

**Future Information:
A Call for Including Instruction on Folksonomy in Schools**

John G. Hendron (johnhendron@gmail.com, www.johnhendron.net)
November 18, 2007 | Version 1.0

Rationale

I believe that being able to organize the digital artifacts at a student's grasp is an important skill. In fact, I would wager to call it a basic literacy skill for living in an information-saturated twenty-first century world. Furthermore, I believe that an education that prepares a student for life in the twenty-first century includes practice in the organization of information using folksonomies—personal, self-created tagging schemes attached to digital files.

This paper outlines some of the influences that gave rise to the term "folksonomy" and the need for social tagging systems. It puts folksonomy into the scope of a definition for information literacy.

Organizing Digital Assets

In the 1980s, with the rise in popularity with personal computing, the paradigm of organizing digital files into "folders" was an attempt to carry the metaphor of the office to the screen.¹ Files to be deleted would be carried to the "Trash." This metaphor worked reasonably well for users familiar with how documents are handled in an office

environment. By the end of the twentieth century however, this paradigm was becoming less-able as computer technology allowed users to store far more information on a hard disk drive than any physical file cabinet could.

The advent of the Web, and shortly thereafter, Web directories such as *Yahoo!*, provided yet another organization scheme built upon hypertext. This was modeled upon the paradigm of the phone book or the library. Yahoo! presented a directory of categories into which the user could "dial-down," getting more specific with each click of the mouse.² The user would eventually land at a list of hyperlinks showing sites of interest. As the number of Web sites increased, the "directory" was replaced by a "search box," where you could enter a title, phrase, or keywords combined with boolean operators to find what you needed across thousands of servers.

On our own computers, we are seeing far more robust search mechanisms that mirror the efficiency of our favorite online search engine. If I commonly use the word "report" in the titles of my word processing documents, I can command the computer to show me all instances of my reports in a few seconds.³

Yet another organizational paradigm is today at play that is useful to

the individual user. It does not rely upon folders or directories.

Combined with a search function, it can be used to look beyond titles and in-between lines of text. This organizational paradigm goes by several names and can be applied to any digital file—not just text. It goes by several names, among them most common, *tagging* and *folksonomy*.

Tagging is frequently an identifier of so-called Web 2.0 tools, including social bookmarking sites (del.icio.us), social photo sharing sites (Flickr.com), blogs (Wordpress.com), and video sharing sites (YouTube.com). In a tagged scheme, each file carries with it data about itself called *metadata*. Metadata can include dates, the identification of who created the data, or a set of keyword tags. On the Flickr website, this might be a photograph *John* took on *November 6, 2006* that deals with *vacation, Cancún, ocean, and fall*. The tags John assigned to this photo help John find his photograph in the future. But unlike a folder scheme, John's tags instantly group his one photo into a larger collection. They join other vacation photos of 2006, other Cancún photos by other photographers, and other ocean pictures taken on November 6, 2006. John's tags enable his photo, then, to belong to any number of diverse "groups" all at the same time. John's tags incidentally help other Flickr users find his photos, too.

The same tagging scheme is at play with the other examples cited

above. At *del.icio.us*, the tags help John find other websites folks have tagged with their own keywords. For blogs, John can visit a site such as *Technorati* and find other bloggers writing about concepts and topics of interest to John based on the tags bloggers are using. More recently, modern operating systems such as Apple's Mac OS X have added tagging capability to the files on John's desktop. Some applications on the desktop are adding metadata to files (Microsoft Word, for one, has added metadata to your documents for years based on your preference settings, such as the date and author). What is new is the user's ability to add their own keywords to the documents, music, photos, and movies they save to their hard drives.

Thomas Van der Wal coined the term *folksonomy* in 2004 to the act of tagging your own data.⁴ The term is also interchanged with "social tagging" and "collaborative tagging" to denote the use of tags by both consumers and producers in the cyberspace arena. What is unique about keyword tags that is reflected in the term "folksonomy" is that the scheme used to organize information is created "by the people" and "à la minute;" just-in-time, typically. Instead of a categorization system designed along the lines of the Dewey system used in libraries by a single, authoritative entity, the folksonomized system relies upon individuals alone and in informal groups.

There are critiques of social tagging. One of the more pronounced drawbacks is the ambiguity inherent in single tag words used by different people, or even the same person, at different times. In a perfect system, John's tagged photo of Cancún in 2006 should appear in searches for "México" but will not, since John did not include a tag for México. If John bookmarks a site about blogs today using *del.icio.us*, he might choose "blog" as the tag. He may also have tagged other, similar sites previously, with "blogging." Or "weblog." Or "MovableType." "Blogger," perhaps. In a similar scenario, John's friend Andy might have tagged his Cancún photos only with "Mexico" and "November" and his blogging bookmarks with "blogs." On the surface, it may appear that folksonomy as a organization system is flawed. Folksonomy is, however, cheap, quick, and easy. Anyone can do it. But few people would do it as well as a centralized, trusted authority.

Some cite the "everyman" approach to folksonomy as a strength.⁵ If users are using the vocabulary and connections to defining key terms themselves (and that make immediate sense to themselves), the system has a high degree of usability for the individual users and those thinking like them. When tags are used in combination with search engines, folksonomy as a scheme takes on further usability. With *del.icio.us* as an example, when John enters his own tags, *del.icio.us*

"autocompletes" the tags and offers suggestions based on the tags already used to categorize that Web link by other users in addition to the tags John has already been using. He can choose all the blogging terms he likes: "blog, blogger, weblog, weblogs," etc., and it is all done quickly by clicking on the most relevant tags to John.

The boolean approach used in search engines comes to play with organizing tagged-content, too. Andy can find other content online by combining tags together ("México" plus "vacation", "México" but not "ocean"). With *del.icio.us* in particular, he can also add tags that assign the website to another user's collection or "cloud" of websites. All of these techniques, combined, allow John, Andy, and whomever else, to find, discover, and make sense of the information their eyes run across.

The question becomes for educators, then, "How important" is this capability of individuals to review, tag, and organize digital content? Today, John's hard disk drive is almost 450 GB in size. In 1998 it was 1 GB. In 1988 it was just 20 MB. More instant messages, movies, music, saved PDF files, and photos will be saved upon it this year. The price of hard disk drives has been decreasing, while disk capacities are increasing.⁶ It is very likely that citizens in the information society today will be able to maintain all of their digital content, as much digital content as they desire, for the remainder of

their lives. Assuming John upgrades to a 750 GB drive in 2008, in that span of 20 years, his digital storage locker increased by 37,500 times. Each hour, day, week, and month more data are being saved to disk. These are data that we may want to access, read, listen to, or watch tomorrow. Multiply one user's personal content with all the content produced by other users online, and the need for an organizational system that looks beyond folders and trash cans is necessary. On a linear scale (and clearly, data capacity has not been increasing on a linear scale), John's storage ability in 2028 would be at 28,125,000,000 MB, or 26 petabytes.

Folksonomy, then, looks like an attractive means of organizing digital content—not only for ourselves, but for each other. The better we become at social, collaborative tagging, the better everyone is for finding relevant, needed content. This is why I feel strongly about teaching students about folksonomy in schools while we wait for a so-called “semantic web.”⁷ The benefit of using a social tagging system to organize digital content benefits the individual student and those with whom this content is shared. Students can even practice folksonomy with everyday objects, including handwritten papers, books, video cassettes, food containers, etc. The future could hold opportunities for tagging beyond digital files, too, if the products we buy are encoded or associated with tag words.

Folksonomy in Praxis

Taxonomies are already an important concept in many curricula. They play a big role in the sciences. Folksonomy, however, is imposing your own classification system upon anything and everything around you. The following list includes binary sets of tags that could be applied to a variety of artifacts, physical or digital.

- like, dislike
- love, hate
- important, trivial
- expensive, cheap
- hot, cold
- sharp, dull
- heavy, light
- rough, smooth

This list includes larger sets, again, that can be used upon a variety of media.

- red, blue, green, yellow, white, black, purple
- *****, ****, ***, **, *
- friend, acquaintance, enemy
- valuable, worthy, worthless
- large, medium, small
- science, languagearts, history, math, arts
- easy, understandable, difficult

The practice of applying tags to the artifacts in one's own world might start with a simple exercise of attaching self-stick notes to physical objects in the classroom. Students in a group could likely generate a number of tags to apply to each object. The experience of agreeing and debating which tags best categorize the object requires collaboration, an important twenty-first century skill. (It is included in several outcomes defined by the *Partnership for 21st Century Skills*, including Learning and Innovation, and Life and Career.⁸) It also engages higher-order thinking skills by requiring evaluative decisions about which sets objects belong.

When working with digital media, students can apply keyword tags to the documents they save, the websites they bookmark, and the photographs they shoot. The practice of applying keywords helps students better the quality of the tags chosen, not only for their own use, but for the use of other students.

Including folksonomy skills in school builds upon the constructivist notion of students creating their own knowledge. As the YouTube video *Information R/evolution* suggests, the systems we use to organize and get-at information are of paramount importance in the digital age.⁹ Social tagging may be the best mechanism we currently have at using the information made available through digital sources.

As you browse through the framework created by the *Partnership for 21st Century Skills*, social tagging impacts a number of elements: critical thinking, communication and collaboration, information literacy, ICT literacy, and productivity as a life/career skill. These elements have been defined by a number of high profile organizations and individuals, including Adobe Systems, the American Association of School Librarians, Educational Testing Service, Ford Motor, Junior Achievement, Microsoft, the National Education Association, and the Sesame Workshop, suggesting their importance for the future of work.

Definitions of Information Literacy

How might we define the term "information literacy?" Let's examine a few different definitions to extract some core attributes. The American Library Association says this: "Information literacy forms the basis for lifelong learning. It is common to all disciplines, to all learning environments, and to all levels of education. It enables learners to master content and extend their investigations, become more self-directed, and assume greater control over their own learning."¹⁰ Oklahoma State University's library says information literacy is "the ability to recognize a need for information, find, evaluate and use that information in whatever format (print index, online database, Internet, etc.) it appears."¹¹ Benedictine University

defines it as such: "The ability to access, evaluate, organize, manipulate, and present information."¹² Davidson University library has their definition: "The competencies and skills students need to locate, retrieve, evaluate, analyze, and use information. These competencies are developed over time and are essential for lifelong learning."¹³ The Wikipedia entry for "information literacy" breaks down the concept into different areas, including: tool literacy, resource literacy, social-structure literacy, research literacy, publishing literacy, emerging technology literacy, and critical literacy.¹⁴

The similarities? *Using information, lifelong learning, locating information, evaluating information...* Clearly, a construct dedicated to creating some order out of the chaos of information has a place within a definition for information literacy. Folksonomy—a social practice of applying keyword tags to information sources and artifacts—hits upon many of the specific aspects cited in the article from the *Wikipedia*. In and outside of the library, our definition for information literacy should evolve to embrace the spirit of the folksonomized Web.

Considering what social tagging does for users of the Web at the start of the twenty-first century, I propose a new definition for information literacy that includes folksonomy as a critical element. Therefore, *information literacy* is "a set of essential skills required to locate, evaluate,

categorize, and organize information for the benefit of one's self and others in a collaborative environment." The categorization and organization of information is what we might identify as new elements in our definition, both requirements and a necessary evolution for a society determined to publish and link an ever-growing collection of information artifacts across distributed computer networks.

Conclusion

Until a better system emerges to organize and categorize the petabytes of information that is magnetically-encoded within our computers and the world's Web servers—something akin to the so-called "semantic web"—the art of folksonomy is something we can all begin using and teaching today. It encourages collaboration among students at higher-order thinking levels, it can be both personally and collectively helpful and relevant, and encourages a lifelong pursuit of productivity in the digital domain.

Further Reading

- Dornfest, Rael and Dan Brickley. *The Power of Metadata* at <http://www.openp2p.com/pub/a/p2p/2001/01/18/metadata.html>
- Lankshear, Colin and Michele Knobel. *Blogging as Participation: The Active Sociality of a New Literacy* at <http://www.geocities.com/c.lankshear/bloggingparticipation.pdf>
- Mathes, Adam. *Folksonomies – Cooperative Classification and Communication Through Shared Metadata* at

- <http://www.adammathes.com/academic/computer-mediated-communication/folksonomies.html>*
- Shirky, Clay. *Many 2 Many—Folksonomy* at *<http://many.corante.com/archives/2004/08/25/folksonomy.php>*
 - Smith, Gene. *Folksonomy: Social Classification* at *http://www.atomiq.org/archives/2004/08/folksonomy_social_classification.html*

Cited Resources

¹ Two sources that follow the evolution of the “desktop” on personal computers of note are the Wikipedia article on the topic, online at *http://en.wikipedia.org/wiki/Desktop_metaphor* but also the evolution of the Apple Lisa’s interface, which was a predecessor to Macintosh: *http://www.folklore.org/StoryView.py?project=Macintosh&story=Busy_Being_Born.txt&topic=User%20Interface&sortOrder=Sort%20by%20Date&detail=medium*

² A version of Yahoo! circa 1995 is available through a screenshot at: *<http://battellemedia.com/archives/000935.php>*

³ Powerful search capabilities have been added to both Microsoft Vista and Mac OS X Leopard. In addition, Google Desktop Search provides similar functionality on either platform.

⁴ see *<http://en.wikipedia.org/wiki/Folksonomy>*

⁵ see *<http://www.adammathes.com/academic/computer-mediated-communication/folksonomies.html>*

⁶ see *<http://www.littletechshoppe.com/ns1625/winchest.html>*

⁷ The semantic web (*<http://www.w3.org/2001/sw/>*) is an effort to codify how a standardized metadata system could work for identifying what content was available on webpages.

⁸ *Partnership for 21st Century Skills* is online at *<http://21stcenturyskills.org/>*

⁹ The video is located at: *<http://www.youtube.com/watch?v=-4CV05HyAbM>*

¹⁰ Definition found at: *<http://lib1.bmcc.cuny.edu/help/glossary.html>*

¹¹ Definition found at: *<http://www.library.okstate.edu/infolit/glossary.htm>*

¹² Definition found at: <http://www.ben.edu/library/help/glossary.htm>

¹³ Definition found at:

<http://www.davidson.edu/administrative/library/refer/jargon.asp>

¹⁴ Article found at: http://en.wikipedia.org/wiki/Information_literacy