Toward an Emerging Principle of Linking Socially-Constructed Metadata

Getaneh Alemu, Brett Stevens, Penny Ross and Jane Chandler

University of Portsmouth, Portsmouth, UK

Abstract

Current discussions on the use of Linked Data for digital libraries cater towards standards-based metadata (expert-created) metadata approaches. The inclusion of socially-constructed (user-created) metadata such as tagging, ratings, reviews and recommendations is very scant or absent. While, the efforts by the world's major national and academic libraries to release their bibliographic data (metadata) is an important step towards the adoption of Linked Data principles, as this paper argues, without metadata richness that would be obtained by including socially-constructed metadata (Web 2.0), the current efforts may result in sub-optimal utilisation of the potential of Linked Data. With the aim of exploring and investigating the potential roles of metadata linking, this paper presents results of 21 in-depth interviews conducted with Library and Information Science (LIS) professionals. A constructivist grounded theory analysis of the interview data resulted in the emergence of four metadata principles, namely the principles of metadata enriching, metadata linking, metadata openness, and metadata filtering. Considering the evidence from data analysis, it was found essential that a mixed metadata approach should be considered, so that socially-constructed metadata approaches augment and enhance standards-based ones through proactive user engagement and a platform of collaboration afforded through the adoption of Web 2.0 technologies. After highlighting on each of these principles, this paper focuses on discussions on the principle of metadata linking.

Keywords: Metadata, Linked Data, Metadata Linking, Web 2.0, Socially-Constructed Metadata, Standards-Based Metadata, Constructivist Grounded Theory Method

Introduction

Most current standards tend to adhere to what is known as the ontologically and objectively true viewpoint, thereby failing to take into account the diversity of cultural, linguistic and local perspectives that characterise library users, a stance for which they have often been criticised (Shirky, 2005; Veltman, 2001; Weinberger, 2007). Libraries have, hitherto, focused too much on standards-based metadata architectures with pre-determined schema (constraints and guidelines) and their associated controlled vocabularies and authority lists (Alemu, Stevens, & Ross, 2012). In their book "The Social Life of Information", Brown and Duguid (2000), contend that institutions in this information age tend to adopt a top-down view of technological determinism, fuelled by the likes of Moore's Law, while ignoring the socio-technical aspects of information technology. Brown and Duguid (2000, p. 1) argue that ignoring the social facets of information leaves out those aspects of information technology use that "lie around the edges, context, background, history, common knowledge and social resources". Libraries, as institutions that utilise information technologies, are not immune from being confronted with these challenges as they endeavour to effectively incorporate the about-ness of information objects (Buckland, 1999). As Buckland (1999) notes, "unfamiliar vocabulary reduces search effectiveness", indicating thus the importance of incorporating the terminologies from the users themselves, who are the raison d'être for libraries.

There are now attempts to incorporate social tagging into metadata systems (Evans, 2009; Farkas, 2007; Guy, 2006; Kroski, 2008; Maness, 2006; Mathes, 2004; Smith, 2008; Udell, 2004; Weinberger, 2005, 2007). Noting

metadata quality concerns, Lu, Park, and Hu (2010) indicate the potential role of social metadata in augmenting expert created metadata such as controlled vocabularies. However, incorporating socially-constructed metadata approaches is not enough. The metadata obtained from users should be linked with both internal and external metadata. To this end, one of the notable current efforts is Linked Data. As the name indicates, Linked Data is a data model that identifies, describes, links and relates structured data elements, analogous to the way relational database systems function, despite the fact that Linked Data is aimed to operate at a Web scale (Alemu, Stevens, Ross, & Chandler, 2012; Berners-Lee, 2009; MacManus, 2009; Shadbolt, 2010; Styles, Ayers, & Shabir, 2008; W3C, 2011; Wallis, 2011; Williams, 2010; Wilson 2010). Linked Data is a meta-model in that it provides a framework to defining, designing, developing and maintaining schemas and vocabularies of any kind and size in a given domain. This in effect means institutions, such as libraries, need not necessarily abandon existing metadata standards (RDA, FRBR and Library of Congress Subject Headings). This is because the adoption of Linked Data can provide an open interactive system, with external links and the ability to make information easily accessible, re-usable and with the possibility of serendipitous discovery of other resources, thus making existing metadata and controlled vocabularies re-usable landscape (Coyle, 2008, 2010; Coyle & Hillmann, 2007).

With regard to this, it is important that international organisations such as the International Federation of Library Associations and Institutions (IFLA) as well as regional standards authorities and libraries should, in general, take the lead and exert concerted effort in initiatives aimed at making contemporary models and standards compatible with Linked Data principles. To optimally benefit this process, libraries should make the most of their knowledge and their expertise in controlled vocabularies such as LCSH, Medical Subject Headings (MeSH), and authority lists. Overall, it is important that library metadata become compatible with Linked Data principles, and as such it is crucial that current models, such as RDA and FRBR, re-conceptualised. Nonetheless, as this paper argues, re-conceptualisation of contemporary metadata principles and adopting Linked Data for existing library metadata (a priori metadata) does not answer the full spectrum of the existing metadata challenges. Thus, it is crucial that current efforts in Linked Data fully consider socially-constructed (post-hoc) metadata.

Currently most metadata discussions cater around the technical implementations of linking such as RDF. In view of current metadata challenges, especially associated with the scarcity of theoretical underpinnings in Library and Information Science supporting metadata functions (Andersen & Skouvig, 2006; Day, 2010; Floridi, 2000; Lawrence, Kehoe, Rieger, Walters, & Kenney, 2000; Lehmann, 2010), it is of paramount importance that technical implementations of linking both standards-based and socially-constructed metadata are underpinned and supported by sound principles. It is with this goal that the principle of metadata linking is explored and developed in this research. Considering the novelty of the paradigm of Web 2.0, and thus the scarcity of existing theories, this paper neither tests a hypothesis nor puts forward a pre-formulated set of research questions. However, using the constructivist grounded theory method the goals of this paper are to:

- Inductively identify concepts and develop the principle of linking socially-constructed metadata;
- Contribute towards the development of a theory of digital library metadata, wherein the principle of metadata linking is a major component;
- Contribute to the development, extension and refinement of contemporary metadata assumptions and principles.

Methodology

Instead of the approach to testing an existing theory (deductive), theory building (inductive) was considered appropriate. It is also important to note that the research problem identification was exploratory rather than pre-formulated. A decision was then made to adopt a research methodology which allows an iterative and evolutionary process of data collection and analysis. Hence, in accordance with the nature of reality in relation to the Web 2.0 paradigm, especially with the acceptance and inclusion of diverse perspectives on matters such as metadata description, it is argued, this paper benefits by adopting an interpretive ontology, a social constructivist epistemology and an inductive methodology. It is indicated that, due to the possible diversity of perspectives that may be held by the potential research participants in this paper, an inductive research methodology would be

appropriate to investigate the problem at hand. Methodologically, the constructivist grounded theory approach affords researchers the opportunity to be part of the research process, through proactive mutual co-construction and reflexivity, during data collection and data analysis.

Thus, adopting a constructivist epistemological approach (Guba & Lincoln, 1989; Guba & Lincoln, 1994) and a constructivist grounded theory method (Charmaz, 2006; Mills, Bonner, & Francis, 2006), in-depth interviews with 21 librarians, LIS researchers, librarians and metadata experts were conducted. The interviewees included expertise from various professional experiences including heads (directors) of metadata and bibliographic services at world-renowned international, national and academic libraries, notable, well-experienced and published researchers, internationally recognised metadata consultants who are also involved in international metadata standards development, experienced faculty members, and practising librarians based at various institutions across the world. Prior to the selection of an interviewee, a prospect's publications as well as on-going research projects were reviewed so as to gather as rich and diverse a view as possible from him/her and also invoke interest in participation in the present study. Interviewees' places of work included the British Library (3), Library of Congress (1), Harvard University (1), University of Portsmouth (3), University of Loughborough (1), Kings College London (1), University of South Australia (1), University of Bologna (1), University of Parma (1), University of Zimbabwe (1), University of North Texas (1), Queensland University of Technology (1), OCLC Online Computer Library Centre (1), University College London (1), Cloud of Data (1), and 2 consultants who are not affiliated to any organisation. The interviews were conducted between January 2012 and December 2012. Prior to each interview, and following selection, introductory contacts were made, via email, in order to obtain the consent of each interviewee as well as to reach bilateral consensus as to the timing and venue of the meeting. Due to geographical dispersion and convenience of availability, 14 interviews were conducted remotely (out of which 12 were via Skype and 2 were over the telephone), while 7 of the interviews were conducted face-to-face. Once each interview is conducted, it is transcribed and stored in NVivo 9 software. The interview data was analysed, employing three stages of analysis: Open Coding, Selective Coding and Theoretical Coding. From the analysis, four overarching Core Categories emerged, namely, Metadata Richness, Linking, Openness and Filtering. Due to space limitations, detailed procedures and processes of the data analysis are not presented in this paper.

Interview Results and Discussions

From the grounded theory data analysis, four overarching metadata principles emerged, namely the principle of metadata enriching, metadata linking, metadata openness, and metadata filtering. The following sections provide a brief highlight on each of these principles and then present the principle of metadata linking in detail.

The four principles emerged from a constructivist grounded theory analysis posit that the socially-constructed (user-created, post-hoc) metadata should augment standards-based (expert-created, a priori) metadata. The adoption of such a mixed metadata approach would result in metadata that is enriched, which in turn would help libraries better accommodate the multitude of perspectives and interpretations of its users, while simultaneously enhancing the find-ability/discover-ability of information objects. The principles state that metadata that is enriched with both standards-based and socially-constructed metadata, cannot be optimally utilised unless the resultant metadata is contextually and semantically linked to both internal and external information sources, thus highlighting the importance of metadata linkage. Furthermore, in order to exploit the full benefits of such linkages, metadata should be made openly accessible, whereby it can be shared, re-used, mixed and matched, thereby precluding metadata duplication. Finally, metadata that is enriched, linked and opened should be presented to the user in a flexible, personalised, and re-configurable interface (Figure 1).
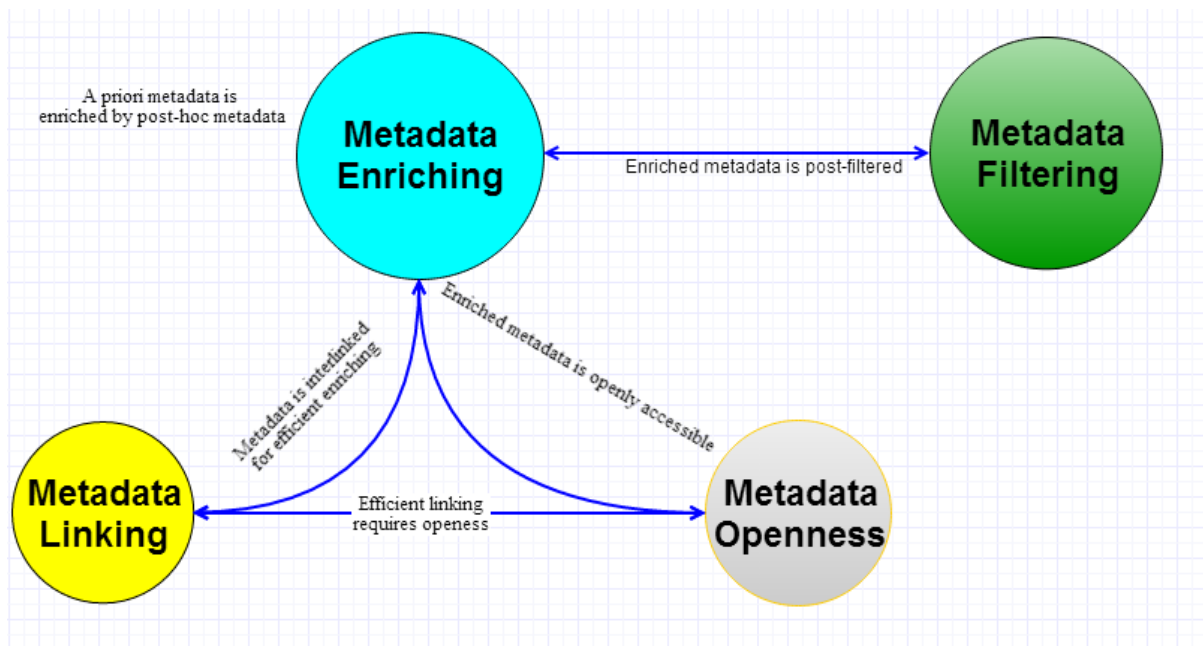
Figure 1: Four Emerging Metadata Principles

The Principle of Metadata Enriching

The results of the grounded theory analysis showed that a mixed metadata approach that combines the strengths of both standards-based and socially-constructed metadata approaches would result in metadata that is enriched. Metadata enriching should be a goal of any metadata system. Conversely, the analysis showed that standards-based metadata approaches alone cannot enrich information objects with metadata that especially reflects the diversity of interpretations that exist in library users. The absence of metadata richness in contemporary metadata approaches can be attributed to the principle of sufficiency and necessity which has been adhered by standards-setting agencies such as IFLA (IFLA, 2009; Svenonius, 2001). The principle of sufficiency and necessity essentially resulted in metadata simplicity. For example such simplicity was pursued by the Dublin Core Metadata Initiative (DCMI), resulting in a minimalist metadata schema that of consisting of just fifteen metadata elements. Whilst Dublin Core serves an appropriate purpose in being a basic descriptive standard for Web-based information, it nonetheless would not be a full-fledge metadata standard (Alemu, Stevens, & Ross, 2012; Lagoze, 2001, 2010; Veltman, 2001) .

This principle of metadata richness is thus emerged. Involving users in the creation of metadata provides a richer metadata environment, within which diverse views are reflected. Full adherence to the principle implies the incorporation of multiple interpretations, if any, pertaining to each information object, including those that originate from users, librarians, authors, and even machines. Such enrichment of information would result in a metadata ecology in which user-generated metadata is likely to be greater in volume than those derived from the other contributors, due to the simple fact that a much larger number of users will make contributions to the system. It is argued that such an increase in metadata richness would enhance the find-ability/discover-ability of information objects.

The Principle of Metadata Openness

One of the key drivers for the success of Web 2.0 approaches is the benefit that can be accrued through openness to data and information (O'Reilly, 2005). This has especially been found crucial for public institutions such as libraries to share their bibliographic information. It has been noted that the key features of Web 2.0 are two-way people-to-people sharing of data, the network effect, collective intelligence, and reusability (Floridi, 2009; Kroski, 2004; O'Reilly, 2005). However, these features cannot be fully harnessed without opening up data and systems. The network as a platform assumes that there is some level of openness to cooperation and sharing

of data (Alexander, 2006; Anderson, 2006, 2010; Shirky, 2005; Tapscott & Williams, 2010; Udell, 2004; Weinberger, 2005, 2007).

Metadata openness is a principle that emerged from the conceptualisation of the data inherent in the perspectives of LIS researchers, lecturers, practising librarians and library users. It is the culmination of the following inter-related processes: the process of metadata creation and the systems that support it, such as metadata interfaces, are designed in such a way that users are able to co-create metadata, along with librarians; Web 2.0 technologies offer both the conceptual and technical foundations that permit the implementation of socially-constructed metadata, such as tagging, rating, reviews, and recommendations; such metadata that enriches legacy library metadata can be effectively harnessed if it is made freely and openly available to external users.

The principle of metadata openness, it was observed, requires an open world assumption, in which any metadata description is valid, and the existence of multiple interpretations (metadata descriptions) are accepted and even encouraged, thus promoting metadata diversity In addition, as one interviewee stated, an open world assumption entails that metadata description is a continuous process of enhancing and augmenting, rather than being a done-state. Thus, instead of a priori complete metadata description, the open world assumption takes the view that an imperfect a priori metadata is continually enhanced by its users as post-hoc metadata.

Metadata openness does not require a priori consensus or agreement on a single set of metadata description. In other words, several interpretations of an information object could mutually exist. The main requirement is that these interpretations are represented in a format that is recognisable and process-able by the user at the other end. For this, metadata should be globally and uniquely identified so to disambiguate, link and re-use it through aggregation and mixing and matching it with other metadata values. As one interviewee noted, the open world assumption is an open thinking approach where "any opinion is valid whether you agree with it or not". However, once every kind of assertion is made about an information object, it is paramount that such metadata values are carefully assembled along with their contextual and semantic significance, thus precluding the 'anything goes' description.

If metadata is uniquely identifiable and structured, opening such metadata may provide an opportunity to create new metadata descriptions by bringing various metadata values from different sources. It is argued that metadata openness enables incremental metadata creation and aggregation, through which metadata richness would be achieved. Overall, metadata openness is an essential principle that should be embraced by libraries, standards-setting agencies and developers, in view of improving the findability and discoverability of information objects, thereby improving the experience of users.

The Principle of Metadata Filtering

According to the findings from the data analysis, library discovery interfaces (OPAC) can be improved through the provision of multiple personalised metadata views where users are able to filter metadata. The kind of filtering suggested is post-hoc rather than a priori filtering. Post-hoc filtering provides users with the flexibility to re-configure and customise their discovery interface displays. They can save search results, bookmark favourites, tag, annotate, leave comments, review and recommend books. The interface should allow users to add and remove information objects into directories of their choice, created on their mini-library catalogue.

The analysis also showed that users should be provided with suggestions during searches. Alternative spelling options (such as "Did You Mean...") should be included. Interviewees recommended that discovery interfaces should enable the integration of disparate information resources, including electronic databases. The legal and ethical aspects of personalisation-associated personal privacy concerns should be strategically addressed at an institutional level.

The inclusion of socially-constructed metadata approaches along with legacy metadata, results in possible proliferation of metadata. The size of metadata grows proportionately with the number of users and the number of collections. However, the enormity of the size of metadata presents a new challenge, i.e., the ability to filter metadata. Without appropriate filtering, metadata enriching becomes more of a problem than a solution. Unless the information displayed to users through discovery interfaces (OPACs) is properly filtered, the sheer volume

of metadata may even hamper the find-ability of information objects, as users will have a hard time navigating through the mass of metadata. This is where the principle of filtering comes to play an important role. Once such a huge amount of metadata is assembled, it should not all be presented to the user. Instead, a filtering process can be instituted, using such criteria as educational level, age, language, personal interests, previous preferences and context of use.

The Principle of Metadata Linking

The notion of metadata linking was thought to be important by interviewees as they discussed the implication of metadata richness in libraries. The increase in the size (diversity and depth) of metadata, unless, cross linked in-house and external information sources, they pointed out, provides little added value. Linkage, therefore, adds value to metadata richness. Interviewees described the importance of linkage in different ways. For example, some interviewees simply mentioned the importance of linking the library OPAC with the biographical information of authors, as found on Wikipedia and other Websites, or linking the library OPAC with authority files, controlled vocabularies, etc., all concurring on the relevance of such links. In other words, metadata richness is partially dependent on the existence of links. Unless there is a link between metadata values, the metadata cannot be mapped to any semantic structure. Thus to optimally benefit from metadata richness, metadata linkage is an important process, and thus is considered as a core category.

At present, as remarked by interviewees, the links to metadata on an OPAC display can only go up to a certain extent and soon reach a dead end, when a data item is not linked any further. For example, current metadata interfaces allow a user to search for a specific book, click on the author's name and see his publications, and, maybe find related/similar books. They, usually, do not offer a seamless link to the author's biographical page or from there to another page, such as one on Wikipedia, Google Books, etc. that cites him, and back to the library listing.

Metadata enriched with contextual and relevant links would allow users to seamlessly navigate between disparate library databases and external information providers, such as other library systems, and search engines. By globally and uniquely identifying entities (such as works, people, places, and events), metadata elements or properties (author, title, subject, relations) and corresponding values (instances), Linked Data offers a multitude of ways for enriching information objects with metadata which should facilitate information discovery and improve the experience of the user in digital libraries.

The metadata that is created and enriched using standards-based and socially-constructed metadata approaches cannot be optimally utilised unless these metadata are contextually and semantically linked to both internal and external data sources. Linking is also considered to provide a level of interoperability. One interviewee believed that metadata linkage bridges inter-disciplinary gaps and enables connections among concepts and objects across boundaries (Interviewee8-Researcher).The principles of Linked Data support the implementation of such kinds of relevant connections. Disparate data sources can be related without putting significant technical effort. The best scenario would be that such connections are seamless. In much the same way that library users do not necessarily know or care how the local and wide area networks are laid out during their use of the library network, future library users would search, find and discover internal and external information services without them being aware of making explicit connections to these latter services.

The metadata richness that resulted from the integration of standards-based and socially-constructed metadata approaches can only be optimally utilised if it is linked in-house (internal linking - within a particular library metadata set) as well as to external metadata sources (external linking). Linking metadata within and outside the library domain is deemed crucial for metadata enriching. A single institution, of whatever size and resources, cannot possibly create metadata richness on its own. Links can thus be used for further enriching, preferably in an open environment than in closed ones (silos). Links are powerful means for creating seamless connections between disparate sources of data. Stand-alone library metadata can be connected to other sources. Links also make library metadata visible. However, the interviewees noted that current library metadata linkage is not well developed (Interviewee12). The interviewees attributed the challenges to both institutional (intellectual property rights and metadata quality concerns) and technological (metadata granularity, structure and formats) factors.

The analysis of interview data discussed hereafter highlights the importance of linking technologies. Interviewees were in particular keen on the use of Linked Data (RDF) for representing and linking library metadata. It is important to note here that interviewees variously identified linking as an essential part of metadata. The interviewees identified limitations of current metadata standards. The interviewees discussed metadata linkage by referring to particular technologies and standards, such as Linked Data, RDF and Semantic Web.

Contemporary Library Standards and Metadata Linking

As the interview data revealed library standards such as MARC did not foresee the need for a linking structure, one that could be easily harnessed by simply pointing to other metadata values. One of the notable metadata standards interviewees cited for failing to adequately enabling a metadata link structure is MARC. Current metadata standards such as MARC consist of structures that do not favour linking among metadata values or making connections to external sources. This was partly due to the fact that MARC was developed in the 1970s, before the invention of the World Wide Web. MARC, one interviewee noted, was an attempt to computerise the contents of card catalogues, hampering thereafter the implementation or introduction of a link structure (Interviewee11). The interviewee recalled that, in its early days of the 1970s, libraries used to display metadata embedded in MARC on character-based dumb-terminals. The interviewees thought that libraries are to be praised for their early recognition of the value of computers and adopting them.

The incompatibility of current library formats to easily cross-link, among metadata values within and to outside libraries is considered a major shortcoming. As one interviewee (Interviewee11) stated, MARC format, rather than creating links between metadata values stored previously, re-creates an entity for each individual entry of metadata that is used for describing information objects. The interviewee illustrated this by providing the following example: "So, if you catalogued, let us say, Hamlet by William Shakespeare, you would get a catalogue card, a typical MARC record that has a lot of information in it, like who the author was, when he was born, when he died, the name of the publisher and other pieces like that. If you then catalogue another book by William Shakespeare, again you would need to duplicate all the same information". Libraries recognise the importance of changing some of their current standards and formats.

However, as one interviewee (Interviewee7) noted, "libraries started creating metadata long before other people did. They already had their concepts developed. Libraries have been creating metadata for 150 years, so libraries tend to have concepts that predate data processing. So the standards were developed to make use of those concepts" (Interviewee7).

One interviewee pointed out the inherent structural challenges as well as the obsolescence of the formats found in contemporary library metadata standards (Interviewee11). The interviewee noted that "the library management community is highly concerned about the way they store and disseminate their data and believe that Linked Data is probably the way forward. The rest of the library community doesn't know, doesn't understand, quite frankly, they wish things didn't change" (Interviewee11). Linking has powerful implications. David Weinberger (Weinberger, personal communication, December 12, 2012) had the following to say:

> "I can imagine a circumstance where a library just doesn't want to provide linking to external sources. But I think that is an unusual circumstance. The content of the library has more value generally when it is linked to other ideas and information. This is not true only for libraries; it is true for all other information providers. It is not so much that information wants to be free; it is that information wants to be linked. It wants to find other information. To cut that off seems to me to want to value the authority of the information over its utility. When you want to know something, almost in every case, you don't want to be finished with it. Sometimes you do but in many cases libraries should and need to stimulate our curiosities to the next piece and hope that next piece also leads us to more curiosity, to guide our pursuit of curiosity. There is no reason why that should occur within the library collection or the library walls and so there seems to me a natural and almost an inevitable desire to pass that boundary. That boundary is artificial. So either you encourage that and make the linking better or pretend that it is not going to happen".

There is recognition now that, as requirements for Web-based scalable metadata formats grow, such library-centric formats are becoming obsolete when it comes to supporting scalable linking. However, there is a need for data convergence, not only among libraries but also across other information sources and data providers, thus the need for adopting new formats becomes all the more relevant. Most interviewees discussed the current status of library metadata formats in comparison to XML and RDF. These new formats were considered more appropriate for representing library metadata. Interviewees acknowledged the importance of re-designing library metadata, in view of the need for linking in-house and with other data sources. The interviewees emphasized the importance of replacing these formats with flexible and scalable ones. However, it was evident that the data represented using these old formats needed to be compatible with the new formats.

One interviewee (Interviewee9) pointed out that library standards have been characterised by linear progression rather than disruptive technologies. The interviewee stated that most of the developments in library catalogues have been rather straight forward. The linear progression, the interviewee believed, has had its own ramifications, as the limitations from older formats were carried forward to new ones. The interviewee further noted the need for leapfrogging to better technical formats, while making the new model backward compatible so that no data is lost.

All interviewees, except for one who thought that it is theoretically sound but technically complex, saw RDF as the potential metadata format that could handle machine process-able and reconfigurable metadata. They noted that RDF provides better granularity than MARC or other formats currently in use in libraries. Scalable metadata creation that begins with making simple assertions was preferred by interviewees to one that relied on the traditional approach of creating complete metadata records using a document-based record structure. The importance of linking to various metadata data sets and mixing metadata using application profiles was considered, by interviewees, as an important component for improving current metadata creation and usage in libraries.

Contemporary library legacy metadata is not enriched with links. The standards and formats used are limited in their handiness for supporting the optimal usage of links. In addition, the concerns for metadata quality, intellectual property issues, and the proprietary nature of library management systems, all hamper metadata linking. As one interviewee alluded, libraries tend to have metadata concepts that predate data processing. This is attributed to an enduring history of library cataloguing and its associated principles. As a result, most interviewees, held the opinion that library metadata has been developed only within library circles. Linking to outside data sources, it was pointed out, has been minimal. Furthermore, library systems were designed to fit library metadata formats. Commercial software vendors are required to base their system design and development around library standards, formats and principles (Interviewee7).

Metadata Linking Technologies

Interviewees expressed their familiarity with the technologies associated with Linked Data, while many alluded to the complexities of the underlying formats such as RDF, OWL and SPARQL. Development of Linked Data in libraries was considered very slow. While there is recognition among the library community that these technologies can be useful for representing library metadata, so that it can be usable in the Web environment, it was also emphasised that Linked Data should not be used for the new technology's or model's sake only. The metadata consultants who worked with library metadata maintained that Linked Data is an appropriate model for scalable library metadata.

The current metadata conversion efforts from contemporary library formats, such as from MARC to RDF, as interviewees observed, were not well thought through nor consensually agreed upon. These efforts are mostly experimental and patchy. The interviewees concurred that the linking infrastructure and the potentials for a scalable metadata model present a good case for the adoption of Linked Data in libraries. However, as the interviewees maintained, libraries are not yet in a position to deal with the technical complexities of RDF and the institutional challenges that it would bring about. In regard to this, it is imperative that a proper distinction is made between the conceptual underpinnings and the technical tools required for metadata linking, so that

libraries can develop the skill sets in-house or use external expertise. At present, the technical complexities seem to have hindered the adoption of the conceptual foundations of Linked Data.

One interviewee acknowledged the efforts, by national libraries and international organisations such as IFLA, in this direction. In contrast, another interviewee criticised current efforts for their lack of coordination and commitment of adequate resources, remarking

> "The work that IFLA is doing to take the standards formats, such as ISBD and UNIMARC, and putting it in RDF is not going to get us where we need to be. It is an interesting exercise, but it isn't really producing [the desired result]. It is much like what we did when we moved from the card to the MARC, where we carried over the same data in the same way, just into a new format. What we need to do is to change the way of thinking about our data. So I kind of understand IFLA doing this, but by doing that I just don't think that moves us very far forward" (Interviewee7).

Making the distinction between the conceptual and technical sides of Linked Data is crucial. In this regard, it is argued, it is important to delineate contemporary library principles and models such as RDA and FRBR vis-à-vis library record format such as MARC. Libraries can leapfrog from current formats to RDF and OWL. However, it is mandatory to re-conceptualize the underlying principles embedded in RDA and FRBR, in order to make these library principles compatible with Linked Data. One interviewee contends that "The question again is whether by doing so we will make enough change. Because RDA, after they started developing it, people were complaining and saying that it was too different, and so, they added back some of the practises, such as alphabetical headings."

As the data analysis showed, in principle, Linked Data is a viable opportunity for libraries to link their metadata with external data providers, which in turn would make library metadata connected to external information sources. However, libraries are not easily going to abandon their principles, which they have developed over hundreds of years. Linked Data principles however do not require libraries to abandon their existing principles and legacy databases. Instead, the question for Linked Data at present is to incorporate four major principles, which would enable libraries to external data sources on a global scale. Linked Data principles start from the most fundamental component of Linked Data, i.e., the use of globally unique URI as names for distinctively denoting such things as information objects, people, places, and events (Berners-Lee, 2009). Other essential principles include an ability to dereference a given URI, using Hypertext Transfer Protocol (HTTP), in order to retrieve relevant information as to what this particular URI refers to; use of data formats, such as RDF/XML, use of vocabulary definition languages, such as RDFS and OWL, and use of query language known as SPARQL; and last but not least is the incorporation of links within and outside the data sets, thereby enriching the data and providing contextual significance (Berners-Lee, 2009).

Links are the most important enabling components of metadata. The interviewees expressed their preferences for the features of links that they would like to see in library metadata:

> "The problem of the Web at the moment is that the link basically says 'See also', it doesn't say anything about the authority of the linker, it doesn't say anything about why they are linking, whereas by using some of the standards of Linked Data, you can make a link, you can describe your authority, you can describe the form of the comment you are making, and there is plenty of scope for disagreement, plenty of scope for interpretation. How a system displays that to the end user is an open question. But all that should be encoded in the description, it is all there" (Interviewee12).

The interviewees tended to gravitate towards particular technologies and formats, such as XML, RDF, Linked Data, OWL, and URI, in their discussion. They described in detail the features and benefits of these technologies. One interviewee compared Linked Data with prior technologies such as relational database systems. The interviewee noted,

> "The power of Linked data is the link. It is the fact that an assertion or an object described in one place, by one person, is linked over the Web to things said elsewhere. A RDBMS is a closed box. It is a very

powerful closed box, but it is very difficult for a relational database to say something about an object and to know it is the same object described by another person in another database" (Interviewee12).

The use of universal URIs, RDF, RDFS and OWL are the main elements of Linked Data technologies. Using Linked Data models, metadata can be represented in simple statements of subject-predicate-object. For example, a metadata statement is fractured in to three (subject-predicate-object) constituent elements such as *"George Orwell → isAuthorOf → Animal Farm"*.

Linked Data, interviews pointed out, can help to break metadata records into a series of atomic metadata statements, which then can help to repurpose and recompose to describe information objects. Linked library metadata, opens opportunities to be mixed with other metadata sources outside the remit of libraries. The metadata should be done using a well-developed approach of linking data elements.

All the four metadata consultants interviewed (Interviewee7; Interviewee9; Interviewee11; and Interviewee12) considered Linked Data an important model to obviate some of the metadata challenges associated with contemporary metadata approaches of creating metadata records as human-readable document. Linked Data promotes the creation of linking incrementally. The incremental link creation can be made not only by one organisation but the new structure allows other organisations to cross-link to the metadata maintained by others.

Current implementations of Linked Data, as mentioned by interviewees, include the British National Bibliography (BNB) – which is published using RDF, OCLC's Virtual International Authority File (VIAF), LCSH, Lexvo, GeoNames, MARC country, and language, Dewey.info, RDF Book Mashup. Interviewees indicated that current developments in using Linked Data for libraries are in its infancy. One interviewee described the current status by stating the problems of absence of agreement on persistent identifiers resulting in weak (limited) connections between data stores. The interviewee described the development of Linked Data for library metadata as very slow (Interviewee7).

Re-conceptualising Library Metadata

The data analysis revealed interviewees' preferences for the creation and representation of metadata that could easily be linked to both internal and external sources. To that end, they underscored the importance of machine process-able, aggregate-able and re-combinable metadata, wherein each metadata value and its associated property (metadata element/field) is uniquely identifiable. For instance, one interviewee dwelt on the importance of decomposing (or atomizing) what had previously been a document-like metadata record structure into a whole series of statements which can provide maximum flexibility. The interviewee noted that the individual statements can then be joined with other statements, thus providing a more complete description of information objects (Interviewee7). This revealed the need for a re-conceptualisation of current metadata formats.

Library agencies at present have been deliberating on finding sustainable solutions to long standing problems of metadata technologies, especially in relation to making library metadata easily interoperable through links. As the interviewees noted, there have been attempts to translate MARC records into MARC/XML, and recently MARC/XML into RDF. At present, as one interviewee observed, some of the world's biggest library institutions, including the Library of Congress, the British Library, the national libraries of France and Germany, and Europeana are playing active roles in this arena (Interviewee4). The re-conceptualisation of metadata representation and linking has several implications on the current use of metadata.

Breaking Metadata Values into Granular Metadata Statements

Links scale if the metadata about an information object is described at a granular or atomic level. As one interviewee (Interviewee11) noted granular description allows metadata to be mixed and recombined. The same interviewee underscored that greater metadata utility emerges from creating simple metadata statements that scale; each statement can then be dissected and linked to another metadata set based on need. Interviewees emphasised the importance of structuring metadata in such a way that it can be linked not only within an institutional or system's boundary but also in the wider Web context.

With such a scalable design of metadata linking, as envisaged by interviewees, an assertion statement can be uniquely identified, and, thus, can be contextualised with other metadata assertions elsewhere (Interviewee12). The interviewee further elaborated that, by using the same identifier, two or more metadata statements can become linked to each other. The same interviewee, however, noted the limitations of the current Web technologies to develop scalable liking at the data level. The interviewee maintained that Linked Data offers better opportunities for providing solutions to current challenges.

Interviewees (such as Interviewee6 and Interviewee11) stated, metadata linking would enable libraries to create cross-links between metadata so that metadata about authors, publishers, subject-vocabularies, and identifiers can be re-used, thus avoiding duplication of effort. As the data analysis showed, in order to allow metadata linking at the atomic (single most common element such as a keyword, or name), the underlying metadata structure plays a significant role. Metadata that is granularly structured facilitates machine process-ability. Library metadata such as that embedded in the traditional record format is stored and exchanged by machines but it predominantly tends to remain document-centric where human users must painstakingly check and use it as opposed to machine process-able metadata that can automatically be re-used and re-purposed.

Unique Metadata Identifiers

The interviewees discussed the issue of uniquely identifying information objects variously, citing the examples of specific instance techniques such as the use of ISBN, URI, and classification numbers. The importance of identification mechanisms that scale and interoperate is considered an important feature of linking library metadata. The role of namespaces to uniquely and globally identify metadata sets and values was emphasised (Interviewee4; Interviewee9). There are currently several international and institutional efforts to develop namespaces for unique identification. However, it was indicated that best practises on the use of global identifiers for metadata is yet to emerge. As one interviewee described the use of an identifier allows recombining various metadata elements and values from disparate sources (Interviewee4). One interviewee described the current status and future prospects of globally identifying metadata as: "If you like to link back to information about a book in the library world, today that is a very difficult thing to do because there is no universally agreed upon identifier for that book" (Interviewee11).

One of the challenges of identity in libraries has been that the usage mechanisms have been predominately used within libraries themselves, and not by non-library service provides, thus making it a challenge to cross-link disparate resources. Current efforts at finding mechanisms for identification (such as the IFLA Namespaces Group, the British Library Data Model, and the Europeana Data Model) focus on standards-based metadata, there is little effort expended to develop identification mechanisms for socially-constructed metadata.

In order for it to be re-usable and shareable, metadata should not only be uniquely identified but the identification mechanism should also be persistent. For this to happen, there should be a concerted co-ordinated effort to develop global identification schemes for metadata elements (properties) and their corresponding values. Due to the apparent enormity of the effort to uniquely identify each metadata value, the efforts should be undertaken both at decentralised and centralised levels. Whilst there are efforts at devising identification mechanisms for standardised metadata, such efforts are not developed to include social metadata.

Uniquely Identifying Socially-Constructed Metadata

It is important that standards agencies and libraries work towards the development of unique identification mechanisms for socially-constructed metadata, so that re-usability and mash-ability of such metadata can be realized. Until such identification mechanisms are found, social metadata will largely remain constrained to the system in which they were created. In addition, as the interviewees mentioned, the provenance of metadata can only be made possible if particular user metadata can be uniquely identified and located, so that the context of its use and who has created it is can be determined. One interviewee noted that "there are many suppliers that are mining text, mining twitter streams, and mining status updates on Facebook to identify concepts, meanings, emotions, sentiments, and things like that, but at the present moment all these systems are very closed. But there is no agreed international identifier at the moment" (Interviewee11). The use of identifiers for social metadata,

as the interview data showed, is not developed very well. It is thus crucial that unique identification mechanisms are developed for socially-constructed metadata.

In connection with the above, if we take tagging, the most structured type of socially-constructed metadata, it has three major components: user, tag and information object. The user is a person who tags an information object. Each of the components can be considered as a class and instances can be identified by a Uniform Resource Identifier (URI) scheme. A URI is a global unique identifier which forms a fundamental part of Linked Data (Allemnag & Hendler, 2008).

Interviews saw benefits of linking library metadata with outside sources such as Amazon. In addition, library metadata can be linked to other library metadata thus avoiding duplication of effort. As one interviewee noted "if the books are not available on the shelf, there is a link to take you off where you would be able to purchase the book" (Interviewee2). Such links may also be integrated to sites such as LibraryThing where users are offered the possibilities of looking for information about a particular book, whether to purchase it or read reviews about it. Linking could enrich the OPAC/discovery interface with more relevant information from outside sources. For example, some users may want to check the biography of the author of a book, which is not recorded in the library metadata but exists on outside sources. Another example, mentioned by interviewees, is book covers which can be copied from external providers, such as Amazon, rather than the library trying to recreate them in-house. Book jackets offer users a visual aid to identify books on shelves and electronic lists. Please note that even through book jackets can be considered a priori metadata (for example as used in Amazon) but as exemplified in LibraryThing, they can be contributed by users, thus making it a constituent of post-hoc metadata.

One interviewee concurred, "I think there is just so much to be done with these metadata and joining it up so that it is not all of us sitting in our little silos of data [which may include the ability to connect with other universities". They also saw the benefits of nation-wide integrated catalogue where one could check what other universities have books and for a user to go and read these books. At present, one could search individual catalogues but it would be more effective and efficient to just search one catalogue. As huge amounts of data are being generated at present and computing power is increasing, they pointed out there is a lot of potential to link, share and utilise this data in the future. Hence, socially-constructed metadata can be uniquely identified and linked so that it can be mixed, matched and re-used. One interviewee mentioned the possibility of sharing tags if it is linked and made available to other users (Interviewee4). Linking metadata and making it available for others to re-use, interviewees noted, adds its return on investment as other libraries and applications may use it for various purposes. One interviewee (Interviewee6) suggested the possible use of library metadata to generate citations automatically.

Benefits of Metadata Linking

Links are a powerful means to improve navigation and browsing experiences by making it easy for users to jump from one information object to another. As one interviewee noted, links provide mechanisms to link from one source to another, to say "this movie is adapted from this book, or this book is based on this other book" (Interviewee7). Links, if consistently designed and implemented, provide scalable means to enrich metadata as different people can create connections between information objects. In other words, the critical mass of links is beneficial to enrich metadata with contextual relations. That is "people will get interested, those people who have read a lot on a topic, will go there and make a couple of connections, somebody else will make a couple of other connections, and then you reach a kind of a critical mass of all these connections" (Interviewee7). Metadata values that are not linked have low visibility by either machines or human users. One interviewee noted that present library metadata has low visibility that affects the library collection's discoverability (Interviewee7). More links, the interviewee noted, is the solution to these challenges.

The benefit of linking is that library resources become more usable. One interviewee noted that "people will be able to find library resources while they are doing whatever kind of information seeking on the Web" (Interviewee7). Attributing Vannevar Bush's idea of the memex, one interviewee cited the importance of the

notion of associative trails or links between documents. Associative trails are one of the great benefits of rich links while also noting that it is also how users get lost navigating and discovering through the links.

> "Because, we want to know how things are connected to one another, one of the ways we learn is by making those connections. So if I am to describe what the inside of an eyeball is like, I would say it is something like jelly, if you've encountered jelly before you know what I'm talking about. If we have digital object of whatever kind, not only that document has some kind of creator, somebody must have made it, somewhere, somehow, but documents themselves belong to a whole Web of other documents. No document exists by itself, not only is it linked to other documents, but to other ideas" (Interviewee8).

As one interviewee described one of the advantages of metadata linking is that it can be enriched with more links in an incremental manner. In addition, linking also means the library should not necessarily create all the metadata description in-house, instead links can be created to external metadata descriptions. For example, a link can be made between a metadata value in an OPAC display to a socially-constructed tag, rating or review in LibraryThing.com (an online collaborative social cataloguing Website).

Interviewees also stated the benefits of metadata linking to aggregate metadata from disparate sources, which in turn is considered to facilitate metadata interoperability, improving cross-searching of information systems.

In contrast to current efforts that focus on incorporating formal and authoritative metadata, most interviewees accentuated the importance of incorporating metadata descriptions created by users. To this end, metadata is aimed to benefit users, it should not be jargonised by language, which might be difficult for users to understand. One interviewee said, "I don't need to have a library science degree to understand the metadata" (Interviewee11). In addition, one interviewee noted that multilingualism aspect of metadata should be considered. Metadata multilingualism was in particular thought to be useful to represent controlled vocabularies such as subject headings (Interviewee11). It is important to note here that the idea of linking metadata values is mainly dependent on the design of metadata identifiers, structure and granularity. Including users, linking to non-library sources, multilingual metadata, harnessing local and international metadata descriptions all seem to indicate to an underlying requirement for metadata diversity.

Interviewees noted the particular relevance of linking library metadata with outside sources such as LibraryThing. One interviewee pointed out that, "if the books are not available on the shelf, there is a link to take you off where you would be able to purchase the book". Such links may also be integrated to sites such as LibraryThing where users are offered the possibilities of looking for information about a particular book whether that is to purchase it or read reviews about it. Linking to external sources such as Amazon or Google Books, as one interviewee (Interviewee5) indicated benefits users to read tables of contents and free chapters so that they can determine the relevancy of particular information objects for their study or research, thus indicating the importance of purposeful linking to such sources. Users may also want to buy information objects rather than borrowing them from the library. One interviewee saw no problem of making such processes easier for library users.

Metadata linking adds value to the experiences of users by providing contextual and relevant connections between resources which traditionally seems to have appeared disconnected. Relationship between concepts in a search result can be linked to subject headings which have structured semantic links between broader, narrower, synonym and related terms. One interviewee (Interviewee4) emphasised that sharing of metadata, crowd sourced tags and the idea of network effect for metadata can only be achieved through the use of effective linking mechanisms. As one interviewee (Interviewee6) noted library metadata should be semantically described and granularity and machine process-ability should be sought so as to utilise such metadata for various purposes other than library cataloguing. One such use case mentioned by the interviewee was citation analysis where similar names of authors and their contributions can be disambiguated so as to also ease the process of citation impact analysis.

Implications for Digital Libraries

From Metadata Simplicity to Metadata Enriching

Libraries should consider socially-constructed metadata approaches as a novel opportunity to re-orient their metadata functions to meet users' needs. In this regard, libraries should be cognisant of the fact that metadata contributors in a socially-constructed metadata approach can possibly be people of expertise in their subject domains and, thereby be valued contributors, for a number of motivations including prestige, altruism or just concern for one's field. In this regard, since librarians are not always experts in subject domains, involving users enables libraries to solicit valuable metadata with regard to the about-ness of information objects. Since post-hoc metadata is created by users themselves, the likelihood to match their information search keywords is much higher than if standardised and formal taxonomies are used.

From Non-Linked to Interlinked Metadata

The principle of metadata linking, which emerged in this research as an important consideration for libraries, would address some of the major limitations of current standards-based metadata approaches. The data analysis revealed that, current library metadata standards are not designed for linking. Thus a lot of future work is needed. The principle of linking can obviate these challenges by providing a granular, machine process-able, re-usable and link-able structure. These benefits can be used in a mixed metadata approach. Thus, the metadata diversity that is resulted from the adoption of socially-constructed metadata approaches can be effectively structured, affixed with provenance and inter-linked. In this regard, current metadata standards and models such as RDA and FRBR should be re-conceptualised, so that it can be retrofitted to the principle of linking.

From Metadata that is Locked (Silos) to Metadata Openness

Embracing the principle of metadata openness helps libraries to look beyond their institutional silos, further benefitting themselves through metadata sharing, re-use, mixing and matching, and integration. Metadata openness significantly reduces the cost of metadata creation by avoiding duplication of effort. However, issues of licensing need to be addressed. Metadata created for a specific purpose could thus be re-purposed and re-used to achieve a completely different purpose in a manner that may not have been foreseen by the original creator. In addition, as the data analyses revealed, metadata openness furthers institutional transparency and accountability. Finally, metadata openness, rather than guarding metadata in institutional silos, increases the returns on investment expended on metadata creation. Hence, libraries should proactively coordinate and implement open data policies and strategies.

Conclusions

The emerging metadata principles (note that this paper specifically focuses on the principle of metadata linking), emerged from empirical grounded theory data analysis technique suggest a paradigm shift in the future of metadata:

- From metadata simplicity to metadata enriching;

- From human-readable and non-linked metadata to granularly structured, uniquely identified, machine process-able and interlinked metadata;

- From metadata that is locked to metadata openness;

- From a single (global) OPAC interface to metadata that can be filtered through re-configurable interfaces.

Among the four emerging principles, the principle of metadata linking posits that in order to optimally benefit from metadata enriching, libraries should aim to harness both in-bound and out-bound links. Effective linking requires a metadata design that takes into account the unique identification of metadata elements and corresponding values. Unique identifiers that persist facilitate metadata sharing and interoperability. Since metadata enriching is provided through the combined use of standards-based (a priori) and socially-constructed (post-hoc) metadata, it is of paramount importance that both strands of metadata are uniquely identified and

maintained. Metadata that is designed with granular structure and metadata that is enriched both with in-bound and out-bound links helps to minimise or avoid duplication of metadata records with-in and among institutions. It brings disparate metadata sources together. It widens opportunities for collaboration and opens vistas for novel uses of metadata that may not have been envisaged by the original metadata creators.

The findings, with regard to the importance of linking metadata, are consistent with emerging web developments, especially with that of Linked Data and the Semantic Web. Whilst the standards-based legacy metadata resulted in read-only (Web 1.0) metadata where users have a passive role of merely searching or browsing through library collections using the library catalogue, socially-constructed metadata approaches result in read-write metadata (Web 2.0) where users are pro-active co-creators as they can participate in the creation and management of metadata. However, in what would be considered as a web 3.0 approach, as the grounded theory data analysis in this research showed, the strengths of both standards-based and socially-constructed metadata approaches can be brought together in a mixed metadata approach, thus metadata that is enriched through socially-constructed metadata can be structured, aggregated and analysed providing a collective metadata intelligence. For all these to take effect, the power of links and relationships is obviously indispensable.

References

Alemu, G., Stevens, B., & Ross, P. (2012). Towards a conceptual framework for user-driven semantic metadata interoperability in digital libraries: A social constructivist approach. *New Library World, 113*(1/2), 38-54.

Alemu, G., Stevens, B., Ross, P., & Chandler, J. (2012). Linked Data for libraries: Benefits of a conceptual shift from library-specific record structures to RDF-based data models. *New Library World, 113*(11/22), 549-570.

Alexander, B. (2006). Web 2.0: A New Wave of Innovation for Teaching and Learning? *EDUCAUSE Review, 42*(2), 32-44.

Allemnag, D., & Hendler, J. (2008). *Semantic Web for the working ontologist: effective modeling in RDFS and OWL*. Amsterdam: Morgan Kaufmann.

Andersen, J., & Skouvig, L. (2006). Knowledge organization: a sociohistorical analysis and critique. *Library Quarterly, 76*(3), 300-322.

Anderson, C. (2006). *The long tail: how endless choice is creating unlimitted demand*. London: Random House Business Books.

Anderson, C. (2010). *Free : how today's smartest businesses profit by giving something for nothing*. London: Random House Business Books.

Berners-Lee, T. (2009). Linked Data. http://www.w3.org/DesignIssues/LinkedData.html

Brown, J. S., & Duguid, P. (2000). *The social life of information*. Boston, Mass: Harvard Business School.

Buckland, M. (1999). *Vocabulary as a Central Concept in Library and Information Science.* Paper presented at the Third International Conference on Conceptions of Library and Information Science Dubrovnik, Croatia.

Charmaz, K. (2006). *Constructing grounded theory: a practical guide through qualitative analysis*. London: SAGE Publications.

Coyle, K. (2008). Meaning, Technology, and the Semantic Web. *Journal of Academic Librarianship, 34*(3), 263-264.

Coyle, K. (2010). Library Data in a Modern Context. *Library Technology Reports, 46*(1), 5-13.

Coyle, K., & Hillmann, D. (2007). Resource Description and Access (RDA): Cataloging Rules for the 20th Century. *D-Lib Magazine, 13*(1/2).

Day, R. (2010). The Self-Imposed Limits of Library and Information Science: Remarks On the Discipline, On the Profession, On the University, and On the State of "Information" in the U.S. at Large Today. *InterActions: UCLA Journal of Education and Information Studies, 6*(2).

Evans, W. (2009). *Building library 3.0: issues in creating a culture of participation*. Oxford: Chandos.

Farkas, M. G. (2007). *Social software in libraries: building collaboration, communication, and community online*. Medford, N.J: Information Today, Inc.

Floridi, L. (2000). On defining library and information science as applied philosophy of information. *Social Epistemology, 16*(1), 37–49.

Guba, G., & Lincoln, S. (1989). *Fourth Generation Evaluation*. London: SAGE Publications.

Guba, G., & Lincoln, S. (1994). Competing paradigms in qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 105-117). Thousand Oaks, CA: SAGE.

Guy, M. (2006). Folksonomies: Tidying up Tags? . *D-Lib Magazine, 12*(1).

Kroski, E. (2008). *Web 2. 0 for librarians and information professionals*. New York: Neal-Schuman.

Lagoze, C. (2001). *Accommodating Simplicity and Complexity in Metadata: Lessons from the Dublin Core Experience* Paper presented at the Seminar on Metadata, Amsterdam.

Lagoze, C. (2010). *Lost Identity: the Assimilation of Digital Libraries into the Web.* (PhD), Cornell University, School of Information Science. Retrieved from http://www.cs.cornell.edu/lagoze/dissertation/CarlLagoze.pdf

Lawrence, G. W., Kehoe, W. R., Rieger, O. Y., Walters, W. H., & Kenney, A. R. (2000). *Risk Management of Digital Information: A File Format Investigation*. Washington, D.C.: Council on Library and Information Resources.

Lehmann, H. (2010). *Grounded Theory and Information Systems: Are We Missing the Point?* Paper presented at the The 43rd Hawaii International Conference on System Sciences, Koloa, Kauai, Hawaii.

Lu, C., Park, J., & Hu, X. (2010). (2010). User tags versus expert-assigned subject terms: A comparison of Librarything tags and Library of Congress subject headings. *Journal of Information Science, 36*(6), 763–779.

MacManus, R. (2009). Understanding the New Web Era: Web 3.0, Linked Data, Semantic Web. *Read Write Web*. http://www.readwriteweb.com/archives/understanding_the_new_web_era_web_30_linked_data_s.php

Maness, J. M. (2006). Library 2.0 Theory: Web 2.0 and Its Implications for Libraries. *Webology, 3*(2).

Mathes, A. (2004). Folksonomies - Cooperative Classification and Communication Through Shared Metadata http://www.adammathes.com/academic/computer-mediated-communication/folksonomies.html

Mills, J., Bonner, A., & Francis, K. (2006). Adopting a constructivist approach to grounded theory: Implications for research design. *International Journal of Nursing Practice, 12*, 8-13.

O'Reilly, T. (2005). What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software. http://oreilly.com/web2/archive/what-is-web-20.html

Shadbolt, N. (2010). *data.gov.uk - The Linked Data Revolution*. Paper presented at the Innovating Through Information Lecture Series, London School of Economics, London. http://eprints.ecs.soton.ac.uk/18787/

Shirky, C. (2005). Ontology is Overrated: Categories, Links, and Tags. *Clay Shirky's Writings About the Internet*. http://www.shirky.com/writings/ontology_overrated.html

Smith, G. (2008). *Tagging : people-powered metadata for the social web*. Berkeley, CA: New Riders.

Styles, R., Ayers, D., & Shabir, N. (2008). *Semantic MARC, MARC21 and the Semantic Web*. Paper presented at the Linked Data on the Web (LDOW2008), Beijing, China. http://events.linkeddata.org/ldow2008/papers/02-styles-ayers-semantic-marc.pdf

Tapscott, D., & Williams, A. D. (2010). *Wikinomics: how mass collaboration changes everything* (Expanded ed.). London: Penguin Books.

Udell, J. (2004). Collaborative knowledge gardening With Flickr and del.icio.us, social networking goes beyond sharing contacts and connections. *InfoWorld*. http://www.infoworld.com/d/developer-world/collaborative-knowledge-gardening-020

Veltman, K. H. (2001). Syntactic and semantic interoperability: New approaches to knowledge

and the semantic web. *New Review of Information Networking, 7*(1), 159-183.

W3C. (2011). Library Linked Data Incubator Group Final Report: W3C.

Wallis, R. (2011). Library of Congress To Boldly Voyage To Linked Data Worlds. Retrieved from http://consulting.talis.com/2011/11/library-of-congress-to-boldly-voyage-to-linked-data-worlds/

Weinberger, D. (2005). Tagging and Why It Matters. http://cyber.law.harvard.edu/sites/cyber.law.harvard.edu/files/07-WhyTaggingMatters.pdf

Weinberger, D. (2007). *Everything is Miscellaneous*. New York: Times books.

Williams, H. (2010). Linked Data and libraries. *Catalogue & Index* (160), 2-5.

Wilson , N. (2010). *Linked Data Prototyping at the British Library*. Paper presented at the Talis Linked Data and Libraries event, British Library, London. http://blogs.talis.com/nodalities/neil-wilson-lightning-talk-%E2%80%93-linked-data-and-libraries-2010