Playfulness in the Archives: Enhancing Digital Collections through Card Sorting

Erin Passehl-Stoddart, Special Collections and Archives, University of Idaho

Author Note:

Erin Passehl-Stoddart is Head of Special Collections and Archives at the University of Idaho.

Correspondence concerning this article should be addressed to:

Erin Passehl-Stoddart
Special Collections and Archives
University of Idaho Library
875 Perimeter Drive MS 2350
Moscow, ID 83844-2350
Contact: estoddart@uidaho.edu

Abstract

Regardless of institutional type or resources, one question facing archives and special collections is how archival collections can be efficiently enhanced with minimal or no original metadata. This issue becomes a focal point when collections are digitized, as metadata is what makes digital collections more accessible and usable. This case study explores the development of a digital collection using card sorting activities and gamification techniques, and analyzes the direct and indirect effects of each strategy, including student employee connections to library learning goals and visual literacy standards.

Keywords: archives and special collections, metadata, digitization, student employees, visual literacy, gamification

Playfulness in the Archives: Enhancing Digital Collections through Card Sorting

Erin Passehl-Stoddart, Special Collections and Archives, University of Idaho

The William Allen Stonebraker Collection[1] was donated to the University of Idaho (UI) Library Special Collections and Archives in 2003. The collection consists of approximately 550 photographic negatives and prints, some located in bound scrapbooks, for a total of 1 cubic foot. W.A. Stonebraker was an amateur photographer, whose images document several important topics in Idaho history from the turn of the 20th century, including remote wilderness areas, early homesteads, and entrepreneurial operations such as dude ranches, pack train operations, mining, big game hunting, wildlife, scenic views, and early aircraft operation. After a patron requested to access the collection from a distance, the University of Idaho Library Digital Initiatives unit scanned the entire collection to add it to the Library's digital collections. Upon creating a spreadsheet with the minimal descriptive information provided with each photograph, the question that emerged was how could we efficiently augment metadata for photographs with little to no original information? My role was to enhance metadata for all 550 images from the Stonebraker collection, with an emphasis on generating geographic location tags for as many images as possible.

Problems with Identifying Trends across Digitized Materials

A few common issues emerged as I looked to create additional metadata for existing images. First, the archival collection contained very little individual, family, or historical information. Only one-third of the images contained identifying description, and it did not match across related photographs. In flipping through the images, staff identified individuals sporadically, and most images lacked geographic locations. Additional research was required to identify people and places, and in turn, better understand the potential importance of the collection.

1. https://www.lib.uidaho.edu/digital/stonebraker

The second problem was the difficulty comparing photographs and metadata across the collection. Digital Lab staff tried sorting the scanned images on computer screens with a Microsoft Excel spreadsheet of the original metadata provided for each image. However, even with dual computer screens, there was not sufficient physical space to compare hundreds of images. Since both the archival collection and digital images were not in any specific type of order, images taken on the same date or location were spread out physically throughout the collection and spreadsheet with different identification suffixes.

The third issue that arose was the more images someone sorted through, the more likely they were to express: "I know I have seen a similar photo before" and "This location looks very familiar." Staff felt as though they were taking two steps back for every connection they made. It became clear that one way to solve these challenges was to remove technology from the equation.

Not Every Digital Project Requires Technology the Entire Time

Many public and academic libraries have embraced the trend of crowdsourcing to enlist volunteers to assist with tagging and generating metadata for digital collections; two successful examples are "What's on the Menu?" (New York Public Library, 2011) and "DIY History" (University of Iowa Library, 2012). However, our unit was not prepared to spend that much time or programming on this project. We decided that a low-tech/low-effort approach utilizing existing infrastructure and staff would work best. Much like the ideology behind the New York Public Library Labs that "...the idea that digitization is the beginning of a whole life cycle, followed by transformations, further processing, and downstream uses, rather than an end point" (as cited in Enis, p. 38), we wanted to generate additional metadata so users could conduct basic search and browse via subjects, location, and date and contribute to public understanding of Idaho's early businesses and rural life. Instead of sourcing this out to crowds online, I created an internal card sort solution that was inexpensive, simple, and utilized analog methods.

Card sorting is a concept to emerge from computer science literature typically associated with organizing and evaluating website architecture that entails providing a group of users with a set of cards with descriptive information that needs to be organized; users sort the cards with similar concepts into piles (Faiks, & Hyland, p. 350). The card sorting activity I used was a cross between website usability card sorting and the board

game *Memory*. The game requires players to take turns trying to pair matching images lying face down, requiring concentration to remember where images were located. Not only did this mimic my experience on the computer, I felt like a game could also be the solution. Converting this activity into a game would make it fun for both students and staff; we could spread out all 550 images at one time for comparison purposes instead of only a handful at a time on a computer screen and find trends and subject similarities much faster; and it would not interfere with the original order or handling of the archival collection.

Resources needed to create our own *Memory*-like card game were minimal: it required a color printer, paper, and a paper cutter. Since the images were already scanned, I printed out the access jpegs at 35 images to one page, 1.5 inch by 1 inch in size, for a total of 16 printed sheets (see Figure 1). Using Microsoft Windows Picture Viewer, we printed the file name/photograph identifier number at the bottom of each image, which was helpful so we could quickly identify the images when making metadata changes in the spreadsheet. It took about an hour to print and slice the images using a paper cutter (see Figure 2).

Over 500 cards with identification numbers were cut out and spread out across a large table in the Digital Lab. At first, it looked like a colorful

Figure 2. Example of image printed with corresponding file identifier. Image courtesy Special Collections and Archives, University of Idaho.

kaleidoscope, and then chaos, as hundreds of images swirled together on the table (see Figure 3). However, within minutes it became easy to see that this was going to help immensely. We could already start to identify images that shared some of the same features, identification that was too difficult to point out when scrolling through images on a computer screen.

Figure 3. View of individual cards printed and spread out on a table. Image courtesy Special Collections and Archives, University of Idaho.

Identifying Trends through Card Sorting

Reactions to a table covered in hundreds of small cards ranged from the curious, "What are you working on?" to incredulous students exclaiming, "This is actually pretty fun!" Similar to traditional crowdsourcing, anyone who walked in, from library staff to student employees, was invited to spend time trying to identify trends and similarities across the collection. This activity could be completed quickly as it did not require a login or learning software and was accessible to all skill levels. It invited different perspectives and levels of expertise to participate; one did not need to have expertise in the life of Stonebraker to make a positive contribution, as card piles could be moved and renamed easily. As participants went through the cards, scrap pieces of paper were used to write down emerging themes and trends that developed naturally. For example, someone would write down winter on a scrap paper, and then start making a pile of images that included something related to winter (snow, activities in snow, etc.). That

Figure 4. Student employee sorting cards by topic. Image courtesy Special Collections and Archives, University of Idaho.
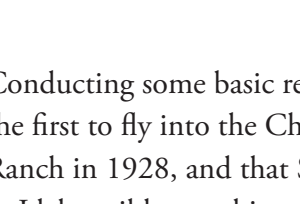
Figure 5. Example of card sort by subject. Image courtesy Special Collections and Archives, University of Idaho.

theme may later break down into smaller categories (snowshoeing, cross-country skiing, exterior buildings in snow, etc.) and the piles would be moved around (see Figure 5). Piles could be pushed around as new cards were analyzed, and cards could be swapped out as new themes emerged or were collapsed. At first, most people were a little intimidated, thinking that they could potentially mess up the order. An advantage of the card sort game was anyone could participate and the activity could take as long or as short as the participant wanted, from a five-minute break between other activities, to working on a specific topic for a half hour.

By sorting the cards by topics, themes, and trends that would directly assist with filling in metadata holes, other categories emerged such as sorting by family, individual, or geographic location. With hundreds of images physically and intellectually spread out in a digital collection, it felt impossible to identify unnamed individuals or locations. However, card sorting allowed us to see where an individual appeared many times; we looked to contextual information in these images to help identify individuals. An example of this was a group of seemingly random images of airplanes. After making a pile of all images related to air travel, it became evident that the same man (pilot Nick Mamer, identified through newspaper articles) occurred in many photos.[2] Because his name was described in the original metadata on one of the prints, but not in the other twelve, we identified the pilot and the year the images were taken.

2. Image comparison online at http://digital.lib.uidaho.edu/cdm/search/collection/stonebraker/searchterm/mamer/field/all/mode/all/conn/and/cosuppress/

	Title	Description	Date	Location	Latitude	Longitude	Identifier	Keywords
Initial metadata		Nick Mamer next to airplane. Man on horseback.	n.a.	n.a.	n.a.	n.a.	PG 26, P055	Aircraft;
Final metadata *(identifies the pilot, date, and location based on newspaper research)*	First Airplane in Chamberlain Basin	Pilot Nick Mamer (of Spokane, Washington) stands next to his airplane in the Stonebraker Ranch meadow. The airplane was a consolidated fleet with the National Guard. A man sits on a horse near the airplane.	08/21/28	Stonebraker Ranch	45.37785	-115.198892	PG 26, P055	military aircraft; horses; pilots;

	Title	Description	Date	Location	Latitude	Longitude	Identifier	Keywords
Initial metadata		Man and woman standing next to airplane. "U.S. Army Consolidated 0-17 NG 28-360" on plane.	n.a.	n.a.	n.a.	n.a.	PG 26, N08	Aircraft;
Final metadata *(identifies family members through other photographs, date and location based on newspaper article)*	National Guard airplane at Stonebraker Ranch	A young boy Adolph "Bill" (left) and mother Golda stand near a National Guard 116th Observation Squadron. The plane reads "U.S. Army Consolidated 0-17 NG 28-360" on its side. The plane is parked in the Chamberlain Basin meadow near the Stonebraker Ranch, piloted by Nick Mamer.	08/21/28	Stonebraker Ranch	45.37785	-115.198892	PG 26, N08	military aircraft; women; children;

	Title	Description	Date	Location	Latitude	Longitude	Identifier	Keywords
Initial metadata		Two men standing in front of porch eating.	n.a.	n.a.	n.a.	n.a.	PG 26, N31	
Final metadata *(identifies men, date, and location based on other photographs)*	Al Stonebraker and Nick Mamer	Pilot Nick Mamer (left) and Al Stonebraker (right) eat from bowls on the porch of Stonebraker Ranch.	08/21/28	Stonebraker Ranch	45.38936	-115.197614	PG 26, N31	military aircraft; women; children;

Conducting some basic research on this pilot, we determined that he was the first to fly into the Chamberlain Basin wilderness near the Stonebraker Ranch in 1928, and that Stonebraker documented this monumental flight in Idaho wilderness history (see Figure 6).

This type of metadata enhancement occurred not just with people, but also with identifying dates and years of events, buildings, and geographic locations. We knew this was important metadata since these are two popular ways for users to search the Library's digital collections per Google Analytics. The geography was important not only in telling the story of W.A. Stonebraker, but as an addition to the Library's Digital Map Room.[3] The Map Room is based on Google Maps and Google Fusion Tables that utilize latitude and longitude coordinates for individual images across digital collections. Wanting the Stonebraker collection to be discoverable in the Map Room meant that we needed to determine as many geographic coordinates as possible. Physically sorting cards meant that we could determine many images overlapped by location, which we could not "see" before.

Direct Effects of Gamification Strategy in Digital Collection Development

While the initial strategy and preparation for the card sort was not centered on gamification or game dynamics, these later had a direct effect on the development of the Stonebraker digital collection. Academic

Figure 6. Metadata descriptions before and after card sorting activity. Image courtesy Special Collections and Archives, University of Idaho.

3. https://www.lib.uidaho.edu/digital/maps/

definitions of gamification mainly focus on video games and motivation; a broad definition of gamification useful for libraries is "the use of game design elements in non-game contexts" (Deterding, p. 12). While the card sort activity was not designed with typical game outcomes in mind (scoring, levels, achievements, etc.), it did lend a sense of competition and playfulness that helped motivate both staff and students that were experiencing fatigue and frustration over trying to identify photographs in the collection. The card sorting activity helped counteract these feelings by increasing staff and student engagement with the metadata process and making it fun. For student employees, the gamification of the metadata process enabled and motivated them to plan and perform simple to complicated tasks towards a goal, as well as exercise a significant level of diligence, creativity and resourcefulness (Kim, p. 466). Other attractive outcomes include inducing strong motivation, spurring a high level of productivity, and simply having fun in a library-related context (Kim, p. 467-468). Overall, the card sort allowed for both staff and students to work through all the photographs at a significantly faster rate while having fun working on a difficult identification process.

Since student employees work in the Digital Lab, it was also important for the card sorting activity to produce and enhance skills that are part of the Library's overall learning goals. Students became engaged with different learning styles, including working individually and collaboratively, and experienced that a process can take different avenues and end with success. Another direct effect was an enhanced set of visual literacy skills. In ACRL's *Visual Literacy Competency Standards for Higher Education*, visual literacy is defined as:

> A set of abilities that enables an individual to effectively find, interpret, evaluate, use, and create images and visual media. Visual literacy skills equip a learner to understand and analyze the contextual, cultural, ethical, aesthetic, intellectual, and technical components involved in the production and use of visual materials. A visually literate individual is both a critical consumer of visual media and a competent contributor to a body of shared knowledge and culture (Hattwig et.al., p. 62).

Specifically, student employees developing metadata for the Stonebraker collection met the following sub points in Standards 3, 4, and 7 (Hattwig et.al., p. 77-78, 82).:

Standard 3: The visually literate student interprets and analyzes the meanings of images and visual media.

> » The visually literate student identifies information relevant to an image's meaning.

> » The visually literate student situates an image in its cultural, social, and historical contexts.

> » The visually literate student validates interpretation and analysis of images through discourse with others.

Standard 4: The visually literate student evaluates images and their sources.

> » The visually literate student evaluates the effectiveness and reliability of images as visual communications.

> » The visually literate student evaluates textual information accompanying images.

> » The visually literate student makes judgments about the reliability and accuracy of image sources.

Standard 7: The visually literate student understands many of the ethical, legal, social, and economic issues surrounding the creation and use of images and visual media, and accesses and uses visual materials ethically.

> » The visually literate student understands many of the ethical, legal, social, and economic issues surrounding images and visual media.

This activity helped increase student's visual literacy and critical thinking skills, including the ability to interpret and analyze images for contextual information and conducting research for source reliability. Students and staff compared notes and talked through how they were determining information from clues found in the photographs. This card sorting technique introduced the standards and sub points found in the Visual Literacy Array[4] based on ACRL's *Visual Literacy Standards*.

Indirect Effects of Gamification Strategy

Although employing a gamification strategy in the development of digital collections yielded many effects internally to the unit and the collection

4. http://www.ala.org/acrl/publications/keeping_up_with/visual_literacy

itself, it also provided indirect results due to the enhanced metadata that was created because of the strategy.

Enhanced Access and Discoverability

One of the benefits of "touching" each image and creating broad subject piles was it later translated into established subject headings that provided increased access and discoverability of the Stonebraker photograph collection. For this collection, we utilized Library of Congress Subject Headings, Getty Thesaurus of Graphic Materials (TGM) and Getty Art and Architecture to add subject headings to individual images. With the use of recognized controlled vocabularies across all images instead of just a small selection with original metadata, the collection can be browsed in a subject cloud using TagCloud software as well as keyword searched, adding another layer of discoverability to the collection that previously was not there.[5] Additionally, adding latitude and longitude coordinates to individual images allows for users to browse the collection by geographic location or map of Idaho.[6] These two features are displayed prominently as additional ways to discover materials in the digital collection.

Enhanced Metadata Tells a Better Story

Before the collection received additional metadata attention, the images were seemingly unconnected to each other, with no links between individuals, family members, businesses, or geographic locations. When archival collections do not have embedded background or historical information before undergoing the digitization process, it can be difficult to describe individual items without an understanding of the whole or story behind the collection. These enhanced descriptions of materials provide a compelling narrative to both staff and students working on collections as well as users of the digital and physical archival collection.

   The collection did not initially provide much biographical information about W.A. Stonebraker. After conducting research using local newspapers and references in books, we determined more aspects of Stonebraker's life that matched the photograph collection. We discovered Stonebraker was one of the original Idaho homesteaders in the Chamberlain Basin area whose businesses had an impact on the local economy. He also helped construct a major route named the Three Blaze Trail which led to Thunder Mountain during the mining boom. His pack train operation is heavily represented in the digital collection, including images of mining towns that

5. https://www.lib.uidaho.edu/digital/stonebraker/subjects.html

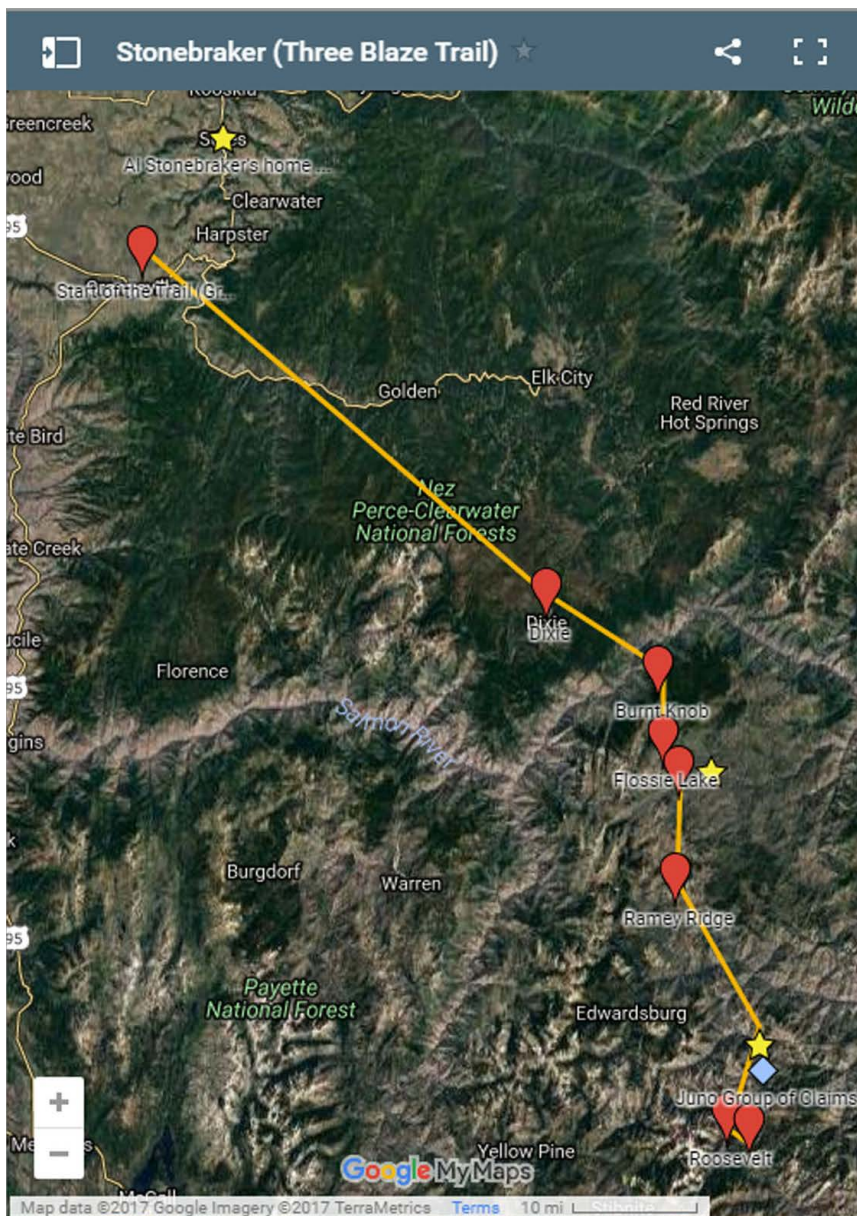6. https://www.lib.uidaho.edu/digital/stonebraker/map.html

Figure 7. Map of Three Blaze Trail created using Google Maps (University of Idaho Library, 2014). Map data © 2017 Google Imagery, © 2017 Terra Metrics.

no longer exist. We determined several properties in the images, including mining properties and a cabin used primarily for big game hunting tours. Stonebraker also captured picturesque views of Idaho's wilderness and the introduction of new modes of transportation to rural Idaho, including railroads, automobiles, and aircraft. Perhaps most significant was the connection that parts of the Stonebraker homestead still stand today, maintained by the Payette National Forest and still only accessible by horseback, foot or air. All in all, by piecing together this rich narrative, the collection became a cohesive story that connects to Idahoans and people studying the American West at the turn of the twentieth century.[7]

Enhanced Metadata Grabs a Larger Audience

What started out as a simple patron request to scan images from the Stonebraker collection became the most interesting digital collection the UI Library released that year. Suddenly the archival collection had quite a story to tell and allowed for a high level of public outreach and engagement with the university, local community, and statewide.

The UI Library issued a press release, resulting in a local TV station interview of our department that aired on the evening news and online (Ryan, 2014). Boise State Public Radio called for a phone interview and published an article online with images from the collection (Wright, 2014). The story was also picked up by a few statewide news stations and newspapers, giving the digital collection excellent coverage. In the weeks after the release, we received emails about how excited people were about the collection. One email from an archaeologist at the United States Forest Service at the Payette National Forest branch offered detailed information about the location of some of the wilderness photographs, as well as corrections to metadata that she and her staff had specific knowledge of, having visited the area many times. Therefore, knowing the collection inside and out gave it more value and a compelling narrative to share with others, which in turn allowed for public engagement.

Conclusion

While the development of the Stonebraker digital collection did not start out with intentional strategies involving crowdsourcing and gamification, it turned into a simple, inexpensive, and effective way to process and create enhanced metadata for over 500 images. This project served as a reminder that analog solutions can sometimes be the answer to digital quandaries and technical difficulties around the development of digital collections. It is easy to get focused on using technological means and forget that sometimes something as simple as a card sorting game can be the answer to the problem. It proved worthwhile in cost effectiveness, student engagement, and an improved digital collection.

In looking to transfer this process to other institutions and collections, this card sorting activity would be useful with large, complex collections where little to no original metadata is provided, but could be used for smaller visual collections as well. The exercise can be easily adapted for different types of collections; the only suggested best practice is to print the unique identifier on each card so that the

image can be traced back to the original metadata. One lesson learned to decrease the amount of time needed to create additional metadata is to sort the original metadata in Microsoft Excel by category instead of applying subject headings and keywords image by image.

Besides instilling a sense of playfulness and fun while learning in the Digital Lab, it was intriguing to see how students engaged with the visual literacy aspects of the activity. Next steps to increase value in these types of exercises and activities is to directly tie them to the ACRL Visual Literacy Standards and demonstrate how it increases student engagement, visual literacy and critical thinking skills, and potentially retention of student employees and interns in the unit. Finally, building in ways to formally assess these measures and standards will help demonstrate the value of the library to student learning in the library and university.

References

Brown, N., Bussert, K., Hattwig, D., & Medaille, A. (2013). Keeping up with… visual literacy. Retrieved from www.ala.org/acrl/publications/keeping_up_with/visual_literacy.

Enis, M. (2015). Wisdom of the crowd. *Library Journal, 140*(11), 38-40. Retrieved from http://lj.libraryjournal.com/2015/07/technology/wisdom-of-the-crowd-digital-collections/#.

Deterding, S., Khaled, R., Nacke, L. E., & Dixon, D. (2011, May). Gamification: Toward a definition. In *CHI 2011 Gamification Workshop Proceedings* (pp. 12-15).

Faiks, A. and Hyland, N. (2000, July). Gaining User Insight: A case study illustrating the card sort technique. *College & Research Libraries, 61*(4), 349–57. Retrieved from http://crl.acrl.org/content/61/4/349.full.pdf+html

Hattwig, D., Bussert, K., Medaille, A., & Burgess, J. (2013). Visual literacy standards in higher education: New opportunities for libraries and student learning. *portal: Libraries and the Academy, 13*(1), 61-89.

Kim, B. (2012). Harnessing the power of game dynamics: Why, how to, and how not to gamify the library experience. *College & Research Libraries News, 73*(8), 465-469. Retrieved from http://crln.acrl.org/content/73/8/465

New York Public Library. (2011). What's on the menu? Retrieved from http://menus.nypl.org/.

Passehl-Stoddart, E. (2015, February). *Match that photo: Embracing analog methods to enhance digital collections.* Presented at the Online Northwest Conference, Corvallis, OR. Retrieved from http://works.bepress.com/erin_passehl/29/.

Passehl-Stoddart, E. (2015, May). *Match that photo: Embracing analog methods to enhance digital collections.* Poster presentation at the Western Round-Up, Denver, CO.

Ryan, J. (2014). The U of I Library's digital initiatives department has 50 digitized collections you can access online. Retrieved from http://klewtv.com/news/local/the-u-of-i-librarys-digital-initiatives-dept-has-50-digitized-collections-you-can-access-online.

University of Idaho Library. The Map Room. (2012).

University of Idaho Library. Stonebraker Photograph Collection. (2014). Retrieved from http://www.lib.uidaho.edu/digital/stonebraker/
University of Iowa Library. (2012). DIY History.

Wright, S. (2014). Idaho's legacy comes to life in 20th century photos. Retrieved