

Minnesota Historical Society

Visualizing History: Using Data Visualization to Transform Primary Sources

Award Amount \$24,813, Total Project Cost \$42,930.18

August 1, 2012 – September 30, 2014

Project Director: Shawn Rounds

Project Summary

The Minnesota Historical Society's Visualizing History project began with two premises. The first is that the 21st-century museum visitor expects access to accurate, original historical evidence, but may not have the time or research skills to discover it themselves in libraries or online catalogs. And second, libraries and archives face a challenge to integrate their content with the world of visual media in a way that does not compromise the complexity of the primary sources they preserve. The challenge of this proposal was to see how data visualization could make complex data sets comprehensible and approachable in a gallery context. Over the course of the project, the project team came to a more complex understanding of the potential and the pitfalls of data visualization, and a more nuanced understanding of the role it plays in libraries, archives, and museums.

Museum visitors, as much as library researchers and users of archives, want access to high-quality, verifiable sources of information. They want to engage with original historical evidence. The typical users of MNHS's primary sources are family historians and older researchers who access our archives in the library's two reading rooms and online. Surveys have found that these users invariably search for answers to the questions they came in with. Gallery visitors, however, are often a different breed: they skew slightly younger, express a desire to be exposed to history, and are open to encountering a wider range of experiences. The history they access in the gallery is based on archives, but the presentation is often disconnected from sources—documents rarely figure into the dynamic experience of the gallery where visitors are free to play, touch, and experiment. Data visualization can help break down these boundaries and maximize archives' accessibility.

At the outset, the project team planned to build three visualizations. The first would be a static graphic; the second, an interactive infographic; and the third, a full visualization. This scaled approach, along with the use of open-source software, allowed us to build on our successes and at the same time model possible approaches to information visualization that could be of use to other museums regardless of size or budget. Our focus was not just on our peers at other institutions, however. We also hoped to demonstrate to our colleagues at MNHS (curators, exhibit designers, and archivists) the potential for data visualization to meet the needs of various cultural heritage audiences.

The impetus for our project emerged from an early data visualization we completed related to one of our historic sites, Fort Snelling. In this visualization, completed in 2011, we digitized several documents related to nineteenth-century mail delivery and supply provisioning at the historic fort to explore the realities of life on the northern frontier. The anecdotal feedback we

received on this small web site convinced us to further explore the possibilities of data visualization as an interpretive tool for our museum.

For our first visualization for the Sparks project, we focused on Civil War soldier Matthew Marvin, a private in the 1st Minnesota Regiment. The data visualization took excerpts from Marvin's wartime diary (the original is the MNHS archives) and explored the toll that the war took on his physical health. This simple visualization was created with the web presentation tool Prezi to demonstrate the possibilities of using readily available, web-based tools to create compelling historical interpretations.

Building on the Prezi work, we next set out to create an interactive timeline. Again relying on Marvin's detail-rich diary, we employed the open-source tool TimelineJS to integrate diary entries, transcriptions, maps, photographs, and objects from the MNHS collections into a narrative structure detailing Marvin's experiences between 1862 and 1863.

For our final visualization, we moved in a different direction, focusing on our extensive newspaper and photograph collection. We created an interactive web page called History Roulette. Visitors to the page make a virtual spin of the roulette wheel, which prompts the program to select a newspaper and a photograph at random from a curated set. The program then presents a mashup of a headline with the photograph, with the result being strange, incongruous, and frequently very funny. When we first launched the program online, we found that Roulette players spun the combinations frequently (web visitors spent 8 minutes on average on the page, far longer than on any other page of our site), finding the process to be, we suspect, somewhat addictive. Visitors can draw others into the experience by sharing their mash-ups through a variety of social media network links

History Roulette was different from the first two visualizations in that it did not aim to educate and inform as the first two did. But like our earlier visualizations, this one served to engage audiences with our collections, to get them to think about the primary source material we collect, and how they might use that material in their own research.

Along with the three visualizations, the team convened several assessments and workgroups in the course of the project. A final workshop, inviting colleagues from the University of Minnesota, was held at the end of the project to discuss and summarize the experience of the project and to talk about the emerging potential for data visualization for a wide variety of applications in libraries, archives, and museums.

Process

Because we looked at the issue of visualization through the lens of the various activities of MNHS, we included a diverse group of professionals on our project team. We asked for and used the staff time of curators, librarians, exhibit designers, information technology specialists, archivists, and educators. The project required staff commitment and cost-share from departments across the institution.

Our aim in using such a large number of our colleagues on this project was not merely to spread the work around. Rather, MNHS is a complex organization encompassing all three functions of libraries, archives, and museums under one roof. We wanted to engage our colleagues, learn from their professional experience, educate them about the potential of data visualization and, as a long-term outcome, incorporate this technology into the processes and products of the Minnesota Historical Society.

As one of the first steps in the project, the team worked with MNHS evaluation coordinator, Sheila Brommel, to define the expected project outcomes. With Brommel's guidance, the team created a logic model for the project. A logic model is a graphical depiction of a project showing the activities of the project and defining their relationship to the project's expected outcomes. The logic model helps answer questions about what the project is and what it is intended to accomplish. Our logic model set out the direction of the Sparks project and provided clarity to the team about the change the project intended to bring about. (A copy of the logic model is included at the end of this report.)

Beyond the advice, evaluation, and collaboration of our colleagues, we also used the services of technical consultants to create the History Roulette visualization. The use of outside programmers was certainly the most expensive and specialized way to build a visualization, but we wanted to demonstrate a range of options in our three products that included both simple DIY techniques to more expensive contracts.

One significant change that took place over the course of the project was a change in content focus for the final visualization. MNHS had been engaged in 2012 in the 150th commemoration of the US-Dakota War. Early plans were to create a visualization that could be used in the gallery exhibit for this war. Ultimately, we shifted focus to our newspaper collection. Over the course of the last four years, newspaper digitization, preservation, and access has emerged as the top priority of the Library and Collections Division of the Minnesota Historical Society. MNHS consulted with the IMLS program officer and received permission to shift the focus of the final visualization away from the US-Dakota War exhibit and instead to create a visualization that could be accessed on touch-screens in our Library lobby exhibit area.

Project Results

As mentioned above, our logic model established several anticipated project outcomes. By the end of the project, we expected to demonstrate that data visualizations of varying levels of complexity could be used in the service of reaching the various audiences of MNHS. We intended that: 1) users of the data visualization would demonstrate increased knowledge of the complexity of the topics interpreted; 2) visitors would display an increased understanding of and willingness to use primary sources; 3) staff would demonstrate increased understanding of data visualization and its potential for use in their activities; and 4) staff would demonstrate increased critical thinking and problem-solving skills.

We initiated the project by surveying MNHS staff to find out what ideas and preconceptions they had about data visualization. We learned that a third of respondents were uncertain about how to

define data visualization or what might be considered an example of one. A majority of staff were interested in creating visualizations to advance their work at MNHS, but most felt they lacked the requisite training and time needed to do the work. They expected the activity to be difficult, requiring skills and software applications that they lacked. When staff did create visualizations, they tended to use common desktop programs like PowerPoint and web pages to represent data visually. After we surveyed staff, we provided an opportunity for them to learn about and discuss data visualization at a brown-bag forum held in March 2013.

We also asked questions about the staff's use of primary sources and their interactions with various audiences in which they used visual aids. Presentations and lectures are common activities among staff at MNHS, even for staff not directly responsible for delivering programs to the public. Most respondents told us they use primary sources for research. They often use such sources to create visual aids in their presentations and believed that sources help them represent the value of MNHS to their audience. The most commonly used source, not surprisingly, was photographs, but staff also reported using a wide variety of document types. The most common way to identify and acquire sources for their presentations was to use the MNHS web catalog, Collections Online, which includes such things as photos, art, maps, audio materials, and objects, although not the finding aids for the archives and library collections (available through a separate online catalog).

By the end of the project, we found a distinct shift validating our expected outcome of deepening staff's understanding of data visualization and its potential for use in their work. At a workshop we convened at the end of the project to discuss the implications of the Sparks work for our institution, staff displayed a clearer understanding of the definition, as well as a greater ability to discuss and analyze the meaning of such presentations for their work.

With a diverse group of professionals in the workshop, suggestions for the use of data visualization were varied and creative. In general, staff saw data visualization on a spectrum with value for different audiences at each point on the spectrum. At one end of the spectrum was the more heavily curated and mediated experiences that take complex data sets and provide innovative visual presentations to make them comprehensive and meaningful. Such curated visualizations become useful interpretive tools, ways of educating students about historical topics, engaging visitors in a gallery, or driving home a point in a web presentation. In many ways, this was where the project began, with an idea of creating these highly curated experiences in order to advance our educational mission.

Over the course of the project, staff began to appreciate data visualization possibilities on the other end of the spectrum, where the data that is created and maintained by MNHS is let loose in the world for others to create visual meaning. Between the time the initial MNHS proposal was submitted in 2011 and work was completed in late 2014, MNHS became an active participant in the Digital Public Library of America (MNHS' public collections catalog will be added to the DPLA in the spring of 2015). The concept of making metadata available on the web with Creative Commons copyright licensing for user-driven mashups and reuse is an exciting idea that moves us away from being a centralized authority of meaning and instead a provisioner of clean data sets for innovative visualizations that can be created by anyone online.

At both ends of the spectrum, the staff emphasized the primary importance of the original data sets, of creating clean, usable, and well-constructed data that could be used by staff, vendors, or the public for visualization. Some staff felt this was the biggest lesson of all in the Sparks project: the importance of good data regardless of the way it was put into the world for education and outreach.

One final insight that emerged from our workshop was the value of data visualization for internal communication and analysis. Staff from the University of Minnesota demonstrated an in-house tool at our workshop that showed visually the geographic range of inter-library loan requests that came to UMN Libraries from across the globe. Although outside the scope of this Sparks project, the team underscored the need for MNHS to pursue internal uses of data visualization for better understanding of the many metrics we collect about our activities and the utility of visualization for communicating our findings to our colleagues.

We surveyed people who tried History Roulette on an interactive touch-screen display located in the Library lobby exhibit area. Our survey caught respondents of all ages, including children who participated with parent/guardian permission. About half of the respondents were experienced users of our newspaper and/or photograph collections, but very few had ever seen History Roulette. In general, respondents indicated that History Roulette was enjoyable and engaging. Specifically:

- The vast majority of respondents (81%) reported having a good (23%), very good (42%), or excellent (16%) experience with History Roulette.
- Half of respondents thought that it was very or extremely engaging way to connect with original documents like photographs and newspaper articles.
- Almost all respondents “played” History Roulette more than once (an average of 3 spins per person) and 28% “spun” the wheel more than five times
- Just under 85% of respondents said that they planned on sharing History Roulette with their friends or family
- Most respondents (68%) chose to further explore primary sources in our collections using the provided links

Recommendations:

1. Start with good data. Our entire staff felt that the most important component of the project was the data. The strength of the data mattered more than the final visualization for all staff.
2. Know your audience. Visualization can serve a wide variety of purposes and it is important to clearly define your purpose and expected outcomes for creating visualizations.
3. Visualization can be as simple or as complex as your needs and budget dictate, with access to a wide variety of tools and an increasingly varied array of data sets on the web.
4. Presenting data visualization in a gallery or display can deepen the visitor’s understanding of and connection to archival materials.

Project Resources

A Place in Time: A Look into Fort Snelling's Past

Documents from the collections of the Minnesota Historical Society, curated and presented to reveal selected details of daily life at the fort in the mid-nineteenth century.

<http://www.historicfortsnelling.org/a-place-in-time/homepage.php>

Matthew Marvin Prezi visualization

Visualization based on Marvin's diary featuring images and transcripts to describe his war wounds during the period 1862-1863.

https://prezi.com/0uzqd24bbzf_/matthew-marvin/

Matthew Marvin's Civil War Diary timeline

Timeline visualization of Marvin's wartime experiences from 1862-1863 based on diary images and transcripts, photos, and items from the MNHS collections.

<http://bit.ly/1BnMUM4>

Minnesota History Roulette

Visualization mashup of historical newspaper headlines and photos from the collections of the Minnesota Historical Society, sharable on a variety of social media networks.

<http://www.mnhs.org/historyroulette>

Related Links

Minnesota Historical Society, Collections Online

<http://greatriversnetwork.org/index.php?brand=cms>

Minnesota Digital Newspaper Hub

<http://newspapers.mnhs.org/web/mhsnews/web/imu.php?request=access>

Digital Public Library of America

<http://dp.la>

University of Minnesota Digital Library Services, Tools + Methodologies

<https://www.lib.umn.edu/digital/dash/tools>

Authors

Jennifer Jones, Director of Division of Library and Collections

Jim Ockuly, Web Project Manager

Shawn Rounds, State Archivist

Visualizing History Logic Model (1.30.12)

