Making XBRL and Linked Data interoperable

Benedikt Kämpgen, Seán O’Riain, Andreas Harth
GLD WG – RDF Data Cube Vocabulary Session
17 May 2012
Outline

- XBRL - eXtensible Business Reporting Language
- Relationship between XBRL and QB
- Use Cases
- Challenges
- Discussion
What is XBRL?

- XBRL allows stakeholders, e.g., technicians and end users, who adopt it as a specification to enhance the creation, exchange, and comparison of business reporting information, e.g., financial statements and non-financial information. [XBRL SPEC]

- Financial information, e.g., yearly balance sheets, in XML

- **Example**: RAYONIER INC had a sales revenue net of 377,515,000 USD from 2010-07-01 to 2010-09-30 [XBRL example]
Why is it interesting?

- XBRL format is more machine-readable than HTML/PDF
- Large amounts of interesting statistics:
  - Since 2009, the U.S. Securities and Exchange Commission forces companies traded at the stock market to publish financial information as XBRL.
  - July 2012: 100% of SEC companies publish XBRL
  - [2012-01-27 13:48:40] @mhausenblas says: Dear @NeelieKroesEU how comes US companies must publish XBRL openly whilst not in the EU? Can we change this, please? /CC @ccbuhr #opendata

However: XBRL challenges regarding interoperability

- e.g., due to extensibility of XBRL, and flexibility of publishers to represent their financial numbers [DEBRECENY]
- Although financial data is published in XBRL, it is still difficult to integrate XBRL data using different taxonomies or taxonomy versions.
- It is still difficult to integrate XBRL with other data sources (Open Data!).
- Linked Data may help here.
Relationship between XBRL and QB (1)

- Based on: XBRL Abstract Model [XBRL AM]
- \texttt{qb:Observation} <-> Fact (e.g., disclosure of RAYONIER INC about sales in third quarter 2010)
- \texttt{qb:DataSet} <-> Fact Collection (e.g., filing with SEC Accession No. 0001193125-10-238973 and CIK no 52827) / Concept (e.g., SalesRevenueNet)

\textbf{qb:Dimensions}
- Period (e.g., from 2010-07-01 to 2010-09-30)
- Entity - Segment (e.g., RAYONIER INC)
- Scenario (e.g., planned)
- TypedAxis Value (e.g., geo location)
- ExplicitAxis Value (e.g., certain product from product catalog)

\textbf{qb:Measure}
- numerator/denominator
- value (e.g., integer, date, string)

\textbf{qb:Attribute}
- unit
Relationship between XBRL and QB (2)

- `qb:DataStructureDefinition` <-> `Concept`
- `qb:Dimension` <-> `Axis`
- `skos:Concept` <-> `Member`

**Relationships between `qb:DataStructureDefinitions` <-> Relationships**

(e.g., Sales Texas „part of“ Sales)
Relationship between XBRL and QB (3)

- Ontology for Digital Financial Report: 

- Ontology for Multidimensional Model: 

Future work:

- How to relate taxonomies and filings? XBRL uses URIs but not in the way Linked Data does.
- Once semantics of XBRL becomes clear, a mapping to Linked Data reusing QB should be possible.
- XBRL community interested in Linked Data, but they need support.
Use Case (1): Publishing XBRL from SEC as Linked Data reusing QB, for data interoperability

- SEC Edgar Linked Data Wrapper: Translates XBRL information from the SEC on-the-fly to RDF and makes it available as Linked Data reusing QB [EDGAR].
- Example: Rayonier filing [EDGAR RAY].
- EDGAR is used in Financial Information Observation System to load the data into a triple store and run OLAP analyses on it [FIOS].
- Allows integration with other data sources (Freebase, DBPedia), or other XBRL data (other taxonomies, taxonomy versions).
Example query for Rayonier Inc Sales from EDGAR [EDGAR RAY]

<table>
<thead>
<tr>
<th>Dtend root level</th>
<th>Dstart root level</th>
<th>RAYONIER INC</th>
<th>FIRSTCITY FINANCIAL CORP</th>
<th>INTERVAL LEISURE GROUP, INC.</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010-09-30</td>
<td>2010-07-01</td>
<td>377515000</td>
<td>998000</td>
<td>100488000</td>
</tr>
<tr>
<td>2011-09-30</td>
<td>2010-07-01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2011-03-31</td>
<td>2010-07-01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-03-31</td>
<td>2010-07-01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2011-06-30</td>
<td>2010-07-01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-06-30</td>
<td>2010-07-01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2011-10-01</td>
<td>2010-07-01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-10-02</td>
<td>2010-07-01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2009-08-30</td>
<td>2010-07-01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2009-09-30</td>
<td>2010-07-01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010-09-30</td>
<td>2010-07-01</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2009-08-30</td>
<td>2010-07-01</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Use Case (2): Transforming statistical Linked Data reusing QB to XBRL

- DERI use case to deliver sustainable IT information as XBRL to the Global Reporting Initiative [UC10].
- Sustainable IT information represented using QB.
Challenges

- Why are these use cases non-trivial, and should be somehow considered in GLD deliverables: Linked Data Cookbook, QB Use Case Document, or even in the core QB vocabulary?
  - Relationships between taxonomies and filings
  - Relationships between qb:DataStructureDefinitions: Aggregations, (more complex) slices, complex measures/formulas.
  - Business rules (Charles Hofmann): How to do consistency checks and automatic inferencing of additional statistics on QB?
    - There have been solutions proposed, but the GLD may give hints.
  - More complicated multidimensional elements
    - E.g., default values for dimensions
  - Extensibility and versioning
Discussion

- Use cases and challenges in- or out-of-scope?
- How to proceed? E.g., one could write a similar document for XBRL SEC data as for [COINS], this could then be linked from an XBRL use case for QB.

Thanks!
References

- [EDGAR RAY] http://edgarwrap.ontologycentral.com/archive/52827/0001193125-10-238973#ds