



U-Content

Doing It With Wikipedia

by Nicholas Tomaiuolo
Instruction Librarian
Central Connecticut State University

The print article appears in the April 2009 issue of *Searcher: The Magazine for Database Professionals*. To purchase a PDF of the full article, go to www.iti-infocentral.com.

An Insider's View on Using Wikipedia to Improve Library Materials Usage

Dreanna Belden, coordinator for grants and development, information technology services, digital projects, University of North Texas Libraries, is the author of "Harnessing Social Networks to Connect With Audiences: If We Build It, Will They Come 2.0?" (*Internet Reference Services Quarterly*, 13(1), 99-111, 2008).

Would you please help clarify for our readers exactly what role you play at the University of North Texas Libraries? How is the topic (i.e., placing links in Wikipedia and other social networking sites) related to what you primarily do there? How does the social networking/library links project fit in with you, your job, your background?

I do a wide variety of work at the UNT Libraries: I write the grants for the digital projects unit, manage relationships with our 95 collaborative partners, handle new partnership inquiries, and create and launch any promotional items or ideas for the Portal to Texas History. I see adding Wikipedia links as a way of increasing awareness and providing access to our digital collections while providing people with relevant, authoritative content.

In the introduction to your *Internet Reference Services Quarterly (IRSQ)* article, you do a good job of convincing librarians that press releases and other mechanisms such as partnerships with like-minded websites aren't optimal ways of driving traffic to library websites, because digital collections, which are often hidden in the deep web, are not as transparent as we'd like. What made you decide to start looking at social networking sites as possible places to help extend the reach of libraries?

The Internet provides the greatest opportunity to grab people at the point that they are looking for information and to direct them to a place that can answer their question or provide more in-depth

resources. I just decided to start experimenting with some different social networking sites to see what was most effective. Some of them proved to be more worthwhile than others.

In your article you often refer to the "staff at UNT." For example, you wrote that, "Staff at UNT noticed that Google searches relevant to the material offered in the Portal to Texas History almost always resulted in a Wikipedia article about the same topic..." and you go on to say that "UNT staff edited over 700 articles in Wikipedia ..."

Well, we are a "we" kind of department in digital projects, but for adding the Wikipedia links it was more like me sitting at my computer at home and adding links! Some people play video games (well I do that too), but for a few months I invested time in adding a lot of links to Wikipedia, especially when we saw the results from my early efforts.

The whole thing really started when I noticed Wikipedia results turning up in my Google searches. Early on when we started the Portal, I searched for websites or forums that I wanted to link to us. I emailed a lot of folks to let them know we were here offering free access to lesson plans and thousands of historical documents, and to let them know they were welcome to link to the Portal. I noticed Wikipedia results showing up in my searches, or if I was researching a topic for a grant. I wondered what would happen if we added some links from Wikipedia into the digital content offered on our site.

How formalized is the project?

It is not formalized at all! Early on I just decided to add some Wikipedia links back into our content. It was fun, so I added sev-

continued on page 2



eral hundred links in the first six months or so when I began. We saw some significant impact on visits to our site through our web traffic software.

How formalized is the commitment?

Well, after the first year of experimenting with this and viewing the great results we had, we added it to my performance agreement. My task is to “explore the use of Social Networking websites for opportunities to increase access to the Libraries’ digital collections.” The Method of Assessment is web traffic analysis, and here are the results for our 2007-08 fiscal year:

Number of visitors coming to UNT digital collections from Wikipedia and GenWeb, for FY 2007–08
Portal to Texas History: 294,000
Congressional Research Service Reports: 131,000
UNT Digital Collections: 42,500

These numbers represent a nearly 100% increase over the previous year.

How is the work divided up?

I’m the only person in the library adding the Wikipedia links, but we reached a “saturation” stage a long time ago, in that I’ve linked to our collections in any way that seems useful and relevant, and just add new collections as they are uploaded. It doesn’t take very much time anymore.

Are there teams (if so, what do the teams do)?

Nope, just me.

Was the real work of the project (placing links in Wikipedia) done systematically or was there a “go for it” feeling driving it (with individuals placing links wherever and whenever they could)?

Definitely a “go for it” type thing, which is pretty much the character in our unit. You try something, it works, you build on it — if it turns out to be a lame idea, that’s ok too. I want to emphasize that I do put a lot of thought into whether linking things is useful or not to most Wikipedia users — and sometimes other editors will disagree or delete entries. But when adding links, I always try to ensure that we really have enough information about a topic to make it worth the user’s time to click through.

How many human hours have been devoted to the project?

Wow, a lot. Early on I spent a significant amount of time doing it, but it takes much less time now. I’m guessing, but total time investment is perhaps 150 hours, with most of that coming up front in the first year. Now all I do is add links to new collections

and materials we upload if a link makes sense. For instance, I already have a link to historic photos from Austin in the “Austin” Wikipedia article. If we uploaded another 1,000 images of Austin, the link is already there, so there is nothing for me to do. At this point I am only adding in links for material that is new and unique or in a new geographical area for us.

I just mentioned that I consider the project at UNT to be on a large scale. Besides this experience and the one reported by Ann Lally at the University of Washington, do you know of other libraries that are really working on doing the same thing on any scale? If so, where are they and what types of collections do they have?

I’m not sure who else might be doing it, but I have talked about it with a number of people in digital libraries and suggested they try it. I know that the “Chronicling America” historic newspaper site at Library of Congress made a decision to not add any Wikipedia links because they know that Wikipedians may interpret these types of links as “spam.” You would think that people editing articles would be looking for authoritative sources to bolster their usefulness! I can’t figure out why people editing at Wikipedia are not doing more to link to digital library content and great resources like Google Books.

In the *IRSQ* you reported that Wikipedia accounts for 48% of referring traffic and Google (surprisingly) accounts for only 16.68%. How have these figures changed since the article was published?

Our traffic has just about doubled since I finalized that article, and we’ve done a lot to open up our content to Google and other search engines in the past two years, so the numbers really reflect those changes. The percentages have flip-flopped. Currently, about 44% of our traffic comes from Google, and only about 14% comes from Wikipedia — but the number of click-throughs from Wikipedia is still increasing slightly every year as we add new links. Our overall traffic has increased dramatically, while the Wikipedia click-throughs have grown less significantly.

You have placed many valuable links in Wikipedia. I’d read about some good resources from the University of North Texas’s “Primary Source Adventures.” One that I was particularly interested in was the Primary Source Adventure called “Iwo Jima: Forgotten Valor.” I wanted to take a look at that resource, but I wanted to do it through Wikipedia and a simple search on the words “Iwo Jima” brought up two possible articles. The UNT primary source site was, appropriately, linked to the second article “Battle of Iwo Jima,” and not the article on the island — this showed that the



UNT Libraries staff have been diligent about putting the links in the correct articles. After I found the “Iwo Jima: Forgotten Valor” link in Wikipedia, I clicked through to the PowerPoint presentation at UNT and became engrossed in a transcript of Charles W. Lindbergh talking about his Air Force training. The transcript came from the University of North Texas’s Charles W. Lindbergh Oral History Collection. All of this history is fascinating. Since the staff at the UNT libraries have edited more than 700 articles, what are some of the other articles that have been edited, and what UNT content do they link to?

Here are a few examples of the types of links we have added:

“Tanning” is a Wikipedia article about processing leather. We created a link to a government publication, “Home Tanning of Leather and Small Fur Skins.” I guess that would be useful to the ultimate do-it-yourselfer.

In the Norman Rockwell article, we provided a link to nine WWII posters that the artist designed.

The Fort Wolters article covers the U.S. Army installation in Mineral Wells, Texas. It served as a training facility for WWII and then served as the Army’s primary helicopter training school for several decades. We linked to 23 volumes of “A Pictorial History of Fort Wolters,” which has thousands of pictures in it.

The Texas City Disaster is still the deadliest industrial accident in U.S. history. We created a link to over 300 photographs we have online that show the extent of the damage and the aftermath.

We’ve also added in a lot of links that are geography-based at the county or city level.

When I read an article in Wikipedia, I noticed that references to the Portal to Texas History, which refer back to the University of North Texas Libraries often have two links. For example, the first link goes to a specific part of the UNT digital collection (the Battle of Iwo Jima article carries a link to “Iwo Jima: Forgotten Valor”) and then an adjacent link appears called “from the Portal to Texas History.”

In my reading of Ann Lally’s article about the University of Washington’s endeavors into library outreach through Wikipedia, she stated that one prominent obstacle to infusing Wikipedia with credible information was that even authoritative links can be construed by Wikipedia editors as “self-promoting.” While I certainly see nothing about a link to the entire Portal to Texas History that indicates self-promotion, this is exactly the type of situation Lally describes as an obstacle to libraries.

Why do you think Wikipedia has permitted UNT to include a general link to the Portal in these articles? Have you had any “talks” with Wikipedia moderators or editors?

Oh yes, and Ann is exactly right about that. I’ve had a few “talks” with Wikipedians, and they often do see this type of linking as spam and “self-promoting.” For some Wikipedians, it doesn’t matter that our site is totally non-commercial and offers authoritative materials totally free for the public to use. I believe their official policy is that any linking to a site you are associated with is not allowed, but I disagree with it and add links anyway. I mean we are a library The fact that most of the links to our site remain on Wikipedia and are not removed tells me that most people find these to be useful information.

As far as putting in two links, I just started formatting them that way and it’s stuck. No one has said anything about it to me.

In the *IRSQ* article, you mention that the UNT Libraries have a presence on MySpace and YouTube. I checked both out. The one at YouTube is a short introduction to the Portal to Texas History — but that seems the only item associated with UNT that I can find. What is the role of MySpace and YouTube (in contrast to Wikipedia) as places to help potential users find your digital content?

I tried the MySpace thing for awhile, but I don’t think it’s the optimal place to connect people with digital collections. It could probably work great for public libraries and teen groups, but I would call our foray on MySpace a failed experiment. As far as YouTube, I’ve just done the one video. We are in the middle of a major site redesign that will launch this spring, and one thing we plan to do is create “help” videos that we will place on YouTube and embed in our site. After trying a number of social networking sites, I would say that Wikipedia is the best way to connect people with digital collections based on our experiences.

The UNT Libraries maintains an online Government Documents collection consisting of 10,200 Congressional Research Service Reports. You’ve managed to embed 160 links within Wikipedia that refer users to these reports. But the Wikipedia articles you mention have widely different foci. For example, you’ve put links to Gov Docs in an article about the spotted owl, teenage pregnancy, and the Gulf War. This is quite different than simply targeting articles in a subject area like Texas history. It seems as though you (or someone) is carefully going through UNT’s digital collections, one item at a time, and then locating the relevant article in Wikipedia. Is this the way it is actually done, or is there a less labor-intensive, automated solution to matching the contents of the collection with Wikipedia articles? If so, can you let us in on it?

Early on, I did go through our content carefully to determine good matches for Wikipedia articles, but I didn’t go through item by item. The CRS Reports do have a browse by subject page, so



that made those a lot easier to conceptualize. For the Portal and our other digital collections, we have a great metadata analysis tool on the back end that Mark Phillips created, which allows you to easily look at the data in different ways. I can look at word clouds of subjects, titles, or other data fields to discover topical areas where we have a lot of content. It's a visual representation of the data, so I can see that we have a lot of materials about a certain person or geographic area. I can then investigate if that would be a good area or topic to link to from Wikipedia.

In the *IRSQ* article you write about many Portal to Texas History resources on the Alamo and that links to the resources have been placed in Wikipedia. Try as I might, I do not see links to the UNT Libraries in either the Wikipedia articles called "Battle of the Alamo" or in the article called the "Alamo."

Did Wikipedia delete the links? Because you wrote: "Schoolchildren in California writing book reports on the Alamo use books, floor plans, and images from the Portal after discovering them through Wikipedia" certainly indicates that links referring back to UNT should appear in the "Battle of the Alamo" Wikipedia article. What do you think?

Yes, it looks like someone deleted the link and many other links that used to be there. You can never count on the fact that what you put up will stay there. "Wikipedia" does not delete the links,

but every article is up to the interpretation of anyone. If some stray Wikipedian decides that they think an external link is irrelevant, BAM! They will delete it without a thought. Here are some examples of links I created that have been deleted:

Austin College article. We have digitized every Austin College yearbook from 1899 to 1950. Every page is there, and the entire collection is searchable. Most people would assume that this is a great resource for people who are interested in Austin College, and it is a noncommercial, free resource. Someone decided that all the external links in this article are a "link farm" and deleted all of them except the one to the official Austin College website. The same thing happened to the Abilene Christian University article and we have all of their yearbooks online from 1916–2007.

When this happens, I don't even try to fight it anymore. I may go back in a few months and add it back in, but I don't confront the person who did it. I found that while some people are reasonable, some just like to make it into a drubbing.

In the "old days," I suppose we just counted on old ways of promoting our digital collections, and were satisfied. What is your overall internet traffic like now?

Visits in the past 14 months — 218,599 from Wikipedia and 360,136 from Google.

An Academic Librarian/Wikipedian Weighs In

Phoebe Ayers is first author of *What is Wikipedia?* (No Starch Press, 2008) and a science librarian at the University of California–Davis. Her professional interests include writing historical science biographies, fact-checking and referencing, teaching how to use wikis, and Wikipedia.

Phoebe, as a Wikipedia editor and a librarian, you must have many insights into the project. Please give me your overall impression and thoughts.

Wikipedia is an extraordinary project: by any measure, the largest reference work to ever be created, undoubtedly the most multilingual, definitely the most free and open, and certainly the most-accessed reference site online today, outranking all other websites except for the most popular search portals and social networking sites. And all this content has been developed by volunteers, working under a handful of basic behavioral and content

guidelines but without any top-down editorial leadership or control. On Wikipedia, the community of editors develops not only content but also, through discussion, trial, and iteration, the standards to which they edit articles. Wikipedia is also relatively new — founded in 2001, the site didn't really take off until 2003, when successive waves of new editors and the content they contributed helped cement the site as a useful reference resource.

One thing I didn't like about Wikipedia when I first began to think about it is that the articles could, potentially, always be under revision and, therefore, weren't necessarily "correct" at any given time. What about that?

What all this means is that the whole site is a work in progress. The openness of Wikipedia is in large part grounded in a belief that not only can everyone contribute something to the endeavor, but also that all content can continue to be improved. Every article is a draft.



Some articles are excellent, having been worked over by dozens of committed writers, editors and researchers who have produced work that rivals the best general encyclopedias; while other articles are simply beginnings, “stubs” that contain little content, or possibly even pieces that don’t comply with Wikipedia content guidelines, either because the authors were unaware of them or because the article is still getting there, waiting for an editor to come along and take the page under his or her wing.

You’re an official Wikipedia editor. Tell me more about editing Wikipedia.

It is this basic task — a volunteer editor taking up an article that looks interesting, or that they are knowledgeable about, or that simply needs some work, and then improving grammar, adding missing content and adding and checking references for accuracy — that makes up the heart of what goes on in Wikipedia. And every task can be further broken down. Many editors specialize (I personally enjoy researching and adding references, even for topics I don’t know much about — I **am** a librarian, after all — and doing some maintenance work, merging articles that are accidental duplicates of one another, or doing copyediting). Other editors enjoy adopting a single article and bringing it up to the highest standards they can; while others keep an eye on new contributions; while still others add those new contributions, trying to reduce the missing articles lists. Still others work primarily on the “back-end” of the community, working on technical development, discussing policies or doing outreach. All of these individuals, whether they fix a typo occasionally or write pages from scratch, are classed as “editors,” and with few exceptions they all have the same privileges — to work on Wikipedia, as long as their changes improve the encyclopedia.

What about my concerns that the articles are never finished, and that so many people have too much input. The old saying goes, “Too many cooks ...”

It’s this unevenness that I think gives most librarians pause. We like to be able to say that a source is good or it isn’t. We can’t do this for Wikipedia. It is not a monolithic source, any more than the community of hundreds of thousands of volunteer editors constitutes a monolithic editorial board. Another common objection is that it’s difficult, and usually impossible, to know whether any particular editor knows what they are talking about when they add content. Editors aren’t vetted for their credentials, the way they are in academic publishing; this is seen as part of the site’s open culture, less a flaw than a prominent success.

What are your recommendations for those of us who want to, as librarians and teachers, evaluate an article?

What anyone who wants to evaluate a Wikipedia article needs to do is go back to the text, to the individual article in question, and ask:

- Is this content accurate — does it match up with what multiple outside reputable references say is true?
- Are those references good quality; are *they* likely to know what they are talking about?
- How’s the writing style? Does it match with Wikipedia’s standards? And, finally,
- Is this the work of only one or of many Wikipedia editors? Can you tell who they are? (this information is provided through the “history” tab).

Note this is a very brief summary of Chapter 4 of my book, “Understanding and evaluating an article” [<http://howwikipedia-works.com/ch04.html>].

Now, where librarians have an exceptional role to play, in my opinion, is in the first two points. We know how to find and check references better than just about anyone. So why don’t we? This peer-review-through-referencing is one of the hardest and most pressing tasks on Wikipedia today, and we are uniquely well-equipped to do so.

I’m not sure if it’s a manifestation of the “If you can’t beat ‘em, join ‘em” or “I want to buttress the encyclopedia’s content” that is beginning to get a foothold. Yet, I know several librarians and professors who won’t even mention the word Wikipedia except in disdain.

The question arises: Why work on Wikipedia at all? Why even bother with a project that doesn’t meet our standards? And my answer is: *because it’s important*. Wikipedia is idealistic, a non-profit organization that aims to do the impossible — collecting all human knowledge in all languages — and present it in an accessible, free, and open fashion. Moreover, our patrons are using it, and they are doing so because it is the best quickly accessible store of basic information online today. Why would we *not* want to make that better? And while we’re at it, we should teach our patrons about it, about how to evaluate articles, about how to cite it properly (hint: the “cite” link on the left-hand sidebar), and other tricks and tips they may not know. Wikipedia is a fabulous laboratory for learning and teaching about research, collaborative work, copyright (and by extension scholarly publishing), and internet content and culture.

It can also be difficult to participate in. As one might expect for a technical project with so many contributors, the site’s culture,



policies, and even technical aspects have become immensely complicated over the years. So while Wikipedia is entirely open, it's not always transparent.

How about wrapping this up by giving *Searcher's* readers some suggestions as to how we can proceed in getting involved with Wikipedia.

Here are some things that librarians can do:

- Learn about the site. Start with <http://en.wikipedia.org/wiki/Wikipedia:About> and then move on to any of the many instructional resources available. (This is why we wrote "How Wikipedia Works.") Your patrons, colleagues, and bosses will and probably do have questions about Wikipedia, and it's not worth ignoring. Then teach the site: For all of the millions of people who use Wikipedia (including every computer-savvy student I've ever met), few of them actually know how it works or what they are looking at.
- Evaluate articles. Are there especially good ones you can use? Especially bad ones that need work? Create a user account (free), then participate in cleanup and evaluation efforts.
- Add sources: As I said above, we're uniquely well equipped to do this. The syntax is tricky but you can follow along with existing code, or the guide here: http://howwikipediaworks.com/ch06s03.html#referencing_styles. Adopt an article about a favorite topic, and then think: How would you reference this article, focusing on references that are accessible to as wide an audience as possible, if you were writing the best general encyclopedia you can imagine? Then do it.
- Check references by other people to make sure that they are as complete and accurate as they can be. Books should have their ISBNs listed. Articles should have links to online versions. And so on.
- Add links to primary sources online or other excellent resources online. Often libraries have developed bibliographies, pathfinders, or digital collections that would be an excellent further reading resource for the readers of an article on that topic. But do be aware of the strict guidelines on spam; don't add links willy-nilly, but think: Would this resource be educational and helpful for the worldwide audience of this article who want to find out more on this topic? If so, add a link under the "external links" or "further reading" sections.