

Increasing the Discoverability of Digital Collections Using Wikipedia

The Pitt Experience

Ed Galloway & Cassandra DellaCorte

Ed Galloway is Head, Archives Service Center, University Library System, University of Pittsburgh, edwardg@pitt.edu

Cassandra DellaCorte is an undergraduate history and communications double major, University of Pittsburgh, cad125@pitt.edu

During the Fall 2013 academic semester at the University of Pittsburgh, two undergraduate history majors performed a Wikipedia internship in the University Library System's Archives Service Center and Special Collections Department. The purpose was to enhance the discoverability of Pitt's digital collections and finding aids by creating links from Wikipedia articles to relevant content held by the library's specialized collection units as well as to generally improve the quality of articles by adding additional information. By editing nearly 100 articles in Wikipedia, the interns developed their own effective strategies to perform this work and learned how to use and edit Wikipedia efficiently, how to navigate library resources effectively, how to decide what types of content would be valuable to add, and how to present new and respectable information. As a result, usage of Pitt's online digitized collections and finding aids appears to have increased.

Introduction

It has become almost second nature to turn to the "all-knowing" Wikipedia to look up a quick fact or learn more about practically any given subject. Students have especially come to rely on Wikipedia as the first source of information to get started on a research project (Kim, Sin, & Yoo-Lee, 2014). If this is where the action is, should not archives and special collections be in that stream of information gathering and dissemination? This is not a new concept since archives and special collections have been actively engaged for many years in editing and/or creating entries in Wikipedia in the hope that the links left behind will be followed to their collections for use and study. Actually, it is more than a hope. Several studies have shown that embedding citations and other external links to

primary sources held by a library do in fact increase traffic to those collections (Combs, 2011). So should we not all be *Wikipedians* if that work can result in more usage of our unique materials? In Fall 2013, a small team of archivists and staff in the University of Pittsburgh's University Library System teamed up with two student interns to jump into the Wikipedia stream.

About the Archives Service Center and Special Collections Department

The Archives Service Center (ASC) and Special Collections Department (SC) are separate units that reside within the University Library System (ULS) at the University of Pittsburgh (Pitt). The ASC comprises the University Archives and the Archives of Industrial Society. While the University Archives collects content focused on the history and life of the university, collections within the Archives of Industrial Society primarily document the growth of a heavy industrial society and contain the records of organizations, businesses, societies, institutions, and individuals from Pittsburgh from the mid-19th through the 20th centuries. The Special Collections Department is the general designation for rare books, literary and historic manuscripts, and specialized collections. Its manuscript collections include the Archives of Scientific Philosophy, the Ford E. and Harriet R. Curtis Theatre Collection, as well as manuscripts and collections of prominent authors or historical figures.

Together, the ASC and SC have nearly 1,000 finding aids online that describe the archival and manuscript collections held by the ULS. The library has been engaged in the digitization of its primary source material for over a decade, which has resulted in websites such as Historic Pittsburgh (digital.library.pitt.edu/pittsburgh), Documenting Pitt (documenting.pitt.edu), Darlington Digital Library (darlington.library.pitt.edu), and many others. To date, the library has digitized over 100,000 unique items, which represent 1.6 million images.

Background

One of the goals of the ULS in Fiscal Year 2013 was to enhance the discoverability of its digital collections. As owner of this goal, the head of the ASC (Galloway) employed a number of strategies to address it, such as publishing LibGuides that describe ULS digital collections as well as creating a presence on Facebook for the ASC and, more recently, Tumblr. In the course of reviewing and analyzing usage statistics of ULS digital collections over the last two years, Galloway learned that much of the traffic volume came to ULS sites from Wikipedia. While he had known this with anecdotal evidence, now he had concrete data proving as much. He found that Wikipedia was the sixth largest provider of traffic to ULS digital collection sites (only the ULS's own collection of servers referred more traffic), and for the ULS Audubon website (audubon.pitt.edu), it was the top referrer. Clearly, the ULS should be working to increase this traffic as a method of enhancing the discoverability of its digital collections.

Environmental Scan

Wikipedia, of course, is not the new kid on the block, having been in formal existence since January 2001. It now contains 30 million articles in 287 languages and represents the sixth most trafficked website in the world ("Wikipedia: About," 2014). Adding content to Wikipedia from the Galleries, Libraries, Archives, and Museums (GLAM) community is also not something new. In fact, some GLAMs have invited Wikipedians (i.e., persons who write and edit the pages for Wikipedia) to be a resident with them for a period of time in order to strengthen the GLAM community's presence in Wikipedia. The literature reflects that in 2007 the University of Washington (UW) Libraries were one of the first to purposely edit Wikipedia articles to increase traffic when they added links from Wikipedia to their digital collections (Lally & Dunford, 2007). Back in 2006, Lally noticed an increase in traffic from

Wikipedia to UW sites and wondered if “adding links from Wikipedia into related areas of our digital collections might be an effective method to integrate our digital collections more fully into the information ‘flow’ of researchers” (Lally, 2009, Implementation of Wikis, para. 1). The UW librarians began with an analysis of subject strengths found in their digital collections and then searched Wikipedia for relevant articles, in which they subsequently added links to their content. As newcomers, the librarians did not register an account with Wikipedia, so all of their edits were done anonymously. As a result of their work, they reported a steady increase in traffic from Wikipedia to their digital collections over a year-long period. Students purportedly edited the entries but the article supplied no information about who they were, how many were involved, or what they learned (Lally, 2009).

In 2008, librarians at Wake Forest University described their effort to edit Wikipedia, while arguing that librarians should be involved in Wikipedia since the online encyclopedia is viewed by some as an unreliable source if it can be edited by anyone. They contend that librarians should be involved as content contributors in order to “assist with the creation and maintenance of a more scholarly environment” (Pressley & McCallum, 2008, Why Librarians Should Contribute, para. 3). They began by establishing a generic Wake Forest user account and wrote two new articles, but they quickly ran into trouble when their account was blocked and the new pages they created were no longer available. As they soon learned, Wikipedia editors frown on self-promotion by an institution and will only allow users to create individual identities to edit. Editing as “Wake Forest” threw up red flags about their intentions. Pressley & McCallum’s article provided very helpful tips on abiding by the proper editing rules and understanding Wikipedia culture, including a discussion on how to properly edit.

Belden (2008), a librarian at the University of North Texas (UNT), described her experience with trying to harness the power of social networks by increasing usage of UNT’s online Portal to Texas History by placing external links to site content from relevant Wikipedia articles. She discussed various strategies to attract potential users to their digital collections and how the ease of use and growing consensus regarding the reliability of Wikipedia made it an obvious choice. She indicated that students and faculty are increasingly mining Wikipedia for the purposes of discovering good references and external links for historical topics of interest. UNT staff edited over 700 articles, which soon accounted for 48% of referring traffic from Wikipedia to their digital collections (Belden, 2008, p. 104).

Work done by the Special Collections Research Center at Syracuse University and reported in 2008 by Combs (2011) was especially helpful to the project team as it embarked on this project. Syracuse’s Special Collections Research Center closely resembles the environment at Pitt in terms of institution size, collection size, and user composition. The ULS project team hoped that its results would mirror those observed at Syracuse and lead to a dramatic increase in traffic from Wikipedia to its online finding aids. Likewise, a similar effort was made by the Archives and Special Collections at Ball State University in 2010. The aim was to “increase[e] the visibility of digital assets in [its] Digital Media Repository,” specifically the Hague Sheet Music collection (Szajewski, 2013, Background, para. 1). Ball State’s work also showed a remarkable increase in page views, results the ULS project team hoped to replicate.

In 2012, librarians wrote about how they were trying to utilize Wikipedia to increase the discoverability of the digitized library resources held within the University of Houston (UH) Digital Library (Elder, Westbrook, & Reilly, 2012). Digital Services staff started promoting their digital collections in 2010 via Flickr, Facebook and Twitter, but soon turned their attention to Wikipedia, spurred on by the success of the University of Washington Libraries. Rather than edit articles, however, the staff at the University of Houston were primarily interested in contributing visual images to Wikipedia to accompany already existing articles. With the help of information science interns, Digital Services staff started by matching their digital collections and images published in the UH Digital Library with existing Wikipedia articles. In order to do this, the staff had to first upload images to Wikimedia Commons (an online repository of free-use media files) where the images can then be used to accompany Wikipedia articles. The authors discussed the procedures and workflow to accomplish this, noting several of the problems they encountered,

such as confronting Wikipedia's non-advertising policy, using correct syntax, assigning proper Creative Commons licenses, and learning how to assign proper categorization terms. Elder et al. (2012) also indicated that the help they received from other editors was invaluable. Usage statistics soon indicated that Wikimedia Commons became the number one referring site to the University of Houston Digital Library collections (Elder et al., 2012, p. 44); part of this was the added bonus of having other Wikipedia editors select UH digital images in Wikimedia Commons for other articles. UH Digital Services staff decided to formalize the project and planned to create a set of easy-to-follow instructions and policies as part of their ongoing Wikipedia program.

Initial Steps and Usage Statistics

As Galloway considered the path to take, he started by offering an internship during the Spring 2013 semester to a graduate student in the School of Information Sciences at Pitt with a concentration in reference services. The student also had an interest in and experience with editing Wikipedia and agreed to create a guidebook to help the ULS get started. Up until that point, staff in both units had not purposely edited and/or created Wikipedia articles with an eye towards sending users to ULS digital collections or finding aids, so the hundred or so articles in Wikipedia that contained outbound links or citations to ULS content were not done by ASC or SC staff. Therefore, designing a manual to guide staff through the process was the first step.

At the same time, using Google Analytics, Galloway conducted an examination of the many Wikipedia articles pointing to ULS digitized content so that staff knew which articles contained links and how often they were followed in any given month. This exercise was important in not only discovering what articles currently contained references to ULS collections, but equally important in knowing what was not there. In other words, knowing the collections' strengths, what avenues to the collections might be expected but were not evident? This twofold strategy would enable staff or future students to know which existing articles to edit (i.e., fact check, enhance readability, improve content), and also inform what potential new articles could be created.

Analyzing the usage statistics of the ULS digital collections surprised Galloway in many ways. As of April 2013, Galloway learned that 298 distinct articles in Wikipedia drove traffic to ULS servers. The most referred page from Wikipedia was the John James Audubon article; the one on *The Birds of America* was a close second. Wondering why? The University of Pittsburgh owns one of only 120 known complete sets of Audubon's famous prints of North American birds. Furthermore, the ULS digitized all 435 double elephant folio plates and made them accessible online with a high-resolution magnification tool. As a result, the site itself is the ULS's single most heavily trafficked site. Other Wikipedia articles that drove traffic to ULS digital collections revolve around the steel and railroad industries, such as the Jones & Laughlin Steel Company or Pittsburgh Railways Company.

Galloway learned that traffic from Wikipedia was responsible for a large part of the digital collection use. Out of over 1,600 referrers to content in Historic Pittsburgh, Wikipedia ranked in sixth place with three ULS referring servers, Facebook, and mobile Facebook being ahead. Within the Darlington Digital Collection, Wikipedia was the top referrer. Similarly, the Documenting Pitt website received the fourth highest referring traffic from Wikipedia entries (only ULS referring servers were ahead).

Student Engagement

In determining the next steps to take, Galloway knew from the literature that some university faculty assigned their students to work inside Wikipedia. One study looked at how students responded to the challenge of writing for a Wikipedia audience as opposed to one person, such as their instructor (Roth, Davis, & Carver, 2013). Although a larger audience can certainly be intimidating, students performed better and took the assignments more

seriously when writing for such an audience than when writing only for an instructor. This knowledge gave Galloway encouragement that Pitt students would behave likewise. Another article by Tsipursky (2013) introduced *class-sourcing* as a similar means for engaging students in their writing assignments. In class-sourcing, teachers give assignments in which students make publicly accessible online digital artifacts (e.g., wikis, websites, blogs, videos, etc.) that they create (Tsipursky, 2013, para. 2). Making students' work public (e.g., in Wikipedia) makes a difference in their level of participation and quality of work. In yet another example, a retired faculty member, who initially authored the Wikipedia article on the War of 1812, challenged his readers to "turn undergraduates into editors" to combat the slow decline of Wikipedia editors (Jensen, 2012, p. 9). In fact, the Wikimedia Foundation, the nonprofit that runs Wikipedia, set up its education program to attract university students as new editors.

With this strategy in mind, in late spring 2013, Galloway reached out to the Department of History at Pitt, where the ULS already had a successful track record of offering internships to undergraduate history major students who worked primarily on processing physical collections to prepare them for research use. This time, however, he invited students to apply to an internship dedicated to working with Wikipedia alone. Two students accepted for the Fall 2013 semester were each obligated to work 150 hours. One was placed at the Archives Service Center and the other within the Special Collections Department.

Galloway then formed the Wikipedia project team, which met regularly over the summer in preparation for the work ahead. In doing so, the team made a fortuitous connection to a local Wikipedian. A team member joined Wikipedia's Pittsburgh Group and reached out for advice. This led the team to Sage Ross, who is employed by the Wikimedia Foundation and responsible for the online training modules associated with Wikipedia's Education Program ("Wikipedia: Training," 2013). Ross introduced these resources and sponsored the course the project team established ("Wikipedia: Courses," 2014). The project team became the instructors for the course while the two students were the class. Learning about and using the Education Program became an ideal way of working within Wikipedia because it gave legitimacy to what the project team was doing. The team presumed, and Ross confirmed, that other Wikipedia editors would be less likely to prevent the project team's work from being published since it was part of a Wikipedia course. Also, working within the course environment allowed the students to work in the *sandbox* (an area in Wikipedia where new content is drafted before it is posted as an independent article), another avenue favorable for attaining legitimacy.

The Pitt Experience

As stated earlier, two students were offered internships: Cassandra DellaCorte in the Archives Service Center and Ariel Hammel in Special Collections. DellaCorte is a sophomore history and communication double major and Hammel is a senior history and anthropology double major, both in the Dietrich School of Arts and Sciences at the University of Pittsburgh. Each reported to a member of the respective departments. DellaCorte and Hammel both enrolled in the ULS course in Wikipedia, identifying themselves as Pitt students who were serving as interns within the ULS and would be specifically editing articles about Pittsburgh and Western Pennsylvania history.

While the project team placed a high value on asking the students to add citations and external links in Wikipedia to ULS online finding aids and digitized content, the team also had a few subsidiary goals. One was to meet Wikipedia's mission to improve the quality of the articles by adding content where necessary and substantially editing an article if called for. In order to sustain the work going forward, the team wanted to learn how to do this the right way by becoming good Wikipedia citizens. There are certainly methods to do this, which the students learned along the way. The other goal was to provide the students with an opportunity to write for a global audience. The Wikipedia experience enhanced their research and writing skills and exposed them to the world of primary source collections. They became adept at searching collections, finding aids, and other library resources.

The project team met several times throughout the semester, but essentially the students worked independently and developed their own practices for editing articles. With this autonomy came a high level of responsibility and trust. The supervisors simply could not review each and every edit made by the students, so they were left to make most judgments for themselves. That is not to say the supervisors were not paying attention or were not available for consultation and feedback. Supervisors often met with the students to discuss articles they were working on, if they were having difficulty finding sources, and any other issues. The students were also encouraged to use the course sandbox when working on any given article. Using Wikipedia's course system helped the students work in standard ways that are consistent with Wikipedia's rules and norms and helped them get started productively without as many of the typical missteps of newcomers.

During the course of the semester, DellaCorte edited 70 existing articles in Wikipedia, and Hammel edited 27 articles. This is not to say that Hammel did less work, because her edits differed dramatically in substance and content from DellaCorte's. Although DellaCorte performed mostly smaller edits, adding just a few facts and a source in most cases, Hammel created new articles and did mostly major overhauls of entries. Nonetheless, the students' goals were twofold: to increase the web presence of ULS materials and to improve the content of Wikipedia. In order to do this, they needed to learn how to

1. Use and edit Wikipedia efficiently.
2. Navigate ULS sources effectively.
3. Decide what types of content would be valuable.
4. Present new, valued information.

In a Wikipedia entry, the goal is to clearly present as much necessary information as possible. Although some topics have articles composed of thousands of words, others are completely covered with just a brief explanation. Unlike a class paper, a Wikipedia entry has no thesis, few transition phrases, and no conclusion. In order to write a successful paper, a student needs to apply only the expectations of their instructor and the conventions of the language, whereas a successful Wikipedia article might be much more elusive.

Wikipedia operates on consensus, just like language conventions, albeit Wikipedia conventions are more open to alterations than the accepted rules of English, for example. However, this consensus applies to all content added. While a student writing a paper need only please a teaching assistant or instructor, a Wikipedia editor must pacify every other editor interested in that entry. Another editor can object to any aspect of new content, from the facts and citations to formatting and word choice. On Wikipedia, a community of thousands figuratively grades the work of a student contributor. It may seem very difficult to add content to Wikipedia without someone else finding it objectionable and removing it, but the closer one adheres to the rules of Wikipedia, the more likely the content will stay. Hammel did run into some difficulty with a particularly aggressive editor but worked through the issues and concerns by using the Talk page (a page on every article where contributors can exchange comments with one another and try to reach consensus before edits are made to the specific article).

Moreover, there are myriad rules to follow, but many of them are more guidelines than laws, and even more are very specific to certain cases. The main rules that applied directly to the ULS Wikipedia project were the policies concerning notability and original research. Students writing content for Wikipedia cannot add their own analysis without violating Wikipedia's No Original Research policy, summarized as follows:

Wikipedia does not publish original thought: all material in Wikipedia must be attributable to a reliable, published source. Articles may not contain any new analysis or synthesis of published material that serves to advance a position not clearly advanced by the sources themselves ("Wikipedia: No Original Research," 2014).

This policy was particularly hard for the students to follow because they were sitting in a building filled with primary sources. Anything produced by the subject of the article (e.g., diaries, personal letters, pamphlets produced by the organization, etc.) cannot be used as a source in an article, so the students had to rely entirely on finding aids or related periodicals. For example, DellaCorte might have known that a person got married from reading a wedding announcement or consulting personal papers, but she had to find a newspaper article or other secondary source with that information in order to post it on Wikipedia. The notability policy states that topics should be of some importance to merit space on Wikipedia. The measure for this requirement is the number of reliable, third-party sources on the subject ("Wikipedia: Notability," 2013). These two policies create an interesting puzzle for a history student attempting to increase the visibility of a body of primary sources. The workaround was to find University of Pittsburgh sources that cover ULS collection topics; however, this was easier said than done and involved a lot of research.

In DellaCorte's case, the first thing she learned about this necessary research process was how little she actually knew. Although she had taken classes that required secondary and tertiary source research, she generally turned to omnibus collections like EBSCO and even general Google searching to find sources, not knowing how many appropriate primary sources are readily available through the university. In addition to the ULS finding aids, DellaCorte learned to use other ULS materials, including Historic Pittsburgh and the University of Pittsburgh Press Digital Editions website (digital.library.pitt.edu/p/pittpress).

Although the finding aids essentially contain an overview of collections of primary sources, they are published by the University of Pittsburgh, making them reputable according to Wikipedia standards. Generally, the historical/biographical sketch section of a finding aid contains information suitable for Wikipedia. For example, an archivist may record the founding members of a corporation as written on the original charter document and then include that information in the finding aid. Then, although DellaCorte did no original research, she could cite that finding aid when she added those names to a Wikipedia entry, and a reader can then find that primary source document via Wikipedia. In the same way, the citations she made lead to digitized documents on Historic Pittsburgh, for example, or references to digitized University of Pittsburgh Press books. By adding information that led back to these online sources, DellaCorte created opportunities for driving more web traffic to ULS content.

Links have to be valid, valuable, and followed in order to make a difference in terms of views per ULS page. Before adding content and the relevant citation, the students had to make sure the subject met the notability requirement and that the source was germane to the entry as a whole. If not, the link or even the whole article could be removed by another user under the auspices of Wikipedia's policies. Due to this possibility, DellaCorte and Hammel had to add only unobjectionable material or risk having their work undone.

The students found that the most impactful way to add meaningful information was to find a popular topic that is poorly covered on Wikipedia. The more important—or "notable" in Wikipedia nomenclature—the subject, the more people read it. More readers mean more potential link-followers. As an example, a major topic DellaCorte worked on was the *Pittsburgh Courier*, which in 1938 was the highest circulating African-American publication in the United States. When DellaCorte began working on the preexisting article, the body of the page contained 236 words. This text had three references and was based on five sources, two of which had only dates (incorporation, reorganization) as relevant information. In two weeks, DellaCorte added over 1,000 words with 38 references, three subsections, and at least nine sources. These sources were two Archives Service Center finding aids and various digitized texts from the University of Pittsburgh Press and Historic Pittsburgh websites. All of this work has remained, and DellaCorte was thanked for her contributions by another Wikipedia editor through the official mechanism on the site.

The most current information DellaCorte added to the *Pittsburgh Courier* article was six years old so not exactly new information. Still the content she compiled from ULS sources was new to Wikipedia, since no one had

taken the time to add it. Although only a few other edits to articles were as major as those concerning the *Pittsburgh Courier*, DellaCorte made an effort to add only content that could help reader understanding. Thus, in the 200 or more edits DellaCorte made over the course of a semester, only one was reverted by another user.

As mentioned earlier, Hammel's edits focused mainly on a few entries in which more extensive work needed to be done. She worked on two pages with high local notability: the Pittsburgh Playhouse and the University of Pittsburgh Stages. Both of these cultural institutions have been in existence for over 100 years. Theatrical organizations that performed here (along with the events at the Carnegie Museums, the Carnegie Music Hall, and the sporting events at Pitt Stadium and Forbes Field) helped to cement Pittsburgh's Oakland neighborhood as a destination separate from the sporting and cultural events of downtown Pittsburgh. Hammel changed the University of Pittsburgh Stages Wikipedia article from one long paragraph into a structured, detailed entry. Hammel benefited not only by her close proximity to the collections of the organizations themselves. She also had access to the archivists who processed and wrote the finding aids she was using as a source. The staff members were able to help guide her to the sections of the finding aids that would include the most relevant information.

Editing Strategies

When asked about her different editing approaches, DellaCorte categorized her editing into three types: general editing, subject editing, and source editing. Wikipedia's Starting Editing pages ("Wikipedia: Starting Editing," 2011) outline the many ways to edit a page. However, the question is not only how to begin, but also where.

Wikipedia encourages new editors to begin with general editing. This type of editing includes everything from checking the grammar and spelling of existing articles to fixing formatting and rewording awkward passages. General editing is great for starting out, as it allows a beginning editor to get familiar with Wikipedia without making major changes. Moreover, assisting with work such as this is unlikely to ruffle the feathers of any other editors, since no content is being added or removed. While this type of work is good practice, it does not necessarily do much for an institution looking to increase web traffic.

Subject editing means adding content related to one subject. An editor would browse both Wikipedia and the holdings of their institution and find places where materials could be applied. Picking a subject to edit is necessary to limit the possibilities, because among the vast number of Wikipedia articles in need of improvement and the institution's materials, an editor would quickly be overwhelmed. Subject editing is both practical and realistic for archives, because archives generally hold information on local history. At the ASC, DellaCorte did a great deal of Pittsburgh-related subject editing. Although she added information on subjects as varied as the Colonial era, buildings, and labor rights movements, all of these topics were, in one way or another, Pittsburgh related. She was able to minimize the time spent choosing new projects by browsing Wikipedia entries related to Pittsburgh and then searching the ULS finding aids for relevant collections.

Source editing is altering Wikipedia articles with references to one source or body of sources. For example, a collection of personal papers could be used to create or improve a biographical Wikipedia article. This is useful if a student editor knows of sources held by the institution that are poorly covered on Wikipedia. Since source editing can be limited to just one document, this is the most focused way to begin work. However, no primary sources should be used as a starting point, since the citation of such materials would constitute "original research" on Wikipedia. Ideally, any sources used will already be published and available to researchers.

An important consideration when an institution begins a Wikipedia project is how much autonomy to give students in regard to their editing. Although source editing is perhaps best for increasing web traffic, it limits students to an extreme degree. For example, if an archive holds many collections on local organizations, genealogy, and defunct businesses, odds are good that few of these topics are notable enough for Wikipedia. Furthermore, the

institution might not hold enough information on a subject to make a quality article based on a particular source, necessitating outside research or a minimal Wikipedia article. A student can probably perform source editing if they are only augmenting existing pages, but when creating new articles, they are more likely to hit a roadblock. When DellaCorte first began at the ASC, she started with source editing, working on creating a new biographical article based on one collection. However, the collection was missing crucial information: how, when, and where the subject of the collection died. Even with supplemental research from non-Pitt sources, DellaCorte was unable to find this information, necessary for even a basic Wikipedia entry. Thus, she spent many hours on work that yielded no results on Wikipedia. This is the danger of source editing, so supervisors and instructors should be cautioned when assigning work on a particular source.

For her work at the ASC, DellaCorte edited Wikipedia entries using a combination of subject and source editing with relative autonomy. She alternated between choosing finding aids of major collections and other Pitt publications to improve Pittsburgh Wikipedia articles. She did diverse work, from general editing, such as creating an *infobox* (Figure 1), to a complete overhaul of the *Pittsburgh Courier* article.

Hammel followed a similar strategy, taking advantage of the Special Collections Department's comparatively small number of manuscript collections to establish a triage system with four different levels. After putting together a list of the material in the department and searching for Wikipedia articles relevant to the collections, she assigned each a status level. These levels ranged from articles with solid Wikipedia entries that needed little to no work, to those that needed general editing, to those articles that needed significant attention, to those that had no related Wikipedia entry. As a result of Hammel's work, the Special Collections Department can confirm that every collection in its possession that already had a related Wikipedia article now has a citation reference to the corresponding finding aid. In her quest to get familiar with the process, Hammel began with general editing before tackling some of the more difficult subject editing. Near the end of her experience, she undertook a source editing project, using the finding aid for the Hervey Allen Papers to significantly add to the Wikipedia entry for his grandfather, Edward Jay Allen, whose manuscripts describe his trip along the Oregon Trail. She also compiled a great deal of information from various sources, including the Curtis Theatre Collection, to create a Wikipedia page list of over 100 Pittsburgh performing arts companies and venues, past and present. As a result of DellaCorte's and Hammel's work, the ULS has a better idea of the different types of Wikipedia editing and will be able to see which types yield the most significant results in terms of web traffic.



Figure 1

Infobox for the Pittsburgh Courier. Infoboxes are templates used in Wikipedia to improve the appearance of articles and to ensure consistency of presentation for articles on similar topics ("Wikipedia: Infobox," 2014).

Student Evaluation

As part of their course requirements, each student maintained a weekly journal in which they noted and reflected upon their progress, challenges, and what they learned throughout the semester. At the end of the semester, they submitted their journals to their respective faculty adviser along with a short paper summarizing their experience. These reports were necessary in order for each student to receive a pass/fail grade. The students were also required to meet several times during the semester with their faculty adviser. The site supervisors in turn submitted a course evaluation supplied by the Department of History, in which they indicated the students' performance on specific assignments, their motivation, initiative, reliability, flexibility, willingness to learn, communication skills, and attendance. The students were not evaluated on their writing ability or the amount of content contributed to Wikipedia; rather, this was considered a learning partnership so that each party could better understand the effort required to be engaged in such an internship.

The culmination of the internship was a joint presentation to ULS library staff by selected project members in November 2013. Galloway, along with DellaCorte, Hammel, and Ross, offered comments on the experience to date, highlighting specific edits and lessons learned. This provided an unusual opportunity for undergraduate students to address a room full of professional librarians, invited faculty, and library administrators. This event honed their presentation skills and afforded an opportunity for public speaking that undergraduates rarely experience.

On reflection, the Department of History saw this as a particularly welcome opportunity that simultaneously familiarized their students with archival collections and practice as well as aspects of digital humanities. As one adviser said in an e-mail,

It helped [the students] to understand the importance of bridging the divide (that still exists in too many cases) between scholarly (archival) and popular (Wikipedia) interpretations of history. Learning to translate and articulate information about the collections and their significance was not only a valuable public service, but it also helped them acquire skills which will be useful as they transition out of the university into professional fields (J. Stoner, personal communication, 2014 January 7).

This adviser thought the students succeeded in all aspects of the internship and would agree to offer it again.

Outcome

The actual impact of this work on the ULS remains to be seen, although preliminary usage statistics seem to suggest an increase in use of the ULS's digital collections and finding aids throughout the Fall 2013 semester and into the Spring 2014 semester. This increase could be due in large part to the activity of the students performing the work. After all, they followed the links they embedded in Wikipedia pages to make sure the links worked properly. However, as only a short time has passed, Galloway knows from reviewing the usage logs that it cannot be just the students who are following these links from Wikipedia into the ULS's digital collections. The students typically worked on an article and then moved on to the next. Taking a look at usage statistics from May 2012 through February 2014, focusing especially on referrals from Wikipedia, Figure 2 reflects referrals to content in the Historic Pittsburgh site from Wikipedia, such as digitized books, maps, and photographs. During the Fall 2013 semester, there was a significant increase in visits to Historic Pittsburgh from Wikipedia. There was also an increase in the number of unique Wikipedia articles that referred traffic to Historic Pittsburgh. In fact, the number almost doubled to a high of 83 articles, and that is in January 2014, after the internship ended. This suggests that the team's effort is being rewarded since the data indicate an increase in the number of users. As already mentioned, the data include page views by the interns themselves, which certainly accounts for some of this activity, yet the data also indicate that multiple users followed links to ULS digital collections from Wikipedia.

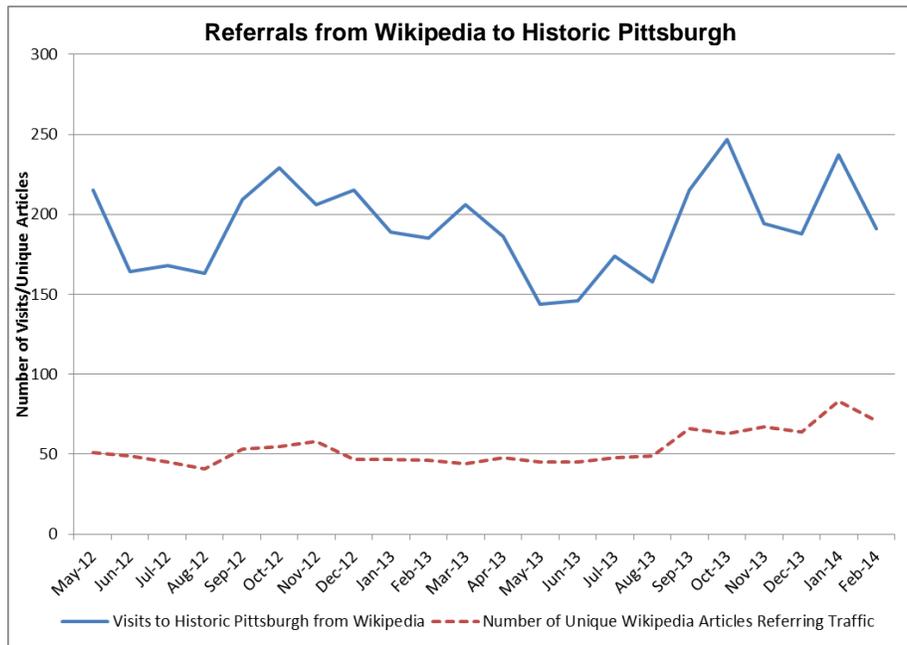


Figure 2
Referrals from Wikipedia to Historic Pittsburgh, May 2012 through February 2014.

This was not the only increase in usage. The majority of work by the interns involved citing and pointing users to ULS finding aids, which describe the archival and manuscript collections in detail and serve as an authorized published source in the eyes of Wikipedia. As Figure 3 indicates, there was a discernible increase in finding aid traffic by Wikipedia visitors, especially during the Fall 2013 semester. Moreover, it can be noted that the number of unique Wikipedia articles from whence this traffic came more than doubled.

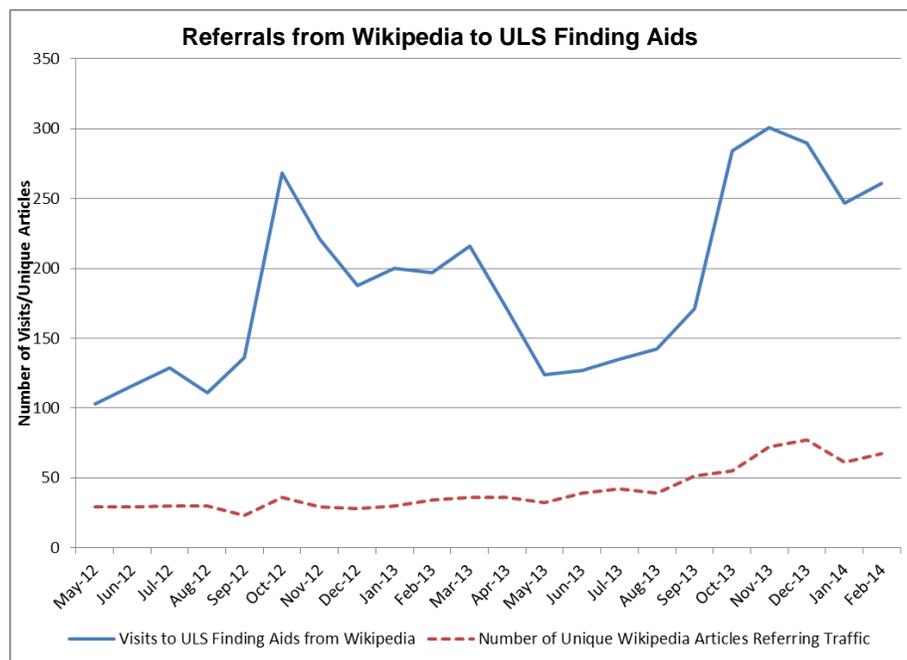


Figure 3
Referrals from Wikipedia to ULS finding aids, May 2012 through February 2014.

It is important to remember that the students did not concentrate on creating new articles; rather, they primarily added to and edited existing articles. So the above increase in traffic from the Wikipedia articles suggests that their edits (and citing sources) dramatically improved the amount of use the finding aids received.

As head of the ASC, Galloway wondered if any of this increase in traffic to the ULS sites would be reflected in e-mail reference statistics. In other words, with more “eyeballs” visiting the online collections and finding aids, are any of these visitors contacting the archives with questions or comments? This cannot be known for certain since each person who contacts the ASC is not asked if they did so because of a Wikipedia article. Nonetheless, as Figure 4 indicates, the Archives Service Center did see a notable increase in e-mail inquiries over the last 24 months in general, and particularly during the Fall 2013 semester. It could very well just be a coincidence, but the data suggests some correlation since it closely mirrors the Wikipedia referral charts.

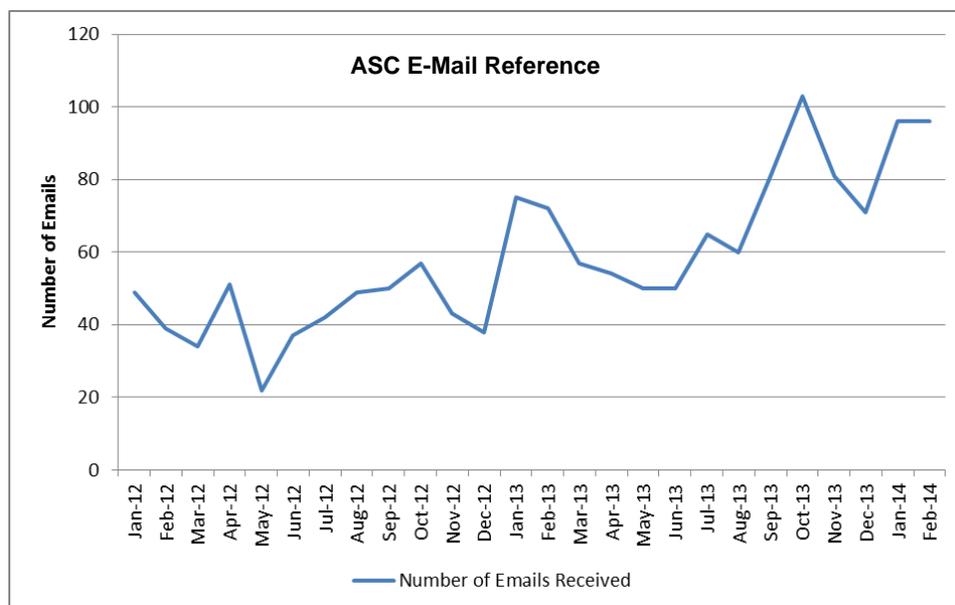


Figure 4
ULS Archives and Special Collections e-mail reference statistics, January 2012 through February 2014.

Future Plans

Now that the ULS’s semester-long experience editing Wikipedia articles is complete, Galloway has plans to keep this initiative going. A new semester has brought a fresh round of students to the Archives Service Center and Special Collections Department, some who have expressed an interest in doing similar work. In addition, DellaCorte will continue to work at the ASC for another semester, concentrating on creating new articles in Wikipedia based on content in the collections. Working with the spreadsheets created by Hammel, Special Collections staff members plan to have future students create stub articles for notable subjects not yet represented in Wikipedia. Staff within the two departments will be trained to learn how to successfully edit articles themselves, so that working in Wikipedia becomes a normal routine when processing collections and writing finding aids. The ULS is excited about this work to increase access to its unique collections.

References

- Belden, D. (2008). Harnessing social networks to connect with audiences: If you build it, will they come 2.0? *Internet Reference Services Quarterly*, 13(1), 99-111. doi:[10.1300/J136v13n01_06](https://doi.org/10.1300/J136v13n01_06)
- Combs, M. (2011). Wikipedia as an access point for manuscript collections. In K. Theimer (Ed.), *A Different kind of web: New connections between archives and our users* (pp. 139-147). Chicago, IL: Society of American Archivists.
- Elder, D., Westbrook, R. N., & Reilly, M. (2012). Wikipedia lover, not a hater: Harnessing Wikipedia to increase the discoverability of library resources. *Journal of Web Librarianship*, 6(1), 32-44. doi:[10.1080/19322909.2012.641808](https://doi.org/10.1080/19322909.2012.641808)
- Jensen, R. (2012). Military history on the electronic frontier: Wikipedia fights the War of 1812. *The Journal of Military History*, 76(4), 523-556.
- Kim, K., Sin, S. J., & Yoo-Lee, E. (2014). Undergraduates' use of social media as information sources. *College & Research Libraries*, 75(4). Retrieved from <http://crl.acrl.org/content/early/2013/02/06/crl13-455.full.pdf.html>
- Lally, A. (2009, May 18). Using Wikipedia to highlight digital collections at the University of Washington. In *The Interactive archivist: Case studies in utilizing web 2.0 to improve the archival experience*. Retrieved from <http://interactivearchivist.archivists.org/case-studies/wikipedia-at-uw>
- Lally, A. M., & Dunford, C. E. (2007). Using Wikipedia to extend digital collections. *D-Lib Magazine*, 13(5/6). doi:[10.1045/may2007-lally](https://doi.org/10.1045/may2007-lally)
- Pressley, L., & McCallum, C. J. (2008). Putting the library in Wikipedia. *Online*, 32(5). Retrieved from http://www.infotoday.com/online/sep08/Pressley_McCallum.shtml
- Roth, A., Davis, R., & Carver, B. (2013). Assigning Wikipedia editing: Triangulation toward understanding university student engagement. *First Monday*, 18(6). doi:[10.5210/fm.v18i6.4340](https://doi.org/10.5210/fm.v18i6.4340)
- Szajewski, M. (2013). Using Wikipedia to enhance the visibility of digitized archival assets. *D-Lib Magazine*, 19(3/4). doi:[10.1045/march2013-szajewski](https://doi.org/10.1045/march2013-szajewski)
- Tsipursky, G. (2013, October 18). Class-sourcing as a teaching strategy. *Inside Higher Ed*. Retrieved from <http://www.insidehighered.com/advice/2013/10/18/class-sourcing-teaching-strategy-essay>
- Wikipedia: About. (2014). In *Wikipedia*. Retrieved February 10, 2014, from <http://en.wikipedia.org/wiki/Wikipedia:About>
- Wikipedia: Courses. (2014). In *Wikipedia*. Retrieved February 10, 2014, from <http://en.wikipedia.org/wiki/Special:Courses>
- Wikipedia: Infobox. (2014). In *Wikipedia*. Retrieved April 27, 2014, from <http://en.wikipedia.org/wiki/Infobox>
- Wikipedia: No original research. (2014). In *Wikipedia*. Retrieved February 10, 2014, from http://en.wikipedia.org/wiki/Wikipedia:No_original_research
- Wikipedia: Notability. (2013). In *Wikipedia*. Retrieved February 10, 2014, from <http://en.wikipedia.org/wiki/Wikipedia:Notability>
- Wikipedia: Starting Editing. (2013). In *Wikipedia*. Retrieved April 28, 2014, from http://en.wikipedia.org/wiki/Help:Starting_editing
- Wikipedia: Training. (2013). In *Wikipedia*. Retrieved February 10, 2014, from <http://en.wikipedia.org/wiki/Wikipedia:Training>