

Angela Evans
S-503: Directed Research

Part One: Definitions and Compare/Contrast

Cataloging provides libraries with a solid and reliable foundation and support for the collections, ensuring through proper and accurate coding of information that the materials the library provides are easily accessible for patrons and staff alike. This is accomplished through a standardized MARC format wherein field 245 will always contain information pertaining to the title of the material, regardless of what library or material format is featured. Because of the thoroughness and regulations in MARC cataloging, libraries around the world are able to share records, improve upon records, and provide users with consistency. No system is perfect however, and aspects of the library catalog are in need of improvement. It's no secret that many patrons find a Google search to yield more accurate results than that of the library's OPAC. In order to better serve patrons, improvements to the records that OPACs pull from are in order. One means of doing this may be the introduction of folksonomy, wherein tagging materials can be performed by anyone, offering a wider variety of terms than just what the cataloging librarian uses. In addition to folksonomy, refining the controlled vocabulary that is used to describe materials to make them more patron-friendly would benefit the library.

The idea of an individual that is untrained in the aspects of cataloging, be it a paraprofessional or civilian, being in charge of something as important as descriptors for a material is somewhat daunting. However that is precisely the goal of a folksonomy, and it isn't as terrible of an idea as it may seem. By allowing other staff members or even the general public assign "tags" to materials, those materials become more searchable. "Tagging is the assignment of a keyword or phrase to a bit of digital data..." (Stephens, 2007, p. 58). The folksonomy itself is described by Stephens as the "group of tags created by the users of a certain collection of data" (Stephens, 2007, p. 58).

By utilizing the option of tagging materials in the library's collection, a variety of phrases, genres, characters, and even various slang or jargon that may not have normally appeared in the catalog record, but because of its inclusion the material has a more thorough record that may appear in a search it otherwise would not have. Arch refers to this potential as bringing "gray literature" to light, literature that, such as a scholarly work, may be difficult to locate through a general search (Arch, 2007, p. 80). With tagging, materials are given a variety of descriptors from multiple individuals. Folksonomies are a way to eliminate the confusion in determining the correct search term, or searching for an item only to come up with no results because a word wasn't included in the item's record, but may still bear a significant role in how much attention the tagging offers.

"...allowing (in)exact word matches, detecting semantically similar keywords, and using representative keywords in a folksonomy as opposed to LCSH, to describe a book, we can significantly reduce, if not eliminate entirely, the relatively high percentage of searches that generate no result or irrelevant results and improve the quality and quantity of the results retrieved for a library query." (Pera, Lund, & Ng, 2009, p. 1394).

In addition to the potential elimination of gray literature as Arch describes, Pera et. al describes the key advantage of folksonomies as a means of providing the most comprehensive access to the materials the library provides. Identifying trademarks of folksonomies are that they are social classifications, essentially user-created metadata (Spiteri, 2006, p. 77), and thus contain no controlled vocabulary.

Cataloging materials requires strict attention to detail. With this in mind, operating on a controlled vocabulary allows for a multitude of terms to trickle down to one meaning through hierarchical relationships. In the first line of his article, Leise summarizes a controlled vocabulary best by describing it as, "A controlled vocabulary is a way to insert an interpretive layer of semantics between the term entered by the user and the underlying database to better represent the original intention of the terms of the user" (Leise, 2012, p. 1). Humans speak in a natural language full of slang, jargon, and shortcuts. In addition to that, for many items there is more than one type or word for an item, such as nicknames or regional dialect differences. In a library setting, the Library of Congress [LC] provides the only subject headings list that is accepted on a worldwide scale. The Library of Congress Subject Headings [LCSH] is an example of controlled vocabulary in that it breaks down the subject categories in a material for a blanket, uniform subject across all records. Through a series of hierarchical breakdowns, broader terms [BT], narrower terms [NT], and other variants such as "USE" and used for [UF], the LCSH listings are easily searched and applied (Leise, 2012).

The function of a controlled vocabulary is to assign meaning and order to what would otherwise have the potential for disorder and chaos. Take Leise's example of pants for instance; the control term is pants, the narrower terms being "men's pants", "women's pants", and "kid's pants", and broken down further from there (2012). The connections made in a controlled vocabulary, or in the case of libraries the LCSH, help to organize and make sense of multiple terms or variables and format them in a way that provides a clean and easy searching experience. As Southwick et. al explains, a standard controlled vocabulary is a list of terms, with or without hierarchical structure, that are controlled by an authoritative organization, such as the Library of Congress, whereas the local controlled vocabulary is a controlled vocabulary in which terms are locally controlled (Southwick, Lampert, & Southwick, 2015, p. 182).

As different as the two systems are, they are simultaneously similar. Folksonomies boast having no connections to a controlled vocabulary allowing the fluidity of natural language impact the tagging and thus promoting a more solid searching experience. While controlled vocabularies on the other hand boast order and organization for a more uniform and cohesive catalog across the board. Between the two are obvious strengths and weaknesses, not only in the overall execution of the systems, but also in various applications of the systems.

Beginning with folksonomies, any librarian can tell you the advantage of the folksonomy without a doubt would be the capability of providing users with a more consistent search result culmination. Little is more frustrating for library patrons than searching what they consider to be the proper terminology and coming up with nothing, because of a lack of proper terminology, misspellings, or homonyms. The patron may assume the library holds no materials on a particular topic, and, depending on the patron's willingness to seek out help, if the patron leaves without resolving their search problem that's one patron that is not benefiting from all their library can offer. Folksonomies would in essence help prevent that by providing a multitude of tagged terms relating to the material. Porter cites an experiment that featured the use of folksonomies in the library's catalogue and concluded with "...a unanimously positive response from users, both in terms of their system's usability and its utility" (Porter, 2011, p. 253). Contrarily, he cited another study of a library that supplemented their catalog with LibraryThing, and also proved positive results, however noted that "folksonomies are only truly effective when large numbers of users have contributed to them (Porter, 2011, p. 253). Porter summarizes the effectiveness of a folksonomy in a library OPAC in the following statement:

“Despite the varied opinions on whether folksonomies are capable of adding value to the user’s experience of the library’s catalogue and resources, there is sufficient evidence in the literature to suggest that they have the potential to improve browsability and discovery of library materials. However, this is dependent on a large number of diverse users contributing to the development of the folksonomy” (Porter, 2011, p. 253).

Disadvantages to folksonomies are oddly the same as their advantages. Lack of semantic and linguistic control, in addition to the potential for inconsistent tagging, “trending” personal tags, and lack of interoperability with multiple languages (Porter, 2011, p. 251). Of these disadvantages, in the event of application of a folksonomy to a library catalogue, regulations could potentially be applied. Though it would somewhat defeat the free-reigning purpose of a folksonomy, it would solve the potentially detrimental disadvantages associated with inconsistencies.

Controlled vocabularies have the advantages of being consistent, frequently revised and updated, and represented by the worldwide accepted authority on controlled vocabularies. Controlled vocabularies in the library field are represented in the Library of Congress Subject Headings, which are consistently revised on a monthly basis (Library of Congress, 2017). For accuracy, and maintaining professionalism, using controlled vocabularies would be in the library’s best interest. Controlled vocabularies, unlike folksonomies, are phonetically sound and can address other issues such as homonyms, synonyms, and even misspellings (Leise, 2012). Eliminating the potential for errors in grammar is one key way that controlled vocabularies have an advantage over folksonomies.

Disadvantages of controlled vocabularies, like folksonomies, lie in the system’s strengths as well. Although controlled vocabularies have strength in their rigid organizational rules and regulations, because of this they are also hindered by not allowing the flow of natural language to occur. In folksonomies, multiple tags can be applied to an item to make it more searchable by use of natural language rather than a language that is meant to aid the information retrieval system. Cost is also a recurring disadvantage mentioned by multiple sources.

While it is apparent that the two systems, folksonomy and controlled vocabulary, are very different in structure, they both serve the same purpose of providing a descriptor or multiple of an item in the hopes of improving information retrieval for patrons and staff. Folksonomies are free-reigning, user-created, largely unfiltered tags with no links whatsoever to controlled vocabularies. Controlled vocabularies however are regulated, professionally-created systems to tag materials in a uniform and understood manner. Some researchers believe that these two systems, as different as they are, can be combined together in a step towards the “Sociosemantic Web” (Harper & Tillett, 2007, p. 65), proving that there is hope for the systems to merge and become the best of both worlds for information seekers on either side of the reference desk. It even seems possible to make the two systems work for each other in the best interest of furthering effective search results citing, “Flickr could convert and makes use of Library of Congress Subject headings to augment both the searching and development of their folksonomies” (Harper & Tillett, 2007, p. 64).

As different as they are similar, folksonomies and controlled vocabularies are essential in effective information retrieval. With the ever-evolving world of the library catalog, the web, and the connection that is thriving between the two, it would not be surprising if the two systems were one day effectively combined for a better searching experience for all.

Part Two: Examples

Folksonomy

In the following table, examples of a folksonomy are portrayed by the title and author of a work being featured, then followed by a portion of the table dedicated to the potential tags that may be associated with that material on sites such as GoodReads, LibraryThing, and so on. Some tags were retrieved from <https://www.goodreads.com> by selecting the “genre” link and seeing how other readers “shelved” the item, which is similar to tagging.

Title	Author
Harry Potter and the Sorcerer’s Stone	J.K. Rowling
Wizard, kid, fantasy, coming of age, fiction, young adult, thriller, paranormal, sci-fi, classic, magic	
Six of Crows	Leigh Bardugo
Fantasy, mystery, lgbtq, romance, action, multiple pov, high-fantasy, dystopian, series, crime, magic	
A Darker Shade of Magic	V.E. Schwab
London, fantasy, magic, lgbt, high-fantasy, historical fiction, action, adventure, parallel worlds, fiction	
The Raven King	Maggie Stiefvater
Magic, urban-fantasy, mythology, tarot, paranormal, romance, young adult, lgbt, friendship, horror	
Illuminae	Amie Kaufman and Jay Kristoff
Science fiction, fantasy, futuristic, dystopian, young adult, trilogy, adventure, action, romance	

Controlled Vocabulary

In the following table, examples of controlled vocabulary via subject headings from the Library of Congress are featured on five new book titles. These subject headings were retrieved from the LOC Authorities site <http://authorities.loc.gov/>. Included also are the LCCN Permalinks.

Title	Author
Astrophysics for People in a Hurry	Neil deGrasse Tyson
HEADING: Physics https://lccn.loc.gov/sh85101653	
Cravings	Chrissy Teigen
HEADING: Cooking https://lccn.loc.gov/sh2010008400	
Good Omens	Neil Gaiman
HEADING: Inspiration Religious texts https://lccn.loc.gov/sh85066720 BT: Supernatural, NT: Inspiration Religious texts	
Undeniable	Bill Nye
HEADING: Evolution (Biology) https://lccn.loc.gov/sh90004042 BT: Evolution, NT: Evolution (Biology)	
Seriously...I’m Kidding	Ellen DeGeneres
HEADING: Comedy https://lccn.loc.gov/sh85028845	

Bibliography

- Arch, X. (2007, Feb). Creating the academic library folksonomy. *College and Research Libraries News*, 68(2), 80-81. doi:<https://doi.org/10.5860/crln.68.2.7748>
- Harper, C. A., & Tillett, B. B. (2007). Library of congress controlled vocabularies and their application to the semantic web. *Cataloging and Classification Quarterly*, 43(3/4), 47-68. doi:https://doi.org/10.1300/J104v43n03_03
- Leise, F. (2012, December 16). *What is a controlled vocabulary?* Retrieved from Boxes and Arrows: <http://boxesandarrows.com/what-is-a-controlled-vocabulary/>
- Library of Congress. (2017). *Controlled Vocabularies*. Retrieved from Library of Congress: <https://www.loc.gov/librarians/controlled-vocabularies/>
- Pera, M., Lund, W., & Ng, Y.-K. (2009, Jul). A sophisticated library search strategy using folksonomies and similarity matching. *Journal of the American Society for Information Science and Technology*, 60(7), 1392-1406. doi:10.1002/asi.21072
- Porter, J. (2011, Aug). Folksonomies in the library: their impact on user experience, and their implications for the work of librarians. *Australian Library Journal*, 60(3), 248-255. doi:<https://doi.org/10.1080/00049670.2011.10722621>
- Southwick, S. B., Lampert, C. K., & Southwick, R. (2015, Jul-Dec). Preparing controlled vocabularies for linked data: Benefits and challenges. *Journal of Library Metadata*, 15(3/4), 177-190. doi:<https://doi.org/10.1080/19386389.2015.1099983>
- Spiteri, L. (2006, Nov). The use of folksonomies in public library catalogues. *Serials Librarian*, 51(2), 75-89. doi:10.1300/J123v51n02_06
- Stephens, M. (2007, Sep/Oct). Tagging and social bookmarking. *Library Technology Reports*, 43(5), 58-61. Retrieved from <https://journals.ala.org/index.php/ltr/article/view/4830/5794>