Libraries and linked data

Ina van der Merwe

Library metadata made the transition from card to online catalogues over 40 years ago, yet the metadata in our online catalogues are no more free to interact on the web today than when they were confined to 3 x 5 inch cards in wooden drawers. What if we could set it free and let bibliographic elements like subjects, creators, dates, places … to name a few, interact independently with data on the web to which they are related?

So why is it not happening? The biggest stumbling block within the library world is the use of the MARC21 format as a coding convention. All our bibliographic records, as well as billions other from libraries all over the world, are coded in this format, a non-web data standard for sharing bibliographic information in a standard human readable form.

You are looking at a simple, yet carefully crafted MARC bibliographic record with content that conforms to controlled vocabularies; yet this format focuses on catalogue records that are independently understandable; encoded in natural language rather than as data with an inability to represent relationships between bibliographic entities.

Fig. 1 MARC bibliographic record

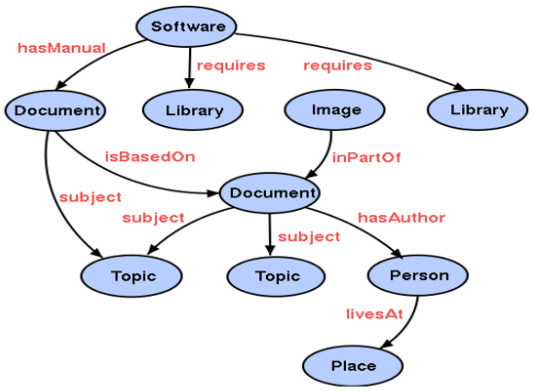
Misc.	082	0	4	823.8|223
Author	100	1		Brontë, Charlotte,|d1816-1855,|eauthor.
Title	245	1	0	Jane Eyre /|cCharlotte Brontë ; with an introduction by Lucy Hughes-Hallett.
Publication Info.	264		1	London :|bDavid Campbell Publishers,|c[2000?]
Description	300			xxxviii, 284 pages ;|c21 cm.
Note	500			Originally published: 3rd ed. 1847.
Note	500			"The Millennium library"--Dust jacket.
Note	504			Includes bibliographical references.
Subject	650		0	Mentally ill women|vFiction.
Subject	650		0	Man-woman relationships|zEngland|vFiction.

Partial record taken from b10138791, 2019, AUT LMS Sierra version 4.1

The bottom line is that MARC data is hidden in the deep web in information silos which cannot be found by search engines unless heavily manipulated into web friendly formats.

The big issue now is how to adapt to fundamental differences in how our data is used without rendering decades of legacy data completely worthless? The answer lies in the vision of the World Wide Web Consortium to build the semantic web, an environment where artificial agents, as well as human readers, are able to acquire, share and reason about information. Other than the World Wide Web, the Semantic Web is a web of structured data, of meaningful things, and it is about making relationship links.

Fig. 2 Semantic Web example



Taken from slideshow: Miller, Eric. 2002. The semantic web and the potential for library leadership

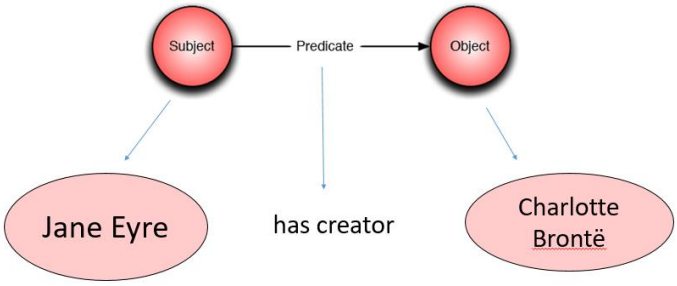
An implementation of these standards that seem to fit well with the legacy metadata produced and maintained by libraries and other cultural institutions, is Linked Data, an open standard for the web. Linked Data is all about the relationships between data.

How does it work?

Linked data is using the RDF standard for sharing data over the web. An RDF statement expresses a relationship between resources using the following structure:

subject – predicate – object; also called a triple

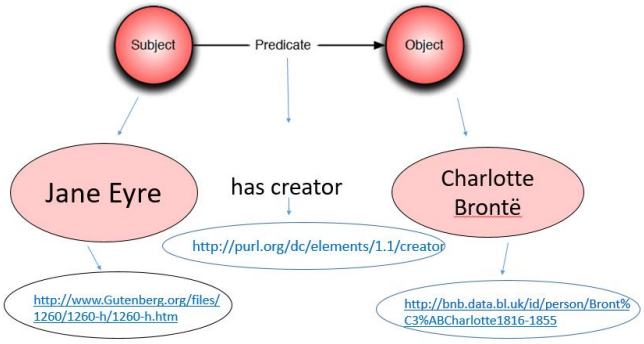
Fig 3. RDF (Resource Description Framework) data model



The subject and the object represent the two resources being related; the predicate represents the nature of their relationship

This standard also makes use of Uniform Resource identifiers, in other words a persistent ID for each element within a triple.

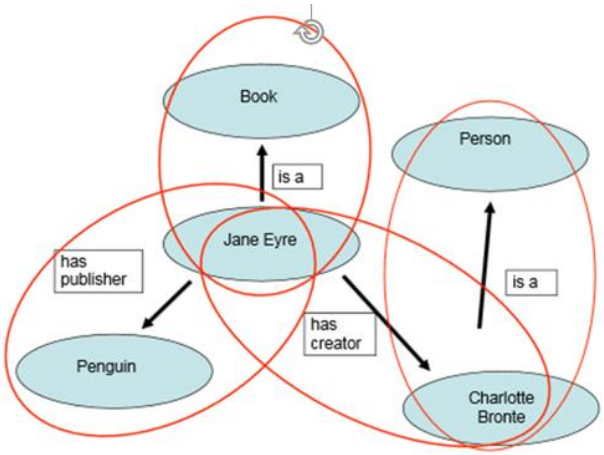
Fig. 4. Uniform Resource Identifiers



Let's refer to all subjects and objects as "things"

Overly simplified it implies that one thing has a relationship with another thing. Each thing has a URI (specific ID), so links have semantic meaning that machines can understand and act upon, and before we know it, all other kinds of relationship are found, i.e. publisher of book, all formats of *Jane Eyre*, all about Charlotte Bronte, etc.

Fig 5. Relationships



and this is a schematic diagram of linkages among metadata pieces, and this illustrates how Linked data can free our information from lockdown catalogue databases and openly share, link and enrich the data beyond libraries to increase our presence on the Web, where most information seekers can nowadays be found.

Fig. 6. Linking Open Data cloud diagram

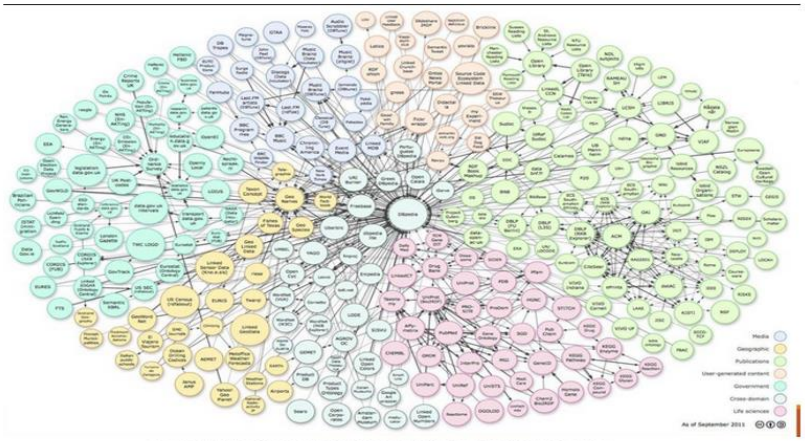


Figure 4. Linking Open Data cloud diagram 2014, by Max Schmachtenberg, Christian Bizer, Anja Jentzsch and Richard Cyganiak. http://lod-cloud.net/

References:

Gonzales, B. M. (2014). Linking Libraries to the Web: Linked Data and the Future of the Bibliographic Record. *Information Technology & Libraries*, *33*(4), 10

HASTINGS, R. (2015). Linked Data in Libraries: Status and Future Direction. *Computers in Libraries*, *35*(9), 12–16.

Rollitt, K. (2014). MARC21 to Bibframe: outcomes, possibilities and new directions. *NZLIMJ*, 55(1),

Schilling, Virginia (2012). "Transforming Library Metadata into Linked Library Data", American Library Association, September 25, 2012. http://www.ala.org/alcts/resources/org/cat/research/linked-data (Accessed March 12, 2019) Document ID: fc4b530d-00bb-4aef-b6d4-21d40f4ae19c

Tillett, Barbara. (2013). Rda and the Semantic Web, linked data environment. *JLIS* 4(1), 139.