

Using these slides

These slides were developed by Owen Stephens (owen@ostephens.com) on behalf of the British Library.

Unless otherwise stated, all images, audio or video content are separate works with their own licence. Such licences are stated where known

This work is licensed under a Creative Commons Attribution 4.0 International License http://creativecommons.org/licenses/by/4.0/.

It is suggested when crediting this work, you include the phrase "Developed by Owen Stephens on behalf of the British Library"



Original proposal at CERN - 1989

Followed up by a more concrete proposal in 1990 which scoped down considerably – talks about 'hypertext pages' What is 'the web' – at it's heart a set of protocols, data standards and identifiers

"The Semantic Web is a *web of data*, in some ways like a global database"

"first step is putting data on the Web in a form that machines can naturally understand, or converting it to that form. This creates what I call a Semantic Web-a web of data that can be processed directly or indirectly by machines"²

I. http://www.w3.org/DesignIssues/Semantic.html

2. Tim Berners-Lee, Weaving the Web. Harper, San Francisco. 1999.

Before 'linked data' came the Semantic Web Very related to all we've looked at today The Semantic Web isn't just about putting data on the web. It is about making links, so that a person or machine can explore the web of data. With linked data, when you have some of it, you can find other, related, data.

I.Use URIs as names for things

2.Use HTTP URIs so that people can look up those names.

3. When someone looks up a URI, provide useful information, using the standards (RDF, SPARQL)

4.Include links to other URIs. so that they can discover more things.

http://www.w3.org/DesignIssues/LinkedData.html

Change of emphasis, although perhaps not of long term vision Lots to talk through here



While most people are familiar with URLs, the concept of a URI is less well known. A URL is a resource locator – if you know the URL, you can locate the resource. A URI is a resource identifier – it simply identifies the resource. In fact, URLs are a special kind of URI – that is any URL is also a URI in that a URL both identifies and locates a resource. So – all URLs are also URIs, but not vice versa. You can read more about <u>URIs on Wikipedia</u>.

Further to this, an 'HTTP URI' is a URL as we are used to using on the web.

This means that the first two principles together basically say you should identify things using web addresses. This sounds reasonably straightforward. Unfortunately there is some quite tricky stuff hidden behind these straightforward principles, which basically come down to the fact that you have to be very careful and clear about what any particular http URI identifies Use URIs as names for things
 Use HTTP URIs so that people can look up those names.

What does this URI identify?

http://www.amazon.co.uk/Pride-Prejudice-Penguin-Classics-Austen/dp/0141439513

For example this URI:

http://www.amazon.co.uk/Pride-Prejudice-Penguin-Classics-Austen/dp/0141439513

Doesn't identify (as you might expect) Pride and Prejudice, but rather identifies the Amazon web page that describes the Penguin Classics edition of Pride and Prejudice. This may seem like splitting hairs, but if you want to start to make statements about things using their identifiers it is very important. I might want to state that the author of Pride and Prejudice is Jane Austen. If I say:

http://www.amazon.co.uk/Pride-Prejudice-Penguin-Classics-Austen/dp/0141439513 is authored by Jane Austen, then strictly I'm saying Jane Austen wrote the web page, rather than the book described by the web page.

3. When someone looks up a URI, provide useful information, using the standards (RDF, SPARQL)

What is RDF?

- Resource Description Framework
 - -"a language for representing information about resources in the World Wide Web"
 - -"RDF can also be used to represent information about things that can be *identified* on the Web, even when they cannot be directly *retrieved* on the Web"
 - -Uses URIs or Literals





Remember the 'just the data' slide from this morning?

Many RDF triples are made up of just three URIs - this means a different way of thinking about the data



In reality would be more complex. For example, the entity that I've labelled 'Jane Eyre' would have a URI, which would then have a name or label



The Power of Three Or should that be 4? The 'Quad' for provenance

4. Include links to other URIs. so that they can discover more things.













Stuff that you don't/can't or won't capture – but someone else will. (Books about places that are mentioned in MO archive?) Network effect – collapse the network – one more node makes the difference Provenance



I used to think the technical aspects of Linked Data were the hard bits – RDF, SPARQL, and a whole load of stuff I haven't mentioned. While there is no doubt that these things are complicated, and complex, I now believe the really difficult bits are the modelling and reuse aspects. I also think that there is an overlap here with the areas where domain experts need to have an understanding of 'computing' concepts, and computing experts need to understand the domain – and this kind of crossover is always difficult.

```
<http://bnb.data.bl.uk/id/concept/lcsh/
Programminglanguages%28Electroniccomputers%29>
skos:inScheme <<u>http://id.loc.gov/authorities/subjects</u>> ;
rdf:type ns0:TopicLCSH ;
rdfs:label "Programming languages (Electronic
computers)" ;
owl:sameAs <<u>http://id.loc.gov/authorities/subjects/
sh85107313</u>> ;
<<u>http://www.bl.uk/schemas/bibliographic/
blterms#TopicLCSH</u>> rdfs:label "Library of Congress
Subject Heading (LCSH)"@en
```

Follow your nose to LCSH data

RDFS (RDF Schema) allows you to define vocabularies for a community

Define a class of objects 'Books'

Define a property 'Creator'

OWL (Web Ontology Language) allows you to put limits of relationships

a Person has one and only one biological mother

OWL allows you to say 'this is the same as that'

This thing that I call 'Mozart' is the same as that thing that you call 'Mozart, Wolfgang Amadeus'



Datamodel for BL BNB as linked data, taken from <u>http://www.bl.uk/bibliographic/pdfs/bldatamodelbook.pdf</u>



The Library of Congress started to publish it's catalogue records on these standard sized cards and by this method could distribute them to other libraries.

This was so successful, that Charles Cutter, who produced a seminal work on building a printed dictionary catalogue quickly had to revise it to take account of the card catalogue. By the time the 4th edition of Cutter's work was published, he prefaced it by saying "any new library would be very foolish not to make its catalog mainly of them [LoC cards]"



By establishing the 'work' (not a FRBR work) raises the possibility of a single 'work' record on the web... Illustration of Bibframe model taken from <u>http://bibframe.org/vocab-model/</u>

BIBFRAME Editor											
Get the code here.	In the meantime, select a Pro	ofile bel	ow to get	starte	d.						
BIBFRAME - Kitchen Sink Profiles	Work - Monograph										
New HeldItem	Authorized access point	Authorized access point					Set				
New Instance											
New Work	Title	Title					Set				
New Work, Instance, & HeldItem	Title variation	Title Entity									
BIBFRAME - Simple Monograph	Author	Person	Organization	Family	Meeting	Jurisdiction					
New Holding New Instance	Editor	Person	Organization	Family	Meeting	Jurisdiction					
New Work	Translator	Person	Organization	Family	Meeting	Jurisdiction					
New monograph (New Work, Instance, and Holding)	Date of work	Date of work				Set					
	Place of creation	Place									
							http://bibframe.org/tools/	<u>'editor/</u>			

MARC 21	MARC 21 Position / Indikator / Unterfeld	Inhalt	RDF- Element	Bemerkung
100	\$0	Personenname (IDN)	dc:creator	wenn Präfix "(DE-588) <idn>" , dann URI <u>http://d-nb.info/gnd/</u><idn></idn></idn>
	\$a	Personenname (Literal)	dc:creator	Literal
110	\$0	Körperschaftsname (IDN)	dc:creator	wenn Präfix "(DE-588) <idn>" , dann URI <u>http://d-nb.info/gnd/</u><idn></idn></idn>
	\$a	Körperschaftsname (Literal)	dc:creator	Literal
111	\$0	Kongressname (IDN)	dc:creator	wenn Präfix "(DE-588) <idn>" , dann URI <u>http://d-nb.info/gnd/</u><idn></idn></idn>
	\$a	Kongressname (Literal)	dc:creator	Literal
700	\$0	Personenname (IDN)	dc:contributor	wenn Präfix "(DE-588) <idn>" , dann URI <u>http://d-nb.info/gnd/</u><idn></idn></idn>
	\$a	Personenname (Literal)	dc:contributor	Literal
710	\$0	Körperschaftsname (IDN)	dc:contributor	wenn Präfix "(DE-588) <idn>" , dann URI <u>http://d-nb.info/gnd/</u><idn></idn></idn>
	\$a	Körperschaftsname (Literal)	dc:contributor	Literal
711	\$0	Kongressname (IDN)	dc:contributor	wenn Präfix "(DE-588) <idn>" , dann URI <u>http://d-nb.info/gnd/</u><idn></idn></idn>
	\$a	Kongressname (Literal)	dc:contributor	Literal

https://wiki.dnb.de/pages/viewpage.action?pageId=68@60017

<hr/> http://viat.org/viat//1388025>								
madsrdf:isIdentifiedByAuthority	<http: 9054114="" authorities="" id.loc.gov="" n="" names=""></http:>							
rdf:type	schema:Person							
schema:name	"Brontë, Charlotte, 1816-1855."							
at.org/viaf/71388025" af.org/viaf/71388025" ama.org/Person" chema.org/author" org/viaf/71388025" chema.org/Person"> rg/viaf/71388025" target="_	new"> <u>http://viaf.org/viaf/71388025</u> >							
	<http: 71388025="" viaf="" viaf.org=""> madsrdf:isIdentifiedByAuthority rdf:type schema:name author" af.org/viaf/71388025" ma.org/Person" hema.org/author" org/viaf/71388025" hema.org/Person"> rg/viaf/71388025" target="</http:>							

http://viaf.org/viaf/71388025/rdf.xml

rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#" xmlns:void="http://rdfs.org/ xmlns:viaf="http://viaf.org/ontology/1.1/#" xmlns:foaf="http://xmlns.com/foaf/0.1/" xmlns:skc xmlns:skos="http://www.w3.org/2004/02/skos/core#" xmlns:rdaGr2="http://rdvocab.info/Elements(xmlns:rdaEnt="http://rdvocab.info/uri/schema/FRBRentitiesRDA/" xml:base="http://viaf.org/"> ▼<foaf:Document rdf:about="http://viaf.org/viaf/71388025/"> <void:inDataset rdf:resource="http://viaf.org/viaf/data"/> <foaf:primaryTopic rdf:resource="http://viaf.org/viaf/71388025"/> </foaf:Document> \rdf:Description rdf:about="http://viaf.org/viaf/71388025"> ▼<!--The primary entity (formerly a viaf:NameAuthorityCluster) --> <rdf:type rdf:resource="http://xmlns.com/foaf/0.1/Person"/> <rdf:type rdf:resource="http://rdvocab.info/uri/schema/FRBRentitiesRDA/Person"/> <foaf:name>Brontë, Charlotte, 1816-1855</foaf:name> <foaf:name>Brontë, Charlotte English novelist, 1816-1855</foaf:name> <foaf:name>Brontë, Charlotte, 1816-1955</foaf:name> <foaf:name>Bronte, Charlotte</foaf:name> <foaf:name>Бронте, Шарлотта, 1816-1855</foaf:name> <foaf:name>Bell, Currer, 1816-1855</foaf:name> <foaf:name>Bolangte, Xialuodi, 1816-1855</foaf:name> <foaf:name>勃朗特, 夏洛蒂, 1816-1855</foaf:name> <foaf:name>Wellesley, Charles, 1816-1855</foaf:name> <foaf:name>Bronte, Charlotte, 1816-1855</foaf:name> <foaf:name>Bronte, Sharlotta, 1816-1855</foaf:name> <foaf:name>Pŭront'e, Syarŭllot'ŭ, 1816-1855</foaf:name> <foaf:name>Po-lang-t'e, Hsia-lo-ti, 1816-1855</foaf:name> <foaf:name>Pirāņțē, Cārlațți, 1816-1855</foaf:name> <foaf:name>Bŭront'e, Syarŭllot'ŭ, 1816-1855</foaf:name> <foaf:name>Bronte, Karlotta, 1816-1855</foaf:name> <foaf:name>Bell, Currer.</foaf:name> <foaf:name>Tree, Captain, 1816-1855</foaf:name> <foaf:name>Douro, Marquess of, 1816-1855</foaf:name> <foaf:name>Бронте, Шарлотта, 1816-1855</foaf:name>

Authorities

- LCSH
- LC NAF
- VIAF
- FAST

...and more

- LC Genre
- Preservation Events
- MARC Relators
- MARC Countries
- MARC Languages
- Dewey.info
- ...







http://data.open.ac.uk/applications/red-author.php?uri=http://data.open.ac.uk/red/person/woolf-virginia

Links http://www.open.ac.uk/Arts/RED/ to Dbpedia

• • •



Every story has a beginning

Entering the web of data

About Full text Resources Timeline Map BOOM!

This is me.

Today, Wednesday, 14 September 2011, I'm honoured to be able to join you here in the luxurious surrounds of the Brighton Savoy Hotel for the 'Indexing See Change' conference. This is an event, a moment in history; we can pinpoint ourselves, this gathering, both in time and in space.

If we do that, if we move outside the moment and position ourselves on a timeline or a map, interesting things start to happen. Connections emerge.

Here we are at number 150, The Esplanade, in Brighton. A bit over a kilometre away is the stately villa, Kamesburgh. For many years Kamesburgh was also known as the Anzac Hostel — a refuge for permanently-incapacitated World War One veterans.

The Anzac Hostel opened on 5 July 1919. Here it is draped in its patriotic finery, from the collections of the Australian War Memorial. According to the caption, the Anzac Hostel was 'a home, not an institute'.

Also amongst the War Memorial's holdings is a wheeled bed that was used at the hostel. This particular bed was apparently occupied by one man, Albert Ward, for forty-three years.

http://wraggelabs.com/shed/presentations/ndf2012/storydata/





http://www.bbc.co.uk/blogs/bbcinternet/2012/04/sports_dynamic_semantic.html



Google Knowledge Graph http://www.google.com/insidesearch/features/search/knowledge.html

Summary

- Things not strings
- Shared identifiers makes it easier to join up disparate resources
- The web allows anyone to say anything about anything
- Approaches to Linked data still developing
- Linked Data is being used in the real world to make things, and likely to be applied in the library sector