

A Review of User-created Metadata in Libraries

Kathryn Miller

LBSC670 Section 0201 Summer II 2012: Organization of Information

August 14, 2012

## A Review of User-created Metadata in Libraries

### **Introduction**

As the internet has become more accessible and websites are more reliant on user-created content, the library science profession has had to take notice. While fixed organizational structures for physical resources are important, digital resources are constantly in flux and can be accessed in multiple, ever-changing locations (West, 2007). How are information professionals supposed to fit these dynamic resources into traditional metadata schemas? One solution is to let end-users add to and change information resource metadata through library catalog Web 2.0 “tagging” features (Majumdar & Shukla, 2008; Porter, 2011; Vander Wal, 2007). From the aggregate of these tags can come a “folksonomy,” which was coined and defined by Thomas Vander Wal as “the result of personal free tagging of information and objects (anything with a URL) for one’s own retrieval...Folksonomy is tagging that works” (Vander Wal, 2007). Basically, if a tagging structure “works” and is implemented properly in a library, then a “user-centric categorical structure” (Porter, 2011) emerges, a folksonomy of tags in non-hierarchical, “related” relationships which other users can search (and retrieve relevant resources), browse, and use as a controlled vocabulary when submitting their own tags. However, there are inherent problems with user-created tags, as is expected when information professionals ask end-users to aid them in creating metadata for information resources (Schwartz, 2008).

This paper will explore the advantages and disadvantages of tagging and folksonomies as discussed in literature, and survey the utilization of user-created tags in LibraryThing, Online Computer Library Center, Inc. (OCLC)’s WorldCat, and two public library catalog interfaces.

## Literature Review

### Advantages of Tagging

Clay Shirky aimed to change the way information professionals organize digital content in his talks “Ontology is Overrated,” given at the O'Reilly ETech conference in March, 2005 and “Folksonomies & Tags: The rise of user-developed classification,” given at the IMCExpo in April, 2005. In an overview of these talks, Shirky summarizes his main argument: information professionals shouldn't be using strategies meant for physical resources when developing metadata schemas and ontologies for digital resources; in the digital world, user-created metadata in the form of tags should replace formal classification structures (Shirky, 2005). Shirky adds that with the indefinite number of users on the web, tagging will “[get] better with scale” (Shirky, 2005); as the amount of users utilizing the system increases, they will start to work together to develop common practices for tagging (i.e., in formatting, use of characters, using popular tags, developing relational relationships, etc.) and a folksonomy will emerge (Shirky, 2005).

While most of the literature isn't so “all or nothing” as Shirky is, it is generally agreed that tagging practices should be implemented into library catalogs in order to keep up with the way users are searching for information (Schwartz, 2008). Tagging allows for more choices in identifying content and is more akin to the user-popular practice of “keyword searching,” whereas traditional cataloging, such as the Library of Congress Subject Headings (LCSH), might limit the information available about a resource (Webb & Nero, 2009; Rolla, 2009). Tagging also increases the findability of resources as users can submit an indefinite number of tags to a resource, which assists with what Majumdar & Shukla (2008) call “lateral searching” (West, 2007; Porter, 2011). That is, instead of just typing in a keyword and getting a page of results, users can see items that share the same tags and search by sub-tags or other relevant tags, thus

increasing the potential for indirectly relevant resources (Majumdar & Shukla, 2008). In an online social environment, users have a broad vocabulary and are willing to start conversations using distinct points of view. Therefore, letting users add to the metadata of information resources through “tagging, tag clouds, folksonomies, and user-driven content descriptions and classifications” (Majumdar & Shukla, 2008) can increase access to and relationships between resources. Additionally, Porter (2011) argues that if technology can search, aggregate, and connect people to these Web 2.0 implementations, then libraries will forever reap the benefits.

### **Disadvantages of Tagging**

While a lot of literature sings the praises of user-created metadata—some almost tongue-in-cheek as they say they are *very* willing to pass over the responsibilities of creating metadata to the user—there are inherent problems when integrating any Web 2.0 features into a system. Most of these problems have to do with quality control. With an indefinite number of users and the ability for resources to be constantly in flux, how can information professionals keep track of it all? Marieke Guy and Emma Tonkin discuss these issues in their paper, “Folksonomies: Tidying up tags?” They argue that current folksonomy systems are imprecise; the tags created by users are usually very personal, inexact, ambiguous, and contain misspellings, plural forms of words, symbols, and synonym/homonym problems (Guy & Tonkin, 2006). These inexact, incorrect, and personal tags create “noise” (West, 2007) in library catalogs, what may be relevant or true to one user might not be relevant or true to the majority of those searching or browsing the tags (Porter, 2011; West, 2007). More than the inconsistent format or personal attributes of tags, Guy & Tonkin (2006) also argue that once a system has a large volume of users, and “popular” tags emerge, users will tend to keep reusing the popular tags, even if they have little to do with the information resource at hand. This is in direct contrast with Clay Shirky’s (2005) argument that a

large volume of users will be able to work together to weed out inconsistent, uncommon tags in favor of more relevant tags that will make up a folksonomy. On the other hand, Rolla (2009) points out that if a system does not have enough users tagging resources, then the few tags that are being submitted are probably not very helpful to other users in their searching activities. While library catalogs have begun to implement tagging into their Web 2.0 features, the lack of semantic control, user propensity to go with tags that are popular rather than those that are specific, truly relevant, or representing a point of view, and reliance on a high volume of users makes it difficult to see the true usefulness in practice of user-created metadata in the form of tagging.

### **Examples of Tagging Implementation**

Using the points brought up by the literature, I have explored two widely used information resource databases and briefly surveyed two public library catalogs.

#### **LibraryThing**

LibraryThing is a Web 2.0 site for personal use; users can sign up for a free account and search through the site's vast listing of books, or add their own records for books, in order to organize and catalog their personal libraries. LibraryThing can also be implemented into an existing Online Public Access Catalog (OPAC) so users of that institution's catalog can get the added benefit of browsing LibraryThing's tags of resources available to them. In turn, the catalog can utilize LibraryThing's more formal classification features, such as the LCSH and Dewey Decimal System Classifications (DDC), in their ontologies.

LibraryThing relies heavily on user-created tags as users more frequently search the tags in order to find relevant resources. This is indicative of the Web 2.0 nature of the site, as users feel more comfortable searching other user-created content to find new books to read, new

genres to enjoy, and ways to connect with other's "libraries." LibraryThing has a "How-to" page on tagging in which they encourage users to add personal tags to information resources as they are organizing their libraries, they give examples of tags such as "favorites" or "books to read" ("LibraryThing Concepts," 2012). One would think the freedom for users to tag anything, anyway they like would hinder the searching process, but LibraryThing has come up with tag combinations and "tagmashes." Tag combinations solve the synonym problem Guy & Tonkin discuss in their paper; LibraryThing users can search, submit, and vote on combinations of tags that are very similar in meaning and subject matter (see Figure 1). This solves having to search separately for "WWII, WW2, and World War Two," since all three mean the same thing ("WikiThing: Tag combining," 2012). Additionally, LibraryThing users can use a sort of Boolean search method when searching tags through "tagmashes," (e.g., if a user wants to find resources tagged "Paranormal" but *not* "Vampires," the way to search would be to type this into the address bar: <http://www.librarything.com/tag/paranormal,+--vampires>, see Figure 2) ("HelpThing: Tagmash," 2012).

As you can see in both Figure 1 and Figure 2, many pages on the LibraryThing interface offer links to relevant tags, tagmashes, and tag combinations. This shows that the site has developed a successful folksonomy, as users can see relationships between resources and are able to easily find indirectly relevant resources. However, there are some "point of view" issues with such a large volume of users. I, for instance, do not believe the *Twilight* series should be tagged "Fantasy" by over 1,600 users so it appears in the same company as *Harry Potter* and *The Lord of the Rings* (see Figure 1). But, that is a matter of opinion, and one that you can share with others through LibraryThing's discussion boards and groups. Through letting users vote on tag combinations, "tagmash" search features that return relevant resources, and the opportunity for

users to discuss tag structures in groups and discussion boards, LibraryThing has created a very thorough folksonomy. This folksonomy has *almost* replaced the formal classifications on the site, as LibraryThing users are more likely to search tags than subject headings.

### **WorldCat**

WorldCat is an online database that lets users search library collections around the world for physical items available to rent and digital content available to download. Similar to LibraryThing, libraries can integrate WorldCat into their catalogs so users can get the added benefit of browsing WorldCat's user-created content. However, WorldCat's tagging system is quite different from LibraryThing. In WorldCat, the tags are merely for personal use and browsing, they cannot be searched nor do the tags directly affect the metadata of the information resource. When WorldCat implemented user tags in 2008 they told Josh Hadro of the *Library Journal* that they were going to wait until a larger volume of users utilized the tagging system before making them searchable; it would simply be an "end user tool" as opposed to interfering or amending the official classification of an item (Hadro, 2008). It has been 4 years and WorldCat still hasn't reached the volume nor the precision in tagging that LibraryThing has, so the tags remain an end-user tool for personal categorization. WorldCat has made a "tag cloud" available on the homepage which contains the most popular user tags (Figure 3) but, compared to the tagging structure of LibraryThing, these "most popular" tags seem oddly specific and personal. Furthermore, the "Fantasy" tag on WorldCat has only been applied to 134 items (as of August 8, 2012), which is in stark contrast to the thousands of items in LibraryThing with a "Fantasy" tag.

Additionally, since WorldCat is used as an OPAC where users can search for items in libraries, all manifestations of a work are listed as individual records (all editions, hardcover, paperback, different languages, etc.) whereas in LibraryThing, each work is listed as one item.

This can make tagging very difficult in WorldCat, as users have to decide which item they are going to tag. This is shown in Figure 4, where there are 22 records for Terry Pratchett's *Good Omens*, none of which have user-created tags. While WorldCat does give users the opportunity to organize their favorite libraries, lists of books, and tag items that have personal meaning for them, in my opinion the site has not successfully implemented the tagging feature in such a way that a folksonomy has the opportunity to emerge.

### **Portland, Maine Public Library Catalog**

The Portland, Maine Public Library has implemented LibraryThing's features into their catalog. Tags are pulled from LibraryThing for each information resource available in the Portland Public Library Catalog. However, users can't search for tags right off the bat. The only way to see the catalog's "Tag Browser" is to find a resource and click on a tag for that resource. Once the Tag Browser is open, the user can search or browse for relevant tags which will give them items available in the Portland Public Library Catalog (see Figure 5). Users can't add their own tags through the Tag Browser, as the catalog is just pulling from tags on LibraryThing. While the library makes it a little difficult to get to the tags, and users are unable to interact directly with their library or other catalog users, it is still a useful tool and gives the user many opportunities to find indirectly relevant resources.

### **District of Columbia (DC) Public Library Catalog**

The District of Columbia (DC) Public Library Catalog has implemented its own tagging system. If a user has a DC library card and account, they can log in to the catalog and submit tags for resources, and all users can search the tags. However, the system has run into the low user count problem Rolla (2009) discussed in his paper. Because only a few users have utilized this feature, tags seem to have been applied at random and are rarely used. For example, on



LibraryThing and WorldCat, *The Lord of the Rings* or *Harry Potter* series hold top spots on the “Fantasy” tag pages; when I searched the “Fantasy” tag in the DC Public Library Catalog, they aren’t even on the first page of results (see Figure 6). The one interesting thing about this Web 2.0 feature is that each tag is divided into sub-topics, so a user can further refine their tag search (see Figure 6). However, low user count, inexact tags, and a greater focus on subject headings makes this feature seem a little irrelevant. With a greater volume of tags and users, a unique, regional folksonomy could emerge, which could have the potential for interesting interactions between DC Public Library information professionals and end-users.

### **Conclusion**

In conclusion, user-created metadata in the form of searchable tags in a Web 2.0 social environment is helpful to libraries in an age when digital content is constantly changing, keyword searching is more common, and users are accustomed to interacting directly with content and other users via the web. While informational professionals might be happy to pass over their indexing responsibilities to the end-user, we have to be wary of the inherent problems with implementing Web 2.0 features into metadata schemas and ontologies. Noise, quality control, and lack of user activity can greatly impede any system that relies too heavily on users to “do the work.” If libraries take a leaf from LibraryThing’s book, per say, and figure out a way to successfully make tag searching return relevant results, and are able to get a high volume of users willing to work together to create a folksonomy, then user-created metadata can greatly aid the searching process, user searching activity will become more social and comfortable, and opportunities for indirectly relevant resources will increase dramatically.

### References

- Guy, M., & Tonkin, E. (2006). Folksonomies: Tidying up tags?. *D-Lib Magazine*, 12(1).  
Retrieved August 8, 2012, from <http://www.dlib.org/dlib/january06/guy/01guy.html>
- Hadro, J. (2008). Tagging added to WorldCat.org. *Library Journal*. Retrieved August 8, 2012,  
from <http://www.libraryjournal.com/article/CA6600525.html>
- DC Public Library. (2012). Online catalog. Retrieved August 8, 2012, from  
<https://catalog.dclibrary.org/vufind/>
- LibraryThing (2012). HelpThing: Tagmash. Retrieved August 8, 2012, from  
<http://www.librarything.com/wiki/index.php/HelpThing:Tagmash>
- LibraryThing. (2012). LibraryThing concepts. Retrieved August 8, 2012, from  
<http://www.librarything.com/concepts>
- LibraryThing. (2012). WikiThing: Tag combining. Retrieved August 8, 2012 from  
[http://www.librarything.com/wiki/index.php/Tag\\_combining](http://www.librarything.com/wiki/index.php/Tag_combining)
- Majumdar, A., & Shukla, A. (2008). *Web 2.0: Implications on library*. Paper presented at the 6th  
International CALIBER 2008 on “From Automation to Transformation”, University of  
Allahabad, Allahabad.
- OCLC Online Computer Library Center, Inc. (2012). WorldCat. Retrieved August 8, 2012, from  
<http://www.worldcat.org/>
- Porter, J. (2011). Folksonomies in the library: Their impact on user experience, and their  
implications for the work of librarians. *The Australian Library Journal*, 60(3), 248-255.
- Portland Public Library. (2012). Portland Public Library catalog. Retrieved August 8, 2012, from  
<http://catalog.portland.lib.me.us/search>

- Rolla, P. J. (2009). User tags versus subject headings: Can user-supplied data improve subject access to library collections? *Library Resources & Technical Services*, 53(3), 174-184.
- Schwartz, C. (2008). Thesauri and facets and tags, oh my! A look at three decades in subject analysis. *Library Trends*, 56(4), 830-842.
- Shirky, C. (2005). Ontology is overrated: Categories, links, and tags. *Clay Shirky's Writings About the Internet*. Retrieved August 8, 2012, from [http://www.shirky.com/writings/ontology\\_overrated.html](http://www.shirky.com/writings/ontology_overrated.html)
- Vander Wal, T. (2007). Folksonomy. Retrieved August 8, 2012, from <http://vanderwal.net/folksonomy.html>
- Webb, P. L., & Nero, M. D. (2009). OPACs in the clouds. *Computers in Libraries*, 29(9), 18-22.
- West, J. (2007). Subject headings 2.0: Folksonomies and tags. *Library Media Connection*, 25(7), 58-59.

**Tag: fantasy**

Includes: fantasy, fantasia, ファンタジー, \*fantasy, fantasy, fancies, fantasy, fantassy, fantasia, fantasia, fantasia, fantasia, fantasia, fic genre: fantasy, Fancies, Fantasy, Fantasy, Fantasy, FANTASY, Fantasy, fanatzy, "Fantasy" - 102 more (what?)

Tag and its aliases used 47,979 times by 43,498 members.

There are open combination/separation votes for this tag

Filter: none (add)

**Most often tagged fantasy**

Using the tag fantasy

bookstopshere (5679), bluetyson (5632), esther\_a (4276), trollsdotter (2816), skullduggery (2702), StormRaven (2612), purpledragon42 (2341), slothman (2302), rt GemotCoenen (1908), roguelibrarian (1903), Rtrace (1897), bookstothesky (1691), AsYouKnow\_Bob (1624), Traste (1620), salimbol (1577), casa\_tali (1561), Jinjifore

**Combine tags?**

**Q:** What is tag combination?

**A:** LibraryThing members can combine closely synonymous tags, such as "wwii" and "World War Two." Combination does not affect individual members' tags, which never change, but can affect tag pages like this and other "global" pages.

**Q:** What are the guidelines?

**A:** Members have evolved **comprehensive guidelines** for what tags should be combined and which shouldn't.

**Q:** How does voting work?

**A:** Because tag combination prompts system-level changes, decisions are now made by votes. See the **guidelines** for how and when votes are closed and combination effected.

**Q:** Can I propose a tag combination now?

**A:** Yes you can.

**Q:** Can I propose to remove a combination?

**A:** Yes you can.

**Proposed combinations**

BarkingMatt has proposed combining this tag and the tag фантасика/fantasy.

**Vote:** Yes | No | Undecided **Current tally:** Yes 8, No 0, Undecided 1

**Related tags** (show numbers)

adventure British children  
 children's children's literature  
 classic classics **discworld**  
 dragons fairy tales **fantasy**  
 fiction graphic novel Harry  
 Potter horror humor  
 literature **magic** Middle Earth  
 mystery mythology **Narnia** novel  
**Pern** Pratchett read romance  
 satire **science fiction** series  
 time travel Tolkien urban fantasy  
**Valdemar** **vampires**  
 werewolves Wheel of Time witches  
 wizards **young adult**

**Related subjects**

Fantasy fiction (1,452)  
 Fantasy (650)  
 Fantastic fiction (497)  
 Science fiction (366)  
 Historical fiction (187)  
 Magic > Fiction (157)  
 Wizards > Fiction (125)  
 England > Fiction (101)  
 Children's stories (95)  
 Magic > Juvenile fiction (92)  
 Humorous stories (88)  
 Fairies > Fiction (87)  
 Large type books (82)  
 Horror tales (77)  
 (show all subjects)

**Related tagmashes**

fantasy, -Tolkien (200)  
 fantasy, -Harry Potter (190)  
 fiction, high fantasy (190)  
 high fantasy, science fiction (190)  
 fantasy, own (197)  
 fantasy fiction, series (196)  
 fantasy, fiction (196)  
 fantasy, series (196)  
 magic, sf (192)

**Combine tags?**

**Q:** What is tag combination?

Figure 1. LibraryThing Fantasy Tag Page (Retrieved August 8, 2012)

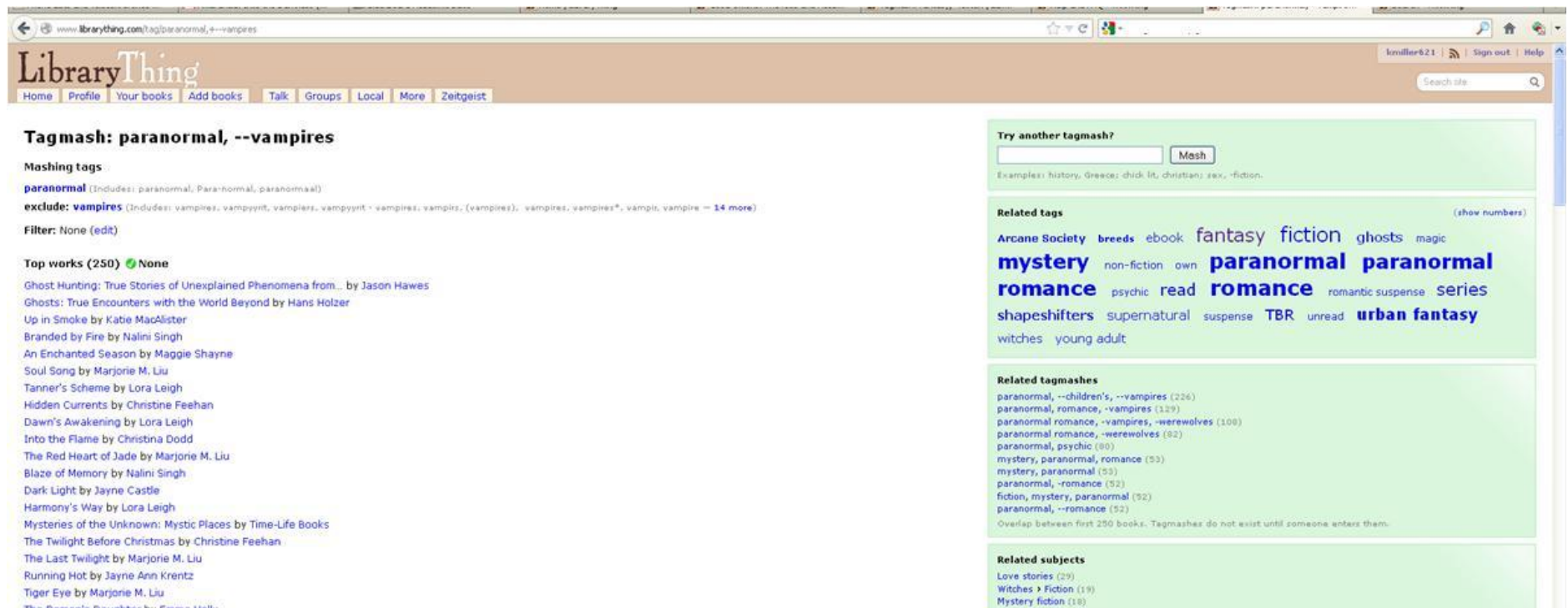


Figure 2. LibraryThing Tagmash Search (Retrieved August 8, 2012)

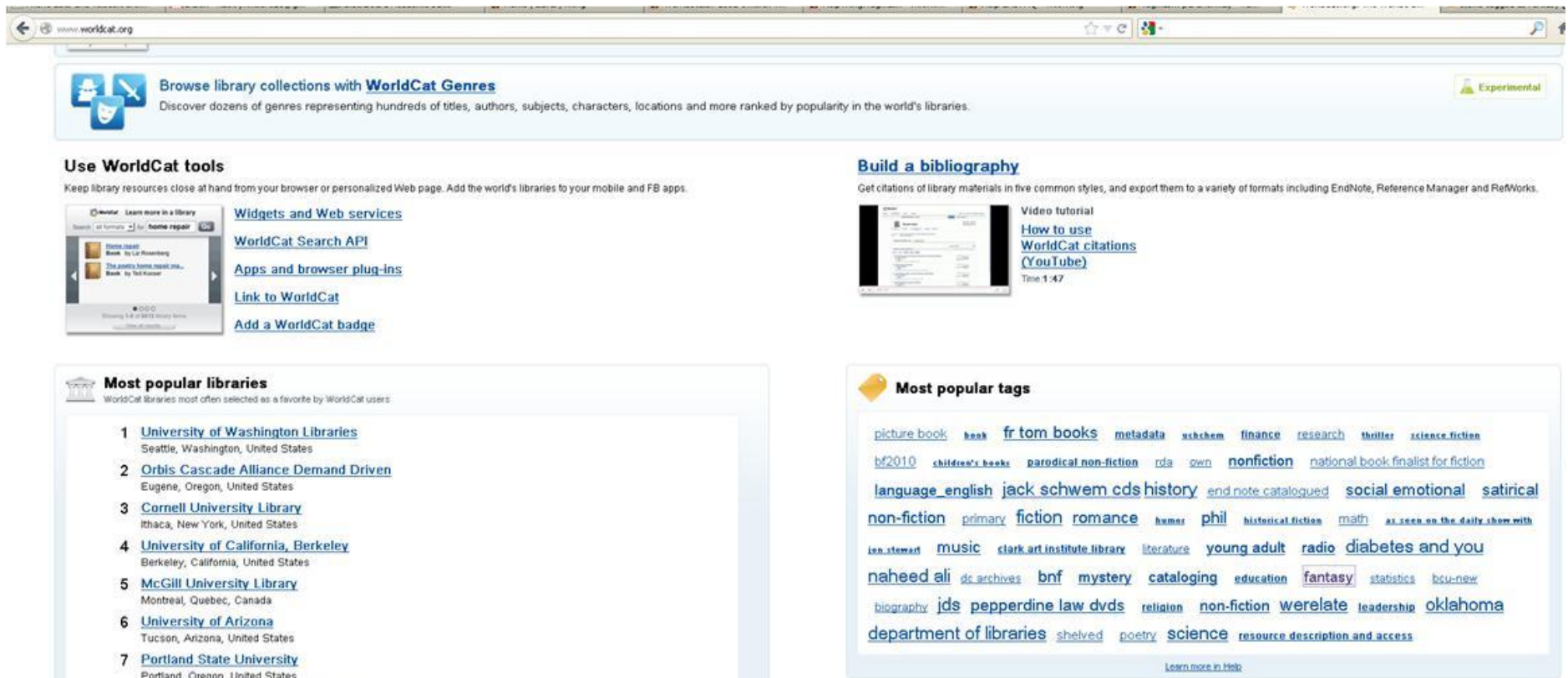


Figure 3. WorldCat.org Most Popular Tags (Retrieved August 8, 2012)

Search results for **"ti:good\_omens\_by\_pratchett" > "Terry Pratchett"**

Results 1-10 of about 22 (1.07 seconds)

Format:  All Formats  Book (22)

Save to: [New List] Save

Sort by: Relevance

**Refine Your Search**

**Author**  
[All Authors](#)  
Terry Pratchett

**Year**  
2007 (5)  
2006 (4)  
1996 (3)  
1991 (5)  
1990 (3)  
[Show more...](#)

**Language**  
English (9)  
German (3)  
Finnish (1)  
French (1)  
Italian (1)  
[Show more...](#)

**Content**  
Fiction (16)  
Non-Fiction (6)

**Topic**  
[Language, Literatures...](#) (9)  
[Computer Science](#) (1)



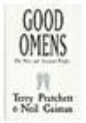

1.  **Good omens**  
by Terry Pratchett, Neil Gaiman  
Book : Fiction  
Language: English  
Publisher: Gollancz, 1990.  
[Editions and formats >](#)
2.  **Good omens : the nice and accurate prophecies of Agnes Nutter, witch**  
by Terry Pratchett, Neil Gaiman  
Book : Fiction  
Language: English  
Publisher: New York : Harpercollins, 2007.  
[Editions and formats >](#)
3.  **Good omens : the nice and accurate prophecies of Agnes Nutter, witch : a novel**  
by Terry Pratchett, Neil Gaiman  
Book : Fiction  
Language: English  
Publisher: London : Gollancz, 2007.  
[Editions and formats >](#)
4.  **Good omens : the nice and accurate prophecies of Agnes Nutter, witch**  
by Terry Pratchett, Neil Gaiman  
Book : Fiction  
Language: English  
Publisher: London : Corgi Books, 1991.  
[Editions and formats >](#)

Figure 4. WorldCat.org Retrieval List for Terry Pratchett's *Good Omens* (Retrieved August 8, 2012)

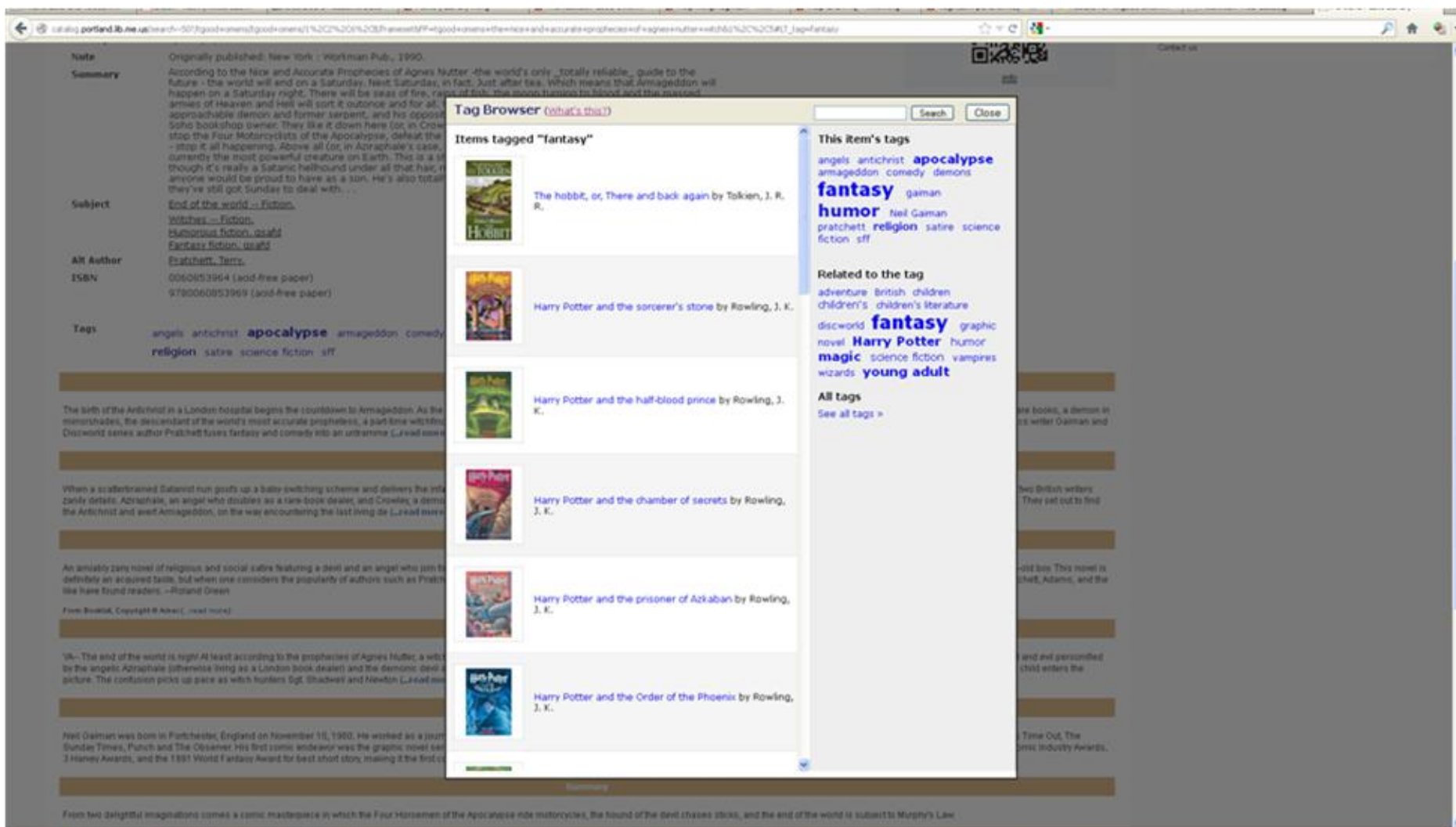


Figure 5. Portland Maine Public Library Catalog Tag Browser – Fantasy Tag Page (Retrieved August 8, 2012)



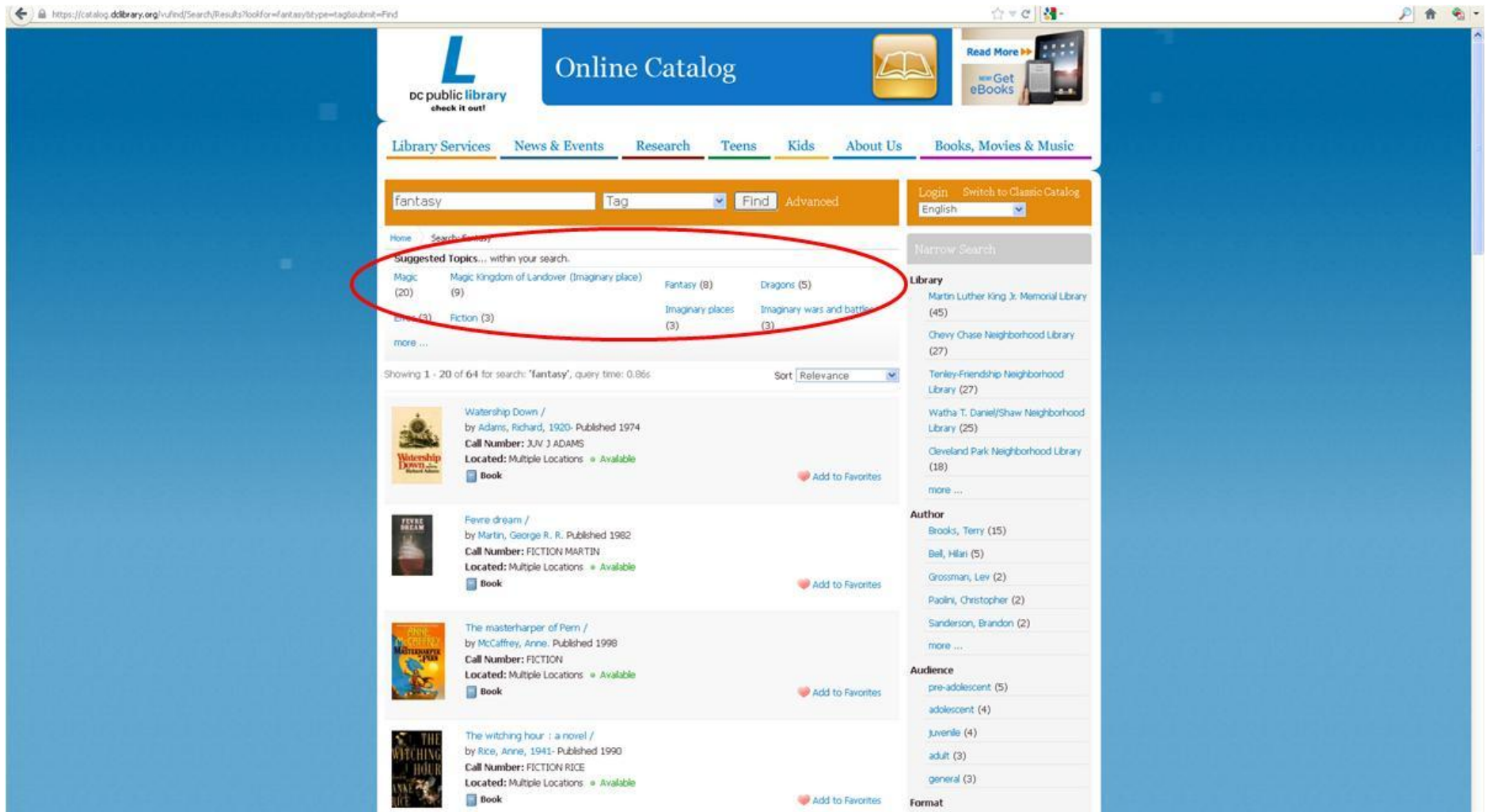


Figure 6. DC Public Library Online Catalog Fantasy Tag Page (Retrieved August 8, 2012)