Serving Ireland’s Authoritative Geospatial Data as Linked Data

Alan Meehan

ADAPT Centre, Trinity College Dublin, Ireland
Contents

• Context

• Requirement Analysis

• Knowledge Representation and Organization

• Linked Data Platform

• Future use of Oracle 12c

• Conclusion
• In 2014, the *Ordnance Survey Ireland* (OSi) delivered a newly developed spatial data storage model known as **Prime2**.

• With Prime2, OSi moved from a traditional map-centric model towards an object-oriented model from which various types of mapping and data services can be produced.

• OSi furthermore aims to adopt Linked Data to enable third parties to explore and consume some of OSi's *authoritative* datasets. *But how? Can Prime2 form the basis for that?*
**Goal:** To lay the foundations of a semantic architecture and Linked Data platform for the OSi taking into account best practices and guidelines in the domain of geospatial information and industry and OSi’s current technology stack.

Starting from the boundaries dataset. These are open and already available on [http://data.gov.ie/](http://data.gov.ie/), but *not* as Linked Data.
Requirements Analysis

- Requirements analysis included engagement with the *Central Statistics Office* and the *Department of Public Expenditure and Reform*.

- Formulation of two use case scenarios from which requirements were distilled:
  1. Accessing the *same features with different geometric representations*, i.e., different generalizations or “resolutions”.
  2. Capturing the *provenance and evolution of features and their geometric representations*. E.g., Statutory Instruments to change boundaries.
Ontologies

- Features and Geometries based on GeoSPARQL
- Provenance using Statute Instruments based on PROV-O
- Static and dynamic boundaries (and their relationships)
- Necessary ontologies developed

Meetings with DPER and CSO on a URI Strategy

- Information Resources vs. Non-Information Resources
- Using Prime 2 GUIDs and a hint of the instance’s nature

- http://data.geohive.ie/page/county/2AE19629144F13A3E0550000000000001
- http://data.geohive.ie/data/county/2AE19629144F13A3E0550000000000001
Mapping the Oracle Prime2 database to RDF with R2RML

Use of (named) graphs to support both use cases

```xml
<#TM50>
  rr:logicalTable [ 
    rr:sqlQuery "SELECT GUID, GEOM FROM COUNTY_50M" ; 
  ] ;
  rr:subjectMap [ 
    rr:class geo:Geometry ; 
    rr:graph <http://data.geohive.ie/graph/50m> ; 
    rr:termType rr:BlankNode ; 
    rr:column "GUID" ; 
  ] ;
  rr:predicateObjectMap [ 
    rr:predicate geo:asWKT ; 
    rr:graph <http://data.geohive.ie/graph/50m> ; 
    rr:objectMap [ 
      rr:column "GEOM" ; 
      rr:datatype geo:wktLiteral ; 
    ] ; 
  ] ;
</#TM50>
```
Knowledge Representation and Organization

Oracle Prime2 Database

R2RML Processor

R2RML Mappings

TRIPLESTORE

<table>
<thead>
<tr>
<th>Graphs Use Case 1</th>
<th>default</th>
<th>Types</th>
<th>Labels</th>
<th>100m resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>default</td>
<td>default</td>
<td>Activities [PROV-O]</td>
<td>Entities [PROV-O]</td>
<td>History of 100m resolution</td>
</tr>
<tr>
<td>50 meters</td>
<td>50 meters</td>
<td>50m resolution</td>
<td></td>
<td>History of 50m resolution</td>
</tr>
<tr>
<td>20 meters</td>
<td>20 meters</td>
<td>20m resolution</td>
<td></td>
<td>History of 20m resolution</td>
</tr>
<tr>
<td>links</td>
<td>With DBpedia</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Triple Pattern Fragments Client Example

Linked Data Fragments client
Enter or choose a SPARQL query below and see then how your browser solves it using only triple pattern fragments.

Choose datasources:
Type a SPARQL query:
SELECT DISTINCT ?subject ?label WHERE {
  { ?subject dcterms:title ?label }
  UNION
  { ?subject rdfs:label ?label }
}
LIMIT 100

...or pick an example query:
Execute query

Query results:
  label: "Williamstown"
  label: "Agha"
  label: "Agha"
  label: "Ballincorrag"
Description of County Dublin linking to its three representations
Default generalization with OSI’s base map.

Different representations
# Download the Data

OSI "boundary" database contains geometrical representations of the boundaries of the administrative units (e.g., county, city, and rural area) of the Republic of Ireland. These are generalized up to 20, 50 and 100 meters. The following table allows boundary data to be downloaded in the RDF Turtle format.

<table>
<thead>
<tr>
<th>Administrative Unit</th>
<th>Boundary Generalisation (in meters)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>20</td>
</tr>
<tr>
<td>Barony</td>
<td>Barony 20m</td>
</tr>
<tr>
<td>Census 2011 Cities and Legal Towns</td>
<td>Cen 11 Cities and Legal Towns 20m</td>
</tr>
<tr>
<td>Census 2011 Electoral Divisions</td>
<td>Cen 11 Electoral Divisions 20m</td>
</tr>
<tr>
<td>Census 2011 Electoral Divisions Links</td>
<td>Cen 11 Electoral Divisions Links</td>
</tr>
<tr>
<td>Census 2011 Settlements</td>
<td>Cen 11 Settlements 20m</td>
</tr>
<tr>
<td>Census 2011 Small Areas Conoral</td>
<td>Cen 11 Small Areas Conoral 20m</td>
</tr>
<tr>
<td>City and County Council</td>
<td>City and County Council 20m</td>
</tr>
<tr>
<td>City Council</td>
<td>City Council 20m</td>
</tr>
<tr>
<td>County</td>
<td>County 20m</td>
</tr>
<tr>
<td>County Council</td>
<td>County Council 20m</td>
</tr>
<tr>
<td>Electoral Division</td>
<td>Electoral Divisions 20m</td>
</tr>
<tr>
<td>Local Electoral Area</td>
<td>Local Elec Area 20m</td>
</tr>
<tr>
<td>Municipal Districts</td>
<td>Municipal Districts 20m</td>
</tr>
<tr>
<td>Parish</td>
<td>Parish 20m</td>
</tr>
<tr>
<td>Rural Area</td>
<td>Rural Area 20m</td>
</tr>
<tr>
<td>Townland</td>
<td>Townland 20m</td>
</tr>
<tr>
<td>Totals</td>
<td>Totals 20m</td>
</tr>
</tbody>
</table>

Links to DBpedia

Full Dump

Datahub: https://datahub.io/dataset/geohive
Future Directions with Oracle 12c

• Use native features of Oracle 12c:

  • R2RML feature
  • Triple store and (Geo) SPARQL processor
  • SPARQL endpoint instead of TPF
  • Access control for closed data
Conclusion and Future Work

We have used OSi’s Prime2 dataset to publish their authoritative geospatial data as Linked Data on the Web by creating R2RML mappings using ontologies that extend GeoSPARQL, and PROV-O.

Future Work:

• Publish authoritative URIs for every geospatial feature in the country (~55 million distinct features – how many triples for each feature?)
Questions?

Contact Email: alan.meehan@adaptcentre.ie