linked Data architecture. They also created the *Linked Data Platform (LDP) Vocabulary*[26], an ontology that provides an informal representation of the concepts and terms as defined in the specification.

Other than this W3C research group, there have been some efforts to take advantage of the best of both approaches – the Linked Data Principles and the REST approach. The most relevant efforts are presented in what follows.

(Alarcon and Wilde 2010) focus on building Restful SPARQL mappings. They claim that a RESTful redesign of SPARQL allows the Semantic Web to evolve in a more decentralized and openly accessible manner than the current style of SPARQL. Their work explores the possibilities and limitations of mapping SPARQL query forms (e. g. SELECT, CONSTRUCT, etc), as well as update forms (INSERT, DELETE, etc) to HTTP methods (GET, PUT, etc), extending the SPARQL and focusing on applying the REST Principles to it (Wilde and Michael Hausenblas 2009).

The main difference when comparing their approach with the LDC Mediator is that the LDC Mediator does not intend to focus on extending the SPARQL in order to create such mappings. The mediator only generates the required SPARQL queries according to the HTTP method requested.

(Stadtmüller and Harth 2012) extended Linked Data with RESTful manipulation possibilities. They proposed a lightweight declarative rule language, with state transition systems a formal grounding, that enables the development of data-driven applications built on the RESTful manipulation of Linked Data resources.

[24] http://www.w3.org/2012/ldp/wiki/Main_Page
[25] https://dvcs.w3.org/hg/ldpwg/raw-file/default/ldp.html
[26] https://dvcs.w3.org/hg/ldpwg/raw-file/default/ldp.ttl

(Speiser and Harth 2011) proposed *Linked Data Services* (LIDS), a general formalized approach for integrating data-providing services with Linked Data. They also present conventions for service access interfaces that conform to Linked Data Principles, and an abstract lightweight service description formalism. They also present algorithms that use LIDS descriptions to automatically create links between services and existing data sets.

3.3 Summary

This chapter presented the related work, grouped as OLAP data approaches and Linked Data with RESTful approaches. The next chapter presents the framework *OLAP2DataCube Catalog On Demand*.