Geographic Information, Linked Data & The Semantic Web

Glen Hart
Head of Research
Ordnance Survey
The daydreams of cat herders
The Prologue

- Linked Data
- Semantic Web
- Implications for GI

- Linked data is part of the Semantic Web

- And ask questions any time!
Based on an idea by Tom Heath
If settlements were data...

Based on an idea by Tom Heath
Linked Data
Some History

“The WorldWideWeb (WWW) project aims to allow all links to be made to any information anywhere. [...] The WWW project was started to allow high energy physicists to share data, news, and documentation. We are very interested in spreading the web to other areas, and having gateway servers for other data. Collaborators welcome!”

—from Tim Berners-Lee's first message
Sharing Data using the Internet

CSV
KML
PDF
<br/>
GML
<br/>
XML
<br/>
XBRL
<br/>
SDMX
<br/>
Statistical Data and Metadata eXchange
The Hampshire Bowman is located in Hampshire, which is a County. It is a Pub.

Has review: “this is a very rural pub that is very popular with bikers. It serves a range of real ales all from the barrel.”

Glen Hart lives in Warsash, which is in Hampshire. He has the birth date of 1959-06-13. His Phone No is 01489 892940.

Is located at Dundridge, which is a Town.

Is within MI5 Database.
RDF

- Resource Description Framework
- Been around for about 10 years
- W3C Standard
- Specialist databases – Triplestores
- SPARQL query language
- It is to the ‘web of data’ what HTML is to the ‘web of documents’
RDF Triple Format

Represents data as “triples”

<subject> <predicate> <object>

The Hampshire Bowman is in Dundridge
The Data Model

subject

predicate

object

subject

predicate

literal

URIs

XML Schema datatypes
Uniform Resource Identifier
http://data.ordnancnesurvey.co.uk/id/12387630

Identify:
- Information Resources
- Non-information Resources
- Properties (relationships or associations)

Let’s face it not great names!
Non-information Resource - Things

- Anything you’re interested in
Properties – Relationship or Associations

- These are URIs that identify predicates, e.g.:
  - Within
  - Same as
  - Etc

- Are they really information resources? (who cares!)
Information Resource

- The information about the thing you’re interested in
- Human readable – Web Page
- Data – RDF

But…you can be interested in “information resources” -> non-information resources!
Non-information to information resource

http://data.ordnancesurvey.co.uk/id/postcodeunit/S0164GU

GET [vocabulary URI]
Accept: application/rdf+xml

303 See Other
Location: [RDF content location]

http://data.ordnancesurvey.co.uk/doc/postcodeunit/S0164GU

GET [RDF content location]
Accept: application/rdf+xml

http://data.ordnancesurvey.co.uk/doc/postcodeunit/S0164GU.rdf

200 OK
<RDF>
URLs for Things & Relationships

- The City of Southampton

http://data.ordnancesurvey.co.uk/id/70000000000037256

- The “touches” relationship

http://data.ordnancesurvey.co.uk/ontology/spatialrelations/.touches
“Southampton touches Hampshire”

http://data.ordnancesurvey.co.uk/id/70000000000037256
http://data.ordnancesurvey.co.uk/ontology/spatialrelations/touches
http://data.ordnancesurvey.co.uk/id/70000000000017765
Data on the Web
Tim Berners-Lee’s Star Awards

- make your stuff available on the web (whatever format)
- make it available as structured data (e.g. excel instead of image scan of a table)
- non-proprietary format (e.g. csv instead of excel)
- use URIs to identify things, so that people can point at your stuff and get back useful information (in RDF)
- (if applicable) link your data to other people’s data to provide context
- Describe the data semantically
  (Ordnance Survey 6 star award)
Linked data is not terrible good at Sums

\[ 2 + 2 = 5 \]
Birmingham City Council

Website: http://www.birmingham.gov.uk
Authority Type: Metropolitan Borough
Address: The Council House, Victoria Square, Birmingham B1 1BB
Telephone: 0121 303 1111
Chief Executive: Stephen Hughes
Population: 1,016,800 (Source: ONS Crown Copyright)
Police Force: West Midlands Police
Pension Fund: West Midlands Pension Fund
1010 signatory: Yes

About this area
Experimental

403,195 dwellings, 977,087 people live in this area, 36.0 years old on average

Religion

- No religion
- Religion not stat
- Muslim
- Jewish
- Hindu
- Sikh
- Any other religion
- Buddhist
- Christian

Demographics
### Boundary-Line™
- **Data type:** Vector
- **Supply format:** ESRI ® Shape
- **Version:** 05/2010

### Code-Point® Open
- **Data type:** Point Data
- **Supply format:** CSV
- **Version:** 08/2010

### 1:50 000 Scale Gazetteer
- **Data type:** Point Data
- **Supply format:** ASCII text, Colon separated
- **Version:** 06/2010
So what?

Linked data Mashup
Artists that were played on specific BBC radio programmes that come from a specified place

Traditional Mashup
Fast food vendors in the US
The City of Birmingham

A description of the resource identified by [http://data.ordnancesurvey.co.uk/id/7000000000000018](http://data.ordnancesurvey.co.uk/id/7000000000000018)

<table>
<thead>
<tr>
<th>Label</th>
<th>The City of Birmingham</th>
</tr>
</thead>
<tbody>
<tr>
<td>Name</td>
<td>Birmingham</td>
</tr>
<tr>
<td>Name</td>
<td>The City of Birmingham</td>
</tr>
<tr>
<td>Type</td>
<td>Metropolitan District</td>
</tr>
<tr>
<td></td>
<td>Borough</td>
</tr>
<tr>
<td>Area</td>
<td>26,778.635</td>
</tr>
<tr>
<td>Area code</td>
<td>MTD</td>
</tr>
<tr>
<td>Census code</td>
<td>000CN</td>
</tr>
<tr>
<td>Official name</td>
<td>The City of Birmingham</td>
</tr>
<tr>
<td>Unit i d</td>
<td>18</td>
</tr>
<tr>
<td>Borders</td>
<td>The District of Lichfield</td>
</tr>
<tr>
<td></td>
<td>The Borough of North Warwickshire</td>
</tr>
<tr>
<td></td>
<td>The County of Warwickshire</td>
</tr>
<tr>
<td></td>
<td>The Borough of Sandwell</td>
</tr>
<tr>
<td></td>
<td>The Borough of Solihull</td>
</tr>
<tr>
<td></td>
<td>The District of Bromsgrove</td>
</tr>
<tr>
<td></td>
<td>The County of Staffordshire</td>
</tr>
<tr>
<td></td>
<td>The County of Worcestershire</td>
</tr>
<tr>
<td></td>
<td>The Borough of Dudley</td>
</tr>
<tr>
<td></td>
<td>The Borough of Walsall</td>
</tr>
</tbody>
</table>


About: **Birmingham**  
An Entity in Data Space: dbpedia.org

Birmingham is a city and metropolitan borough in the West Midlands county of England. Birmingham is the most populous of the English Core Cities, with a population of 1,006,500 (2006 estimate).

<table>
<thead>
<tr>
<th>Property</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>dbpedia-owl:Place/areaTotal</td>
<td>267.7700</td>
</tr>
<tr>
<td>dbpedia-owl:Place/elevation</td>
<td>140</td>
</tr>
<tr>
<td>dbpedia-owl:Place/nickname</td>
<td>&quot;Brum&quot;, &quot;Brummagem&quot;, &quot;Second City&quot;, &quot;City of a thousand trades&quot;, &quot;Workshop of the World&quot;</td>
</tr>
<tr>
<td>dbpedia-owl:PopulatedPlace/areaCode</td>
<td>0121</td>
</tr>
</tbody>
</table>
| dbpedia-owl:PopulatedPlace/establishedTitle | City  
  Founded  
  Municipal borough |
| dbpedia-owl:PopulatedPlace/foundingDate | 1838-01-01 00:00:00 (xsd:date)  
  1889-01-01 00:00:00 (xsd:date) |
| dbpedia-owl:PopulatedPlace/leaderName  | dbpedia:Conservative_Party_(UK)  
  dbpedia:Liberal_Democrats  
  dbpedia:Mike_Whitby  
  dbpedia:Government_of_Birmingham#Birmingham_City_Council |
| dbpedia-owl:PopulatedPlace/leaderTitle | Council Control  
  Council Leader  
  Deputy Lord Mayor  
  Governing body  
  Lord Mayor |
| dbpedia-owl:PopulatedPlace/motto       | Forward                                                               |
| dbpedia-owl:PopulatedPlace/populationAsOf | 2005-01-01 00:00:00 (xsd:date)                                        |
| dbpedia-owl:PopulatedPlace/populationDensity | 3739                                                               |
| dbpedia-owl:PopulatedPlace/populationTotal | 0 (xsd:integer)  
  67 (xsd:integer)  
  2284093 (xsd:integer) |
Broadcast are an electronic music band, based in the Kings Heath part of Birmingham, England. Original members were Tristram Cary (vocals and guitar), Conor O'Brien (bass). Various drummers have played with the band, including Keith York, Phil Jenkins, Jeremy Barnes, Steve Perkins and Nick Jenkins.

<table>
<thead>
<tr>
<th>Property</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>dbpedia-owl:Artist/genre</td>
<td>dbpedia:Dream_pop</td>
</tr>
<tr>
<td></td>
<td>dbpedia:Electronic_music</td>
</tr>
<tr>
<td>dbpedia-owl:Artist/label</td>
<td>dbpedia:Duophonic_Records</td>
</tr>
<tr>
<td></td>
<td>dbpedia:Warp_Records</td>
</tr>
<tr>
<td></td>
<td>dbpedia:Tommy_Boy_Records</td>
</tr>
<tr>
<td>dbpedia-owl:MusicalArtist/background</td>
<td>group_or_band</td>
</tr>
<tr>
<td>dbpedia-owl:Person/activeYearsEndDate</td>
<td>1995-01-01 00:00:00 (xsd:date)</td>
</tr>
<tr>
<td>dbpedia-owl:Person/activeYearsStartDate</td>
<td>1995-01-01 00:00:00 (xsd:date)</td>
</tr>
<tr>
<td>dbpedia-owl:Person/homeTown</td>
<td>dbpedia:Birmingham</td>
</tr>
<tr>
<td></td>
<td>dbpedia:England</td>
</tr>
<tr>
<td>dbpedia-owl:activeYearsEndDate</td>
<td>1995-01-01 00:00:00 (xsd:date)</td>
</tr>
<tr>
<td>dbpedia-owl:activeYearsStartDate</td>
<td>1995-01-01 00:00:00 (xsd:date)</td>
</tr>
<tr>
<td>dbpedia-owl:background</td>
<td>group_or_band</td>
</tr>
<tr>
<td>dbpedia-owl:genre</td>
<td>dbpedia:Dream_pop</td>
</tr>
<tr>
<td></td>
<td>dbpedia:Electronic_music</td>
</tr>
<tr>
<td>dbpedia-owl:homeTown</td>
<td>dbpedia:Birmingham</td>
</tr>
<tr>
<td></td>
<td>dbpedia:England</td>
</tr>
<tr>
<td>dbpedia-owl:label</td>
<td>dbpedia:Duophonic_Records</td>
</tr>
<tr>
<td></td>
<td>dbpedia:Warp_Records</td>
</tr>
<tr>
<td></td>
<td>dbpedia:Tommy_Boy_Records</td>
</tr>
</tbody>
</table>
Broadcast

UK electronic music band on Warp Records

Broadcast are an electronic music band, based in the Kings Heath part of Birmingham, England. They recorded albums for Warp Records and Boo jay's Records. Their debut album was 1998's 'The Fool's哲学', and their second album was released in 2000.
07/07/2009

**AVAILABILITY:**
Sorry, this programme is not available to listen again. (why?)

Last broadcast on Tue, 7 Jul 2009, 21:00 on BBC 6 Music

Gideon presents more concerts and sessions from the BBC archives, featuring live tracks from Tahiti 80 and The Ruts plus sessions from OMD and Broadcast.

**TRACKLIST**

- reg wale group — fruity flute / farmhouse kitchen
- The Pastels/Tenniscoats — Vivid Youth
- Sid Presley Experience — Hup 2 3 4 BBC Session 31/05/1984
- **Iron & Wine** — love vigilantes
- **Man or Astro-man?** — Rovers BBC Session 21/12/1993

**More details**

A PROGRAMME FROM
- Gideon Coe

WEBSITE
Go to the homepage for 'Gideon Coe'

DURATION
180 minutes

**More like this**
Find related BBC 6 Music programmes.

CATEGORIES
GEHRE:
- **Music** > **Classic Pop & Rock**
So what?

- Links are:  
  - Preserved  
  - Published  
  - Reusable  
- Data structure problem reduced  
- Amount of code required reduced
<sameAs> interlinking the Web of Data

The Web of Data has many equivalent URIs. This service helps you to find co-references between different data sets. Enter a known URI, or use Sindice to search first.

Equivalent URIs for http://data.ordnancesurvey.co.uk/id/70000000000000018 –


Show 21 more

rdf+xml · n3 · json · text
Linked Data Principles

- Name Things with URI
- Use HTTP URI so they can be looked up on the Web
- Provide some useful information in open standard formats
- Link to other things...

Tim Berners-Lee
Linked Data - Summary

- The Linked Data Web does what it says on the tin
- The simple data model reduces the data structure problem
- Links are as important as the data
- The links you make can be reused by others
The Semantic Web
Semantic Web

- In 1999, Tim Berners-Lee, ‘invents’ the Semantic Web
- Content marked up to indicate meaning

In the Semantic Web Ontologies are used to describe what things are.

Ontology definition:
Southampton is a kind of City…
A Failed Vision?

- The Semantic Web still isn’t really here after 10 years
- Data was in the wrong format
- Lack of tools
- But....
Semantic Web Reborn?

- Increasing recognition that Semantic Web technologies have an important role to play in data integration
- So today the Semantic Web is seen as more about joining up data in the “Deep Web”
- Less about intelligently marking up web pages.
Hidden knowledge

- Concealed in the database design
- In application code
- In documentation (or not!)
- In peoples’ heads

- Ontologies can help expose this knowledge
Ontologies

Dictionaries are like watches, the worst is better than none and the best cannot be expected to go quite true.

Samuel Johnson

- Repositories of knowledge
- Independent from the data
- Formal descriptions
- In computer science normally based on first order logic
- Semantic Web: OWL or RDFS
Where Ontologies Fit

Application

Knowledge

Ontology

Relational Database

Triple Store
OWL

- OWL – Web Ontology Language
- Represented as triples (subject predicate Object – sound familiar?)
- E.g.:
  - Every Mayfly is a kind of Insect.
  - Every School has a Building.

- RDFS is (sort of) a simpler version of OWL
A Geography Question

Is Nottingham in No

Why?
How Ontologies Work

Know more than is in the data

Every Place has a Footprint.
Every Place has a Purpose.

Every City is a kind of Place.
Every City contains Buildings.
Every City contains Roads.
Every City has purpose Provision of Housing.

Every University has purpose Education.

University of Nottingham is a University.

Nottingham is a City.
Nottingham contains the University of Nottingham.
Geographic Ontologies

Features
- City
- Road
- River
- Building
- Field
- Orchard...

Land Use
- Agriculture
- Market Gardening...
- Arable farming
- Manufacturing
- Education

Land Cover
- Rock
- Scree
- Rough Grass
- Trees
- ...

Every Orchard is a kind of Enclosed Land.
Every Orchard contains Fruit Trees.
Every Orchard has purpose Market Gardening.
Relationships (Properties)

- Main ones:
  - Subsumption
  - Topology
  - Mereology
  - Function (purpose and use)
Spatial Relations RCC8
Regional Connection Calculus 8
Ordnance Survey Topologic Relations

- Disjoint
- Touches
- Partially overlaps
- Within
- Contains
- equal
Principles

- No God’s Eye view or Super all encompassing Ontology
- Lots of light weight interconnected ontologies
- Ontologies are a network of statements
  - Interconnected within
  - and between ontologies
Semantic Web Summary

- The Semantic Web is slow in coming
- Needs both data and ontologies
- Separates knowledge from application and data
- Enables you to know more than is in the data
- Support for multiple World views
Implications for GIS

- Technical

- Cultural

the united kingdom vs the united states: CULTURE CLASH!!!

I’ll learn how to play cricket if you’ll learn how to drink watered down beer. DEAL
Implications for GI - Technical

- Less focus on geometry
- Relationships become much more important
- Topology & Mereology
  - Next to
  - Within
  - Part of…
Cultural Changes

- Where we believe we “fit” as a community
- The way we develop standards
Where does GI fit?

80% of all organisational Information

GI
The rest
80% of all organisations

A 1986 brochure published by the Municipality of Burnaby, British Columbia¹

Local Authorities deal with lots of “local” data

But what about a pharmaceutical company? Or a bank?

¹Muki Hakley: (http://povesham.wordpress.com/2010/02/22/the-source-of-the-assertion-that-80-of-all-organisational-information-is-geographic/).
So is GI unimportant?

- The point is not that GI is unimportant – it clearly is very important

- But that we should not over estimate it’s importance – this promotes a geocentric view of the world
Standards

- We depend on standards to communicate between ourselves
- E.g. GML, WMS, WFS
- SDIs are built on them

- But they aren’t known to many outside the GI Community
We are an island race

GIS Island,
Home to the GI Community

Wish you were here?
It’s not just us!
A Sticky view of GI

Geography is glue, and no one looks at the glue
Semantic Web & Linked Data can help

- Adopt Linked Data as the means to exchange information

- Describe our data with ontologies

- Be part of a worldwide Data Infrastructure not an SDI
We can help develop the Semantic Web

- We can use our knowledge to modify the web standards
- We can develop spatial ontologies
Conclusion

- There’s a lot of stuff going on!
- Linked Data is here and growing
  The wider semantic web is beginning to stir again
- The GI Community needs to embrace these activities but we need to recognise the cultural changes that are required?
- We can bring our knowledge to the Semantic Web
Thank you, Questions?

The map is definitive, reality is frequently inaccurate