Official Statistics meets the Semantic Web: How SDMX and RDF can live together

Raffaella Maria Aracri, Stefano De Francisci, Andrea Pagano, Monica Scannapieco

Istituto Nazionale di Statistica
Istat, Italy
Objective

- Designing and implementing translation from SDMX data into RDF data
  - Format and Model translation
  - Integration of SDMX dissemination architecture with Semantic Web technologies

[Part of Eurostat: “Horizontal and vertical integration: implementing technical and statistical standards in ESS”]
Agenda

- Background:
  - RDF, RDF Schema, RDF QB
  - SDMX Reference Infrastructure (SDMX-RI) and SDMXMLSource.NET

- Methods and Results:
  - Modeling: from SDMX to RDF Data Cube Vocabulary
  - Adding structure and data format to SdmxSource.NET
  - Testing

- Conclusions
BACKGROUND
Semantic Web Standards: Data Models

- World Wide Web Consortium (W3C) data models

- RDF (Resource Description Framework): RDF 1.1 Suite, W3C Recommendation February 2014

- RDF-QB: RDF Data Cube Vocabulary, W3C Recommendation January 2014, to represent multidimensional data structures
SDMX-RI and NSI Web Services

Off-line mapping of NSI’s data

- .NET and Java implementation
- Open Source.

SDMX.Source.NET as the implementation of SDMX code source on which RI tools are built.

Web Services to make data available as generic or structured format according to client requirement.
METHODS & RESULTS
The problem

Given a Client request to Web Services of the SDMX-RI…

returning an RDF Structure and Data
Mapping SDMX data model to RDF-QB data model

- The translation step from SDMX to RDF is not only a “format” translation, but it involves a “model” translation.

RDF Data Cube Model

- qb:DataStructureDefinition
  - qb:structure
    - qb:DataSet
      - qb:dataSet
        - qb:Observation
  - qb:componentnt
  - qb:ComponentSpecification
    - qb:componentProperty
      - qb:DimensionProperty
      - qb:AttributeProperty
      - qb:MeasureProperty
      - qb:CodedProperty
Example: RDF Data Cube Model - Data Structure Definition “Environment and energy – Air”

- `qb:component`
- `qb:dimension`
- `qb:attribute`
- `qb:measure`

```xml
<rdf:type rdf:resource="purl.org/linked-data/sdmx#DataStructureDefinition"/>
```

```xml
<qb:component rdf:resource="datiopen.istat.it/property/USO_ENERGIA"/>
```

```xml
<qb:component rdf:resource="datiopen.istat.it/property/IND_TYPE"/>
```

```xml
<qb:component rdf:resource="datiopen.istat.it/property/REF_AREA"/>
```

```xml
<qb:component rdf:resource="datiopen.istat.it/property/TIME_PERIOD"/>
```

```xml
<qb:component rdf:resource="datiopen.istat.it/property/OBS_VALUE"/>
```

```xml
<qb:component rdf:resource="datiopen.istat.it/property/UNIT_MEAS"/>
```

```xml
<qb:component rdf:resource="datiopen.istat.it/property/TIME_FORMAT"/>
```

```xml
<qb:component rdf:resource="datiopen.istat.it/property/OBS_VALUE"/>
```

```xml
<qb:component rdf:resource="datiopen.istat.it/property/UNIT_MEAS"/>
```

```xml
<qb:component rdf:resource="datiopen.istat.it/property/TIME_FORMAT"/>
```

```xml
<qb:component rdf:resource="datiopen.istat.it/property/OBS_VALUE"/>
```

```xml
<qb:component rdf:resource="datiopen.istat.it/property/UNIT_MEAS"/>
```

```xml
<qb:component rdf:resource="datiopen.istat.it/property/TIME_FORMAT"/>
```

```xml
<qb:component rdf:resource="datiopen.istat.it/property/OBS_VALUE"/>
```

```xml
<qb:component rdf:resource="datiopen.istat.it/property/UNIT_MEAS"/>
```

```xml
<qb:component rdf:resource="datiopen.istat.it/property/TIME_FORMAT"/>
```

```xml
<qb:component rdf:resource="datiopen.istat.it/property/OBS_VALUE"/>
```

```xml
<qb:component rdf:resource="datiopen.istat.it/property/UNIT_MEAS"/>
```

```xml
<qb:component rdf:resource="datiopen.istat.it/property/TIME_FORMAT"/>
```

```xml
<qb:component rdf:resource="datiopen.istat.it/property/OBS_VALUE"/>
```

```xml
<qb:component rdf:resource="datiopen.istat.it/property/UNIT_MEAS"/>
```

```xml
<qb:component rdf:resource="datiopen.istat.it/property/TIME_FORMAT"/>
```

```xml
<qb:component rdf:resource="datiopen.istat.it/property/OBS_VALUE"/>
```

```xml
<qb:component rdf:resource="datiopen.istat.it/property/UNIT_MEAS"/>
```

```xml
<qb:component rdf:resource="datiopen.istat.it/property/TIME_FORMAT"/>
```

```xml
<qb:component rdf:resource="datiopen.istat.it/property/OBS_VALUE"/>
```

```xml
<qb:component rdf:resource="datiopen.istat.it/property/UNIT_MEAS"/>
```

```xml
<qb:component rdf:resource="datiopen.istat.it/property/TIME_FORMAT"/>
```

```xml
<qb:component rdf:resource="datiopen.istat.it/property/OBS_VALUE"/>
```

```xml
<qb:component rdf:resource="datiopen.istat.it/property/UNIT_MEAS"/>
```

```xml
<qb:component rdf:resource="datiopen.istat.it/property/TIME_FORMAT"/>
```

```xml
<qb:component rdf:resource="datiopen.istat.it/property/OBS_VALUE"/>
```

```xml
<qb:component rdf:resource="datiopen.istat.it/property/UNIT_MEAS"/>
```

```xml
<qb:component rdf:resource="datiopen.istat.it/property/TIME_FORMAT"/>
```

```xml
<qb:component rdf:resource="datiopen.istat.it/property/OBS_VALUE"/>
```

```xml
<qb:component rdf:resource="datiopen.istat.it/property/UNIT_MEAS"/>
```

```xml
<qb:component rdf:resource="datiopen.istat.it/property/TIME_FORMAT"/>
```

```xml
<qb:component rdf:resource="datiopen.istat.it/property/OBS_VALUE"/>
```

```xml
<qb:component rdf:resource="datiopen.istat.it/property/UNIT_MEAS"/>
```

```xml
<qb:component rdf:resource="datiopen.istat.it/property/TIME_FORMAT"/>
```

```xml
<qb:component rdf:resource="datiopen.istat.it/property/OBS_VALUE"/>
```

```xml
<qb:component rdf:resource="datiopen.istat.it/property/UNIT_MEAS"/>
```

```xml
<qb:component rdf:resource="datiopen.istat.it/property/TIME_FORMAT"/>
```

```xml
<qb:component rdf:resource="datiopen.istat.it/property/OBS_VALUE"/>
```

```xml
<qb:component rdf:resource="datiopen.istat.it/property/UNIT_MEAS"/>
Example: RDF Data Cube Model - Observation of DataSet “Environment and energy – Air”

```
rdf:type purl.org/linked-data/cube#Observation
datiopen.istat.it/dataset/DS_ENERGY_ITC41/ITC41/EE_DOM_XAB/1/A/2000

property:REF_AREA
datiopen.istat.it/code/1.0/CL_FREQ/A
property:FREQ
datiopen.istat.it/code/1.0/CL_USO_ENERGIA/1
property:USO-ENERGIAENERGIA
datiopen.istat.it/code/1.0/CL_AMBIENTE_INDICATOR/EE_DOM_XAB
property:IND_TYPE
datiopen.istat.it/code/1.0/CL_PROV_LOMB/ITC41
property:REF_AREA

1194
```

Monica Scannapieco, NTTS, Brussels - 10-12 March 2015
Implementation of the RDF «writer»

- Leverage SdmxSource.NET source code where possible
- Specific ‘add-on’ to be included in future SDMXSource.Net releases
Testing

- To validate the results we leverage the “Sdmx-RI Test Client”.
- The Test Client is a tool to test the SDMX Reference Infrastructure building blocks and to expose/browse the dissemination environment of an NSI.
CONCLUSIONS
Conclusions

- Translating SDMX into RDF is possible!
- Working solution has been implemented
- Integration of the solution with Linked Open Data dissemination channel based on a SPARQL Endpoint: datiopen.istat.it
Questions?