

## **Perspectives of LIS Academics and Post Graduates on Standards-based and Socially-Constructed Metadata Approaches**

**Getaneh Alemu, Brett Stevens and Penny Ross**  
University of Portsmouth, Portsmouth, UK

### **Abstract**

With the aim of developing a conceptual metadata framework that takes into account user-driven metadata, this paper presents the perspectives of library and information science (LIS) academics and postgraduates on optimality of standards-based versus socially-constructed metadata approaches. Based on analysis of data collected through eleven in-depth interviews, three broad metadata approaches are identified: standards-based, socially-constructed and mixed-metadata approaches. In relation to standards-based approaches, the paper discusses the obsolescence of the OPAC in comparison to contemporary web search engines and the limitations of standards in addressing the changing user needs, expectations and evolving vocabularies. The paper also discusses the roles ascribed to Web 2.0 and Linked Data technologies in libraries, and the importance of enriching information objects with metadata that better conveys the various perspectives of users. The paper suggests a mixed-metadata approach that includes the strengths of the standards-based metadata approaches and the social space of metadata that comes from socially-constructed metadata approaches.

**Keywords:** Metadata, OPAC, Web 2.0, Linked Data, Socially-constructed metadata

### **1. Introduction**

As the size of collections in digital libraries continues to grow, contemporary standards-based metadata approaches fail to scale in enriching the ever increasing amount of information objects with appropriate metadata (Lagoze, 2010; Lagoze & Patzke, 2011; Mathes, 2004; Shirky, 2005; Veltman, 2001; Weinberger, 2007). This is due, first of all, to the fact that current metadata standards are mainly concerned about the physical characteristics of information objects (Day, 1997 as cited by Wright, 2007, p. 86). Secondly, librarians may not have the required expertise in specialised domains (e.g., Egyptology) in order to adequately describe the semantic aspects of information objects (Lagoze, 2010; Shirky, 2005). Thirdly, as the size of their collections grows, librarians will increasingly find it difficult to describe every digital object (Shirky, 2005). In this connection, Lagoze and Patzke (2011, p. 375) contend that “well established and well-known bibliographic standards, undeniably useful for traditional library cataloguing of a range of resource types, have been shown to be too complex and not scalable for digital resources”. Fourthly, as

Shirky (2005), Veltman (2001), and Weinberger (2007) argue, human beings are highly unlikely to agree on a singular, authoritative and hierarchical classification of objects.

Unfortunately, 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, an outcome for which they have been often criticised (Shirky, 2005; Veltman, 2001; Weinberger, 2007). As Shirky (2005) argues, the art of classifying and categorising information objects in the digital world using traditional categorisation schemes is a forced limitation imposed from the habits of organising physical objects in the physical world. Thus, Shirky (2005) and Weinberger (2007) argue that conventional library classification systems, such as the Dewey Decimal (DDC) and Library of Congress (LC) classification systems, are severely limited in their ability to organise and make digital content accessible and findable for the diverse group of digital library users. In addition, it is also argued that the cost of creating and maintaining taxonomies is a major constraint (Barbosa, 2008; Shirky, 2005). In his oft-cited article, Shirky (2005) contends that “well-managed, well-groomed organizational schemes get worse with scale, [...] because the costs of supporting such schemes at large volumes are prohibitive”. Perhaps the most important of these limitations, which is also the concern of this paper, is that in standards-based metadata approaches “the vocabulary of the information professional/librarian may be at odds with the language of content creators and users, obscuring the very information the taxonomy should reveal” (Barbosa, 2008).

To overcome these limitations, an alternative approach has evolved with the advent of social media and Web 2.0 technologies (O'Reilly, 2005). A collaborative metadata approach, known as “folksonomy” (Vander Wal, 2007), has emerged and is currently in widespread use in popular web applications such as Wikipedia, Flickr, LibraryThing, Delicious and Technorati (Farkas, 2007; Maness, 2006; Pressley, 2005; Smith, 2008; West, 2007; Yi & Chan, 2009). Vander Wal (2007) defines folksonomy as “the result of personal, free tagging of information and objects (anything with a URL) for one's own retrieval. The tagging is done in a social environment (usually shared and open to others)”.

In this paper, the term socially-constructed metadata is adapted as an approach that caters for not only social tagging, but also for the incorporation of user reviews, comments, highlights, ratings, and recommendations. For example, some social-bookmarking applications (e.g., diigo.com) and e-book readers (such as Amazon's Kindle and Apple's iPad) provide users, among other services, with features to highlight a part of a text, leave comments on margin (sticky) notes or write critical reviews about some of the works offered on readers. In principle, service providers could glean relevant information from such user-contributed metadata.

In a socially-constructed metadata paradigm, users not only search/browse and access content but also participate in its production and description. Put simply, this paradigm provides users of online systems (such as Flickr) with easy to use, flexible, collaborative and personalised tools to proactively tag, rate, link, review, comment, highlight and recommend content without the need for adhering to the strictures and tenets of standards-based metadata approaches (Kroski, 2005; Mathes, 2004; Smith, 2008; Udell, 2004). Discussions and debates on the notion of users/consumers/readers as contributors and co-producers of content and services goes back before the advent of the web. Foreseeing a read-write web, which has now become a reality, technology futurist Alvin Toffler coined the term “prosumer” in the early 1980s (Toffler, 1980), wherein he predicted that the delineation between producers and consumers of information objects would blur and that the two roles would converge. The implications of this convergence have been widely discussed in the literature (Anderson, 2006, 2009; Benkler, 2006; Surowiecki, 2005; Tapscott & Williams, 2010). The “prosumerism” phenomenon is enabled by what O’Reilly calls an “architecture of participation”, which is one of the core competencies of Web 2.0 technologies along with harnessing collective intelligence, remixability and mashability of data, radical user trust and data ownership (O’Reilly, 2005). This architecture lowers the barrier to entry for new contributors. Furthermore, central to this architecture of participation is the importance of embracing openness to share and collaborate (Alexander, 2006; Anderson, 2006; Shirky, 2005; Tapscott & Williams, 2010; Udell, 2004; Weinberger, 2005, 2007). Web 2.0, as O’Reilly (2005) and Miller (2005) contend, has a lot to do with attitude and culture than technology. In other words, Web 2.0 is not a new invention per se (Lagoze, 2010); rather it is defined by the participatory and collaborative culture that is built around Berners-Lee’s web (also called Web 1.0). In their bestselling book “Wikinomics”, Tapscott and Williams (2010, p.10) contend that “due to the deep changes in technology, demographics, business, and the world, we are entering a new age where people participate in the economy like never before”. The authors identify openness as one of the major principles of mass collaboration along with peering (self organised peer networks of contributors), sharing, and acting globally (expand the possibilities to tap a much larger pool of talent) (Tapscott & Williams, 2010, pp. 20-30).

One of the notable implications of Web 2.0 is the notion of user-generated content – a phenomenon that has stormed the scene, as exemplified by the rapid evolution and growing popularity of Wikipedia. Despite this explosive development, the phenomenon is not wholly uncontroversial. In what has become a widely cited article that appeared in the journal *Nature*, entitled “Internet encyclopaedias go head to head” (Giles, 2005), a total of 50 entries that appeared in both *Encyclopaedia Britannica* and Wikipedia on topics in the realm of science were compared. The pairs of entries were sent for blind review by experts, wherein the latter were asked to identify factual errors, critical

omissions and misleading statements in each pair. Forty-two of the entries were examined by the reviewers and the results revealed that “the difference in accuracy [between the two] was not particularly great: the average science entry in Wikipedia contained around four inaccuracies; in [Encyclopaedia] Britannica, about three,” indicating that malicious entries are “not rules but exceptions”. As pointed out in the study, the novel features offered by social media, whereby anyone, anywhere can add and edit entries, has given Wikipedia a competitive edge over Britannica. Furthermore, the study goes even further, arguing that Wikipedia has the advantage, as it benefits from having a wide spectrum of contributors, an increasing volume of entries, more frequent updates, and an in-built mechanism for resolving disputes. This study, however, did not go unchallenged (Britannica Inc., 2006). In a rather complete rebuttal of the study, Britannica Inc. (2006, p. 7) refuted the findings, stating that: “the study was so poorly carried out and its findings so error-laden that it was completely without merit.” Whereas it is not the aim of this paper to compare Wikipedia with Encyclopaedia Britannica, what is evident is the implication of social media and socially-constructed metadata approaches cannot be simply ignored, or as Lagoze (2010, p. 37) advises, “the participatory nature of Web 2.0 should not be dismissed as just a popular phenomenon [or a fad]”.

Despite several criticisms, particularly in relation to a lack of structure as well as an absence of editorial quality, authority and credibility (see Gorman, 2005; Keen, 2007), the actual and potential benefits of Web 2.0 technologies in general, and socially-constructed metadata in particular, have been convincingly and persuasively demonstrated (Anderson, 2006; Casey & Savastinuk, 2006; Casey & Savastinuk, 2007; Evans, 2009; Guy, 2006; Kroski, 2005; Lagoze, 2010; Maness, 2006; Mathes, 2004; Miller, 2005; Pressley, 2005; Shirky, 2005, 2008; Smith, 2008; Weinberger, 2005, 2007). Whilst some efforts have been made by libraries to introduce aspects of Web 2.0 services (Casey & Savastinuk, 2006; Maness, 2006; Miller, 2005), their application in metadata creation and utilisation is still limited (Evans, 2009; Lagoze, 2010; Pressley, 2005). Evans (2009, p. 10) notes that “Web 2.0 is a dated term, but its implications are not yet fully realised, especially for librarianship” perhaps because, as Lagoze argues, there exists a conceptual incompatibility between traditional library models and the emergent Web 2.0 approaches (Lagoze, 2010, p. 73). In connection with this, the scarcity of theories in library and information science (Andersen & Skouvig, 2006; Day, 2010; Floridi, 2000; Lehmann, 2010) implies that the problem could partially be attributed to the absence of a conceptual metadata framework that could serve as a basis for a better understanding of the possible uses of Web 2.0 services in libraries. Since metadata constitutes a central part of digital libraries (Anderson, Delve, Pinchbeck, & Alemu, 2009; Chan & Zeng, 2006; Day, 2003a, 2003b; Duval, Hodgins, Sutton, & Weibel, 2002; Nilsson, 2010; Pressley, 2005), it is of paramount importance that the choice of metadata approaches be underpinned by a sound theoretical framework. This framework must, of course, take into consideration recent theoretical and technological developments, such as the shift towards user-generated content, the spread of

social tagging practices, adoption of social networking applications (O'Reilly, 2005), as well as the move towards the acceptance of disparate points of views and negotiated meanings regarding digital objects (Weinberger, 2007).

## **2. Methodology**

The design and deployment of the Online Public Access Catalogues (OPAC) seem to favour an objectivist perspective (Shirky, 2005; Weinberger, 2007), whereas the proliferation of Web 2.0 applications, such as the advent of socially-constructed metadata approaches, seems to follow an interpretive philosophical paradigm. A social constructivist epistemological perspective is considered appropriate to explore issues of socially-constructed metadata approaches. According to Duffy and Jonassen (1992), social constructivism posits that “meaning is imposed on the world by us, rather than existing in the world independently of us. There are many ways to structure the world, and there are many meanings or perspectives for any event or concept.” This is contrary to the objectivist view that “truth and meaning reside in their objects independently of any consciousness” (Crotty, 1998). The nature of knowledge in social constructivism focuses on “individual reconstructions coalescing around consensus” thus promoting shared and negotiated meaning (Guba & Lincoln, 1994).

Adopting such a constructivist epistemology has also implications in the methodology and methods of research. As Guba and Lincoln (1994), Charmaz (2006) and Mills, Bonner and Francis (2006) contend, in a constructivist approach, the investigator is not an objective observer but co-creator of meaning along with the participants. Whilst objectivists accentuate the need to eliminate researcher bias by controlling confounding factors, such as researcher's assumptions and points of views, constructivists encourage an active engagement and co-construction of knowledge. This paper follows an inductive approach and by that it tests no hypothesis.

Intensive interviewing technique was chosen so as to enable the researcher to ask for more detail, delve into an issue, go back and forth among important points and request for more explanation (Charmaz, 2006). The interviews mostly consisted of open-ended questions. According to Charmaz (2006) “the in-depth nature of an intensive interview fosters eliciting each interviewee's interpretation of his or her experience.” Prior to each interview, 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. Interviews were made purposefully informal, in order to encourage dynamic participation, on the part of interviewees, in the ensuing discussions, which were expected to constitute a significant portion of the meeting.

A total of eleven in-depth interviews were conducted — with two lecturers, one PhD student and eight MSc students in Digital Library Learning (DILL), all of whom were associated, in one capacity or another, with the EU-funded,

Erasmus Mundus programme offered by the Oslo University College (Norway), University of Parma (Italy) and Tallinn University (Estonia). The selection of interviewees was essentially purposive. Countries of origins of interviewees included Bangladesh, China, Denmark, Estonia, Ethiopia, Italy, Turkey, Uganda, Venezuela and Zimbabwe. Findings are discussed under the emergent themes presented below.

### 3. Interview Results and Discussions

#### 3.1. Standards-based Metadata Approaches

All eleven interviewees acknowledged the prolificacy of standards. Some even repeated the often cited adage that *“the good thing about standards is that there are so many you can choose from.”* Under these circumstances, making the correct choice, when selecting a standard, as well as ensuring interoperability between digital libraries is bound to be a daunting task, although interviewees are not currently involved in these decisions. One interviewee expressed the opinion that *“libraries should base their [selection] decisions on the type of resources [that they have] and the subjects they are describing.”* During the discussion, it became apparent that, for this interviewee, interoperability was a much sought after issue, even though it was a complicated one. Nevertheless, the same interviewee conceded that established standards such as Dublin Core, MARC, and LC should still be maintained by libraries, as they are the basis for fulfilling various library functions.

Interviewees have alluded to the complexity of MARC vis-a-vis the simplicity of Dublin Core, noting that this relative ease-of-use comes at the expense of metadata richness. It was also pointed out that some standards, such as MARC, have been unnecessarily stretched too wide, in order to enable them to accommodate certain other genres of information objects. In particular, one interviewee had this to say:

*“We now have a novel information landscape, but a standard that is anachronistic, whilst most OPACs lack interactivity, as they are essentially static. Currently, libraries are so fixated on such archaic standards that they look a little outdated when compared to search engines, such as Google. MARC is a complicated but rich, standard. In contrast, Dublin Core is simple but lacks descriptiveness and richness.”*

Most interviewees are unanimous in stating that they find the OPAC out-dated, especially in comparison with most currently popular web search engines. In support of this view they point out for example, the fact that most OPACs do not seem to have an alternative spelling option. The absence of this and similar, seemingly simple, features makes the OPAC less user-friendly. In addition, most OPACs do not have features that enable users to rate, comment, review, or share resources with others, a serious shortcoming, in view of most users' needs. Most interviewees asserted that they rarely use the library's OPAC. One interviewee noted that:

*“I’m not quite sure if I’m using the OPAC quite extensively as a catalogue [...] because somehow it seems to me that I’ve created my own personal library on my computer. I even like to call it my desktop library”.*

Likewise, another interviewee expressed:

*“The OPACs have a big problem in that they were instituted just to replace card catalogues [and] it didn’t evolve adequately enough with developments within the world of the Web. Hence the catalogues and other library instruments always tend to be a little behind when compared to search engines”.*

However, one interviewee described the OPAC as the “biggest innovation for libraries that has ever happened,” although further suggesting that libraries are changing too slowly in trying to cope with users’ novel needs and expectations. The interviewee noted in particular that current cataloguing systems follow “a refined ways of cataloguing [oriented to assuring metadata quality]. Mistakes, for instance, are not allowed in catalogue searches. But, in reality, users make mistakes.” According to this interviewee, the search engine Google reflects reality better than those libraries that stick to strict cataloguing rules.

In a discussion on how the OPAC is performing in the current information landscape, one interviewee remarked that libraries are trying to catch-up, but: *“When things are no longer cool anymore, then librarians start doing them, [such as] implementing Web 2.0 features in the OPAC.”* According to this interviewee, most OPACs lack interactivity and are mainly static. Thus as a whole, as the responses from interviewees revealed, their preferences for Google-like search interfaces, in which a user is allowed to search with any combination of keywords. This is in agreement with a report prepared by Calhoun (2006, p. 5), who suggested that *“large and growing number of students and scholars routinely bypass library catalogues in favour of other discovery tools, and the catalogue represents a shrinking proportion of the universe of scholarly information”*. Further, Choy (2011, p. 65) argues that library catalogues are becoming inconvenient as they require users to accurately remember and submit bibliographic details such as author, title, ISBN or other details during the information retrieval process.

Another major limitation of OPAC systems has been singled out by Borgman (1996), who asserts that *“most end users of online catalogues are perpetual novices who lack the requisite conceptual knowledge for searching”*. Borgman (1996, p. 501) expounds on what she calls is an obvious disconnect between the assumptions made in the design of OPAC systems and users’ mental models. As Borgman points out these systems require users to formulate their queries in strict syntaxes but, more often than not, users face difficulties in translating their questions in to queries as required by the system. Similar problems are reported by Fifarek (2007) who contends that OPACs *“function like they were still*

*running on mainframes, with their contents just one step away from the 3 by 5 inch [catalogue] cards*". Evans (2009, p. 14) provides examples of OPAC systems which are unforgiving for those users that are not attuned to the niceties and particulars of the bibliographic details of the book. In a recent study that investigated the search behaviours of customers in a physical book shop, Buchanan and McKay (2011, p. 277) reported that book buyers more often than not opt for "simple information seeking strategies and rudimentary query formulations". According to them, most users' search behaviour is characterized by the employment of unconventional metadata, such as the colour of a book (e.g., "I'm after the book by Andrew Marr - the thick blue one?" or approximation of author names (e.g., "I think the author name finished with 'Ishky'"), as well as the submission of incomplete titles (Buchanan & McKay, 2011, p.272). Furthermore, as reported in the study, subtle cultural cues in connection with a book have been found to be more pertinent to discoverability than its accurate bibliographic description. The responses from interviewees and also the literature reviewed indicate the importance of giving due consideration to the varying user preferences and information search behaviours. In this connection, Bates (1989, p. 421) observes the disconnect between search techniques used by users and what is expected of them by designers of database systems.

From the above, it is evident that users' formulations of queries in search systems appear to be more nuanced, naive and deeply influenced by cultural and social cues. Buchanan and McKay (2011) and Choy (2011) contrasted this with the rather 'crisp' and accurate bibliographic description of information objects that are to be found in OPAC systems. Their findings clearly demonstrate the importance of reconceptualising current metadata systems in light of changing user needs, expectations and evolving vocabularies. Overall, since metadata constitutes an important function of libraries (Anderson, et al., 2009; Chan & Zeng, 2006; Day, 2003a, 2003b; Duval, et al., 2002; Nilsson, 2010; Pressley, 2005), getting the metadata right is a fundamental prerequisite for the achievement of their broader goals. The problems described above with regard to the current status of the OPAC can partly be alleviated by the inclusion of web 2.0 technologies such as socially-constructed metadata approaches to augment existing standards-based metadata systems.

### **3.2. Socially-Constructed Metadata Approaches**

Responding to the use of socially-constructed metadata approaches in libraries, interviewees are wary of the lack of control and structure in Web 2.0 applications, including tagging. In connection with this, one interviewee reflected on how some Web 2.0 technologies come and go. The interviewee stressed the need for questioning the purpose served by employing a specific technology before adopting its use in library functions. The same interviewee cited the example of Second Life and how libraries had adopted it just to be part of Second Life, and hence librarians started creating their own avatars without addressing the rationale behind its use. Elaborating on experience, the



interviewee observed that, currently, the use of Second Life in libraries has diminished. Another interviewee also added that they do not foresee Web 2.0 technologies replacing the old systems of information organisation. Many interviewees pointed out that libraries should adopt Web 2.0 technologies, using them in parallel with their metadata standards schemes so that the two serve complementarily. Thus, they advise that libraries need not necessarily discontinue legacy systems or abandon current categorisation systems, as doing so would mean removing those users who utilise them.

In answering the question of how interviewees perceived the socially-constructed metadata approach, as compared to the standards-based one, one interviewee replied that they were not particularly a “*big fan of the folksonomy and social web stuff.*” The same interviewee expressed their belief in the superiority of a more structured system and the roles played by librarians in creating such structured systems. According to the said interviewee, they found it difficult to abandon their library-oriented view of the world, which, they added, might be biased. From the conversation, it became evident that, the interviewee did not totally dismiss folksonomies, as they later stated the probable usefulness of employing tagging as a compliment to legacy metadata usage. The interviewee concluded by observing that the social approach of tagging could be utilised side-by-side with structured metadata created by librarians. Nevertheless, the interviewee advised, the two types of metadata should be kept separate and be maintained at different levels, in such a manner that a user can switch from one approach to the other, as the need arises.

In a discussion focussing on categorisation and classification systems, one interviewee gave details of the techniques they use for organising their personal music and games collections on their own computer. For organising the songs, they rely on common, widely-used tags such as the name of the singer, the title of the song, the name of the group who plays it, the year it was issued, etc., which, they stated, did not pose any particular difficulties for them. The interviewee, however, mentioned their problem when trying to make sense of the differences among the various music genres defined by publishers. This is because, as the interviewee noted, music genres are very subjective, similar to an art collection. The problem, the interviewee indicated, arose from the fact that most genres reflected the requirements of American music radio stations; hence, some genres were irrelevant for their purposes. The interviewee pointed out “*genres such as: Adult contemporary, Album oriented rock, Soft rock, etc are not meaningful for me. You may know what hard rock is, but soft rock is not that comprehensible.*” As a result, the interviewee had had to re-organise the genres and rename some of them, based on information obtained from Wikipedia, which, according to the interviewee, has a system that better reflects their needs.

Another interviewee agreed that tagging is a good thing but expressed concerns regarding the absence of control, citing synonym and homonym ambiguities as

problems afflicting tagging systems. For example, the interviewee selected the term Torino, which may refer to a city, a football club or the car company in Italy – thereby creating ambiguities. According to the same interviewee, tags, by themselves, are not enough as they should be used with the more structured and organised metadata created by librarians. For this interviewee, the claim that tags should not have any structure and control does not make a lot of sense. The said interviewee noted that current tagging systems, such as LibraryThing, do seem to have some form of a structure, as for example the description of a book consists of standard metadata elements such as title, author(s), and publisher.

Another interviewee pointed out that Web 2.0 technologies such as social tagging bring similar information together and provide information that is very practical for and widely known by the lay community. They observed that libraries traditionally offered highly formalised knowledge, such as text books, but for daily usage, they said, they preferred using web search engines. A deeper grasp of the concepts and terms employed in a given topic of interest offered them the opportunity to coin keywords, for use in these search engines. Another interviewee remarked that they took a “softer” approach towards Web 2.0, in that they saw it as a useful approach for augmenting library functionality. Elaborating, the same interviewee spoke thus:

*“If we are talking [about the fact] that the library is meant for users, I think we should not only consider taking users’ needs and wants into account, but they should also be involved when systems intended to serve are designed and developed”.*

Similarly, one interviewee also expressed the view that Web 2.0 (social media) is important for practical applications. They cited, as an example, the world of gardening, about which, they said, they would be interested in discovering information that is expressed in terminologies of day-to-day usage and laymen’s terms, rather than in scholastic ones as found in books from libraries. One good example, they observed, was the use of standard metric units in formal text books, whilst most of these metrics are rarely employed in everyday language.

Contrary to the above, another interviewee pointed out that the current usage of Web 2.0 technologies in libraries has not been well thought out. As an example, they singled out the current usage of Facebook by libraries, whereby both users and librarians make postings. However, as the number of postings grows, searching and/or browsing through them become increasingly cumbersome. The problem, they opined, was that Facebook had not been designed to serve as a search engine. In addition, at present, most of libraries’ postings on Facebook are limited to such basic facts as opening hours and location. It is imperative that libraries employ Web 2.0 technologies proactively and strategically, instead of “adopting technology for technology’s sake”, they concluded.

Most interviewees indicated the importance of some control and structure in Web 2.0 applications. According to one interviewee, even Wikipedia, arguably

considered the freest and most democratic approach, has some basic rules and structure underlying it. The said interviewee maintained that which makes Wikipedia different is the fact that it has no scope. Nor does it have a rule stipulating “*let us create these numbers of pages for these numbers of topics*”. Wikipedians, they asserted, act according to the following maxim:

*“Let us put everything out. If you want to add voice, you can do so. If you want to make the voice louder, you can do that too. If the voice stays [hushed], because nobody is contributing, that is ok. It is so heterogeneous.”*

As the said interviewee explained, Wikipedia has aims, to at least present a neutral point of view. This, according to them, is one of Wikipedia’s few rules. Additionally, one should be able to convey other, even contrary points of views. Furthermore, Wikipedia also watches out for offensive language and imposes some writing style. Hence, according to them, overall, Wikipedia “*is not completely devoid of rules and there is always a structure.*” As further expressed by them, the case of Wikipedia demonstrates that a user-driven contribution can be allowed within a structure, albeit a minimal one.

### **3.3. Towards Mixed Metadata Approaches**

In the metadata literature there are three broad positions with regard to the use of Web 2.0 technologies within libraries. Firstly, there are those that advocate a fully socially-constructed metadata approach. Authors like Shirky (2005) and Weinberger (2007) argue that current standards-based metadata systems are broken and hence, they suggest, should be substituted by an open, democratic, and socially-constructed metadata approach. The digital world, Shirky (2005) argues, is a radical break with the print and analogue paradigm as a single information object can now be categorised in an infinite number of places using hyperlinks. Shirky’s argument about the inherited limitation of categorisation from the physical world is also echoed by Weinberger (2007). Supporting the use of Web 2.0, Weinberger (2007) argues that, adopting the Aristotelian principle of “*carving nature at its joints*” through the use of taxonomies and controlled vocabularies to an increasingly chaotic and complex digital information landscape is a futile effort. Furthermore, as Shirky (2005) contends, standards-based categorisation systems (including Yahoo categories, Dewey Decimal and Library of Congress classification systems) are best fit for physical libraries, and goes on to state that pre-defined categories can only work for domains that have stable categories, limited collections, and clear edges. In addition, he argues such systems can only be functional in domains where there are expert librarians in the specific domain, who are able to describe the content for an already expert user bases. Both Weinberger (2007) and Shirky (2005) agree that, in the digital information landscape and in domains where there exists huge collections of information objects and where users come from diverse backgrounds, employing diverse numbers of terminologies in their search for information, pre-determined categorisation systems (standards-based

approaches) are very likely to fail to scale and become unwieldy. Consequently, they recommend the use of Web 2.0 and support the notion of individual users contributing content and metadata (Weinberger, 2007).

In contrast, the second position suggests the limitations of socially-constructed content and metadata approaches and contends that relying on a system where structure, authority and editorial quality are absent is likely to result on a sub-standard and chaotic information landscape. For example, Andrew Keen, in *"The Cult of the Amateur"*, (2007) dismisses the notion of user generated content and uncontrolled social media as amateurish and stresses on the need for control and regulation by an established authority. He argues that *"what the Web 2.0 revolution is really delivering is superficial observations of the world around us rather than deep analysis, shrill opinion rather than considered judgment"* and the consequence of this uncontrolled content, as Keen warns, is chaos and disappearance of truth (Keen, 2007, p.16). Keen argues that the unedited, uncontrolled and unfiltered content from Web 2.0 threatens the fabric of traditional quality controlled, edited, and guarded media. In a similar vein, Michael Gorman (2005), former president of the American Library Association, stated that the Web 2.0 in general and blogs in particular lack credibility.)

The third, middle-of-the-road, position recognises inherent limitations in both approaches and contends that finding a middle ground could rectify some of these limitations. Authors such as Gruber (2007), Smith (2008), Veltman (2001) and Wright (2007) argue that a balance should be struck between the standards-based and socially-constructed metadata approaches. As Gruber (2007) argues socially-constructed metadata approaches are one-dimensional, plagued with inconsistency and lack of organisation, whilst, standards-based metadata approaches are forced upon users and fail to represent the users' world views. Gruber (2007) suggests that ontologies should be developed to capitalise on the best of both worlds. At a conceptual level, Morville (2005) and Wright (2007), both information architects, contend that the two metadata approaches can productively co-exist. Morville (2005, p. 139) argues that *"ontologies, taxonomies, and folksonomies are not mutually exclusive"*. However, a closer look at Morville's arguments reveals a tendency to suggest that socially-constructed metadata has its contextual place solely in the blogosphere and social media environments and not entirely in portals and digital libraries. This is also evidenced by his choice of terminologies, such as *"mob indexing"* for what is otherwise termed as social classification or folksonomy, elsewhere in the metadata literature. In his book *"Ambient Findability"*, Morville (2005) contextualises the theory of *"Pace Layering"* from Stewart Brand's *"How Buildings Learn"* and argues that *"taxonomies and ontologies provide a solid semantic network that connects interface to infrastructure"* whilst folksonomies are overlaid on the taxonomic metadata infrastructure providing it with the fast-moving and volatile vocabularies of users (see also Campbell & Fast, 2006; Smith, 2008). Both Wright and Morville, renowned information architects, approach the issue of folksonomies with caution, arguing against the view that

folksonomies are fundamentally new ways of organising information. Others such as Campbell and Fast (2006) emphasise the importance of embracing Web 2.0 categorisation techniques. They contend “*we ignore ‘mob indexing’ at our peril; by refusing to change our methods or our approaches in the face of collaborative tagging systems, we run a serious risk of allowing information architects to become anachronisms in emerging information environments*”.

In connection with the above, one interviewee stated that the issue of using standardised approaches (such as hierarchies and categories) versus Web 2.0 technologies is more of a philosophical nature than a technological one. In support of this, the interviewee cited the literary work known as the “*Divine Comedy*”, by Dante, where the organisation of the poem reflects the theoretical (philosophical) framework of medieval European society. According to the said interviewee, the work is a complete summary of all the medieval beliefs and church teachings extant then. Furthermore, the division of the poems is well thought out, each category having 33 divisions, which, along with the introduction, brings the total number of categories to 100. They then contrasted this with the Dewey Decimal Classification system. The same interviewee remarked that both Dewey and Dante represented cultural frameworks of their societies and that they were correct in their own way. He noted that the situation now is entirely different “*because there are too many traditions altogether and we don’t believe any more in a rigid, [monolithic], structure. We [do] believe in change.*”

As noted by almost all interviewees, there exists some bias in current classification systems and standards. This is in agreement with Van House’s (2005) assertion that:

*“Classification systems and categories carry their history within them, including the politics of the time and place in which they are created, and the participants in the decision making”.*

Hence, it is important that information organisation systems reflect the diversity of users’ perspectives and interpretations of information objects that have been deposited in digital libraries. To this end, as Van House (2005) advises, the philosophical assumptions that underlie standards or categorization systems should not “*valorise*” one view whilst disparaging others. In their oft-cited book, “*Sorting Things Out: Classification and Its Consequences*”, Bowker and Star (1999) contend that “*each [classification system] and each category valorizes some point of view and silences another*” which they admit is inescapable but nonetheless such limitation should be taken into account and its ultimate consequences should be carefully weighed. To this end, the decisions to choose one metadata approach over another or the choice of standards or perhaps the absence thereof should be undergirded with sound theoretical foundation.

Overall, as the responses of the interviewees and a review of the literature demonstrate, the implication of Web 2.0 technologies for library metadata

functions merits a closer examination. This paper, whilst concurring with the mixed metadata approach, as recommended by Morville and Wright, advocates that socially-constructed metadata should have a viable, substantive, place in digital libraries. To adopt mixed-metadata approaches, integrating the two approaches remains a big challenge. In this regard, it is important to explore the opportunities presented by Linked Data and its associated technologies such as the Resource Description Framework (RDF) and the Web Ontology Language (OWL).

### **3.4. Slow Adoption of Linked Data in Libraries**

Another question posed to interviewees was the role of Linked Data for metadata functions in libraries. After acknowledging being aware of the visions of Semantic Web technologies, many of the interviewees confessed to being unaware of any Semantic Web (also referred as Linked Data) application appropriate for use in digital libraries. One interviewee expressed their belief that Linked Data would offer solutions to the problems that they had previously highlighted during the interview, including cataloguing and storage of information in a manner that can be searched semantically. Within this context, the interviewee foresaw “*Linked Data*” becoming an effective method of metadata representation, at the most discrete and atomic data level. They continued saying that, “*You just describe at the meaning level and then create associations. Such meaning will lie not with the object itself but on the context of the associations. The OPAC should be more robust and allow natural language searches.*” Similarly, another interviewee characterized the Semantic Web as: “*A system built on meaningful relationships between topics. The world is modelled in words. And then you can see the relationships between the words and what kinds of relationships there are. It is a good technology but I don’t see libraries using it. And I don’t currently see a proper Semantic Web*”.

Another interviewee stated that they have heard a lot about the Semantic Web but hasn’t come across any real instance of its application in libraries. Yet another interviewee described the Semantic Web as a system in which one resource describes another but he says resources. They added “*to tell you frankly I [have] never used any Semantic Web application. But I think from what they say, it would be quite a difficult task*”. One of the interviewees portrayed the Semantic Web as an “*an awesome idea*” but expressed their reservation as to whether they would at all need agents proposed by Berners-Lee, Hendler, & Lassila, (2001), who had envisaged a role for semantic agents in facilitating such tasks as coffee making and fixing appointments. Most of the promises of the Semantic Web, according to a recent review by Powell, Black and Collins (2011), “*have yet to be fulfilled*”. It is perhaps worth mentioning here that Thomas Gruber, who is mainly known for his oft-cited definition of the term ontology (i.e., “*an explicit specification of a conceptualization*” (Gruber, 1993)), has for so long been advocating the role of semantic agents not from an artificial intelligence point of view but from ontology-based metadata specification of concepts, objects and its relationships. It is also important to note here that

Gruber is also behind the development of the Siri software, a semantically intelligent virtual personal assistant, which uses Semantic Web technologies.

Most interviewees acknowledged the potential use of Semantic Web technologies for information organisation and access. In connection with this, one interviewee called attention to the technical complexity of Semantic Web technologies. Linked Data, the same interviewee observed, is an important part of the Semantic Web. This interviewee also predicted that Semantic Web will be more popular in the coming five years. The said interviewee's prediction doesn't seem to be farfetched, given the fact that these technologies have promised to provide concrete solutions for metadata representation and utilisation as early as 2003 (Day, 2003b) and subsequently in 2008 (Rothenberg, 2008). Another interviewee concurred, stating that they understood the Semantic Web as a concept, but stressed that the technologies associated with it should be taught in computer science departments. They foresaw applications ensuing from a wider deployment of these technologies, for which development and acquisition of new skills, by both students and users is a prerequisite.

It is important to note here is that, reading through its origins, it is evident that the Semantic Web has also been about metadata. In January 6, 1997, Tim Berners-Lee wrote a proposal entitled "*Metadata Architecture*" and defined metadata as "*machine understandable information about web resources or other things*" (Berners-Lee, 1997). It is also apparent that the use of the word metadata in the author's initial proposal refers to what later came to be known the Resource Description Framework (RDF), which was subsequently approved by W3C in 2004. Berners-Lee argues in favour of maintaining the centrality of metadata. He points out that metadata about one document can occur within itself, within a separate document, or may even be transferred accompanying the document; metadata can describe metadata (in other words metadata about metadata); and that different things may be asserted of the same thing and may stay independently or in combination (Berners-Lee, 1997). His concern was mainly of metadata to describe web-resources. He was also primarily interested in defining machine process-able metadata for data and documents, providing meaning and context through typed relations; which is tied to his vision of the Semantic Web as described in Berners-Lee, Hendler, & Lassila (2001).

It is worth noting that as recently as in 2011, some national and regional initiatives such as the British Library, the National Library of France, and The Europeana Digital Library have announced their plans to open their legacy bibliographic records as Linked Data (Helen, 2010). A report commissioned by the World Wide Web Consortium (W3C) is published in October 2011 (W3C, 2011). The report acknowledges the low uptake of Linked Data in libraries whilst providing recommendations for libraries to embrace the Linked Data principles which, among other things, includes the use of technologies such as URI, RDF, SPARQL, and OWL. The report states the importance of making

bibliographic library data openly and freely accessible in a form that is “*shareable, extensible, and easily re-usable*” (W3C, 2011).

The Semantic Web has not made significant strides in the library domain, nonetheless the potential role of Semantic Web technologies such as the Resource Description Framework (RDF), RDFS (RDF-Schema), and Web Ontology Language (OWL) for metadata encoding, representation and sharing is clearly evident (Day, 2000, 2003a, 2003b; Nilsson, 2010; Rothenberg, 2008). It has been demonstrated that RDF’s simple data model enables the creation of semantic links among information resources. An RDF schema adds vocabularies – such as Class, SubClass, Domain, and Range – to enable a more meaningful representation of resources. By extending RDFS with yet additional vocabularies, OWL allows the definition of additional semantic constructs, such as equivalency, inverse and cardinality relations and constraints (Allemnag & Hendler, 2008; W3C, 2004). One of the defining features of the RDF model is its ability to identify resources and metadata attributes (relations) uniquely and globally using Uniform Resource Identifiers (URIs). The use of URIs for metadata element names, labels, and relations, according to Nilsson (2010), helps to avoid naming and identification conflicts in the use of elements. This is also suggested by (Alemu, Stevens, & Ross, 2012), Day (Day, 2000, 2003a, 2003b) and Rothenberg (2008). Unfortunately, although there happen to be several academic papers and technical specifications regarding RDF, RDFS, SPARQL (SPARQL Protocol and RDF Query Language), and OWL, there are, up until now, no viable Semantic Web related metadata solutions in widespread use. In a more positive note, Gruber (Gruber, 2007, 2008) argues that Web 2.0 and the Semantic Web (Web 3.0) are complementary and hence can be leveraged to provide unique and enriching user experiences.

### **3.5. Conclusions**

The interviewees characterised the current state of the OPAC as being anachronistic, especially when compared with contemporary search engines. It has also been reported in the literature that there is a disconnect between the designs of OPAC systems and the searching behaviour of users. Interoperability problems among disparate digital libraries, arising from the proliferation of metadata standards, have also been brought to light by several interviewees in the domain. It has also been suggested that existing standards-based metadata systems be re-evaluated, especially in light of socially-constructed metadata approaches. Whilst, it has been remarked that these two metadata approaches do not stand in opposition to each other, interviewees have recommended that, libraries should embrace Web 2.0 technologies strategically, rather than adopting it for the sake of the technology. There is consensus among them that, in as far as they would like to see the adoption of the Linked Data, there is still a lack of applications that serves libraries. It is, however, anticipated that, once Linked Data is implemented and adopted widely, the Linked Data has a lot to



offer for metadata representation and exchange. It has also become apparent that the primary challenge facing librarians today is, to first of all place users at the centre of all information organisation decisions, and to that effect, ensure that metadata systems reflect the world view of their users in terms of representing various perspectives and interpretations.

In view of all the above, in this paper it is argued that there is a need to reconceptualise current metadata systems in light of changing user needs, expectations and evolving vocabularies. Hence, it is deemed essential that the creation and utilisation of metadata be underpinned by sound theoretical frameworks. A conceptual metadata framework that caters for the inclusion of socially-constructed metadata approaches is therefore proposed. It is imperative that such a metadata framework takes into consideration the conceptual foundations of current Web 2.0 technologies as well as recent developments in the Linked Data . These should include, but not be limited to, Bush's concept of associative trails (Bush, 1945), Berners-Lee's Linked Data (Berners-Lee, 1997; Berners-Lee, 1998; Berners-Lee, et al., 2001), O'Reilly's collective intelligence (O'Reilly, 2005) and Anderson's long tail (Anderson, 2006). In the context of a socially-constructed metadata approach, especially relevant is Paul Otlet's vision of "*the social space of documents*", where after reviewing the efforts of Melville Dewey's DDC and Panizzi's cataloguing scheme, Otlet came to conclude that such tools only guide the reader as far as the location of the book but not to the contents within and also relationships between documents (Wright, 2007). Otlet thus envisioned a system called the "*réseau*"- a tool to create semantic links between documents and keep track of the annotations made by readers, eventually forming new trails of documents, which he calls "*the book about the book*" (Wright, 2007). Thus, in conclusion, with a view to improving current library metadata functions (such as the OPAC), Paul Otlet's vision of "*réseau*" matches the interviewees' expectations and should be implemented and widely adopted. In this connection, structured metadata should not only consist of the physical description of information objects (such as author, title, ISBN, subject, format, etc) but also incorporate elements describing its socio-cultural facets (user tags, comments, reviews, links, ratings (likes and dislikes), recommendations). In other words, the representation of the social space of metadata should be considered as equally important as the recording of standardised and objectivistic metadata elements that have hitherto been used to characterise the physical characteristics of information objects.

## 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.
- 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 Unlimited Demand*. London: Random House Business Books.
- Anderson, C. (2009). *The Longer Long Tail : How Endless Choice Is Creating Unlimited Demand* (Expanded & updated ed.). London: Random House Business Books.
- Anderson, D., Delve, J., Pinchbeck, D., & Alemu, G. A. (2009). *Preservation Metadata Standards for Emulation Access Platforms*. Portsmouth, UK: KEEP Project, European Commission.
- Barbosa, D. (2008). Taxonomy Folksonomy Cookbook Available from [http://solutions.dowjones.com/cookbook/ebook\\_sla2008/cookbookebook.pdf](http://solutions.dowjones.com/cookbook/ebook_sla2008/cookbookebook.pdf)
- Bates, M. J. (1989). The Design of Browsing and Berrypicking Techniques for the Online Search Interface. *Online Review*, 13(5), 407-424.
- Benkler, Y. (2006). *The Wealth of Networks: How Social Production Transforms Markets and Freedom*. New Haven [Conn.]: Yale University Press.
- Berners-Lee, T. (1997). Metadata Architecture. Retrieved from <http://www.w3.org/DesignIssues/Metadata.html>
- Berners-Lee, T. (1998). What the Semantic Web Can Represent. Retrieved from <http://www.w3.org/DesignIssues/RDFnot.html>
- Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The Semantic Web: A New Form of Web Content That Is Meaningful to Computers Will Unleash a Revolution of New Possibilities. *The Scientific American*, 284(5), 34-43
- Borgman, C. L. (1996). Why Are Online Catalogs Still Hard to Use? *Journal of the American Society for Information Science*, 47(7), 493-503.
- Bowker, G., & Star, S. L. (1999). *Sorting Things Out: Classification and Its Consequences*. Cambridge, Mass: MIT Press.
- Britannica Inc. (2006). Fatally Flawed: Refuting the Recent Study on Encyclopedic Accuracy by the Journal Nature. Retrieved from [http://corporate.britannica.com/britannica\\_nature\\_response.pdf](http://corporate.britannica.com/britannica_nature_response.pdf)
- Buchanan, G., & McKay, D. (2011). *In the Bookshop: Examining Popular Search Strategies*. Paper presented at the JCDL-2011, 11th annual international ACM/IEEE joint conference on Digital libraries, Ottawa, Canada.
- Bush, V. (1945). As We May Think. *The Atlantic Monthly*. Retrieved from <http://www.theatlantic.com/magazine/archive/1969/12/as-we-may-think/38812/>
- Calhoun, K. (2006). *The Changing Nature of the Catalog and Its Integration with Other Discovery Tools: Final Report Prepared for the Library of Congress*. Ithaca, N.Y: Cornell University Library.
- Campbell, D. G., & Fast, K. V. (2006). *From Pace Layering to Resilience Theory: The Complex Implications of Tagging for Information Architecture*. Paper presented at the The 7th Information Architecture Summit, Vancouver, British Columbia, Canada.
- Casey, M. E., & Savastinuk, L. C. (2006). Web 2.0: Service for the Next-Generation Library. *Library Journal*. Retrieved from <http://www.libraryjournal.com/article/CA6365200.html>
- Casey, M. E., & Savastinuk, L. C. (2007). *Library 2. 0: A Guide to Participatory Library Service*. Medford, N.J: Information Today, Inc.
- Chan, L. M., & Zeng, M. L. (2006). Metadata Interoperability and Standardization – a Study of Methodology Part I :Achieving Interoperability at the Schema Level. *D-Lib Magazine*, 12(6).

- Charmaz, K. (2006). *Constructing Grounded Theory: A Practical Guide through Qualitative Analysis*. London: SAGE Publications.
- Choy, F. C. (2011). From Library Stacks to Library-in-a-Pocket: Will Users Be Around? *Library Management*, 32(1/2), 62-72.
- Crotty, M. (1998). *The Foundations of Social Research: Meaning and Perspective in the Research Process*. London: SAGE Publications.
- Day, M. (2000). Resource Discovery, Interoperability and Digital Preservation: Some Aspects of Current Metadata Research and Development. *VINE*, 36(117), 35-48.
- Day, M. (2003a). *Integrating Metadata Schema Registries with Digital Preservation Systems to Support Interoperability: A Proposal*. Paper presented at the International Conference on Dublin Core and Metadata Applications, Seattle.
- Day, M. (2003b). *Preservation Metadata Initiatives: Practicality, Sustainability, and Interoperability*. Paper presented at the ERPANET Training Seminar on Metadata in Digital Preservation. Retrieved from <http://www.ukoln.ac.uk/preservation/publications/erpanet-marburg/day-paper.pdf>
- 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).
- Duffy, T. M., & Jonassen, D. H. (Eds.). (1992). *Constructivism and the Technology of Instruction: A Conversation*. Hillsdale, New Jersey: Lawrence Erlbaum Associates, Inc., Publishers.
- Duval, E., Hodgins, W., Sutton, S., & Weibel, S. L. (2002). Metadata Principles and Practicalities. *D-Lib Magazine*, 8(4).
- 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.
- Fifarek, A. (2007). The Birth of Catalog 2.0: Innovative Interfaces' Encore Discovery Platform. *Library Hi-Tech News* (5), 13-15.
- Floridi, L. (2000). On Defining Library and Information Science as Applied Philosophy of Information. *Social Epistemology*, 16(1), 37-49.
- Giles, J. (2005). Internet Encyclopaedias Go Head to Head. *Nature*, 438. Retrieved from <http://www.nature.com/nature/journal/v438/n7070/pdf/438900a.pdf>
- Gorman, M. (2005). Backtalk: Revenge of the Blog People! . *Library Journal*.
- Gruber, T. (1993). Toward Principles for the Design of Ontologies Used for Knowledge Sharing. *International Journal Human-Computer Studies*, 43, 907-928.
- Gruber, T. (2007). Ontology of Folsonomy: A Mash-up of Apples and Oranges. *International Journal on Semantic Web & Information Systems*, 3(2).
- Gruber, T. (2008). Collective Knowledge Systems: Where the Social Web Meets the Semantic Web. *Journal of Web Semantics: Science, Services and Agents on the World Wide Web*, 6(1), 4-13.
- Guba, G., & Lincoln, S. (1994). Competing Paradigms in Qualitative Research. Chapter 6 in N.K. Denzin & Y.S. Lincoln (Eds) Sage. 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).
- Helen, W. (2010). Linked Data and Libraries. *Catalogue & Index*(160), 2-5.

- Keen, A. (2007). *The Cult of the Amateur: How Today's Internet Is Killing Our Culture and Assaulting Our Economy*: Nicholas Brealey Publishing.
- Kroski, E. (2005). The Hive Mind: Folksonomies and User-Based Tagging. <http://infotangle.blogspot.com/2005/12/07/the-hive-mind-folksonomies-and-user-based-tagging/>
- Lagoze, C. (2010). *Lost Identity: The Assimilation of Digital Libraries into the Web*. Cornell University, School of Information Science.
- Lagoze, C., & Patzke, K. (2011). *A Research Agenda for Data Curation Cyberinfrastructure*. Paper presented at the The 11th annual international ACM/IEEE joint conference on Digital libraries - JCDL '11, Ottawa, Canada.
- 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.
- 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 Retrieved from <http://www.adammathes.com/academic/computer-mediated-communication/folksonomies.html>
- Miller, P. (2005). Web 2.0: Building the New Library *Ariadne*, 45.
- 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.
- Morville, P. (2005). *Ambient Findability*. Sebastopol, CA: O'Reilly.
- Nilsson, M. (2010). *From Interoperability to Harmonization in Metadata Standardization: Designing an Evolvable Framework for Metadata Harmonization*. Unpublished PhD, School of Computer Science and Communication, KTH, Stockholm.
- O'Reilly, T. (2005). What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software. Retrieved from <http://oreilly.com/web2/archive/what-is-web-20.html>
- Powell, J. E., Black, K., & Collins, L. M. (2011). A Semantic Registry for Digital Library Collections and Services. *D-Lib Magazine*, 17(11/12).
- Pressley, L. (2005). Folksonomies, Tags, and User-Created Metadata: A Truly Emerging Technological Trend Unpublished Seminar on Emerging Technological Trends in Libraries. Reynolds Library, Wake Forest University.
- Shirky, C. (2005). Ontology Is Overrated: Categories, Links, and Tags. *Clay Shirky's Writings About the Internet*. Retrieved from [http://www.shirky.com/writings/ontology\\_overrated.html](http://www.shirky.com/writings/ontology_overrated.html)
- Shirky, C. (2008). *Here Comes Everybody : The Power of Organizing without Organizations*. London: Penguin Books.
- Smith, G. (2008). *Tagging : People-Powered Metadata for the Social Web*. Berkeley, CA: New Riders.
- Surowiecki, J. (2005). *The Wisdom of Crowds: Why the Many Are Smarter Than the Few*. London: Abacus.
- Tapscott, D., & Williams, A. D. (2010). *Wikinomics: How Mass Collaboration Changes Everything* (Expanded ed.). London: Penguin Books.
- Toffler, A. (1980). *The Third Wave*. New York: Bantam Books.
- Udell, J. (2004). Collaborative Knowledge Gardening with Flickr and Del.Icio.Us, Social Networking Goes Beyond Sharing Contacts and Connections. *InfoWorld*.

- Retrieved from <http://www.infoworld.com/d/developer-world/collaborative-knowledge-gardening-020>
- Vander Wal, T. (2007). Folksonomy: Coinage and Definition. Retrieved from <http://www.vanderwal.net/folksonomy.html>
- 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. (2004). Owl Web Ontology Language Overview. from <http://www.w3.org/TR/owl-features/>
- W3C. (2011). *Library Linked Data Incubator Group Final Report*: W3C.
- Weinberger, D. (2005). Tagging and Why It Matters. Retrieved from <http://cyber.law.harvard.edu/sites/cyber.law.harvard.edu/files/07-WhyTaggingMatters.pdf>
- Weinberger, D. (2007). *Everything Is Miscellaneous*. New York: Times books.
- West, J. (2007). Subject Headings 2.0: Folksonomies and Tags. *Library Media Connection*, 25(7), 58-59.
- Wright, A. (2007). *Glut: Mastering Information through the Ages*. Ithaca: Corenell University Press.
- Yi, K., & Chan, L. M. (2009). Linking Folksonomy to Library of Congress Subject Headings: An Exploratory Study. *Journal of Documentation*, 65(6), 872-900.

#### **About the authors**

Getaneh Alemu is a PhD student in the School of Creative Technologies, and Dr Brett Stevens (Principal Lecturer, School of Creative Technologies), and Dr Penny Ross (Senior Lecturer, School of Computing), are his PhD supervisors, all at the University of Portsmouth, Portsmouth, UK. Getaneh Alemu is the corresponding author and can be contacted at: [Getaneh.Alemu@port.ac.uk](mailto:Getaneh.Alemu@port.ac.uk)