5 The LDC Mediator

This chapter presents the architecture of the LDC Mediator. Section 5.1 argues for the advantages of using the REST Principles for consuming linked data cubes. Section 5.2 presents the components of the LDC Mediator. Section 5.3 describes the implementation of the mediator, including the platform, language, and tools used to build the mediator. This section also presents each package and its main java classes. Finally, Section 5.4 illustrates, through one of these Brazilian datasets collected, the process of consuming data cube metadata, as well as the process of consuming data cube observations.

5.1 Combining REST and Linked Data

The REST Principles are in a sense related to the Linked Data Principles in terms of the concept of resource:

- Resources have a unique identifier (URI) that must be dereferenceable through HTTP;
- Resources are interlinked, and, by following those links, new resources can be discovered.

However, the Linked Data Principles are only concerned with the publication and retrieval of data. The access to large collections of RDF data is usually done through a SPARQL endpoint. RESTful Web services in turn have no endpoints. They only provide a collection of resource URIs and operations to access and change the state of the resources. REST requires a uniform interface, a set of methods with known semantics that access and change the state of the resources. These methods are external to the resources, which can have multiple representations, and are invoked by HTTL messages sent to the server.

Therefore, in this context it is intuitive to combine Linked Data with REST Principles to allow data cube access and manipulation (Stadtmüller and Harth 2012). The RDF model is also flexible enough to support multiple representations, which capture the resource's current state (for instance XML or JSON).

5.2 The LDC Mediator components

Figure 21 shows the architecture of the mediator, which mediates access to the underlying statistical relational databases and exposes data cubes to the applications through a RESTful API.



Figure 21: LDC Mediator components

The LDC Mediator has the following components: a LDC RESTful API, a Mediation Engine, a Metadata service and a RDB-to-RDF Converter. The *LDC RESTful API* is the component of the mediator that exposes the data cubes through a Web service implemented using the REST Principles (Fielding and Taylor 2002). The *Mediator Engine* is the component that receives a request from a Client Application and chooses how to process the request, depending on its type. The *Metadata Service* is the module responsible for retrieving data cube metadata requested by the Mediator Engine. These metadata include the cube dimensions,

the measure and also the connection information used to access the underlying relational database. The RDB-to-RDF Converter is the component used to generate the RDF triples that represent the data cube observations from the relational databases.

5.3 Implementation of the LDC Mediator

The LDC Mediator was developed using the following platforms, languages and tools:

- Windows 7 Starter SP1 32 bits
- Java Plataform, Enterprise Edition 7 SDK²⁷ (version 1.7.0 25) •
- NetBeans IDE 7.2.1²⁸
- My Sql Server 5.5²⁹
- Toad for MySql 6.3 Freeware ³⁰
- Virtuoso open source edition ³¹
- DB2Triples RDB2RDF Antidot implementation³²
- Jersey RESTful Web Services in Java³³ •
- Astah Professional 6.6.4/41775³⁴

In addition to the documentation of each of these frameworks, some documents were essential in the implementation of the LDC Mediator:

- Stack Overflow, a question and answer Web site for programmers³⁵
- GUJ forum for Java programmers³⁶ •
- The guide to install and configure Virtuoso on Windows³⁷ •
- Examples of using Jena to manipulate RDF graphs³⁸
- "My first mapping from RDB to RDF with R2RML", Ivan Herman³⁹

²⁷http://www.oracle.com/technetwork/java/javaee/javaee7sdk-install-1957708.html 28 https://netbeans.org/

²⁹ http://dev.mysql.com/downloads/mysql/

³⁰ http://www.quest.com/toad-for-mysql/

³¹ http://virtuoso.openlinksw.com/dataspace/doc/dav/wiki/Main/VOSDownload

³² https://github.com/antidot/db2triples

³³ https://jersey.java.net/

³⁴ http://astah.net/editions/community

³⁵ http://stackoverflow.com/

³⁶ http://www.guj.com.br/forums/list.java

³⁷ http://virtuoso.openlinksw.com/dataspace/doc/dav/wiki/Main/VOSUsageWindows

³⁸ http://virtuoso.openlinksw.com/dataspace/doc/dav/wiki/Main/VirtJenaProvider

The LDC Mediator architecture is divided into three packages, according to its components (see Figure 21): the Mediator Engine package, the Metadata Service package and the RDB-to-RDF package. These packages are detailed in the following sections (the test package, the exception handling, and other auxiliary classes were omitted).

5.3.1. The Mediator Engine package

The Mediator Engine package contains classes that directly deal with client requests, as well as classes that represent the REST resources. The main class of this package is the MediatorService.java class which has attributes and methods used to interact with the other packages. It also has two classes that represent a data cube resource (DataCubeResource.java) and a list of data cubes resources (DataCubesListResource.java), as indicated in Figure 22.



Figure 22: Mediator Engine package

As previously mentioned in Section 2.6.2, data cube metadata is divided into three parts: the dataset definition, the structure definition and the measure(s) definition. In order to simplify cube metadata consumption, the MediatorService.java class has three independent methods: the first method (findDataSetMetadata) gets the dataset definition part of the cube definition; the second method (findDSDMetadata) gets the data structure definition (DSD) and the third method (findMeasureMetadata) gets the metadata of the measure(s).

In order to represent the REST resources, this package uses the Java API for RESTful Services, called JAX-RS⁴⁰. The reference implementation for this

³⁹ <u>http://ivan-herman.name/2010/11/02/my-first-mapping-from-rdb-to-rdf-using-r2rml/</u>

⁴⁰ http://jax-rs-spec.java.net

specification is an open source framework known as Jersey⁴¹, which basically contains a REST server and a REST client. On the server side, Jersey uses a servlet which scans predefined classes to identify RESTful resources. Table 4 presents the most important JAX-RS annotations and their descriptions.

JAX-RS annotation	Description
@PATH(path)	Sets the path to base URL + /path.
@POST	Indicates that the following method will answer to a HTTP POST request
@GET	Indicates that the following method will answer to a HTTP GET request
<pre>@Produces(MediaType.TEXT_PLAIN [, more-types])</pre>	Defines which MIME type is delivered by a method annotated with @GET. In the example text ("text/plain") is produced. Other examples would be "application/xml" or "application/json".
@Consumes(type [, more-types])	Defines which MIME type is consumed by this method.
@PathParam	Used to inject values from the URL into a method parameter. This way you inject for example the ID of a resource into the method to get the correct object.

Table 4: The JAX-RS most important annotations

The DataCubeResource.java and the DataCubeListResource.java classes represent the predefined data classes (resources) that the Jersey scans automatically. Figure 23 shows the annotations used to implement the first resource when a client application requests a data cube metadata, for instance HTTP GET http://host:port/LDC_Mediator/dataCube/dataset-residents. The MediaType.APPLICATION_OCTET_STREAM type (line 39) is a generic mime type for non-text data and is part of the HTTP content negotiation (see Section 2.2) from the server to the clients.

```
30 🖯 /**
      * REST Web Service
31
32
      * @author Livia Ruback
33
34
      */
35
     @Path("dataCube/{uriCube}")
36
     public class DataCubeResource {
37
38
         @GET
         @Produces(MediaType.APPLICATION OCTET STREAM)
39
40 ±
         public Response getDataCube(@PathParam("uriCube") String uriCube
```



⁴¹ http://jersey.java.net/

5.3.2. The Metadata Service package

The Metadata Service package, shown in Figure 24, includes the class that connects directly to the Catalogue stored in the Virtuoso server (CatalogueConnection.java). This package also contains the auxiliary class (SparglQueries.java) that builds the SPARQL queries needed to access the data cube metadata stored in the Catalogue. This class is also composed of some static attributes (omitted from the figure), which represent all vocabulary namespaces used to build the queries.



Figure 24: Metadata Service package

The main method of CatalogueConnection.java is executeSparqlQuery(String sparqlQuery, VirtGraph set), presented in Figure 25. It uses classes from the Jena framework (Query.java, QueryFactory.java and ResultSet.java) and from the Virtuoso Jena Provider framework (VirtGraph.java, VirtuosoQueryExecution.java, VirtuosoQueryExecutionFactory.java). It receives a generic SPARQL query as parameter and a graph of triples, it returns a list of results and it is always used when a SPARQL query is executed against the Catalogue.

32	public ResultSet executeSparqlQuery(String sparqlQuery, VirtGraph set) {
33	try{
34	<pre>Query sparql = QueryFactory.create(sparqlQuery);</pre>
35	
36	VirtuosoQueryExecution vge = VirtuosoQueryExecutionFactory.create(spargl, set);
37	
38	ResultSet results = (ResultSet) vge.execSelect();
39	
40	return results;
41	}
42	catch(Exception ex) {
43	<pre>System.out.println("Error on method executeSparqlQuery: " + ex.getMessage());</pre>
44	return null;
45	}
46	

Figure 25: The generic method to execute a SPARQL

5.3.3. The RDB-to-RDF package

The RDB-to-RDF package includes the classes used to convert the data cubes stored in the relational databases to RDF triples using R2RML mapping files. The two main classes of the package are: the DataCubeTriplification.java and the DB2TriplesParameters.java, shown in Figure 26.



Figure 26: RDB-to-RDF Converter package

In order to read and interpret R2RML mapping files, the DataCubeTriplification.java employs the db2triples library⁴², a software tool for extracting data from relational databases and loading data into RDF. It implements R2RML and Direct Mapping standards defined by the RDB2RDF working group⁴³. This library accepts some parameters to configure the type of RDB-to-RDF conversion, listed in Table 5.

Parameter	Description
-b	Database name
-d	Driver to use (default:com.mysql.jdbc.Driver)
-f	Force loading of existing repository (without remove data)
-i	Base URI (default:http://foo.example/DB/)
-1	Database URL (default :jdbc:mysql://localhost/)
-m	Mandatory conversion mode, 'r2rml' for R2RML and 'dm' for Direct Mapping

 ⁴² <u>http://www.w3.org/2001/sw/wiki/Db2triples</u>
 ⁴³ <u>http://www.w3.org/2001/sw/rdb2rdf/</u>

-n	Native store output directory
-0	Output RDF filename (default : output)
-р	Database password
-q	Transformed graph output file (optional if sparql option is not specified, default : sparql_output otherwise)
-r	R2RML config file used to convert relational database into RDF terms.
-S	Sparql transform request file (optional)
-t	RDF syntax output format ('RDFXML', 'N3', 'NTRIPLES' or 'TURTLE')
-u	Database user name
-V	Version of norm to use (1 = Working Draft 20 September 2011 (default), 2 = Working Draft 23 March 2011)

These DB2Triples parameters are represented in the DB2TriplesParameters.java class by the static attributes that are in fact used to triplify the data cubes: -m (conversion mode, "r2rml"), -u (user), -p (password), -b (database name), -t (RDF syntax output format) and -r (the path of the R2RML mapping file). The conversion mode parameter is used to inform the RDB-to-RDF conversion approach: the Direct Mapping (see section 2.7.1) or the R2RML Mapping (see section 2.7.2).

5.4 Consuming data cubes with the LDC Mediator

This section provides an example that illustrates how the LDC Mediator works.

5.4.1. Datasets collected

A small data set from SIDRA (Sistema IBGE de Recuperação automática)⁴⁴ was collected to test the LDC Mediator. SIDRA is a Brazilian Government Repository which stores demographic census data. Initially, three datasets were extracted:

⁴⁴ http://www.sidra.ibge.gov.br/

- [1] Foreigners living in Brasil⁴⁵, which is broken down by condition (naturalized or not), gender, living situation (urbana or rural) and country of birth containing 828 observations;
- [2] Residents of Brazil⁴⁶, which is broken down by race, sex, living situation, age group and age(year) containing 872 observations;
- [3] Residents of Brazil by Religion⁴⁷, broken down by religion, year and region of Brazil encompassing 864 observations.

These three datasets were extracted in CSV format and transformed into three different relational star schemas, each one containing one fact table related to their dimensions respectively. Figure 27 shows the star schema related to second dataset: residents of Brazil. Section 5.4.3 presents the data cube description related to this star schema.



Figure 27: A residents star schema

⁴⁵ <u>http://www.sidra.ibge.gov.br/bda/tabela/listabl.asp?z=cd&o=2&i=P&c=2093</u>

⁴⁶ http://www.sidra.ibge.gov.br/bda/tabela/listabl.asp?z=cd&o=2&i=P&c=2093

⁴⁷ http://www.sidra.ibge.gov.br/bda/tabela/listabl.asp?z=cd&o=2&i=P&c=137

5.4.2. The RESTful Web service definition

The LDC Mediator provides access to the data cubes through a RESTful Web service, following the REST Principles. The RESTful HTTP methods implemented by the component are shown in Table 6. The first HTTP method offers a list of all cubes stored in the Catalogue. The second HTTP method provides the data cube metadata. Finally, the third HTTP method retrieves all the individual observations related to the cube.

HTTP Method	URI Template
GET	http://host:port/LDC_Mediator/dataCubes
GET	http://host:port/LDC_Mediator/dataCube/{uriDataCube}
GET	http://host:port/LDC_Mediator/observations/{uriDataCube}
	HTTP Method GET GET GET

Table 6: LDC RESTful API methods

By requesting the first HTTP method (GET http://host:port/LDC_Mediator/dataCubes) the client receives a list containing all cubes available and their descriptions, depicted in Figure 28. The list also contains links to the other HTTP methods in order to request cube metadata and cube observations. The second and the third methods will be exemplified in sections 5.4.3 and 5.4.4 respectively.

Data Cubes List			
			URI Cube 🌣
http://purl.org/GovDataCube/resources/dataset- residents	The number of residents in Brazil. Dimensions: Race, Sex, Age, Area and Year.	ه	٩
http://purl.org/GovDataCube/resources/dataset- foreigners	The number of foreigners in Brazil. Dimensions: Country, Sex, Region and Year.	ه	A
http://purl.org/GovDataCube/resources/dataset- devotees	The number of devotees in Brazil. Dimensions: Religion, State and Year.	ه	٩

Figure 28: List of cubes

5.4.3. Requesting a cube metadata

The second HTTP method requests the cube metadata and is shown in Flow 1 of Figure 21 through the interaction between the components *Mediator Engine* and *Metadata Service*. Following the REST Principles, the HTTP request GET http://host:port/LDC_Mediator/dataCube/dataset-residents returns the state of the resource "data cube residents", which in this case is the metadata of the cube, presented in Figure 29.

1	Annefix odf.	<pre>chttp://www.w3.org/1999/02/22-rdf-syntax-ns#></pre>
2	Oprefix rdfs:	<pre><htp: 01="" 2000="" rdf-schema#="" www.w3.org="">.</htp:></pre>
3	Oprefix xsd:	<http: 2001="" www.w3.org="" xmischema#="">.</http:>
4	Oprefix ob:	<pre><http: cube#="" linked-data="" nurl.org=""></http:></pre>
5	Aprefix ex-property:	http://purlog/GovDataCube/properties/
6	Appefix ex-propercy:	http://purl.org/GovDataCube/presources/>
7	Oppefix sdmy_attribute:	http://purl.org/lipked_data/cdmx/200/attribute#\
8	Oppefix sdmx-accribuce.	<pre>//purl.org/linked-data/sdmx/2009/measure#>.</pre>
0	wprerix sumx-measure.	<pre>chech.//puri.org/iinkeu-uaca/sumx/2005/measure#/.</pre>
10		
11		
12	# Cube Residents	
12	#	
14		
15	ex-resource:dataset-resi	dents a db:DataSet:
16	rdfs:label "Number o	f residents":
17	rdfs:comment "The nu	mber of residents in Brazil. Dimensions: Race, Sex, Age, Area and Year.":
18	ab:structure_ex-reso	unce: dsd-residents
	deliberaceare ex rese	
19		
20	ex-resource:dsd-resident	s a gb:DataStructureDefinition;
19 20 21	ex-resource:dsd-resident # The dimensions	s a qb:DataStructureDefinition;
19 20 21 22	ex-resource:dsd-resident # The dimensions ab:component [ab:dim	s a qb:DataStructureDefinition;
19 20 21 22 23	ex-resource:dsd-resident # The dimensions qb:component [qb:dim qb:component [qb:dim	s a qb:DataStructureDefinition; ension ex-resource:dimRace]; ension ex-resource:dimSex];
19 20 21 22 23 24	ex-resource:dsd-resident # The dimensions qb:component [qb:dim qb:component [qb:dim qb:component [qb:dim	s a qb:DataStructureDefinition; ension ex-resource:dimRace]; ension ex-resource:dimSex]; ension ex-resource:dimAge];
19 20 21 22 23 24 25	ex-resource:dsd-resident # The dimensions qb:component [qb:dim qb:component [qb:dim qb:component [qb:dim qb:component [qb:dim	s a qb:DataStructureDefinition; ension ex-resource:dimRace]; ension ex-resource:dimSex]; ension ex-resource:dimAge]; ension ex-resource:dimYear];
19 20 21 22 23 24 25 26	ex-resource:dsd-resident # The dimensions qb:component [qb:dim qb:component [qb:dim qb:component [qb:dim qb:component [qb:dim qb:component [qb:dim	<pre>s a qb:DataStructureDefinition; ension ex-resource:dimRace]; ension ex-resource:dimSex]; ension ex-resource:dimAge]; ension ex-resource:dimArea]; ension ex-resource:dimArea];</pre>
19 20 21 22 23 24 25 26 27	ex-resource:dsd-resident # The dimensions qb:component [qb:dim qb:component [qb:dim qb:component [qb:dim qb:component [qb:dim qb:component [qb:dim	s a qb:DataStructureDefinition; ension ex-resource:dimRace]; ension ex-resource:dimSex]; ension ex-resource:dimAge]; ension ex-resource:dimYear]; ension ex-resource:dimArea];
19 20 21 22 23 24 25 26 27 28	ex-resource:dsd-resident # The dimensions qb:component [qb:dim qb:component [qb:dim qb:component [qb:dim qb:component [qb:dim qb:component [qb:dim # The measure	<pre>s a qb:DataStructureDefinition; ension ex-resource:dimRace]; ension ex-resource:dimSex]; ension ex-resource:dimAge]; ension ex-resource:dimYear]; ension ex-resource:dimArea];</pre>
19 20 21 22 23 24 25 26 27 28 29	ex-resource:dsd-resident # The dimensions qb:component [qb:dim qb:component [qb:dim qb:component [qb:dim qb:component [qb:dim # The measure qb:component [qb:mea	<pre>s a qb:DataStructureDefinition; ension ex-resource:dimRace]; ension ex-resource:dimSex]; ension ex-resource:dimYgar]; ension ex-resource:dimYgar]; ension ex-resource:dimYarea]; sure ex-property:numberResidents];</pre>
19 20 21 22 23 24 25 26 27 28 29 30	<pre>ex-resource:dsd-resident # The dimensions qb:component [qb:dim qb:component [qb:dim qb:component [qb:dim qb:component [qb:dim qb:component [qb:dim # The measure qb:component [qb:mea # The attributes</pre>	<pre>s a qb:DataStructureDefinition; ension ex-resource:dimRace]; ension ex-resource:dimSex]; ension ex-resource:dimAge]; ension ex-resource:dimYear]; ension ex-resource:dimArea]; sure ex-property:numberResidents];</pre>
19 20 21 22 23 24 25 26 27 28 29 30 31	<pre>ex-resource:dsd-resident # The dimensions qb:component [qb:dim qb:component [qb:dim qb:component [qb:dim qb:component [qb:dim qb:component [qb:dim # The measure qb:component [qb:mea # The attributes gb:component [qb:att</pre>	<pre>s a qb:DataStructureDefinition; ension ex-resource:dimRace]; ension ex-resource:dimAge]; ension ex-resource:dimAge]; ension ex-resource:dimYear]; ension ex-resource:dimArea]; esure ex-property:numberResidents]; ribute sdmx-attribute:unitMeasure; qb:componentAttachment qb:DataSet;].</pre>
19 20 21 22 23 24 25 26 27 28 29 30 31 32	<pre>ex-resource:dsd-resident # The dimensions qb:component [qb:dim qb:component [qb:dim qb:component [qb:dim qb:component [qb:dim qb:component [qb:dim # The measure qb:component [qb:meas # The attributes qb:component [qb:att</pre>	<pre>s a qb:DataStructureDefinition; ension ex-resource:dimRace]; ension ex-resource:dimSex]; ension ex-resource:dimAge]; ension ex-resource:dimYear]; ension ex-resource:dimArea]; sure ex-property:numberResidents]; ribute sdmx-attribute:unitMeasure; qb:componentAttachment qb:DataSet;].</pre>
19 20 21 22 23 24 25 26 27 28 29 30 31 32 33	<pre>ex-resource:dsd-resident # The dimensions qb:component [qb:dim qb:component [qb:dim qb:component [qb:dim qb:component [qb:dim qb:component [qb:dim # The measure qb:component [qb:mea # The attributes qb:component [qb:att ex-property:numberReside</pre>	<pre>s a qb:DataStructureDefinition; eension ex-resource:dimRace]; eension ex-resource:dimSex]; eension ex-resource:dimYear]; eension ex-resource:dimYear]; eension ex-resource:dimArea]; sure ex-property:numberResidents]; ribute sdmx-attribute:unitMeasure; qb:componentAttachment qb:DataSet;]. nts a rdf:Property, qb:MeasureProperty;</pre>
19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34	<pre>ex-resource:dsd-resident # The dimensions qb:component [qb:dim qb:component [qb:dim qb:component [qb:dim qb:component [qb:dim qb:component [qb:dim # The measure qb:component [qb:mea # The attributes qb:component [qb:att ex-property:numberReside rdfs:label "number component</pre>	<pre>s a qb:DataStructureDefinition; mension ex-resource:dimRace]; mension ex-resource:dimSex]; mension ex-resource:dimAge]; mension ex-resource:dimArea]; mension ex-resource:dimArea]; musure ex-property:numberResidents]; ribute sdmx-attribute:unitMeasure; qb:componentAttachment qb:DataSet;]. mts a rdf:Property, qb:MeasureProperty; f residents";</pre>
19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35	<pre>ex-resource:dsd-resident # The dimensions qb:component [qb:dim qb:component [qb:dim qb:component [qb:dim qb:component [qb:dim qb:component [qb:dim # The measure qb:component [qb:mea # The attributes qb:component [qb:att ex-property:numberReside rdfs:label "number co rdfs:subPropertvOf s</pre>	<pre>s a qb:DataStructureDefinition; ension ex-resource:dimRace]; ension ex-resource:dimSex]; ension ex-resource:dimAge]; ension ex-resource:dimYear]; ension ex-resource:dimArea]; sure ex-property:numberResidents]; ribute sdmx-attribute:unitMeasure; qb:componentAttachment qb:DataSet;]. ents a rdf:Property, qb:MeasureProperty; f residents"; dmx-measure:obsValue;</pre>

Figure 29: Return of the cube metadata request

In order to simplify metadata consumption, the *Metadata Service* divides the metadata into three parts (the three figure frames) and groups the results to return the complete cube metadata (see Section 2.6.2). The first part (lines 15-18) creates a resource of type qb:DataSet (line 15) and represents the entire data set. The second part (lines 20-31) defines the structure of one or more datasets (a qb:DataStructureDefinition resource type). It also defines the dimensions, attributes, and measures of the structure. Figure 6 reveals that the components of a data structure definition can be of the following types: qb:DimensionProperty, qb:AttributeProperty or qb:MeasureProperty. The third part (lines 33-36) defines the measure corresponding to the phenomenon being observed that will give the value of each observation.

There are two approaches to finding the predicates and objects related to a specific resource. To exemplify each approach, the first part of the cube metadata presented in Figure 29 is considered. The first approach is a SPARQL query,

which explicitly informs a group pattern that must match the graph stored in Virtuoso triplestore⁴⁸, shown in Figure 30.

```
1 PREFIX gb:
                           <http://purl.org/linked-data/cube#>
 2 PREFIX rdfs:
                          <http://www.w3.org/2000/01/rdf-schema#>
                           <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
 3 PREFIX rdf:
 4 PREFIX ex-resource: <a href="http://purl.org/GovDataCube/resources/">http://purl.org/GovDataCube/resources/</a>>
 5
 6 select ?comment ?label ?dsd
7 where {
    ex-resource:dataset-residents rdf:type gb:DataSet.
8
    ex-resource:dataset-residents rdfs:label ?label.
9
10 ex-resource:dataset-residents rdfs:comment ?comment.
11 ex-resource:dataset-residents gb:structure ?dsd.
12 }
```

Figure 30: SPARQL Query to return a metadata cube subset

Alternatively, instead of a static SPARQL query, one can dynamically search starting from a certain resource (identified by its URI), as presented in Figure 31. The piece of code shown in Figure 31 uses the Jena Framework⁴⁹, an API for reading, processing and writing RDF data, among other purposes.

63	public Model findDataSetMetadataPart1 (Model model, String uriDataSet) {
64	<pre>Model modelOutput = ModelFactory.createDefaultModel();</pre>
65	
66	Resource resource = model.getResource(uriDataSet);
67	<pre>StmtIterator iter = resource.listProperties();</pre>
68	<pre>while (iter.hasNext()) {</pre>
69	<pre>Statement nextStatement = iter.nextStatement();</pre>
70	// triple's property
71	Property property = nextStatement.getPredicate();
72	
73	// triple's object
74	RDFNode object = nextStatement.getObject();
75	
76	<pre>// adding triples to model</pre>
77	<pre>modelOutput.add(resource, property, object);</pre>
78	}
79	return modelOutput;
80	L }

Figure 31: Searching dynamically triples related to a resource

In Jena⁵⁰, the class used to represent a single triple is Statement. The methods for extracting the other components of a statement starting from the subject (represented by a Resource class element) are: getPredicate() that returns a Property and getObject() that returns a RDFNode.

⁴⁸ http://virtuoso.openlinksw.com/

⁴⁹ http://jena.apache.org/

⁵⁰ http://jena.apache.org/documentation/rdf/

In most cases, the Metadata Service implements the dynamic approach to find other resources related to a resource to take advantage of the flexibility of the RDF model.

5.4.4. Requesting cube observations

The third HTTP method

GET http://host:port/LDC_Mediator/observations/{idCube} retrieves all the individual observations related to the cube. It has the unique identifier {idCube} and is shown in Flow 2 of Figure 24 through the interaction between the *Mediator Engine* and *RDB-to-RDF Converter* components.

Using the same residents cube example, the *Mediator Engine* receives the GET http://host:port/LDC_Mediator/observations/dataset-residents request and finds the R2RML mapping matching with this cube, which is depicted in Figure 32.

The URI of the triple map is

ex-resource:TriplesMapFactResidents (line 12). Lines 14 to 24 create a view that maps each result row to a number of RDF triples, which are defined by predicate object maps. These, in turn, consist of predicate maps and object maps (or referencing object maps). Figure 32 presents only two dimensions (the others were omitted).

```
<http://purl.org/GovDataCube/properties/>.
 1 @prefix ex-property:
 2 @prefix ex-resource:
                           <http://purl.org/GovDataCube/resources/>.
 3 @prefix ex-class:
                           <http://purl.org/GovDataCube/classes/>.
 4 @prefix sdmx-attribute: <http://purl.org/linked-data/sdmx/2009/attribute#>.
                           <http://purl.org/linked-data/cube#>.
 5 @prefix qb:
 6 @prefix rr:
                           <http://www.w3.org/ns/r2rml#>.
                           <http://www.w3.org/1999/02/22-rdf-syntax-ns#>.
 7 @prefix rdf:
 8
9 # R2RML mapping to generate the residents cube observations from the fact table
10 ex-resource:TriplesMapFactResidents a rr:TriplesMap;
       d2rq:dataStorage ex-resource:databaseResidents;
11
       rr:logicalTable [ rr:sqlQuery ""
12
13
                               SELECT DISTINCT
14
                                  idPerson,
15
                                 idArea.
                                 idTime,
16
17
                                 numberResidents
18
                                FROM
                                 factResidents """; ];
19
20
        rr:subjectMap [
21
           rr:template "http://purl.org/GovDataCube/resources/Observations/
                                   {idPerson}_{idArea}_{idTime}_{numberResidents}";
22
23
           rr:class qb:Observation;
24
      ];
       rr:predicateObjectMap [
25
26
           rr:predicate qb:dataSet;
27
           rr:objectMap [rr:constant ex-resource:dataset-residents];
28
       1;
29
       rr:predicateObjectMap [
30
           rr:predicate ex-resource:dimSex;
31
           rr:objectMap
32
               rr:parentTriplesMap ex-resource:TriplesMapSex;
33
               rr:joinCondition [
                   rr:child "idPerson";
34
35
                   rr:parent "idPerson";
36
               ];
37
           ];
38
       1;
       rr:predicateObjectMap [
39
40
           rr:predicate ex-resource:dimAge;
41
           rr:objectMap [
42
               rr:parentTriplesMap ex-resource:TriplesMapAge;
43
               rr:joinCondition [
                   rr:child "idPerson";
44
                   rr:parent "idPerson";
45
46
               ];
47
           ];
48
       ]; {other dimensions ommited}
49
       rr:predicateObjectMap [
50
           rr:predicate ex-property:numberResidents;
51
           rr:objectMap [rr:column "numberResidents"];
52
      ];
53
       rr:predicateObjectMap [
54
           rr:predicate sdmx-attribute:unitMeasure;
55
           rr:objectMap [rr:constant <http://dbpedia.org/ontology/Person>];
56
      ].
```

Figure 32: R2RML mapping file to the residents cube

Figure 33 shows an observation (1 out of 6480 observations) generated from the R2RML file described in Figure 32. Nine triples were generated for each observation. The first triple indicates that the resource is of the qb:Observation type (line 11). The second triple links the observation with its corresponding dataset (line 12). The third triple informs the class that qualifies the interpretation of the observed value, http://dbpedia.org/ontology/Person (line 13). This class belongs to an external dataset (DBpedia) and enriches the data cube representation. The fourth triple presents the measure of the observation, exproperty:numberResidents (see the measure definition in Figure 29) and its value (line 14). Finally, the other triples link this observation to the dimension values (lines 21 to 34).

```
1 @prefix xsd:
                         <http://www.w3.org/2001/XMLSchema#>.
 2 @prefix qb:
                         <http://purl.org/linked-data/cube#>.
 3 @prefix ex-class:
                         <http://purl.org/GovDataCube/classes/>.
 4 @prefix ex-property:
                         <http://purl.org/GovDataCube/properties/>
                         <http://purl.org/GovDataCube/resources/>
 5 @prefix ex-resource:
 6 @prefix ex-observation: <http://purl.org/GovDataCube/Observations/>
 7 @prefix sdmx-dimension: <http://purl.org/linked-data/sdmx/2009/dimension#>.
 8 Oprefix sdmx-attribute: <http://purl.org/linked-data/sdmx/2009/attribute#>.
 9 @prefix dbpedia-ont:
                         <http://dbpedia.org/ontology/>.
10
11 ex-observation:201_1_2_6931 a qb:Observation
12
      ab:dataSet ex-resource:dataset-residents :
      sdmx-attribute:unitMeasure dbpedia-ont:Person ;
13
      ex-property:numberResidents "6931"^^xsd:integer ;
14
15
      ex-resource:dimArea <http://purl.org/GovDataCube/resources/Area/03e5d717932176817f3cf0a794094c34>;
16
      ex-resource:dimYear <http://purl.org/GovDataCube/resources/Year/08f90c1a417155361a5c4b8d297e0d78> ;
17
      ex-resource:dimAge <http://purl.org/GovDataCube/resources/Age/8d02c4cafd598c790c2555cdac8088c3> ;
      ex-resource:dimRace <http://purl.org/GovDataCube/resources/Race/8ec4466b9f817546ff3cafa8341cf356> ;
18
19
      ex-resource:dimSex <http://purl.org/GovDataCube/resources/Sex/97e8578f2aab557cf86014abb07db01a> .
20
21 <http://purl.org/GovDataCube/resources/Area/03e5d717932176817f3cf0a794094c34> a ex-class:Area ;
22
      <http://www.w3.org/2000/01/rdf-schema#label> "Urbana"@pt
23
24 <http://purl.org/GovDataCube/resources/Year/08f90c1a417155361a5c4b8d297e0d78> a ex-class:Year
25
      26
27 <http://purl.org/GovDataCube/resources/Age/8d02c4cafd598c790c2555cdac8088c3> a ex-class:Age ;
      <http://www.w3.org/2000/01/rdf-schema#label> "70 a 79 anos"@pt .
28
29
30 <http://purl.org/GovDataCube/resources/Race/8ec4466b9f817546ff3cafa8341cf356> a ex-class:Race ;
31
      <http://www.w3.org/2000/01/rdf-schema#label> "Indígena"@pt .
32
33 <http://purl.org/GovDataCube/resources/Sex/97e8578f2aab557cf86014abb07db01a> a ex-class:Sex ;
34
      <http://www.w3.org/2000/01/rdf-schema#label> "Mulheres"@pt .
```

Figure 33: Example of an observation

5.5 Summary

This chapter presented the architecture of the LDC Mediator. The components of the mediator were presented: the Mediator Engine, which provides a RESTful interface to access the data cubes; the Metadata Service, which connects with the Catalogue and manages the building and execution of the SPARQL queries; and the RDB-to-RDF Converter, which converts the star schemas stored in relational databases into RDF triples using R2RML mapping files. This chapter also discussed the implementation aspects, including the technologies, IDE's and frameworks used to build the LDC Mediator; as well as the packages created for each of these three components and their main classes.

Finally, an example was presented to illustrate data cube consumption, including cube metadata request and a cube observations request.