

For the Common Good: The Library of Congress Flickr Pilot Project



October 30, 2008

Michelle Springer, Beth Dulabahn, Phil Michel, Barbara Natanson, David Reser,
David Woodward, and Helena Zinkham

Table of Contents

Executive Summary	iii
Background.....	1
Development of a Pilot	2
Challenges to Launch.....	5
Photo Selection and Preparation	6
Technical Issues	8
Resources	9
Final Phase	12
Outcomes	14
Increasing Awareness	14
Gaining a Better Understanding of Social Tagging and Community Input.....	17
Tags.....	18
Comments	25
Notes	32
Gaining Experience Participating in Web 2.0 Communities	32
Recommendations and Conclusion.....	33
Appendix A: Examples of Flickr Interactions	37
Appendix B: Strategic Goals and the Flickr Project.....	40
Appendix C: Media Coverage	44
Appendix D: Catalog Records for FSA/OWI and Bain.....	46
Appendix E: Traffic to the Library’s Web Site	47

Note: Photo on the cover is “Women Enlisting - England,” Bain News Service, publisher, [between 1910 and 1915]. http://www.flickr.com/photos/library_of_congress/2163930482/

Executive Summary

The Library of Congress, like many cultural heritage organizations, faces a number of challenges as it seeks to increase discovery and use of its collections. A major concern is making historical and special format materials easier to find in order to be useful for educational and other pursuits. At the same time, resources are limited to provide detailed descriptions and historical context for the many thousands of items in research collections. The Library also faces competition for the attention of an online community that has ever-expanding choices of where to pursue its interests.

One solution worth exploring is to participate directly in existing Web 2.0 communities that offer social networking functionality. Reaching out to unknown as well as known audiences can attract more people to comment, share, and interact with libraries. Taking collections to where people are already engaged in community conversations might also encourage visits to a library's Web site where the full wealth of resources are available.

To begin to address these issues, staff in the Office of Strategic Initiatives (OSI) approached the Library Services Prints and Photographs Division (P&P) in early 2007. A small team developed a pilot project that could be rapidly implemented with limited resources. The formal Library of Congress strategic goals to expand outreach and improve the user experience shaped the primary objectives:

- increase awareness by sharing photographs from the Library's collections with people who enjoy images but might not visit the Library's own Web site,
- gain a better understanding of how social tagging and community input could benefit both the Library and users of the collections, and
- gain experience participating in the emergent Web communities that would be interested in the kinds of materials in the Library's collections.

Once the popular photosharing Web site Flickr was selected as a venue that would meet the Library's requirements, the pilot team contacted Flickr to discuss its available rights statement options--none of which was appropriate for the Library's content. These discussions began the collaboration that resulted in the launch of *The Commons* (www.flickr.com/commons), a designated area of Flickr where cultural heritage institutions can share photographs that have no known copyright restrictions to increase awareness of their collections. Flickr members are invited to engage with Commons collections by describing the items through tags or comments. A growing number of libraries, museums,

and archives, intrigued by the possibilities of this model, have followed the Library's lead and launched accounts within the Commons framework.

Two collections of historical photographs were made public on a Library account on the Flickr photosharing site in January 2008. The response from Flickr members and observers of the pilot was overwhelmingly positive and beneficial. The following statistics attest to the popularity and impact of the pilot:

- As of October 23, 2008, there have been 10.4 million views of the photos on Flickr.
- 79% of the 4,615 photos have been made a "favorite" (i.e., are incorporated into personal Flickr collections).
- Over 15,000 Flickr members have chosen to make the Library of Congress a "contact," creating a photostream of Library images on their own accounts.
- For Bain images placed on Flickr, views/downloads rose approximately 60% for the period January-May 2008, compared to the same time period in 2007. Views/downloads of FSA/OWI image files placed on Flickr rose approximately 13%.
- 7,166 comments were left on 2,873 photos by 2,562 unique Flickr accounts.
- 67,176 tags were added by 2,518 unique Flickr accounts.
- 4,548 of the 4,615 photos have at least one community-provided tag.
- Less than 25 instances of user-generated content were removed as inappropriate.
- More than 500 Prints and Photographs Online Catalog (PPOC) records have been enhanced with new information provided by the Flickr Community.
- Average monthly visits to all PPOC Web pages rose 20% over the five month period of January-May 2008, compared to the same period in 2007.

This project significantly increased the reach of Library content and demonstrated the many kinds of creative interactions that are possible when people can access collections within their own Web communities. The contribution of additional information to thousands of photographs was invaluable. Performance measures documented in this report illustrate how the project has been successful in achieving the objectives and desired outcomes of the Library's strategic goals. The Flickr project increases awareness of the Library and its collections; sparks creative interaction with collections; provides LC staff with experience with social tagging and Web 2.0 community input; and provides leadership to cultural heritage and government communities. The specific strategic goals are listed in Appendix B.

Concerns about loss of control over content will continue to be discussed but can also be mitigated. Community practices and forums like the new Flickr Commons, where cultural organizations can now offer collections, help reduce the risks. Pilots like the Flickr project

provide practical experience and concrete data on Web 2.0 risks and rewards, and help staff learn to operate in less formal environments that enhance recognition for a library's valuable cultural roles.

As the Library considers new strategies to make its resources available, discoverable, and useful, pilots of this type are essential for learning how best to engage audiences with library collections in ways that benefit the public at large. The Flickr team recommends that this experiment in Web 2.0 become an ongoing program with expanded involvement in Flickr Commons and other appropriate social networking opportunities for non-photographic collections. The benefits appear to far outweigh the costs and risks.

Key Links for More Information

- The Commons <http://www.flickr.com/commons/>
- Library of Congress Photos on Flickr: Frequently Asked Questions
http://www.loc.gov/rr/print/flickr_pilot_faq.html
- Library of Congress Photos on Flickr (announcing the pilot and linked off the Library's Home page) http://www.loc.gov/rr/print/flickr_pilot.html
- The Library of Congress Profile on Flickr
http://www.flickr.com/people/library_of_congress/
- The Library of Congress Flickr Account
[http://www.flickr.com/photos/library_of_congress.](http://www.flickr.com/photos/library_of_congress)

Background

“The Library of Congress is asking the public for help.” This message exploded on the Web in the week following the launch of the surprisingly popular Library of Congress Flickr pilot project on January 16, 2008. What value did the Library perceive could be gained by joining a social media photosharing site? Like any cultural heritage institution, the Library of Congress is always seeking to broaden the awareness of the resources that it collects, preserves, and makes accessible to the public to inspire, educate, and illuminate. The Library, a pioneer in the digitization of its collections, recognizes the power of the Web to enhance access and expose these resources to the world.

As the Library seeks to reach new audiences and increase discovery and use of our unique collections, we face a number of challenges: a limit to institutional resources to provide detailed descriptions, historical context, and transcriptions of the thousands of items in large collections; a need to make the materials in those collections easily retrievable and accessible; competition for the attention of an online community that has ever expanding choices of where to pursue their interests, and a technical infrastructure that does not easily allow users to comment, share, and interact with content in the manner offered by popular social networking sites.

We entered this pilot with a number of questions about how Web 2.0 environments might help the Library meet its challenges. Could Web users contribute useful information, knowledge, and energy for the Library? Could the Library tap the knowledge and energy of the user community to augment its own efforts? As users become increasingly accustomed to tagging online content for their own purposes, would they be interested in contributing information for community benefit? What’s the quality of the information gained through crowdsourcing? Would the pay-off justify the Library’s investments in such an effort?

Attempting to address these questions led to the launch of the Library of Congress (LC) Flickr Pilot project at http://www.flickr.com/photos/library_of_congress.¹ In the months since launch, this project has allowed the Library to

- encourage discovery of historical material by presenting it in a venue that heretofore

¹ For additional information see: “Library of Congress Photos on Flickr: Frequently Asked Questions,” at http://www.loc.gov/rr/print/flickr_pilot_faq.html ; “Library of Congress Photos on Flickr.” at http://www.loc.gov/rr/print/flickr_pilot.html and the Library’s Flickr profile page at http://www.flickr.com/people/library_of_congress/. The Library of Congress Flickr Account http://www.flickr.com/photos/library_of_congress.

- had been focused on contemporary, personal materials and increase engagement with those materials;
- increase learning and stimulate an educational interest in history;
 - foster the development of personal relationships with these materials as they are shared, identified as favorites, and re-used in new creative ways;
 - stimulate communication, not just between members of the public and the Library, but also with each other, using Library material as a catalyst to spark those conversations;
 - tap into the expertise in the communities of interest on the Web and elicit contributions that provide a richer experience for the users of the collections;
 - collect user-centric, relevant terms that have the potential to increase retrieval of items in the Library's collection;
 - learn by doing; and
 - surface issues related to presentation and engagement that can inform decisions about the materials presented on the Library's own Web site.

The pilot resulted in many positive yet unplanned outcomes: Flickr members' willingness to expend high levels of effort on history detective work; the unprompted sourcing of new information through links to newspaper archives and highly specialized Web sites; the outpouring of appreciation for having started this project and the positive reception for old un-retouched photos; the desire of other institutions to launch similar efforts; the rapid speed of news about the project within the first few days of reaching the blogosphere; and the speed with which new tags and comments continue to be added following our weekly upload of new photos. What follows outlines how we approached the development of the pilot, the outcomes of this experiment, and what we've learned.

Development of a Pilot

In early 2007, discussions began among a small group of staff in the Library's Office of Strategic Initiatives (OSI) about exploring user-generated content in connection with LC's digital resources. Various alternatives were weighed, and photographs were proposed as good candidates for an initial effort. Why photographs as a first step? Photographs have the advantage of being interesting to a wide variety of audiences who can potentially add useful information. They can be appreciated on many different levels regardless of a viewer's native language or expertise and can be quickly absorbed at a basic level without an extended time investment. In addition, the Library has large collections of photographs that could benefit from added description.

OSI approached the Library Services' Prints and Photographs Division (P&P), which was receptive to the idea of a pilot, not only to expand the visibility of the Library of Congress' photo collections and to test user-generated content, but also to increase the public's awareness of the photographic collections held by libraries generally. Soliciting help from volunteer researchers was not a new concept for P&P, which for years has benefited from information provided by volunteer experts and researchers. In 2005, for example, the Pictorial History Committee of the Society for American Baseball Research offered to review all of the baseball-related images within the George Grantham Bain News Service Collection to identify and provide more complete or corrected information about players, teams, stadiums, events, and dates. The information provided by the Committee allowed the Library to update caption information for more than 2,100 photographs in the collection.

Our small team² gathered to discuss strategies for a pilot and explore options. It was soon realized that a pilot to open even a test bed of Library of Congress photographs to public tagging on LC's own site faced high technical and logistical hurdles. Various proprietary and open source solutions were discussed, but it was evident that any in-house solution would not be something we could do quickly, and we wanted to get started. In addition, there was a desire to keep initial expenditures and levels of effort low until analysis and assessments indicated that a more substantial resource investment would be warranted. The Library of Congress is a U.S. federal government agency, and policies and procedures for allowing user-generated content on its Web pages added another layer of complexity.

The 2007 Computers in Libraries conference brought to our attention successful projects by the National Library of Australia and other cultural institutions on the photosharing Web site Flickr (www.flickr.com). We also knew of instances where members of the public, on their own initiative without Library of Congress involvement, had downloaded selected content from the Library's Web site and uploaded it to their personal accounts on various social networking sites, not always with accurate (or any) statements of attribution or permissions status. We were also aware that a handful of federal agencies were experimenting with multi-channel distribution strategies using social media Web sites in addition to offering content

² Initially the team comprised Phil Michel, Barbara Natanson, and Helena Zinkham (Prints and Photographs Division); Beth Dulabahn, Michelle Springer, and Justin Thorp (Office of Strategic Initiatives), David Reser (Cataloging Policy and Support Office), and Erik Delfino (Technology Policy Directorate). David Woodward from Information Technology Services subsequently joined the team to provide programming expertise. Hope O'Keeffe (Office of General Counsel) and Maria Pallante (Copyright Office) provided legal and copyright expertise; Matt Raymond and John Sayers managed public communication strategies; Joe Pagano analyzed LC Web statistics. Colleen Candrl, from the Catholic University of America library school, assisted the team throughout the spring as an OSI intern.

via their own sites.

The seeds were planted to look at established social media sites focused on images, where members were both passionate about photographs and well-versed in tagging. This would expose the Library's collections to a diverse pool of the public and potentially gain an audience interested in interacting with those collections, as well as ensuring that authentic information about the collections was present and accurately sourced, along with links to enable additional discovery. In this context, three objectives supporting the Library's strategic goals to increase Library outreach and improve the user experience, were formalized for a pilot. The relevant strategic goals are listed in Appendix B. The objectives were to:

- increase awareness by sharing photographs from the Library's collections with people who enjoy images but might not visit the Library's own Web site;
- gain a better understanding of how social tagging and community input could benefit both the Library and users of the collections;
- gain experience participating in the emergent Web communities that would be interested in the kinds of materials in the Library's collections.

Aside from the obvious requirements to support tagging and photosharing, the optimal site would be one that offered more than image hosting; it would also provide a connection to an image-focused community. Once the team decided to use an external service, a set of principles guided the project and the selection of venue:

- Content used in the pilot must be already available on the Library's Web site. This offered several advantages, including increasing the likelihood that any unknown permissions issues would have surfaced in the years during which the Library had been offering this content on the Web. For materials already on the Library's Web site, the Library and the public could be confident that the material would remain accessible regardless of the long-term viability or future access policies of the non-Library site.
- Any agreement to place content on a non-LC site must be non-exclusive.
- Access to the Library's content must be free (i.e., without cost).
- Notwithstanding advertisements that might appear on various search results pages, no advertising should appear on the Library's account page in close proximity to the Library's content. An option to control or entirely exclude advertising on the account was preferred.
- The Library should be clearly identified as the source of the content. Pages with Library content should be branded both graphically and through account naming

conventions.

- The site must allow the Library to remove and moderate user-supplied content to ensure that inappropriate material, if found, could be promptly removed (or in a pre-moderation model, not be posted).
- The material uploaded by the Library must be clearly distinguishable from any non-Library generated content.
- It must be possible to accurately convey the copyright status of the material.

Flickr was identified as a venue that would meet our requirements and in which we could observe these principles, save one. Flickr had a vibrant and large photosharing community (gaining in name-recognition and market share in mid-2007); it was built to allow tagging, comments, and notes; the community conversations were focused on photos and photography; and, it offered APIs that allowed for batch loading of photos and extraction of community provided information.³ The single impediment to using Flickr was that the licensing statement options they provided were not suitable to describe the copyright status of the LC material.

Challenges to Launch

None of the available license options offered by Flickr was appropriate for a cultural institution that was not the originator or creator of a work. Like most collecting institutions, the Library of Congress contains valuable archives of photographs for which there are no known copyright restrictions. In contrast to a photographer or other rights holder, however, the Library does not own copyright for its materials and has no authority to waive copyright protection or to transfer or license exclusive rights. Lack of evidence of any rights holder is not the same thing as stating affirmatively that something is in the public domain.

The pilot team contacted managers at Flickr, detailing the goals of the pilot and raising concerns related to the licensing options. These discussions produced the “No known copyright restrictions” licensing statement programmed by Flickr staff to appear on the Library’s account and began the collaboration that resulted in the launch of The Commons.⁴ Developing a new rights statement was

³ For more information on Flickr, see <http://www.flickr.com/help/faq/>.

⁴ The Commons is a Flickr initiative in which cultural heritage institutions share photographs that have no known copyright restrictions as a way to increase awareness of their collections. Flickr members are invited to engage with Commons collections by describing the items through tags or comments. A growing number of libraries and museums, intrigued by the possibilities of this model, have followed the Library’s lead and launched accounts within the Commons framework. For more information on The Commons, see: <http://www.flickr.com/commons/>; more information on the rights statement can be found at <http://www.flickr.com/commons/usage/>.

not an initial project goal. But as we worked through this issue with Flickr staff, we had an opportunity to propose a new model that would allow for any collecting institution to contribute items while stating the rights information that they believed, but could not warrant, to be true.

Photo Selection and Preparation

As possible collections were nominated for consideration, two principles emerged: include a mix of images with varying amounts of metadata to gauge the kind of tags and notes users were interested in contributing; and, try to appeal to a number of audiences by covering a breadth of subjects, locations, and time periods.

Approximately 1,600 color photographs in the Farm Security Administration/Office of War Information (FSA/OWI) collection have proven popular with diverse communities, both national and international. In addition to representing excellent photographic work from the 1930s-40s, the FSA/OWI photos consistently surprise and delight viewers unaware of the existence of high quality color images from that period. The catalog records maintained for these images have fairly complete captions and include topical and geographic subject headings added by Library catalogers in 1984-1985.

Complementing this color collection with a black-and-white collection was an attractive option, and the George Grantham Bain News Service collection, digitized in 2005-2007, was deemed a good match. It featured sports events, theater, celebrities, crime, strikes, disasters, political activities, women's suffrage, and other newsworthy events from the early 1900s. Most of the images have very little subject information beyond abbreviated key words in the titles and are good candidates for additional research and description. To balance the FSA/OWI collection, a sequential run of approximately 1,500 photos was selected for inclusion from the approximately 40,000 glass negative images that have been digitized. We released the two sets with 3,000 images all at once. We recognized that uploading such a large number of photos at once would be a departure from typical photosharing behavior, but we wanted to show the diversity of the collection content, possibly encourage repeat viewing, and provide a range of choices for our potential crowd of participants.

Once the target collections were identified, we had to decide how to present them. Although we could have organized them into various multiple sets by subject at least within the FSA/OWI collection (e.g., "women in wartime," "military aircraft," and "farming"), there were compelling arguments against that approach. We didn't want to pre-determine how the two collections would be experienced (we were asking the community to describe that via tagging, after all), and as noted, the Bain collection had little subject information so the level of effort would have been prohibitive. We

decided to leave the original order of the collections in tact with two exceptions. As the large set sizes meant that the photos at the beginning and the end of the photostream for each set would be viewed the most, we pre-selected 20 interesting images to “book-end” the color set. In addition, the few news photos that show graphic violence or nudity were not uploaded to Flickr to ensure that we complied with the Flickr Community guidelines for its “safe” search filter, (recommending “Content suitable for a global, public audience” and children).⁵

Because the Library could control the elements of the records accompanying the photographs, this project provided a place to experiment and modify the display and sequence of catalog record elements. We also decided to add links to Library Web pages, because these records would display independently of and could be exported far beyond the context of the LC Web site.

Ultimately, modifications to MARC records included adding notes about the sources of the original (historical) titles and adding lot numbers to an item-group linking field (MARC field 773) to allow for parsing the FSA/OWI set into smaller batches of related images as needed. Three links to Library Web pages were embedded into each of the FSA/OWI and Bain records:

- the URL (“handle”) of the Prints and Photographs Online Catalog (PPOC) highest resolution TIFF file, identified as “persistent URL” (rather than “digital ID”),
- the URL (“handle”) of the Prints and Photographs Reading Room Home page, added to the repository information (MARC field 852), and
- the URL (“handle”) of a page with general information about the FSA/OWI or Bain collections to connect viewers to contextual information about each image (MARC field 545). It was initially added just to Flickr project records, but is now being added to all PPOC records.

The Minaret software used by P&P to manage these two collections is flexible and could accommodate these global updates without a high level of effort. On Flickr, selected elements of the PPOC MARC records were loaded as descriptions with user-friendly labels. We could have turned the subject terms in the MARC records into Flickr tags, but the idea was discarded in an effort to reduce any unintended influence on the tagging suggestions of the Flickr community.

⁵ <http://www.flickr.com/help/filters/#258>

Technical Issues

It was determined early in the technical investigation stage that purchase of a Flickr Pro Account (currently \$24.95 US annually) would expand the options we would have on the account, such as loading an unlimited number of sets and collections with unlimited bandwidth, removing limitations on the number of photos viewable in our photostream, and allow uploading photos up to 20MB each. Finally, as a Flickr Pro Account holder, the Library would have access to daily aggregated statistics of the page views, visits, referrers, and community activity on the account.

Flickr provides their account holders with a step-by-step process to manually upload photos and descriptive information. Several requirements of the pilot made manual upload methods inadequate. We needed:

- an automated method to derive Flickr display titles and descriptions from existing MARC records using prescribed mapping
- to use existing image files that were already associated with the descriptive information
- a repeatable process that would allow ongoing uploads of additional photo batches to the appropriate set and accommodate changes (such as the need to overwrite titles and descriptions for photos already in our Flickr account after updating the source records with new information from Flickr members)
- to associate the photos and activity in our Flickr account with the corresponding source photos in the Library's collections in order to support subsequent updates, and
- to be able to do all this with batches of up to 1,500 photographs.

A customized upload application was required. Fortunately, Flickr allows access to its Web services through a public application programming interface (API) in a variety of programming languages. Additionally, third party developers have created software development kits that ease the use of the API. Using the flickrj toolkit for java and marc4j (a MARC record toolkit), an Information Technology Services (ITS) staff member developed a specialized application to upload photos and implement the MARC-to-title-and-description mapping.

Using identifying information from the catalog record and Flickr's machine tags, the upload application was able to store the association between each Flickr photo and its source photo. Machine tags use a special syntax to define extra information about a tag and take this form: <namespace>:<predicate>=<value>. One machine tag was added containing an identifier for

the photo (based on its digital identifier/handle), and a second machine tag put the identifier in the context of the Dublin Core namespace (defining the identifier type used to express the first machine tag). Here is an example of a machine tag combination:

```
dc:identifier=http://hdl.loc.gov/loc.pnp/ggbain.11317
xmlns:dc=http://purl.org/dc/elements/1.1/
```

To facilitate moderation and detailed analysis of community activity, the Library developed another set of applications to harvest the comments, tags, and notes generated by the Flickr membership. Using Flickr APIs, this information is downloaded in a format suitable for relational databases and spreadsheets. Other custom applications allow viewing of harvested comments relating LC and Flickr IDs. These technical solutions made it efficient to manage the new information being provided by the Flickr community and to analyze the kinds of tags and comments we received.

Resources

Compared to the impact of the Flickr pilot, Library investments were relatively minor. No staff members were ever assigned to work full time on this project; work was performed in addition to other ongoing priorities, and deadlines were adjusted accordingly. A number of factors affected our resource allocation, and experience at other institutions may vary depending on their circumstances. The technical work in our case was fairly straightforward. For example, the team had access to in-house IT expertise and experience with the image and record files, and a large body of images was already digitized with metadata available to optimize for Web display.

Investigation of technical requirements and how to meet them began during the summer of 2007. Technical requirements were primarily addressed by two team members, one with programming expertise and one with curatorial expertise. Some of this work is discussed in the *Technical Issues* section of this report and includes determining how to link content in Flickr to the Library source content via machine tags, formatting those tags, loading the featured images for the sets, and checking account settings.

Non-recurring Resource Allocation

Approximately 222 hours of technical programming work was performed over a six-month period. Broken into discrete tasks, this included:

- 24 hours investigating Flickr API and helper kits and establishing upload requirements, input format, and display mappings
- 80 hours of designing, coding, and initial testing of upload, delete, and update

- programs
- 60 hours of design, coding, and testing activity reporting programs and databases
 - 8 hours of uploading the initial batch of photos
 - 50 hours of testing, completing the record preparation (this included completing any lingering image-related issues and image cleanup if needed), and ensuring that no images were included in the pilot set that would be inconsistent with Flickr's community guidelines.

Once Flickr set the date by which they would perform the necessary programming to allow a new rights statement to appear on the Library's account, the implementation team began to meet in earnest each week. Approximate staff time for pre-launch preparation included:

- 20 hours of team meetings and discussions (8 staff members @ 20 meetings = 160 total staff hours) on non-technical tasks
- 10 hours of drafting and coding various descriptive materials to appear on Flickr and loc.gov. This included composing the account profile and set descriptions, identifying what were likely to be the frequently asked questions and drafting the answers, designing the page on the Library's site describing the pilot, working with the Office of Communications on the new Library brand mark, which debuted publicly with this project.
- 6 hours of developing procedures to work with expected user-generated content (e.g., moderation of content added to the account, reviewing likely group requests, handling Flickrmail inquiries and responses, etc.).
- 45 hours of communications work including arranging internal briefings and external publicity.

Sporadic dialogs between Flickr and the Library stakeholders took place throughout 2007 to discuss various issues related to the proposed project. The time line for this phase was more a reflection of the challenge to coordinate the busy work and travel schedules of key individuals (both at Flickr and the Library) than of the complexities of the collaboration. This phase, involving approximately 70 hours of staff time, included:

- meetings of project team leads with Library legal and copyright counsel to establish the parameters under which such a project could operate
- conference calls made between Flickr and Library staff enumerating and resolving the copyright statement issues
- Library General Counsel attorneys working with Flickr legal staff on the language for a modified Yahoo Terms of Service Agreement that would reflect the indemnification limitations on federal agencies.

Recurring Resource Allocations

External expenditures consisted only of a \$24.95 Flickr Pro Account, necessary to allow the Library to upload an unlimited number of photos. This will be an annual expense as long as the project continues.

Since January, more than 15 P&P staff members (almost half the staff including reference service, cataloging, and digital library perspectives) have participated in LC's Flickr account to gain experience interacting with Web 2.0 communities. The popularity of the project immediately after launch resulted in an unexpectedly high amount of user activity, strongly impacting the time and personnel needed to moderate the user-generated content. All new user-generated content is monitored frequently for spam and offensive material. A miniscule amount (fewer than 25 instances) of the user-generated content has required removal. This includes the rare occasions when a tag deemed inappropriate by multiple Flickr users was called to our attention in the comments or by Flickrmail.

A greater investment of time is spent to verify and incorporate new information about the photos. In total, these activities include (on average):

- 2 hours a week to check the account activity for spam or obscene material. This includes a check of the master list of all tags, performed once a week. Two team members in the Prints and Photographs Division (P&P) rotate this moderation responsibility
- 15-20 hours a week spent by a P&P staff member (or his backup) on all aspects of community activity: reviewing and verifying provided sources to substantiate new information, determining suitability for addition to PPOC records, responding to questions from the community placed in the comments, and evaluating requests to add pictures to groups
- 10 minutes per photo description to modify the Library of Congress source records in PPOC as a result of information submitted by Flickr viewers

Project coordination activities can take up to 10 hours a week. The 7-member team continues to meet one hour a week as needed to discuss issues, such as new functionality introduced to Commons members and requests for information. Time investment on project management tasks such as compiling statistics (beyond what is provided by Flickr), tracking use of LC Flickr photos in blogs and other Web pages, communicating with Flickr executives, keeping team members and other interested Library staff informed of new Commons developments, preparing briefings, etc., takes about 2 hours a week on average.

Uploading a new batch of 50 Bain photos each Friday requires 45-60 minutes a week for P&P to prepare the photos and records; it takes roughly the same amount of time for those photos and records to be batch uploaded to Flickr by Information Technology Services. Photos from new collections might take more time depending on idiosyncrasies of the collections selected.

Final Phase

As the pilot phase wraps up, exploration of Flickr community interests continues. To commemorate the November 11th armistice ending World War I, the Library is participating in the first synchronized uploading by Commons members of photos on the same subject. The Library will offer a set of 15 photographs with a World War I theme from a collection of more than 4,000 panoramic photos. In addition, we are preparing a new set using the Library's Photochrom Print Collection to try alternating with the news photos. The colored images in this collection look like photographs but are actually ink-based photolithographs that feature travel views of Europe, the Middle East, and North America created in the 1890s-1910s. When color photography was still rare, demand for these color images was high, and they seem likely to appeal to new audiences one hundred years later.

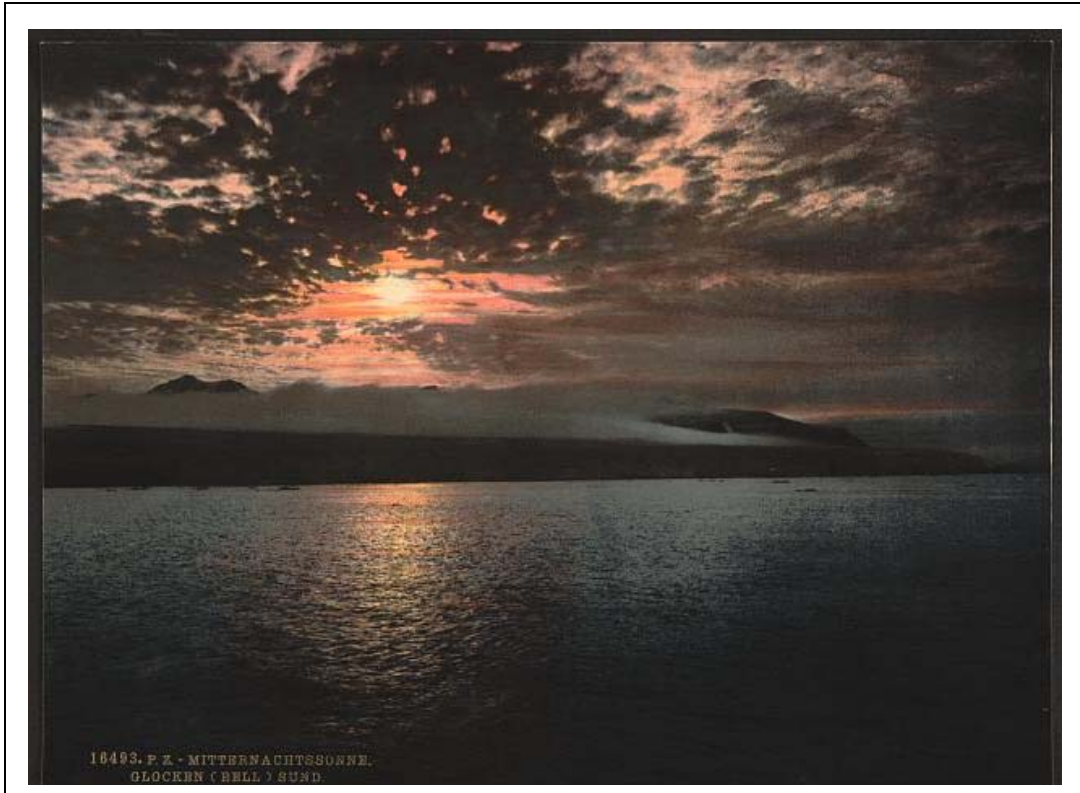


Figure 1. Midnight Sun, Bell Sound, Norway, approximately 1900. Photochrom Print Collection.⁶

⁶ <http://hdl.loc.gov/loc.pnp/ppmsc.06255>

Outcomes

Increasing Awareness

Ironically, some of the email and comments we received after launch noted how pleased people were to see the Library “begin” to digitize their photographs and make them available online. In fact, the Library of Congress Prints and Photographs Online Catalog (PPOC) at <http://www.loc.gov/rr/print/catalog.html> contains more than one million digitized images that have been available on the Library’s Web site for years. Feedback of this nature suggests that as a result of this project the Library is reaching new audiences—people who did not or could not find this material on our own site, and people who never thought to look here.

In the first 24 hours after launch, Flickr reported 1.1 million total views on our account; a little over a week later, the account had received 3.6 million page views and 1.9 million total visits. That included over 2 million views of the photos, and over 1 million views of the photostream. By early October, LC photos were averaging approximately 500,000 views a month and had crossed the 10 million mark in total views and the 6 million mark for visits. Interestingly, 82% of this traffic was referred from within Flickr; only 3% came from search engines. Users continue to offer substantive comments on individual photos, which engages fellow Flickr members and increases interest in our photos. Following the typical Flickr pattern of use, we’ve noticed over time that Mondays regularly show an uptick in the number of views. Daily views for the month of August ranged from 11,119 to 48,585 and averaged 18,263 per day. These spikes can be dramatic, moving from 11,000 views on a Sunday to 41,000 views the next day.

The decision to publicize this pilot solely via the Library and Flickr blogs rather than by the usual method of a press release tested a new model for getting the word out on Library initiatives. The reaction by the blogosphere was astonishing and resulted in thousands of blog posts picking up the story, prompting coverage in the mainstream media: newspapers, magazines, online news services, even television and radio began to cover the pilot (see Appendix C for a bibliography of the coverage). Most posts linked to the Flickr and Library of Congress blogs, which unexpectedly translated into significant visibility for the Library’s blog, in existence for less than a year at the time of launch.

Another unexpected outcome was the relatively high visibility the Flickr account images receive on the major search engines. When relevant keywords are searched, the images

placed on Flickr have greater weight and are returned higher in search engine results than their counterparts in the Library's PPOC. A Google search for the baseball player Germany Schaefer produces the LC Flickr account photo "*Germany Schaefer, Washington AL (baseball)*"⁷ as one of the top 5 results on the first Google page; the PPOC version of the same photo does not appear until the 5th page of search results. The major search engines index Flickr tags, comments, and notes, giving the Flickr versions extra weight in keyword queries.

The surprise and pleasure at the positive reception to the Library's pilot has been well documented (see Appendix C). Why were we so welcomed so warmly? Why did people tag our photos? Was it a type of altruism bred by media sharing sites like Flickr? Was it the fact that we specifically asked for help to tag our photos? Did some Flickr members approach the task as a game? Does releasing public content with no known restrictions create a sense of democratic access or increase the sense of public ownership and shared stewardship for public cultural heritage resources? Certainly the following factors contributed to the unanticipated explosion of interest in the account:

- We appear to have tapped into the Web community's altruistic substratum by asking people for help. Taggers tag for a variety of different reasons, and this diversity is part of what makes Flickr photo collections valuable to a wide membership base. The original Flickr blog post and text announcing the Commons ("This is for the good of humanity, dude!!") struck just the right chord. People wanted to participate and liked being asked to contribute.
- We satisfied a desire for high-quality content without copyright restrictions. Web 2.0 is all about sharing. Providing a rich pool of images that users can easily add to their blogs, download, and re-use in a variety of creative ways satisfied that voracious appetite for unrestricted content.
- The Library was the first institution participating in The Commons, and Flickr's publicity about their new initiative usually included links to the Library's account.
- A venerable institution like the Library of Congress participating (and seemingly conversant) in a popular Web 2.0 space was unexpected and attracted attention.
- The photos are wonderful. We knew from experience that these photos, particularly the color FSA/OWI images, were popular, not just with American but also international audiences.
- The strategy to publicize this project via the blogosphere exponentially increased "word of mouth" traffic. People read the Flickr blog. As the public became aware of

⁷ http://www.flickr.com/photos/library_of_congress/2162646403/

the project, they fueled a grass-roots viral marketing campaign.

In another measure of outreach success, 15,000 Flickr members have chosen to make the Library of Congress a “contact.” Flickr members select other members as “contacts” in order to set up a network of photostreams they’d like to visit often, to have an ‘address book’ of Flickr members that they want to keep in touch with without having to search for or remember the account name, and to easily view new photos loaded by contacts.

We are more than satisfied that we have met our first goal of this pilot, which was to increase awareness of the Library’s photo collections.

The primary Flickr community activity is simply looking at the photos, but Flickr offers additional ways for users to engage. The second most popular community activity has been bookmarking LC’s photos as “favorites,” which allows account holders to select and view Flickr photos they like as a single set on their own account, without having to click back to the original accounts to view them. 79% of the 4,615 photos currently in the Library of Congress account have been incorporated into one or more of these personal collections of favorite images, juxtaposing our historical images with selected contemporary photos in personal photo albums. Similarly, over the course of 8 months the Library has accepted over 200 invitations from Flickr public group administrators. (Flickr members set up groups to gather images from different photo streams for viewing or discussion. Groups can either be public, public (invite only), or completely private.) We’ve allowed images from our account to appear in public groups designated safe for all ages, gathered on diverse themes (like “houses with towers” or “I love grain elevators”). These special subject groups add another way in which Flickr viewers may serendipitously come across our photos. (For a discussion of issues related to groups, see the *Gaining Experience in Participating in Web 2.0 Communities* section below.)

An early decision was made to enable the “blog this” functionality, which has facilitated the inclusion of LC’s photos into personal blogs and Web pages and further increased the visibility of specific pictures. Some meaning can be lost when images are separated from their provenance and collection context. They potentially lose the associations to creators, subjects, and provenance-related images that shape their meaning because the full descriptions rarely travel with the images into new venues. On the other hand, using the “blog this” feature automatically provides a citation and link back to the source photos on Flickr where the fuller description is available. We are able to track usage of this type through various Internet search strategies. Initially photos were used to illustrate posts announcing the discovery of the Library’s account on Flickr. As the novelty of our account

has worn off, we now find these historical photos used for a variety of creative purposes—a photo of two World War II nurses illustrate a blog post on giving blood; a half-built skyscraper announces a trip to New York City; a picture of a 2-story high stack of paper outside a 1940s paper mill adds a humorous take on a judicial order regarding access to data. Creating a new life for these wonderful old photos has been a direct result of providing these pictures in a venue that makes it easy to share and integrate, mix past and present, and we hope, contributes to a better informed public.

Referral Traffic to the Library's Web Site

One of the questions we are frequently asked is to what extent did this project raise or lower traffic to the Library's Web site? Visits to LC Web pages originating from Flickr in the first five months of 2008 rose over 2,000% -- probably an indicator that traffic to the Library's Web site originating from Flickr was relatively low prior to pilot launch. Of greater interest is the data that indicates that in 2008 the Flickr pilot had a positive effect on the traffic to the Prints and Photographs Online Catalog (PPOC), both in terms of the total catalog and specifically for the two collections placed on Flickr. We looked at traffic to the collections in total, as well as traffic to the unique items placed on Flickr. While other factors may have contributed to the rise in the use of the catalog (such as the announcement of a new Abraham Lincoln inauguration photo in January 2008), the Flickr pilot was the most likely variable that could account for these increases.

Full details of the analysis can be found in Appendix E, but here are the highlights:

- Average monthly visits to all PPOC Web pages rose 20% over the five month period for the months of January through May 2008 compared to the same period in 2007.
- Total views/downloads of image files of the color FSA/OWI collection items that were placed on Flickr rose approximately 13% for the period January through May 2008, when compared to the same time period in 2007.
- Total views/downloads of image files of the specific Bain collection items that were placed on Flickr rose approximately 60% for the period January through May 2008, when compared to the same time period in 2007. We think that rate may be higher for Bain photos than for FSA/OWI photos because we are continuing to load Bain photos each week, there are more of them, and more history detective work is leading people to the higher resolution images on the Library's Web site.

Gaining a Better Understanding of Social Tagging and Community Input

Flickr provides three ways for users to react to posted photos: tags, free-form comments, and

notes applied to the pictures themselves. As discussed above, the major search engines index this user-generated content, so they provide extra weighting in keyword queries and their value for discovery extends beyond the Flickr universe. Following our weekly upload of photos, avid contributors add new tags and comments with surprising speed; it's not unusual to see new tags appear within minutes of uploading new images.

Tags

One of the easiest ways for Flickr members to interact with the photographs is to contribute tags. In the Flickr context, tags are “keywords or labels that you add to a photo to make it easier to find later.”⁸ Flickr limits the number of tags that can be added to each photo to 75. The tag characters can be in any alphabet: Roman, Cyrillic, Arabic, Chinese, etc. Tags are typically entered with all lowercase letters, but uppercase letters can be used. Tags can be single words or characters, or can be multi-word phrases such as “telephone wires,” “Lancaster County,” and “john vachon.” No distinctions are made between tags that identify creators, places, time periods, genre, format, subject (of and about), or general associations (reminds me of ...). Flickr tags, like comments and notes, are indexed by the major search engines so their value to increasing discovery and retrieval extends beyond the Flickr Web site.

At the time of original upload, each photo included only 3 tags: one ‘regular’ tag (“Library of Congress”) and two machine tags to correlate the LC and Flickr photographs through identification numbers (see *Technical Issues* above). All other tags on the photos were added by the community. Anyone can see the entire set of tags on our account at http://www.flickr.com/photos/library_of_congress/alltags/. A tag cloud of the 150 most popular tags is available at http://www.flickr.com/photos/library_of_congress/tags/. It's important to note that for the purposes of this pilot, we took a very “hands off” approach to the tags, other than to check for blatantly inappropriate content. On a few occasions, Flickr members commented (or sent us Flickrmail) that they found a tag offensive and asked for its removal. There were exceptionally few tags that fell below a level of civil discourse appropriate to such an online forum—a true credit to the Flickr community. We did not correct spelling or syntax, or remove tags that appeared to be of little or no value to anyone but the tagger. We did delete or spell-correct tags that members had copied from the LC provided description, when additional information from Flickr members revealed that those words had misspellings or other errors. Some tags that would not typically be provided by Library catalogers are obviously more important in a social media site like Flickr, where

⁸“Help/FAQ/Tags” on <http://www.flickr.com/help/tags/>.

photos may be sought out for reuse based on such attributes as predominate colors.

As mentioned previously, the high level of public interest in the account was unexpected. What was especially surprising, however, was that an analysis of the 59,193 tags added through May 13, 2008 revealed that 40% of these tags were added by a small group of 10 “power taggers” (defined as taggers who provided over 3,000 tags each). Some people *really* like to tag; one account holder was responsible for over 5,000 tags. A few statistics provide a sense of the general level of activity. As of October 23, 2008:

- There were 67,176 community added tags (as compared to 10.4 million views);
- 14,472 (21%) were unique⁹ tags;
- 4,548 of the 4,615 photos have at least one community-provided tag;
- 9 photos had reached 75 tags (and people commented that they were disappointed they could not add more tags on those photos)
- 2,518 unique taggers had participated
- More than 500 Flickr members supplied only one tag.

Using the Flickr API to download the entire set of tags added to our account, we were able to analyze the types of tags added by the Flickr community through May 13, 2008.¹⁰ Given the amount of data, we decided to draw representative samples from each of the two sets of photographs. Beginning with the start of the photostream, every 10th photo was selected for analysis until the target number of 100 sample records (50 from each set) was reached. Photos earlier in the photostream are more heavily tagged so this provided a mix of heavily and lightly tagged photos. The FSA/OWI and Bain sets were analyzed separately to determine whether the types of tags added to the sets were different in any way, and to understand if the difference in the fullness of the LC-supplied descriptions had any impact on tagging behavior. (The MARC catalog records for the FSA/OWI photos included significantly more subject/genre terms than those for the Bain photos. See Appendix D for a comparison). The analysis categories assigned to the tags are:

I. LC description-based (words copied from the Library-provided record): e.g., titles, names, subjects, etc.

II. New descriptive words (words not present in the Library-provided description):

- **Place:** e.g., cities, counties, countries, natural feature names
- **Format** (physical characteristics of the original photos). *Sample tags: LF, large format, black and white, bw, transparencies, glass plate*
- **Photographic technique.** *Sample tags: shallow depth of field*

⁹ The unique tag count treats as a single tag terms that vary only in spacing between characters or use of uppercase and lowercase letters.

¹⁰ Analysis performed by Colleen Candrl.

- **Time period.** *Sample tags: wartime, WWII, 1912*
- **Creator name:** e.g., photographer's name

III. New subject words (words not present in the Library-provided description):

- **Image** (items seen in the image itself). *Sample tags: cables, trees, apples, windows, hat, yellow*
- **Associations/symbolism** (phrases and slogans evoked by the image). *Sample tags: Rosie the riveter, Norman Rockwell, We can do it!*
- **Commentary** (revealing the tagger's value judgments). *Sample tags: Sunday best, proud, dapper, vintage.*
- **Transcription** (transcribing words found in items such as signs, posters, etc., within the photo)
- **Topic** (terms that convey the topic of the photo). *Sample tags: architecture, navy, baseball, story*
- **Humor** (tags intended to be humorous rather than descriptive) *Sample tags: UFO, flying saucer*

IV. Emotional/aesthetic responses: (personal reactions of the tagger). *Sample tags: wow, pretty, ugly, controversial*

V. Personal knowledge/research (tags that could only have been added based on knowledge or research by the tagger, and that could not have been gleaned solely from the description provided or examination of the photo): For example, the tag *murder* used on a portrait of someone who was later murdered or tags added for the specific county when that information was not part of the description.

VI. Machine tags (added by the community not Library-supplied): e.g., geotags and Iconclass tags

VII. Variant forms (representing terms already tagged but in a different form, such as synonyms (e.g., *WW2, WWII, World War II, worldwarii*) or plural/singular differences (e.g., *transparency/transparencies*))

VIII. Foreign language (tags in foreign languages/scripts, whether they are translations of English-language tags, or new tags)

IX. Miscellaneous (tags that are not readily understood, that provide corrections to LC descriptions or to other taggers (e.g., *notpeaches*), or tags later removed)

Assigning tags to categories was challenging, and some tags did not fall into a single category. We decided up to three type categories for each tag. For example, when the primary tag type was "LC description-based," a secondary type was also selected so that conclusions could be drawn about the kinds of "LC description-based" words that Flickr members selected to place in the tags. One category was always considered secondary-- "Variants" by definition were always a different representation of information already found in another tag type category.

FSA/OWI Sample Analysis

Total tags in sample: 991; Tags per photo: 19.8

Two category types covered almost 80% of the tags supplied by the Flickr community in the FSA/OWI sample set: tags for what was seen in the images and tags repeated from the LC description. Although some of the additional categories beyond these two are well-represented in the secondary tag types, a large percentage of those came from Library descriptions as well (for example, all 70 of the secondary tags for creator names were from LC descriptions, 88% of the place names were from LC descriptions, 79% of image tags were from LC descriptions, and 98% of all time period tags were from LC descriptions).

Table 1

FSA/OWI Tag Types	Primary Type	Secondary Type	Additional Type
LC description-based	450 (45%)	0	0
Image	336 (34%)	28	16
Commentary	61 (6%)	9	3
Personal knowledge/research	40 (4%)	0	2
Topic	26 (3%)	26	9
Transcription	15 (1%)	2	4
Place	14 (1%)	161	9
Associations/symbolism	13 (1%)	0	0
Foreign language	8 (<1%)	0	0
Format	7 (<1%)	2	0
Miscellaneous	5 (<1%)	3	0
Emotional/aesthetic response	4 (<1%)	7	1
Time period	3 (<1%)	46	0
Photographic technique	2 (<1%)	1	0
Humor	1 (<1%)	0	0
Creator name	0 (0%)	70	0
Variant format	0 (0%)	69	113
Total tags in sample	991	424	157

Bain Sample Analysis

Total tags in sample: 1163; Tags per photo: 23.3

As with FSA/OWI, the two top categories were for items seen in the image and repetition of LC description information. However, several other categories are represented in larger numbers than in the FSA/OWI sample. The “image” category overtaking “LC description-based” can be explained by the fact that the Bain descriptions from Library catalog records lacked most of the detail found in the FSA/OWI descriptions (i.e., there was less data to repeat). The fact that there were more tags per photo and the presence of additional tag types might reflect Flickr tagger efforts to recognize the general lack of information and “step up to the plate” to add information that the Library itself had not provided. The larger number of foreign language tags (in several non-Latin scripts) reflects the international diversity found in the Bain photos. Like FSA/OWI, a large percentage of the secondary tag types were also repetitions of LC description information (e.g., 100% of the “Time period” secondary tags and 51% of the “Place” tags.)

Table 2

Bain Tag Types	Primary Type	Secondary Type	Additional Type
Image	292 (25%)	11	0
LC description-based	271 (23%)	0	0
Place	151 (13%)	29	0
Commentary	117 (10%)	1	0
Foreign language	87 (7%)	4	0
Topic	62 (5%)	12	1
Personal knowledge/research	50 (4%)	2	0
Associations/symbolism	29 (2%)	3	0
Time period	24 (2%)	53	0
Format	22 (2%)	6	0
Miscellaneous	11 (<1%)	1	1
Transcription	14 (<1%)	2	0
Photographic technique	6 (<1%)	0	1
Emotional/aesthetic response	4 (<1%)	2	0
Variant format	0 (0%)	147	14
Creator name	0 (0%)	2	0
Humor	0 (0%)	0	0
Total tags in sample	1163	275	17

A number of other observations can be made either from the sample analysis or the gut reactions of team members who examined the tags:

- The issues commonly cited in comparisons of social tagging vs. assignment of controlled vocabulary terms are represented in the LC photos on Flickr. For example, we found some typographical errors (albeit in impressively low numbers), intended word mergers to overcome system syntax requirements (real or perceived), unintended de-linking of multi-word phrases and terms, overly broad terms that may not help to home in on a desired subset as the corpus on Flickr continues to grow (e.g., *USA*), inability to account for synonyms and homonyms.
- Variant terms are generally not supplied by the same tagger, with the exception of some “power taggers” (defined in this report as individual taggers who provided more than 3,000 tags), who consciously provided variants (e.g., *WWII*, *WW2*, *WWii*) for concepts lacking a consistent vocabulary.
- Although the LC-supplied descriptions had a heavy influence on Flickr members, it did not appear that taggers mined the community-supplied comments with any regularity. Given the richness of information found in the comments (see the *Comments* section below), this is a little disappointing. It is also clear that some members who provided significant numbers of comments on LC photos did not add tags.
- While many of the tags represent terms that might help Flickr searchers find photographs, a number of tags seem to have value primarily to the original tagger (e.g., *6*, *22*, *7th*, *XX*). There were also some tags that are helpful to a certain degree, but are not likely to be used for searching by other Flickr members (e.g., “*not peaches*,” a tag left in response to a “*peaches*” tag).
- Given the emphasis of “place” in user-supplied tags, “geotagging” (specifying the location at which the image was taken using special tags called geotags) should be a popular activity, but the number of geotags remains significantly lower than 1%. The lack of an easy mechanism for taggers (other than the photo owner) to supply geotags within Flickr probably contributes to this dearth. Flickr has indicated interest in changing the functionality.
- Do the community-supplied tags work to drive searchers/browsers to LC’s photos? The lack of a mechanism to easily correlate searches to community-supplied tags remains a problem when trying to answer this question. We know from blog posts

and emails that some users access the photos directly from the photostream—in fact, they anxiously await the weekly load of new photos. What is not as clear, however, is the degree to which discovery is based on searches of the tags, descriptions, or groups other than for searchers who begin their searches outside of Flickr (through Google referrals, for example).

Future Tag Analysis Interests

A few suggestions for next steps with tags include:

1. Examine the relationship between the tags and the keyword searches that resulted in traffic to our photos. As part of the account statistics reported to each Pro Account holder, data is provided on the referring Web sites that linked to individual photos in our Account. This data also shows what keywords were searched that resulted in views of/links to our photos. Unfortunately, at this time this data is not downloadable through an API, and the manner in which it is provided makes its extraction for analysis prohibitively labor intensive. It's hoped that this data will be downloadable in the future, as a detailed analysis would yield insights on the keywords and tags that are useful in finding our photos.
2. Compare tags used by Flickr members against terms/references found in vocabulary lists used primarily to describe photos at LC like *Thesaurus for Graphic Materials* (TGM) or Library of Congress Subject Headings (LCSH). An example of a non-LCSH/non-TGM tag is "Rosie the Riveter." This tag was applied by users 73 times (LC had assigned "Women—employment" and "World War, 1939-1945"). Incorporating popular concepts or variants into our controlled vocabularies might be a way to derive benefit from this kind of user-generated data.
3. Consider bringing tags into LC's search environment, in some manner. The Powerhouse Museum has begun ingesting Flickr tags into their catalog along with the other tags and is marking them as originating from Flickr. They note on their site that they do not verify the 'accuracy' of these keywords. LC does not have a mechanism for easily incorporating these tags. Given the high proportion of tags based on information already in the LC records, the benefit of harvesting all the tags seems limited. But a few options could be explored:
 - Populate bibliographic records with tags, clearly labeled as uncontrolled index terms (e.g., MARC 653 fields) or notes. Some evaluation and weeding of terms would

probably be necessary. It is likely that we would want to exclude tags in non-Latin scripts (MARC records can contain such scripts, but the need to place such terms in a separate field (MARC 880) with accurate script indicators may prove to be more trouble than it is worth).

- If it proves too difficult to populate bibliographic records, an alternative could be to harvest tags from Flickr and park them on LC Web pages that would be crawled by the Library's search engine (possibly using sitemap technology). Hits on terms might then re-direct the Web site searcher to the appropriate presentation in PPOC. This would be the most complicated approach to implement and should be approached as a pilot.

4. As mentioned previously, we have rarely altered the tags on photos in the Flickr account. Should LC staff remove some of the less useful tags (e.g., those of dubious appropriateness or those that have been debunked by comments)? Should LC staff alter existing tags to assist searchers (e.g., correct typos)? Should LC staff monitoring valuable comments that provide new details assign additional tags based on the comments? Should LC delete some categories of tags on 'full' photos (75 tags assigned) to provide room for more relevant tags? We might also consider doing something with tags that other Flickr members have asked to have removed as inaccurate, such as "Dirigibles" and "Zeppelins" on a photo of a barrage balloon. Because these terms could still be useful as entry vocabulary for non-specialists, it might be useful to qualify misused terms, e.g., "Dirigible (similar)" or "Dirigible (related to)."

Comments

When publicizing the pilot, we noted that many of our old photos came to us with very little description and that additional description would be appreciated. The wealth of interaction and engagement that has taken place within the comments section has resulted in immediate benefits both for the Library and users of the collections. By the end of October 2008, 2,562 unique Flickr accounts had added 7,166 comments on 2,873 photos (and the two set pages). A core group of about 20 "power commenters" returns regularly and provides corrected place names, more precise dates, event names, and fuller names for individuals previously identified only by surname. These Flickr "history detectives" contribute historical information, frequently supporting the information they offer by adding links to the New York Times archive, Wikipedia, and highly specialized Web sites devoted to specific relevant areas of interest such as military aviation, railroads, and sports history. Identification

of these highly specialized Web sites may have lasting benefits as research resources for the future.

It is particularly gratifying to see Flickr members provide all kinds of connections between the past and the present through discussions of personal histories including memories of farming practices, grandparents' lives, women's roles in World War II, and the changing landscape of local neighborhoods. As an example of the memories evoked by these images, see the reminiscences posted about the *Sylvia Sweets Tea Room*,¹¹ especially the detailed and moving account from the restaurant owner's family.

Past and present are juxtaposed as individuals take a photo of the same location and then post their picture in the comments with accompanying information on how the location has changed or remains the same.¹² These 'then and now' debates between members take place over locations; changes in purpose and architecture over time are noted with accompanying links to Mapquest and Google Earth providing exact addresses.

Links and samples of the various types of exchanges occurring in the comments, (as well as notes and tags) are available in Appendix A. Two examples of extended interactions follow. Conversations are not uncommon in which multiple members contribute tags and comments over the course of several months. Flickr members added helpful tags to a photo called *Weavers at Work*¹³ to indicate that the women are blind, that they are making rugs at the New York Association for the Blind, and much more. Another user pointed out a "Byron of New York" photographer's logo which provided a fundamental new piece of contextual meaning. We joined the conversation to encourage future viewers to check out the almost 200 photos in the Byron Collection at the Museum of the City of NY that show related activities at the Blind Association. Four months later, yet another Flickr member called attention to the music device in the right corner asking "Is that a violano-virtuoso?" and adding a link to a Wikipedia illustration as encouragement for others to join in the answer (Figures 2-5).

¹¹ http://www.flickr.com/photos/library_of_congress/2178249475.

¹² See "Lowell, Mass., street (LOC)" at http://www.flickr.com/photos/library_of_congress/2179042198/ and http://www.flickr.com/photos/library_of_congress/2179041336/ as examples of "then and now" interactions.

¹³ http://www.flickr.com/photos/library_of_congress/2163450764/



Figure 2. “Weavers at Work”

Byron (Firm : New York, N.Y.), photographer.

Weavers at work

[between 1910 and 1915]

1 negative : glass ; 5 x 7 in. or smaller.

Notes:

Title from unverified data provided by the Bain News Service on the negatives or caption cards.

Photo shows blind women weaving what appear to be rugs. (Source: Flickr Commons project, 2008)

Forms part of: George Grantham Bain Collection (Library of Congress).

Format: Glass negatives.

Rights Info: No known restrictions on publication.

Repository: Library of Congress, Prints and Photographs Division, Washington, D.C. 20540 USA, hdl.loc.gov/loc.pnp/pp.print

General information about the Bain Collection is available at hdl.loc.gov/loc.pnp/pp.gqbain

Persistent URL: hdl.loc.gov/loc.pnp/gqbain.09146

Call Number: LC-B2- 2191-8

Figure 3. Updated LC Descriptive Information in Flickr for “Weavers at Work”

freezejeans pro says:

This is indeed awesome...like peering back through a time machine.
 Posted 8 months ago. ([permalink](#))

LANCERIKA says:

Beautiful vintage clothing*~*
 Posted 8 months ago. ([permalink](#))

byronphoto pro says:

This photo was taken by my Great Grandfather Percy Byron. The Byron Company NY has a large collection at [mcny.org](#) over 22,000 prints.
 Posted 8 months ago. ([permalink](#))

Sarah61489 says:

I agree...it does look like they are weaving rugs.
 Posted 5 months ago. ([permalink](#))

The Library of Congress pro says:

byronphoto: Thanks for making this good connection between Byron and Bain. We'll add Byron's name as the photographer along with his title for the image.

Here's the direct link to view the Byron Collection at Museum of the City of New York: [museumofnyc.doetech.net/voyager.cfm](#) (Includes about 200 more photos of the New York Blind Association.)
 Posted 4 months ago. ([permalink](#))

krseward says:

Any ideas about the cabinet at the far right?

My odd guess is an automatic music device that runs on a music roll (like a player piano or a [Violano-Virtuoso](#)).

It may not be such an odd guess as the NY Blind Assn. pics in the Byron Collection depicts music classes, dances and dance classes, plays. Maybe if this picture was at the NYBA, such a music machine would be a more likely thing. Maybe.

Anyone know?

Figure 4. Comments on “Weavers at Work”

Tags

































-  Library of Congress
-  weavers
-  women
-  5x7
-  loom
-  work
-  industry
-  workers
-  weaving
-  textiles
-  labor
-  NY
-  Association
-  for
-  Blind
-  byron
-  Shuttle
-  1910-1915
-  George Grantham Bain Collection
-  United States South
-  USA
-  southern
-  history
-  B&W
-  disabled
-  disability
-  The N.Y. Association For The Blind
-  N.Y. Association For The Blind
-  new york
-  floor rugs
-  hair bow
-  Gibson Girl hairstyle

Figure 5. Tags Added to “Weavers at Work”



Figure 6: “North American's P-51 Mustang Fighter is in service with Britain's Royal Air Force, N[orth] A[merican] Aviation, Inc., Inglewood, Calif.”¹⁴

An example of studied expertise in a subject area is revealed a conversation over an aircraft image that took place among several commenters over the course of two months. Figure 6, “North American's P-51 Mustang Fighter is in service with Britain's Royal Air Force, N[orth] A[merican] Aviation, Inc., Inglewood, Calif.”¹⁵ sparked a discussion of whether this is in fact a P-51 or an airacobra aircraft. Flickr member bootload said, “The version you see is probably a P-51 B/C with a Rolls Royce Merlin engine. It could also be a Mk1 (Mustang I/P-51) which went into combat on 12 May '42. The bubble cockpit was introduced later in as a solution to the "poor rearward view. A quick solution was the "bulbous" Malcolm Hood and later replaced with a teardrop canopy so widely associated with later P-51's. Another commenter, antwharrington, replied, “It's a P51A, with the Allison V-1710 engine. The 'bump' on top of the engine is the air intake, Merlin versions had the intake under the nose rather than on top. Allison-powered versions also had a three rather than four-blade prop and a shallower radiator scoop on the belly, but neither of those is easy to spot from this photo.” Two months later, a third member, fjcpictures added, “If memory serves, it was the

¹⁴ http://www.flickr.com/photos/library_of_congress/2179916794/.

¹⁵ http://www.flickr.com/photos/library_of_congress/2179916794/.

Brits who got the idea of using the RR Merlin in the P51 because they felt the Allison wasn't fast enough. One of their best ideas.” Two other members later weighed in on the merits of using the Rolls Royce Merlin engine.¹⁶

After verification by Library staff, the corrections and additions to captions and titles, and the identification of individuals being provided in this manner by the Flickr community is being slowly incorporated into PPOC records and then loaded back to Flickr records to keep them in synch. Changes began to appear the week of February 25th. By August, more than 500 records in PPOC had been enhanced with new information and cited the Flickr Commons project as the source of the information that was changed or added. A photo once simply captioned, “*Reid Funeral*”¹⁷ is now more fully described with the note: “Photo shows the crowd gathered outside of the Cathedral of St. John the Divine during New York City funeral of Whitelaw Reid, American Ambassador to Great Britain. (Source: Flickr Commons project, 2008).” Other records might contain a new note field, such as “Current title devised by Library staff based on information provided by the source: Flickr Commons project, 2008.”

By the end of July the team determined that it would benefit users of the Library’s catalog to alert them to corresponding Library Flickr pages. Though we are steadily enhancing PPOC records with new data over time, new user-generated information may be added to Flickr records at any moment as they are discovered by Flickr members. In addition, the personal experiences and reminiscences provided in the conversations between Flickr members in many cases enrich the appreciation of the photos, and that information best resides in its original context on Flickr. By August we added the appropriate Flickr URL to the “additional version available” field (MARC field 530) in relevant FSA/OWI and Bain PPOC records along with language suggesting that additional information might be available through the Flickr Commons project.

¹⁶ Another example of detailed P51 expertise by other commenters: “*P-51 "Mustang" fighter in flight, Inglewood, Calif. The "Mustang", built by North American Aviation, Incorporated, is the only American-built fighter used by the Royal Air Force of Great Britain.*”

http://www.flickr.com/photos/library_of_congress/2179075547/
¹⁷ http://www.flickr.com/photos/library_of_congress/2515741281/.

Click on picture for larger image, full item, or more versions.



[How to obtain copies of this item](#)

TITLE: Weavers at work

CALL NUMBER: LC-B2- 2191-8[P&P]

REPRODUCTION NUMBER: LC-DIG-[ggbain-09146](#) (digital file from original neg.)

RIGHTS INFORMATION: No known restrictions on publication.

SUMMARY: Photo shows blind women weaving what appear to be rugs. (Source: Flickr Commons project, 2008)

MEDIUM: 1 negative : glass ; 5 x 7 in. or smaller.

CREATED/PUBLISHED: [between 1910 and 1915]

CREATOR:

[Byron \(Firm : New York, N.Y.\) photographer.](#)

RELATED NAMES:

[Bain News Service](#), publisher.

NOTES:

Forms part of: George Grantham Bain Collection (Library of Congress).

Title from unverified data provided by the Bain News Service on the negatives or caption cards.

General information about the Bain Collection is available at <http://hdl.loc.gov/loc.pnp/pp.ggbain>

Additional information about this photograph might be available through the Flickr Commons project at http://www.flickr.com/photos/library_of_congress/2163450764

Temp. note: Batch two loaded.

Temp. note: Batch x.

FORMAT:

[Glass negatives.](#)

OTHER TITLE:

[The N.Y. Association for the Blind.](#)

REPOSITORY: Library of Congress Prints and Photographs Division Washington, D.C. 20540 USA <http://hdl.loc.gov/loc.pnp/pp.print>

DIGITAL ID: ([digital file from original neg.](#)) [ggbain 09146](#) <http://hdl.loc.gov/loc.pnp/ggbain.09146>

CONTROL #: [ggb2004009146](#)

View the [MARC Record](#) for this item.

Figure 7. Updated Library of Congress PPOC Record

Notes (Annotations left directly on an image)

Notes, when used by Flickr members with serious purpose, have proven to be a useful way to focus on specific items within a photo, such as the identity of persons in crowd scenes or words on signs and placards. For example, on “*Parade of Olympic Athletes*”¹⁸ one Flickr member looked at the larger image on the Library’s Web site and then provided notes transcribing the tiny text identifying each athlete on the signs on each automobile. Notes are also a means of adding graffiti-type messages and smart-aleck humor to the images. A proliferation of notes can be visually jarring, obscuring an image with nested squares and rectangles. Fortunately, mousing away from the image makes the notes disappear. Although some useful conversations have taken place between members in the notes, this is rare. More common is using a note to highlight something in the image and ask a question (or make a joke) about what it is. Of the three ways for users to interact with the photos on Flickr, the team considered this method the most problematic. We have received requests to turn off this option.

Gaining Experience Participating in Web 2.0 Communities

The Flickr pilot helped both the team and P&P staff gain experience with Web 2.0 conventions and learn more about the photos through interactions with Flickr members via the comments and Flickrmail. A Web 2.0 persona is expected to be less formal than official face of the institution—we wrote the LC profile in a light tone and strove to sound personal in the responses added to the string of everyone else’s comments. These conversations between LC and Flickr users have differed from a standard one-on-one reference desk exchange in that they can take place over months and occur among multiple speakers. Staff monitoring the account have become more familiar with the rhythm of Web 2.0 exchanges and the value, in many cases, of letting Flickr members converse with each other rather than injecting Library comments or rejoinders immediately. At the same time, we have wanted to encourage repeat visits and so have felt it’s important to provide feedback and convey appreciation when new data is provided (such as thanking commenters for providing corrections and alerting them to our plan to fix the source data). P&P staff members have developed a “cookbook” to aid rotating staff on the types of interactions they are likely to see and options for responding, with encouragement to use their own voices.

¹⁸ http://www.flickr.com/photos/library_of_congress/233283282/

As we've moved forward, we've tried to strike a balance between the participation levels expected of community members and the limitations on resources available to allocate to the pilot. We've proceeded cautiously where accepting invitations that would add links from our account to other individual or group accounts. We have not, for example, followed the typical Flickr practice of reciprocating contacts—i.e., making contacts of everyone who made the Library a contact. We have accepted requests to add our photos to groups when we can confirm that they are public,¹⁹ which is one of the moderation tasks undertaken by P&P staff (See the Resources section). The need to be judicious in adding links from our accounts is one of the ways we differ from a standard individual account. Despite our informal tone on Flickr, we recognize that we are representing an institution in a venue designed for individuals.

Recommendations and Conclusion

Ten months into the pilot, the question looms whether to move from pilot project to program. Performance measures documented in this report illustrate how the project has been successful in achieving the objectives and desired outcomes of the Library's strategic goals. The Flickr project increases awareness of the Library and its collections; sparks creative interaction with collections; provides LC staff with experience with social tagging and Web 2.0 community input; and provides leadership to cultural heritage and government communities. The specific strategic goals are listed in Appendix B.

Participating in Flickr can open many new avenues for making the Library's visual collections useful as well as available. Additional resources would be needed to maximize the opportunities offered by Flickr and other Web 2.0 communities. But the overwhelmingly positive response to the digitized historical photographs in the Library's Flickr account suggests that participation in The Commons should continue. To date, ten institutions²⁰ from six nations are sharing selections from their photo archives and inviting the public to contribute information within the Commons' specially designated Flickr space, and more are expected to follow in the coming months. Reports so far indicate that they are pleased with the additional exposure and the quality of the tagging their collections are receiving.²¹ Flickr

¹⁹ See <http://www.flickr.com/help/groups/> for more information on groups.

²⁰ Commons members include: The Powerhouse Museum, The Brooklyn Museum, The Smithsonian Institution, The Bibliothèque de Toulouse, George Eastman House, Biblioteca de Arte da Fundação Calouste Gulbenkian, The National Media Museum, The National Maritime Museum, The State Library of New South Wales, The Library of Virginia, The Musée McCord Museum, and the Nationaal Archief.

²¹ Chen, Seb. "Commons on Flickr - a report, some concepts and a FAQ - the first 3 months from the Powerhouse Museum," <http://www.powerhousemuseum.com/dmsblog/index.php/2008/07/21/commons-on-flickr-a-report-some-concepts-and-an-faq-the-first-3-months-from-the-powerhouse-museum/>.

is now providing Commons users with the ability to limit their keyword searches to Commons collections and to view a tag cloud for the top 150 tags ascribed to photos in The Commons.

The following outline covers the range of possible options for future participation (note: not all are mutually exclusive).

- (1) Walk away. Discontinue adding new photos and turn off interaction options to remove the need for moderation of photos already in Flickr.
 - Pro: Annual savings of 1 FTE salary
 - Con: Loss of good will and the opportunity to collect valuable information.

- (2) Continue “as is”--add 50 photos/week and moderate account
 - Pro: Modest expense to expand to 1.5 FTE from current 1 FTE (shared by OSI and LS among 20 staff). Additional .5 FTE needed to keep up with the amount of user-generated content on a growing account—both in moderation and in changes to the catalog records (both in Flickr and PPOC).
 - Con: Loss of opportunity to engage even more people with Library’s visual collections. Risk of losing attention from a Web 2.0 community that expects new and different content and interaction as often as possible.

- (3) Add new photo collections from P&P
 - Pro: Attract new audiences with more diverse subjects, e.g., a selected set of treasures; black-and-white Great Depression photos (FSA/OWI); Civil War photos; child labor collection (Lewis Hine); color travel views of Europe; and Russian Empire (Prokudin-Gorskii).
 - Con: Cost of .5 FTE additional staff (.25 FTE to prep, load, and update larger quantity of metadata and files and .25 FTE to moderate an account with more numerous photos).

- (4) Contribute photos and video content from many areas of LC
 - Pro: Attract new audiences with more diverse subjects, e.g., American Folklife and Veterans History collections, exhibits such as the MacDowell Colony.
 - Con: Cost of 1 FTE additional staff. (.25 FTE to prep and load more metadata and files, .25 FTE to moderate expanded account, and .5 FTE for project management and coordination among multiple divisions).

- (5) Increase analysis efforts
 - Pro: Learn more about how the Web 2.0 interactions and networks are leading people to the Flickr photos. Analyzing the search queries could improve understanding of how people seek images and lead to improvements in LC cataloging practices. See the *Tags* section for detailed options with tag analysis.

Con: Cost of .5 FTE and library school intern. Flickr members may not want to be studied. Flickr may not be able to provide the search queries and referral statistics.

(6) Create a virtual reference center or reading room for photos

Pro: Increase awareness and use of LC collections by responding to comments proactively and organizing our Flickr photos by subject to make them even easier to find and interact with. For example, refer people to related resources that might be of interest even when they haven't asked a specific question; geotag our photos; draw attention to mystery photos that need special assistance to decipher; acknowledge volunteer contributions formally to encourage continued involvement.

Con: Cost of 1 FTE additional staff. (.25 FTE to prep and load more metadata and files, .5 FTE to moderate expanded account).

(7) Showcase LC visitor experiences

Pro: Create and administer a group where people could post photos taken while visiting the Library; build a virtual community of Library friends and attract new visitors to Capitol Hill because of what they see in Flickr.

Con: Cost of .5 FTE additional staff to moderate the account. Organizational administrative responsibility could be unclear.

At the start of the pilot, critics pointed out several risks often expressed as questions. Experience so far has not borne out their concerns. The skeptics wondered: Would the public conversation contribute to a better understanding of the photos or would fan mail, false memories, fake facts, and uncivil discourse obscure knowledge? Would a public-commercial partnership undermine the Library's reputation for impartiality? Would the Library lose control of its collections? Would library catalogs and catalogers become obsolete? Would the need to moderate and respond to comments overwhelm all other work? Would history be dumbed-down? Would photographs be disrespected or exploited? Would entire collections be welcome or would selection of safe content border on censorship of historical information?

Since the Library first launched its account the public has allayed many of the misgivings by lauding the rapid access to interesting photographs that could be enjoyed and used without restriction. News media complimented the Library for making publicly held information widely and freely available and also praised our openness to participatory cataloging. Fellow cultural heritage organizations quickly began to join Flickr's Commons because 'taking the pictures to the people' resulted in reaching large new audiences.

Increasing the ability to engage and connect with photos increases the sense of ownership

and respect that people feel for these photos. As the Library considers strategies, technologies, and solutions to making our resources available, discoverable, and useful, lessons learned from this project provide guideposts to the type of experience that people would like to have with our collections, as well as informing Library decision-makers on the benefits of distribution beyond the walls of the loc.gov Web site. Using Flickr features and functionalities gave us experience with search and display capabilities currently beyond LC systems. We gained a deeper understanding of how users want to interact with our collections. Engaging with the public in Flickr also helped to identify how to manage the challenges of Web 2.0 participation: how to interact with users in ways that are less formal without diminishing the reputation of the institution; how to reconcile the inevitable loss of control over content with the recognition that we can significantly increase the reach of that content if people can access and interact with it in the communities in which they participate. Continuing the Flickr project would inform this ongoing conversation at the Library. It should come as no surprise, then, that the Flickr team recommends that this experiment in Web 2.0 cease to be characterized as a pilot and evolve to an expanded involvement in this growing community (and other appropriate social networking opportunities that may arise) as resources permit. The benefits appear to far outweigh the costs and risks.

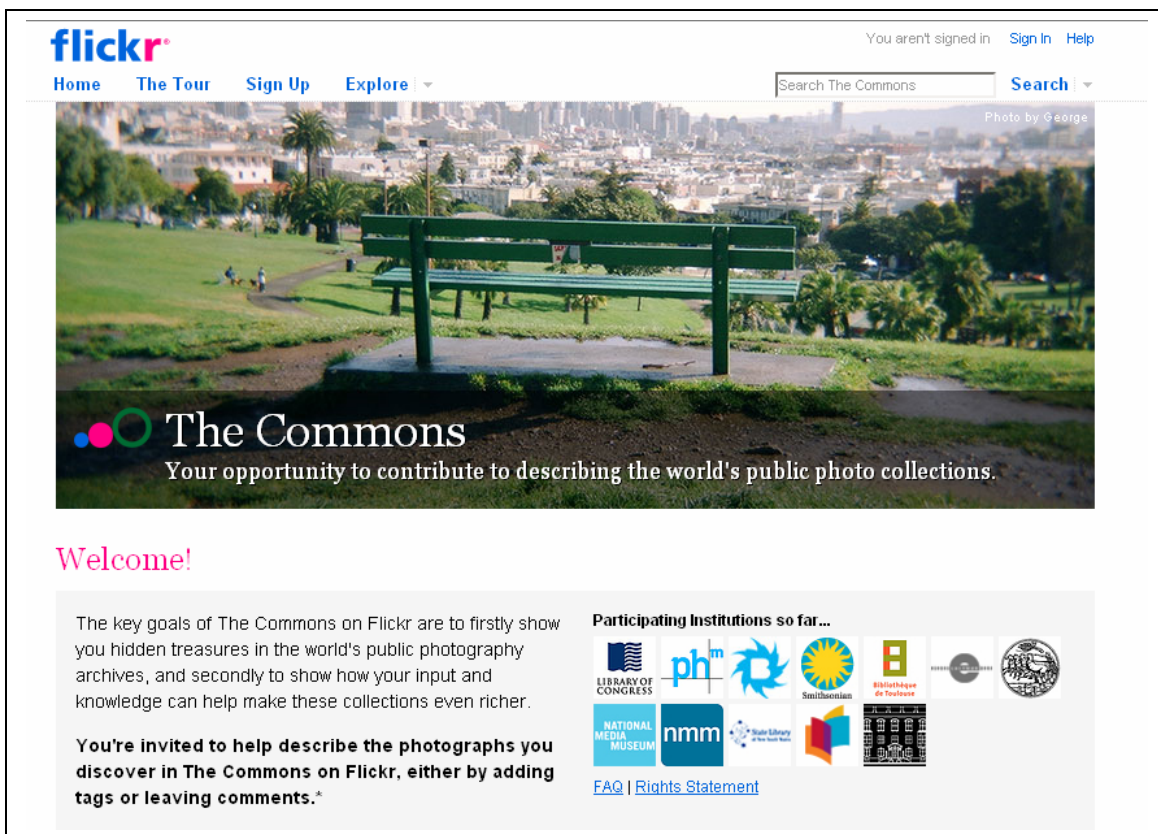


Figure 8. The Commons in Flickr

Appendix A: Examples of Flickr Interactions

One of the best ways to understand Web 2.0 is to look at the interactions in comments, notes, and tags. The following examples are categorized to indicate the inventive and impressive range of engagement with historical images. People are ...

- **Contributing specialist knowledge**

This photo of cars parked next to a huge haystack prompted people to comment on the shape of the stack, the reasons to store so much hay uncovered, and the sparseness of old tires.

http://www.flickr.com/photos/library_of_congress/2178331373/

In this photo of a parade of Olympic Athletes, many notes provide the text of the signs on the cars and on the windows. The debate over the location ended with a photo from another library that confirms the location. Another member corrected assumptions about a 'trolley car' in the photo to clarify that it's really a cable car on the Broadway line of the Metropolitan Street Railway.

http://www.flickr.com/photos/library_of_congress/2332832827/

For this photo of women weavers, a Flickr user pointed out a photographer's logo (bottom left): "Byron of New York," which provided a fundamental new piece of information and connections to many related photos. Four months later, the music device in the right corner caught attention with a new question, "Is that a violano-virtuoso?" and a link to a Wikipedia illustration was added to encourage others join in the answer.

http://www.flickr.com/photos/library_of_congress/2163450764/

- **Adding value through links, comparison, notes**

This rather plain photo of a barn from news photos taken between 1910 and 1915 sent Flickr participants fishing in the *New York Times* archive to link the image to the event alluded to in the title ("*Jones Barn where dynamite was found*"). An additional commenter provided links to archival holdings that could shed further light on the story.

http://www.flickr.com/photos/library_of_congress/2163452890/

Flickr members have sent in photographs of how a place looks today, making then-and-now comparisons possible. A mill in Lowell, Mass.:

http://www.flickr.com/photos/library_of_congress/2179041336/

<http://www.flickr.com/photos/pearlson/2209255147/>

Railroad cars and factory buildings in Lawrence, Mass.:

http://www.flickr.com/photos/library_of_congress/2178249693/

<http://www.flickr.com/photos/bgd73/2208349678/>

<http://www.flickr.com/photos/prybylphotos/1208753606/>

- **Sparking memory and conversations about history**

This photo was enriched with town history. The original FSA/OWI agency caption was “Industrial buildings and a town in Mass., possibly Brockton.” In the comments, one account holder noted that she had spoken to her father who had worked in factories in Brockton. He identified this site as a shoe factory. Other members confirmed the address.

http://www.flickr.com/photos/library_of_congress/2179042924/

Here’s an example of an account holder whose grandmother was a real life riveter for American Aviation in California. She placed a drawing of her grandmother in the comments section.

http://www.flickr.com/photos/library_of_congress/2179930812/

- **Sparking creativity and memory**

Flickr members learn about the photographs by imitating or recreating the scene today. Brownpau and his wife re-enacted the photo taken in front of Union Station in Washington, DC.

http://www.flickr.com/photos/library_of_congress/2179918784/

Painting a plane:

http://www.flickr.com/photos/library_of_congress/2178436353/

<http://www.flickr.com/photos/ataricharm/2232599895/>

- **Looking from all over the world and reflecting on related experiences**

Flickr member Catmandu identifies herself as an ignorant Australian who wonders what you do at a Bull Moose convention. Another user points to a source of information, and she comes back with interesting reflections on women’s suffrage in various countries.

http://www.flickr.com/photos/library_of_congress/2163955672/

- **Being prompted to new awareness of history through the photos and the conversations and investigations they spark**

After seeing a 1943 photo showing an African American woman working on a dive bomber, Flickr member Kenorland observed “Okay so they never taught us that black women worked as Rosie. What a crime of omission! So cool to see this.”

http://www.flickr.com/photos/library_of_congress/2179038448/

A photo of Japanese Americans at Manzanar generated a long conversation about old photos going beyond nostalgia, why would people smile in such a circumstance, and more.

http://www.flickr.com/photos/library_of_congress/2179117431/

- **Looking closely**

Commenters found plenty of information to question in the photo and in the original caption, including the fact that the photo shows home plate, not second base.

http://www.flickr.com/photos/library_of_congress/2163461044/

Careful observers noted the stunt photography aspect of this auto polo scene, and also linked the image to sources showing that the sport itself was real.

http://www.flickr.com/photos/library_of_congress/2477625961/

- **Helping each other to understand the images and data we provided, and the historical context of the original descriptions**

In this case, one commenter questioned the use of the word “girl” in the caption.

Flickr member Caffeinatedlibrarian helped point out that the word was in the original caption.

http://www.flickr.com/photos/library_of_congress/2179922510/

- **Offering visual humor**

The photos inspire wit. This commenter clearly saw something different in a 1942 photo of a Tennessee Valley Authority facility in Alabama. The comment: “I think this was an early conceptualization of the Internet.”

http://www.flickr.com/photos/library_of_congress/2179139231/

Appendix B: Strategic Goals and the Flickr Project

The Flickr Project serves numerous strategic goals at the Library of Congress. The mapping below highlights four project features that increase awareness of the Library and its collections; spark creative interaction with collections; provide LC staff with experience with social tagging and Web 2.0 community input; and provides leadership opportunities to cultural heritage and government communities. The Strategic Goals quoted below are from the *Library of Congress Strategic Plan, Fiscal Years 2008-2013*,²² the *Office of Strategic Initiatives Strategic Plan, Fiscal Years 2008-2013 (OSI)*,²³ and the *Library Services Strategic Goals, Strategic Objectives, Performance Goals, Activities, and Performance Measures (LS)*.²⁴

I. The Flickr Pilot increases awareness of cultural heritage and educational resources by sharing photographs from the Library’s collections with people who enjoy images but might not visit the Library’s own Web site. Taking the pictures to the people attracts new audiences and maximizes the use of visual collections.

Library Outreach Goal: Increase awareness of the value and utility of the Library. (...increasing and maximizing use of the Library’s resources and services by the Congress and the public will make it possible to improve learning and creativity.)

Outcome 2: Maximized use of the Library.

Strategy: Ensure that the Library’s resources and services make full use of changing technologies.

Performance Goal: Collaborative programs with external entities to increase use of specific Library programs, resources and services.

LS Goal 3: Deepen the general understanding of American cultural, intellectual, and social life and of other peoples and nations.

Strategic Objective: 3.A.: Present our collections to new and broader audiences.

OSI Goal: Enhancing access and usage of the Library’s collections and services.

Strategic Objective: Increased use of and seamless access to digital content and services.

²² http://www.loc.gov/about/strategicplan/2008-2013/StrategicPlan07-Contents_1.pdf

²³ http://www.loc.gov/staff/osi/about/OSI_StrategicPlan.pdf

²⁴ <http://www.loc.gov/staff/lis/planning/080608LSGoalObjGoalAct.pdf>

Outcome: Increased usage and awareness of digital content and services by target user communities.

Outcome: Expanded channels for delivering content and services.

II. The Flickr Pilot sparks creative interaction with LC's collections.

Library Customer Goal: Improve our customers' experiences in seamlessly finding and using Library resources. (...improving the customers' ability to access what they need quickly and in the most seamless way, but also ... informing public discourse and understanding about the world in which we live.)

Outcome 1: Increased use of the Library's digital resources to promote knowledge and better understanding.

Library Outreach Goal: Increase awareness of the value and utility of the Library.

(...increasing and maximizing use of the Library's resources and services by the Congress and the public will make it possible to improve learning and creativity.)

Outcome 3: Improved quality of life-long learning and creativity through use of Library services and resources.

Library Content Goal: Expand and preserve in accessible form a unified and universal body of knowledge and creativity. (...stimulating creativity, generating content, and developing descriptive data to ensure that quality collections are available to Congress, the nation, and the world.)

Outcome 4: Increased creative and intellectual output that contributes to the body of knowledge available to the Congress and other constituencies.

LS Goal 2: Provide the most effective methods for connecting the Library user to our collections.

Strategic Objective 2.B: Adopt technology that makes the collections more accessible to users.

Performance goal: 2.B.4: Employ new technologies to deliver collections to users.

OSI Goal: Enhancing access and usage of the Library's collections and services.

Strategic Objective: Increase use of and seamless access to digital content and services.

Outcome: Increased usage and awareness of digital content and services by target user communities.

III. The Flickr Pilot helps LC staff gain a better understanding of how social tagging and community input could benefit both the Library and users of the collections. We can explore the potential of user-generated content and tap expert community contributions to describe the content of special collections that we have the resources to digitize but not subject index.

Library Customer Goal: Improve our customers' experiences in seamlessly finding and using Library resources. (...improving the customers' ability to access what they need quickly and in the most seamless way, but also ... informing public discourse and understanding about the world in which we live.)

Outcome: Improved customer's ability to get what they need, when they need it, with minimal effort.

LS Goal 2: Provide the most effective methods for connecting the Library user to our collections.

Strategic Objective: Determine the bibliographic description framework needed in the digital era.

Performance Goal 2.A.2: Increase the accessibility of the Library's special collections through bibliographic description or finding aids.

OSI Goal: Enhancing access and usage of the Library's collections and services.

Strategic Objective: Increase use of and seamless access to digital content and services.

Outcome: Increased usage and awareness of digital content and services by target user communities.

IV. The Flickr Pilot can lead the way for other institutions. Collaborating directly with Flickr on the pilot proposal and licensing issues resulted in a new specially designated area of Flickr (The Commons) for collecting institutions to contribute items and state the rights information that they believed, but could not warrant, to be true.

Library Content Goal: Expand and preserve in accessible form a unified and universal body of knowledge and creativity. (...stimulating creativity, generating content, and developing descriptive data to ensure that quality collections are available to Congress, the nation, and the world.)

Outcome 4: Increased creative and intellectual output that contributes to the body of knowledge available to the Congress and other constituencies.

LS Goal 4: Provide leadership for the library community.

Strategic Objective 4C: Develop new tools and standards for librarianship.

Performance Goal: 4.C.1.: Lead the development of national and international standards that promise to enhance library services ... allows wide sharing of content and data.

OSI Goal: Enhancing access and usage of the Library's collections and services.

Strategic Objective: Increase use of and seamless access to digital content and services.

Outcome: Expanded channels for delivery content and services.

Strategy: Collaborate with units across the institution, external technology and content partners, and other Federal agencies to target and reach the Library's core online audiences.

Appendix C: Media Coverage

Blogs, Books, Magazines, Newspapers, Web Sites

- Braiker, Brian. "Flickr Helps the Library of Congress." *Newsweek.com* (24 Jan. 2008).
<http://www.newsweek.com/id/103198/Newsweek>.
- Catone, Josh. "Library of Congress Teams with Flickr." *ReadWriteWeb.com* (16 Jan. 2008).
http://www.readwriteweb.com/archives/library_of_congress_teams_with_flickr.php
- Chan, Wade-Hahn. "Library of Congress Tests Web 2.0 Photo Archive." *FCW.Com* (24 Mar. 2008).
<http://www.fcw.com/online/news/152025-1.html>.
- Riley, Duncan. "Flickr Takes Historical Imagery To The Masses." *TechCrunch.com* (16 Jan. 2008).
<http://www.techcrunch.com/2008/01/16/flickr-takes-historical-imagery-to-the-masses/>.
- Fost, Dan. "Killer Statue- Psyched about the Site!" *New York Times* (12 Mar. 2008).
<http://www.nytimes.com/2008/03/12/arts/artsspecial/12social.html>.
- Gilbertson, Scott. "Flickr Taps User Tags to Organize Library of Congress Images." *Wired.com* (16 Jan. 2008). <http://blog.wired.com/monkeybites/2008/01/flickr-taps-use.html>.
- Glen, Joshua. "Marching off to..." *Boston Globe* (3 Feb. 2008): p. C4.
- Hunt, Tara. [Telephone interview for "The Whuffie Factor: Using the Power of Social Networks to Build Your Business."] (17 Apr. 2008).
- Jardin, Xeni. "Rosie the Riveter: one of many finds in that LoC Flickr set." *BoingBoing.net* (18 Jan. 2008). <http://www.boingboing.net/2008/01/18/rosie-the-riveter-on.html>
- Kaplan, Rebecca. "Flickr, Library of Congress find something in 'Common'." *USA Today* (28 Jul. 2008): p. 4D. http://www.usatoday.com/news/education/2008-07-27-flickr-commons_N.htm.
- Kramer-Smyth, Jeanne. SAA 2008: Revealing Collections at the Web's Surface. Spell-Bound blog post, Sept. 2, 2008, <http://www.spellboundblog.com/2008/09/02/saa2008-revealing-archival-collections-at-the-webs-surface-session-102/>
- Lieberman, Michael. "The Library of Congress Hits the Jackpot on Flickr." *Seattle Post-Intelligencer Reader Blogs* (21 Jan. 2008).
http://blog.seattlepi.nwsourc.com/bookpatrol/archives/129927.asp?from=blog_last3.
- Mihm, Stephen. "Everyone's a historian now: How the Internet - and you - will make history deeper, richer, and more accurate." *Boston Globe* (25 May 2008).
http://www.boston.com/bostonglobe/ideas/articles/2008/05/25/everyones_a_historian_now/.
- Sarah. "Color photos from the 1930's & 40's." *Signal vs. Noise (37signals.com)* (4 Apr. 2008).
<http://www.37signals.com/svn/posts/965-color-photos-from-the-1930s-40s>.
- Seligson, Joelle. "Web 2.0 Partnership Opens Doors for Public Collections." *AVISO* [American Association of Museums newsletter] (1 Mar. 2008).

<http://newsmanager.commpartners.com/aama/issues/2008-03-01/>.

Shankland, Stephen. "Flickr to host Library of Congress photos." *CNet News* (16 Jan. 2008) (article and podcast) http://news.cnet.com/8301-13580_3-9852087-39.html?part=rss&subj=news&tag=2547-1_3-0-5.

"My Friend Flickr: A Match Made in Photo Heaven." *Techmeme.com* (16 Jan. 2008).
<http://www.techmeme.com/080116/p76#a080116p76>

Terdiman, Daniel. "Flickr brings tagging to vintage images." *CNet News* (17 Jan. 2008).
http://news.cnet.com/Flickr-brings-tagging-to-vintage-images/2100-1025_3-6226687.html.

"Uncle Sam Needs You." *Newsweek* (4 February 2008); p. 4.

Yehle, Emily. "Flickr Offers Surfers Peek at Library's Photo Collection." *Roll Call* (17 Jan. 2008).
http://www.rollcall.com/issues/53_80/news/21640-1.html?type=pf.

Podcasts, Radio, and Television

Boyd, Clark. "WTP 186: Taliban Threatens Mobile Masts, LOC Does Flickr, and Blogging WWI Redux." *PRI's The World: Technology*, [a BBC World Service news program] (19 Mar. 2008).
<http://www.theworld.org/pod/tech/WTP-podcast186.mp3>.

Falls, Sarah. "Interview #6: Helena Zinkham and Michelle Springer: Library of Congress, Prints and Photographs Division, Flickr pilot project." *ArtTECHtonic* [a virtual "session" for ARLIS-NA 2008 annual conference] (21 Apr. 2008). <http://arttechtomic.wordpress.com/2008/04/20/interview-6-helena-zinkham-and-michelle-springer-library-of-congressflickr-project/>.

Inskeep, Steve. "Library of Congress Looks for Help on Photo Labels." *Morning Edition*, National Public Radio (21 Jan. 2008). <http://www.npr.org/templates/story/story.php?storyId=18279042>.

Nurenberg, Gary. "Library of Congress Gets Online Help." *WUSA9.com*. (26 Feb. 2008).
<http://www.wusa9.com/news/local/story.aspx?storyid=68985>. [Broadcast on WUSA 9].

"Photo Share Podcast #21: Flowers, Prizes and The LOC." *photosharepodcast.com* (21 May 2008).
<http://www.photosharepodcast.com/2008/05/photo-share-podcast-21-flowers-prizes-and-the-loc/>.
And "Photo Share Podcast #22: Billboards and News." *photosharepodcast.com* (27 May, 2008).
<http://www.photosharepodcast.com/2008/05/photo-share-podcast-22-billboards-and-news/>.

Scobel, Robert. "Getting USA's Treasured Images Onto Flickr." *Scobelizer/ Fast TV.com* (October 2, 2008). <http://www.fastcompany.tv/video/getting-usa-s-treasured-images-onto-flickr>.

Scobel, Robert. "Visiting the Library of Congress and Meeting the Flickr Team." (17 Sept. 2008).
Scobelizer/ Fast TV.com. <http://scobleizer.com/2008/09/17/visiting-the-library-of-congress-and-meeting-the-flickr-team/>.

Weinberg, David. "Knowledge, Noise and the End of Information." *Smithsonian Institution Libraries*. (September 16, 2008). http://www.sil.si.edu/lectures_40th_Weinberger.html.

Society of American Archivists. "American Archives Month October 2008 Public Relations Kit." (October 2008); p. 12. http://www.archivists.org/archivesmonth/AAM2008/2008_AAM_Kit.pdf.

Appendix D: Catalog Records for FSA/OWI and Bain

The catalog records supplied by the Prints and Photographs Division for use as descriptions in Flickr for FSA/OWI photos had considerably more information than the records supplied for the Bain photos. A detailed comparison of the types of information in the Library-provided metadata can assist tag analysis. The information elements in the records were encoded in the MARC21 format.

	FSA/OWI (N=1,616 records)	Bain (N=1,500 records)
Name headings		
MARC 100/700 (personal)	1,557	0
MARC 110/710 (corporate)	0	1,500 ²⁵
MARC 111/711 (meeting)	0	0
Names as subjects		
MARC 600 (personal)	4	1
MARC 610 (corporate)	827 ²⁶	0
MARC 611 (meeting)	21 ²⁷	0
Topical subjects		
MARC 650	4725 ²⁸	3
Geographic subjects		
MARC 651	100	0
Uncontrolled terms		
MARC 653	0	902 ²⁹
Genre/Form		
MARC 655	73 ³⁰	1,500 ³¹

²⁵ All for Bain News Service

²⁶ Most for railroads, US military branches, companies

²⁷ Most for events, e.g., Vermont State Fair

²⁸ 22% from LCSH (virtually all for World Wars, ethnic groups (e.g., Norwegian Americans), or holidays); 77% from LC Thesaurus of Graphic Materials.

²⁹ Most for places or types of sports

³⁰ All from TGM

³¹ All for glass negatives, from TGM

Appendix E: Traffic to the Library's Web Site

Team members are frequently asked what effect this project had on traffic to the Library's Web site. Did it increase traffic or perhaps reduce traffic? We performed several metric assessments that confirm that the Flickr pilot had a positive effect on the traffic to the Prints and Photographs Online Catalog (PPOC), measured by views, visits, and downloads,³² both in terms of the total catalog and specifically for the two collections placed on Flickr. In addition, data supports that new audiences are coming to the LC Web site as a result of this pilot. Various factors may have contributed to the rise in the use of the catalog (a new Lincoln photo in the LC collection was announced in January 2008, for example), but it is highly likely that the Flickr pilot was a strong factor in these increases.

The construction of the URLs of the various pages and files found within the Prints and Photograph Online Catalog (PPOC) complicated the metrics analysis process. The PPOC is constructed of various Web pages:

- "About the collection" information
- Item level bibliographic information
- Information on how to order copies
- Digital item display pages (from which higher resolution JPEG and uncompressed archival TIFF files can be downloaded), and
- Search and search results pages.

An effort was made to count views and visits to each digital item display page, the bibliographic information page and the obtain copies page, and to exclude search results pages from the Web pages counted.

In order to determine effects of the pilot, we analyzed the following data for the months of January through May in 2008 and the same period in 2007:

- Web page views and visits to the digital item display pages, the bibliographic information pages and the "obtain copies" pages (excluding the "about the collection" pages and search results pages) for the entire PPOC catalog
- Total views/downloads of TIFF and JPEG files for the entire PPOC catalog
- Unique TIFF and JPEG files viewed/downloaded from the entire PPOC catalog
- Total views/downloads of TIFF and JPEG files for the entire FSA/OWI color collection

³² A page view refers to an individual viewing a single Web page. A site visit refers to an individual accessing any page, or group of pages, on the Library's site prior to leaving for another site.

- Unique TIFF and JPEG files viewed/downloaded from the entire FSA/OWI color collection
- Total views/downloads of TIFF and JPEG files from the entire Bain collection, the subset of Bain photos loaded onto Flickr, and those not loaded onto Flickr
- Unique TIFF and JPEG files views/downloaded from the entire Bain collection, the subset of Bain photos loaded onto Flickr, and those not loaded onto Flickr.
- Visits to the Library’s Web site originating from the Flickr.com domain
- Page views and visits to the five Library Web pages most frequently cited in our Flickr profile page and in various blog posts about the project.

Percentage of total PPOC views and visits

In order to measure the effect of the Flickr pilot on the entire Prints and Photographs catalog (PPOC), we compared month over month Web page views and site visits within the PPOC catalog for the period January through May of 2007 and 2008.³³ Analysis indicated that the percentage of Web pages viewed within the entire PPOC catalog rose for each of the first five months in 2008 when compared to the same month in 2007 (see tables below). The monthly views of PPOC Web pages for the months of January through May 2008 compared to the same period in 2007 showed an average monthly increase of 12% over the five month period. The percentage change of visits told a similar story-- an average monthly increase of 20% over the five month period.

Percentage Change of Page Views by Month, 2008 over 2007		Percentage Change of Visits by Month, 2008 over 2007	
January	15.2%	January	27.0%
February	14.5%	February	21.5%
March	6.6%	March	8.0%
April	10.6%	April	23.2%
May	12.9%	May	24.3%
Average rise	12.0%	Average rise	20.8%

Views/Downloads of FSA/OWI color and Bain collection TIFF and JPEG files

We were interested in the answer to this question: For the pictures we’d uploaded to Flickr, did views/downloads of the TIFF and JPEG files from the FSA/OWI and Bain collections rise or fall? The answer to this query was more straightforward for the color FSA/OWI

³³ An effort was made to exclude search results pages from the PPOC HTML pages and only count the digital item display, the bibliographic information pages and the “obtain copies” pages (not the “about the collection” pages). Our metrics counting was based on a unique string within the URLs of PPOC pages.

photos, as the entire collection (approximately 1,600 photos) was uploaded onto Flickr. The analysis of Bain collection data had an added element of complexity-- only a relatively small set of Bain photos was loaded onto Flickr (2,100 out of approximately 44,000). For the Bain collection, we needed two analyses: an assessment of the effect on the specific photos that were uploaded and an additional analysis on the collection as a whole, to see if there had been any supplementary rise or fall in downloads of the non-Flickr Bain photos. We also examined unique views/downloads vs. total downloads to help determine if some files were downloaded multiple times, while other files were only viewed/downloaded once.

Analysis of the data reveals that downloads rose on every parameter for both collections in 2008 when compared to 2007.

FSA/OWI Color

- Total views/downloads of TIFF and JPEG files of the color FSA/OWI collection rose 13.6% for the period January through May 2008, when compared to the same time period in 2007. Downloads of unique files rose 4.6%.

Bain

- Total views/downloads of TIFF and JPEG files across the entire Bain collection rose 34.3% in 2008 for the period January through May, when compared to the same time period in 2007. Downloads of unique files rose 30.6%.
- Total views/downloads of TIFF and JPEG files of the Bain collection items that were placed on Flickr rose 58.9% for the period January through May 2008, when compared to the same time period in 2007. Downloads of unique files rose 51.9%.
- Total views/downloads of TIFF and JPEG files of the Bain collection items that were not placed on Flickr rose 30% in 2008 for the period January through May, when compared to the same time period in 2007. Downloads of unique files rose 27.6%.
 - In the same time period the 2,100 photos that were added to Flickr in 2008 experienced a 58.3% rise in instances of downloads in 2008.
 - It's interesting to note that the unique files that were viewed/downloaded in 2007 that were not placed on Flickr in 2008 actually saw a 30% drop in views/downloads in 2008, while total views/downloads of unique files not on Flickr rose 30%. This indicates that different files were being viewed/downloaded in 2008 than were downloaded in 2007.

New traffic to LC Web pages from Flickr

We also know that as a result of the pilot, traffic from Flickr to the Library's Web site

increased dramatically. That conclusion is based on measuring visits that originated from the Flickr domain in the first five months of 2008 and comparing those numbers to the same time period in 2007. There was an average rise of 2,205% with the majority of the traffic coming in January and February (see table below).

Percentage Change of traffic coming from Flickr by Month, 2008 over 2007	
January	5,483%
February	3,162%
March	743%
April	1,119%
May	521%
Average rise	2,205%