Linked Open Data Visualisation

Conal Tuohy conaltuohy.com/about-me/

Linked Open Data Visualisation

"Linked Open Data"

- What even is it?
- Why would you want to use it?
- How do you make it?
- How do you visualise it?

"Linked Open Data" (LOD)

Term defined by Tim Berners-Lee in 2006

Linked Open Data =

- data available on the web
- (ideally) as **RDF** and **SPARQL**
- (ideally) including links to other people's data
- under an open licence

www.w3.org/DesignIssues/LinkedData.html

Resource Description Framework (RDF) A web-based technology for representing knowledge

- An initiative of the WWW Consortium
- Breaks down barriers between data silos
- Many open source and commercial tools
- Radically open and interoperable
- Models knowledge as a network or "graph"

What is an RDF "graph"?

Not *this* kind of graph

This kind of graph





RDF Statements ("triples")



Uniform Resource Identifiers (URIs)

ex:Mammal

"Cat"

kind of

label

ex:Cat

- Subject: http://example.com/Cat
- Predicate: http://www.w3.org/2000/01/rdf-schema#subClassOf
- Object: http://example.com/Mammal
- Subject:http://example.com/CatPredicate:http://www.w3.org/2000/01/rdf-schema#labelObject"Cat"

RDF Syntaxes

- Don't be put off!
- There are heaps of different ways to write RDF triples
 - N-Triples, Turtle, RDF/XML, JSON-LD, Notation3, TriX, TriG, etc. etc.
- But they all mean the same thing!
- Why so many?! Why?!!

Linked Data identifiers are also links

- In "Linked Data" those identifiers have to work in a browser!
- Every HTTP URI identifies something (a "resource")
 - A "resource" can be anything at all; a web page, a person, an historical period...
- Every HTTP URI is also a hyperlink: following the link should return information about that resource which it identifies
- i.e. the web of Linked Data is hypertext, like the HTML web.

Linked Open Data Visualisation

What does it mean for visualisation?

- Calls for some specific technical knowledge & software skills (RDF, SPARQL, etc.)
- LOD simplifies accessing and integrating data from different sources
- SPARQL makes it easy to select from, and analyse the data
- It's natural to visualise the data as graphs (networks)
- ... but other forms of visualisation also possible

LODLAM

Linked Open Data in Libraries Archives & Museums

LODLAM acronym coined 2010 by Jon VossA growing international community of practiceLODLAM Summits

- 2011 San Francisco
- 2013 Montreal
- 2015 Sydney

LODLAM a Linked Data buzzword

- A loose organisation / movement: LODLAM.net #LODLAM
- Representing cultural heritage knowledge in RDF graphs
- Shared conceptual models of cultural heritage
- The interlinked "cloud" of cultural heritage knowledge

Demo: LODLive and Museum Victoria LOD



type

E22_Man-Made_Object Resource

P3.1_physicalDescription

This is a glazed earthenware teacup which has been reconstructed. It is decorated with a blue or black vine and leaf design around outside and inside of the cup which is known as 'Moss Rose' pattern.

P3.1_objectSummary

This reconstructed cup was excavated at the Commonwealth Block site between 1988 and 2003. There is a matching saucer that was found with it. The pattern is known as 'Moss Rose' and was made between 1850 and 1851 by Charles Meigh, Son & Pankhurst in Hanley, Staffordshire, England. Homewares. Numerous crockery pieces were found all over the Little Lon site. Crockery gives us a glimpse of everyday life in Melbourne in the 1880s. In the houses around Little Lon, residents used decorated crockery. Most pieces were cheap earthenware or stoneware, yet provided colour and cheer. Only a few could afford to buy matching sets, and most china was probably acquired second-hand. Some were once expensive pieces. Householders mixed and matched their crockery from the great range of mass-produced designs available. 'Blue and white' and the 'willow' pattern, was the most popular choice and was produced by English potteries from 1790.

Similar to the Structured Query Language (SQL) used in relational databases. Different kinds of questions can be asked:

- What are the names of all the kinds of mammals you know of?
- Do you know of any mammal called "Fido"?
- What is everything you know about Fido?

Logical questions:

What are the names of all the kinds of mammals you know of?

prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>
prefix ex: <http://example.com/>

select ?name where {
 ?mammal a ex:Mammal.
 ?mammal rdfs:label ?name.

name	
Cat	
Dog	

True or false questions, e.g.

Do you know of a mammal named "Fido"?

prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> ask {

?mammal a <http://example.com/Mammal>.

?mammal rdfs:label "Fido".



Open-ended questions, e.g. What do you know about Fido?

describe <http://example.com/Fido>

SPARQL a key Linked Data analysis tool

SPARQL makes the RDF learning curve worth it

- Select data from the graph that match certain criteria
- Extract data as sub-graphs, or re-arrange the graph
- Or convert the graph data into a table
- Compute statistics

Demo: PROVisualizer Javascript app using D3.js



Digital Daisy Bates



Cytoscape network visualisation tool



org.cytoscape.command-executor-impl started

Cytoscape network visualisation tool

- Cytoscape is an open source software platform for visualizing molecular interaction networks and biological pathways...
- Although Cytoscape was originally designed for biological research, now it is a general platform for complex network analysis and visualization.

http://www.cytoscape.org/what_is_cytoscape.html

Cytoscape network visualisation tool

- Cytoscape has a bunch of "Apps" for dealing with SPARQL databases
- Or you can just import tables in CSV format

Aj	pp Manager + ×
Install Apps Currently Installed Chec	k for Updates
Download Site: Cytoscape App Store	Manage Sites
Search:	
network dynamics (6) functional module detered gene regulation (5) network manipulation Pathway finding (4) network learning (4) Connection sub-graph networks (4) microarrays (4) simulation (4) linked data (4) R (4) SPARQL (4) export (4) attribute generation (4) centrality analysis (4) molecular structure (4)	General SPARQL SemScape Vital AI Graph Visual WikiDataScape
Install from File	View on App Store Install
	Close

Cytoscape visualisation of Bates vocabularies



Summary

- Graph visualisation offers valuable insights into your data
- There are LOD datasets out there waiting to be visualised
- If your source data isn't yet LOD you can convert it (and Linked Data is a great way to publish and share your data)
- SPARQL is a wonderfully simple and powerful tool

Questions?