

6-3-2018

A. Introduction to the Collection

Robert W. Lankston
University of Montana, Missoula

Let us know how access to this document benefits you.

Follow this and additional works at: https://scholarworks.umt.edu/mmp_metadata_research

Recommended Citation

Lankston, Robert W., "A. Introduction to the Collection" (2018). *Mining the Internet for Metadata*. 1.
https://scholarworks.umt.edu/mmp_metadata_research/1

This Book is brought to you for free and open access by the Mansfield Library at ScholarWorks at University of Montana. It has been accepted for inclusion in Mining the Internet for Metadata by an authorized administrator of ScholarWorks at University of Montana. For more information, please contact scholarworks@mso.umt.edu.

Mining the Internet for Metadata

Introduction

My entry into the world of digital archiving was not something that I ever imagined for myself. As a senior in college, I took a course in folklore. The course required a term project that consisted of making a collection for the university archive. I collected stories in the folk tradition, i.e., told to me orally, from my geology classmates. The stories related to the vicinity of Whitehall, MT, where Indiana University conducted geology courses in the field during the summers. Maybe the assignment in that folklore course lit an ember that glowed for a long time before actually burning.

In 2006, I took my retirement benefit from an international petroleum company. My wife, Marian, and I moved from the big city to Missoula, where we had lived during my graduate school days 35 years earlier. Marian took a job at the Maureen and Mike Mansfield Library on the University of Montana (UM) campus. I had some part time consulting work for several years for the company from which I had retired, and I had opportunities to teach geophysics at UM for a few semesters.

My reconnection with the UM Department of Geosciences occurred, initially, through my interest in archiving physical documents that had survived in the Geosciences Department since 1970 when a seismic survey of Flathead Lake was conducted. The K. Ross Toole Archives at the UM library agreed to take custody of the paper documents and to store them appropriately. Prior to delivering the documents to the Archives, I had each item scanned. The scans then became the starting point for a digital archive.

In developing the digital archive of the [1970 seismic survey](#) on the ScholarWorks platform at the UM library, I did a tremendous amount of internet research because the 1970 project had been documented in only one publication, and that publication gave little information on the equipment that was used and the procedures that were followed. That research actually continues to a certain degree as people discover the ScholarWorks website and offer feedback on what I have placed in the digital collection.

As I mentioned above, Marian was working during that time at the UM library, and one of her tasks was cataloging images of photographs in the UM photo collection that were being scanned and placed into the Montana Memory Project ([MMP](#)) digital collection. One day in early 2014, Marian sent me a simple request based on an image that she was cataloging. My email archive does not go back that far, but some of the early questions about photographs that Marian posed were of the form, “when was Sentinel High School built”, “which Missoula church is this”, and “based on the cars, what is the approximate date of this photo?” The questions were posed with the hope that with some research, aspects of the metadata for the respective photos could be completed or even improved.

While questions such as those above seem straightforward enough, finding the answers can lead in unexpected directions. For example, in those cases above, the date for the construction of the school was found easily enough, but the photo was not of Sentinel High School under construction, and the church was not in Missoula. The stories of the research trails for these and other cases can be found in the research accounts on this site.

Those occasional inquiries from Marian in early 2014 grew to a steady stream of inquiries, and my relationship with the UM library as a volunteer was formalized. Since starting on this effort, I have commented on 300 or more photos, and I have looked at, perhaps, twice that many more in the MMP collection. As of this writing, the UM library has loaded more than 7000 scanned photos from its collection onto the MMP site. The potential for much more metadata research would seem to exist.

My Objectives for the Research

As I started to respond to specific requests from the UM library staff, I realized that the questions tended to be in just a few areas. The questions related to providing one or two elements of the metadata for the respective images. In general, missing or questionable metadata entries resulted from transcription of information available with the original physical photograph into the digital archive, i.e., not usually from data entry errors.

Date and Description are the main fields that I study. Edits in these fields might trigger changes in other metadata fields such as Geographic Coverage and Time Period Represented.

The Description field is fairly open in that it could contain any or all of the following: the location of the scene, the people in the scene, the names of vehicles and machinery in the scene, or information about an activity in the scene.

Probably, the most common question that I try to answer relates to the date of the photo. The date is elusive. If the original photo is not dated in some way in the physical archive from which it was scanned, the date must be resolved to the degree possible with research. In the case in one story in this collection, my research not only found the exact date but also bracketed the time of day to within an hour or two. In another case, I was able to narrow the date to one of two days. These were exceptional cases.

Generally, if I can provide any information for the Date field, the date becomes Circa ####, which generally indicates a date range of 3 years, or the date field becomes a Before #### or After ####, as appropriate. These approximate dates are usually based on automobiles in the photograph. Stars on an American flag that is in the photo may help to bracket a time range. The style of women's clothes can help to bracket dates, also. Men's clothes are not quite as distinctive in this regard.

Perhaps as often, or being a close second after date, is a request for the location of a photo. A vintage car parked in front of some trees or that church presumed to be in Missoula but with nothing in the background of the scene can be hard to locate. In other cases, some background feature, e.g., a mountain behind the main subject of the photo, allows the location to be defined fairly accurately. In the case of city houses or commercial buildings that have been destroyed by fire or torn down for some reason, old maps might be able to help in setting the location. I am successful in defining the location of some scenes, but the locations of many are still unknown.

After date and location, the next most common request is "who are the people in the photograph?" For example, a photo may show a well-known person, such as Franklin D. Roosevelt, in a group of people. The Title and Description metadata fields may name the president but not the other people, some of whom might have been important figures

themselves. A person doing research on one of those other people in the photo might be happy to get a link to the MMP image.

Occasionally, I get a question about machinery. Railway locomotives and rolling stock, aircraft, farming and ranching implements and tools, construction equipment, logging equipment, and mining and oil field equipment are common objects of machinery inquiries. In a view of machinery, the request from the library may inquire as to what is happening in the photo. If such an explicit request is not made and I have a sense of what the activity is, I will provide that in my response to the library staff.

In my responses to the library, I will often note features in the image that are not the central subject, but, like the group of people in the Roosevelt photo mentioned above, those other elements in the photo may be of some significance to one researcher another. One example of this was an aerial photograph of a lumber mill in Missoula. The mill was in the center of the frame, but in the background, one could see another lumber mill, the sugar beet processing plant, and a drive-in movie theatre. All of those businesses in the photo are gone now. Not in that aerial view was the Russell Street Bridge over the Clark Fork River. The original bridge was completed in the late 1950's. The observation that the bridge was not in place would have helped to bracket the date of the photo. In this case, however, the date, April 30, 1955, was already in the Date field of the metadata.

In summary, I try to provide information on the date the original photo was taken, the location of the scene, the names of any people in the scene, and the names and uses of machinery in the scene. My responses to the library staff cover the full spectrum from a much usable information to no information, as you might imagine.

As I saw the pattern of requests that I was getting, i.e., dates, locations, people, and so forth, I began recognizing metadata issues on my own when I would open an MMP photo as part of my research on another photo. In such cases, I would note the image's serial number and return to study it when the time was right.

My Sources of Information

As the title of this collection suggests, my main source of information is the internet. An internet search will usually turn up some textual material in Wikipedia. In researching images of automobiles and locomotives, I will often choose the images option for the links returned by the search engine. I can search on "1951 Hudson", for example, and see many photos of 1951 Hudson cars ranging from rusting away hulks to shiny restorations. I can often find a view of an automobile in which the automobile is oriented in a way that is similar to the orientation of the car in the image on which I am doing research. This helps in resolving subtle differences in features such as grille, tail lights, and others that might vary from one model year to the next.

I use a variety of resources for resolving locations. Google Earth, Google Maps, and Google Street Views are useful. Google aerial views are useful if I am looking at roof outlines and attempting to correlate them to building outlines in Sanborn Fire Maps. Some Sanborn maps are online, but, occasionally, I will look at the physical map books that are available in Archives at the UM library.

Google Earth is helpful in attempting to replicate the viewing angle of an old photograph. Google Earth and Google Maps provide a Street View feature. Of course, the views were made

within the last few years. Sometimes, however, the Street View will be clear enough to simulate the view of the photo being studied. I call using Street View “driving my Virtual Car”. Do you remember the question above about the church in Missoula? That was found during a drive of my Virtual Car right to the spot where the photographer originally stood. That trip in the Virtual Car is described in one of the accounts on this site.

Google Earth ground level and aerial views provide my “Virtual Aircraft”. I use my Virtual Airplane for rapidly flying along above roadways, railroads, and rivers, and I use my Virtual Helicopter when I need to hover and descend for a closer look at the ground.

The Sanborn Fire Maps are an amazing resource for resolving building locations and street names from the late 19th century until well into the 20th century. The maps show building outlines and general construction style, i.e., frame versus brick. Porches on the buildings are also indicated. I have often used an aerial Google view to look at the roof of a 100+ year old house and see the exact same roof outline as in an old Sanborn map. The Sanborn drawings will also capture a turret on a house, and that can be a distinctive feature when trying to specify the location of a house or other building in a photo. The Sanborn maps show the locations of fire hydrants. A feature in a photo of the Eddy, Hammond and Co. store in Missoula can be correlated to the location of a fire hydrant in an 1880’s Sanborn map.

Another map source that I find useful is the US Geological Survey (USGS) collection of historical topographic maps. A map in that collection played a role in the story that evolved from the question “when was Sentinel High School built?” Important features on the old USGS topographic maps are the locations of highways before the interstate system was built and the locations of railways. An old map of a certain area may show a railway of interest. Many times, however, no old map of the area of interest is available. In a modern version of the map of the area of interest, however, the map may show a path labeled as an abandoned railroad grade where the railroad once ran. In some cases, a Google Earth aerial view can complement the topographic map in determining the location of abandoned railways.

Finally, the Montana Memory Project collection itself can be a source of information. Sometimes, a particular photographer took more than one photo of a scene. In responding to a request about one photo, I might come across another one in the photographer’s series of shots, and that other photo might have a valuable clue. For example, the metadata itself may be more complete, e.g., the exact date is given. The second scene might be from a slightly different angle such that a recognizable topographic feature or a building is in view.

An important point to make here is that I just do research. I provide input to the UM library faculty and staff, and they sort out how much of my input is of value for inclusion in the metadata for the respective images. Updates to the metadata are made by the library personnel.

The Stories

The accounts of my research published on this site attempt to give specific examples of the sources that I use to try to improve the MMP metadata entries. In a few cases, the research specifically relating to metadata turned up a sidebar story that I include. One of these relates to a young girl who lived a few miles northwest of Helena in 1890 and who received a free subscription to a children’s magazine of the day. She happened to be at the right place at the

right time. The place was a railroad junction on the H&N railroad. The railroad was actually the original target of my research, not the girl's good fortune.

This collection is not intended to be static. As research on those 7000+ images continues, new accounts will be published and some existing accounts may be edited. This collection of stories about the research is not intended to be the final word on the respective photos. Feedback from viewers that could provide additional data is welcome. Such feedback can be directed to me, and the MMP website itself has instructions for how to provide feedback. To be sure, if you find broken links, please contact me, and I will endeavor to repair them.

Each account in this collection includes the link to the image in MMP that is at the heart of the story and the identification numbers of the image in the MMP collection. Other links to MMP, the USGS, university and high school yearbooks, and so forth may also appear. In reading the research accounts, you might find using a computer with two display screens helpful.

6/3/2018