Beyond Crowdsourcing: Working With Donors, Student Fieldworkers, and Community Scholars to Improve Cultural Heritage Collection Metadata

Andrea Payant
*Utah State University*

Becky Skeen
*Utah State University*

Anna-Maria Arnljots
*Utah State University*

Randy Williams
*Utah State University*

Follow this and additional works at: https://digitalcommons.usu.edu/lib_pubs

Part of the Library and Information Science Commons

**Recommended Citation**

This Article is brought to you for free and open access by the Libraries at DigitalCommons@USU. It has been accepted for inclusion in Library Faculty & Staff Publications by an authorized administrator of DigitalCommons@USU. For more information, please contact digitalcommons@usu.edu.
Beyond crowdsourcing: Working with donors, student fieldworkers, and community scholars to improve cultural heritage collection metadata

Andrea Payant

is the Metadata Librarian for Utah State University Libraries. Her research interests include metadata quality benchmarks, outreach for technical services, assessment methods for cataloguing and metadata services, research data management, and optimising metadata creation based on user search behaviours. Her work has been published in a variety of books and journals and she has presented at numerous conferences.

Utah State University Libraries, 3000 Old Main Hill, Logan UT 84322-3000, USA
E-mail: andrea.payant@usu.edu

Becky Skeen

is the Special Collections Cataloging Librarian for Utah State University Libraries. Her research interests include outreach for technical services, assessment of special collections cataloguing training and professional development, assessment methods for cataloguing and metadata services, and optimising cataloguing records based on user search behaviours. She has presented at numerous conferences and contributed to a number of research articles.

Utah State University Libraries, 3000 Old Main Hill, Logan UT 84322-3000, USA
E-mail: becky.skeen@usu.edu

Anna-Maria Arnljots

is a cataloging and metadata services assistant at Utah State University Libraries. Her responsibilities include creating Dublin Core metadata and helping to maintain digital collections and exhibits. She holds an MA in American studies with an emphasis in folklore, and an MS in technical writing, both from Utah State University. She has presented on cataloguing and folklore at a number of conferences.

Utah State University Libraries, 3000 Old Main Hill, Logan UT 84322-3000, USA
E-mail: anna-maria.arnljots@usu.edu

Randy Williams

is an emeritus associate librarian at Utah State University Libraries. Her research interests include diversity, equity and inclusion, and she previously received a Utah Humanities’ Human Ties award for her work on the university’s Latinx Voices Project. She is also Archival Liaison for the American Folklore Society, a board member of the Folklore Society of Utah, the Vice President of Cache Refugee and Immigrant Connection, and the 2019–20 Chair of the Utah Humanities Board of Directors.

Utah State University Libraries, 3000 Old Main Hill, Logan UT 84322-3000, USA
Tel: +1 435 797 3493; E-mail: randy.williams@usu.edu

Abstract

Utah State University Libraries (USU) employs community-based crowdsourcing metadata practices that provide in-depth, collaborative strategies that go beyond more commonly used collecting methods. Improved metadata quality is a result of working closely with donors, community scholars, students, and the public to describe cultural heritage collections fairly and thoroughly. This article provides an overview of successful ways to collaborate with those outside of traditional library units to create more diverse, equitable and inclusive descriptions of archival resources.

KEYWORDS: metadata, crowdsourcing metadata, ethical cataloguing, ethical fieldwork, community-based fieldwork

Introduction

Utah State University Libraries (USU) uses unique crowdsourcing approaches, as well as the more traditional methods, to describe cultural heritage materials in Special Collections and Archives and to improve overall metadata quality for
digital collections. Efforts include outsourcing, organising community events, conducting metadata interviews and coordinating library ethnographic description. The disciplined skills of cataloguing and metadata services (CMS) and special collections and archives (SCA), when partnered together, facilitate effective metadata crowdsourcing. These units shape their practices to leverage shared knowledge to focus on deeply embedding community scholars and/or collection donors in the process of metadata creation. Their expert voices are considered the most valuable for providing accurate, diverse, equitable and inclusive descriptive content.

**Related work**
Recent scholarly works about metadata crowdsourcing largely reflect positively on the practice. The majority of these works are evaluative in nature and present case studies for specific projects. Other common themes include the prominent use of online or social media platforms as a method for involving broad ranges of participants to generate metadata. In ‘Crowdsourcing our Cultural Heritage’¹ Ridge shares a compilation of various case studies that highlights the benefits of crowdsourcing as well as essays that discuss the subject on a more theoretical level. Additional authors have written similar case studies that illustrate the usefulness of crowdsourcing and the potential uses for online tools with their respective projects. Causer and Terras² share details of a large-scale transcription initiative that successfully implemented crowdsourcing to increase access to primary source materials. Noordegraf³ remarks on the value of tagging archival audiovisual content and how participatory knowledge gained through crowdsourcing complements authoritative knowledge offered by metadata professionals. Paraschakis and Friberger’s⁴ work involves the creation and use of a social media game revolving around art collection to produce archival metadata, which, after analysis and testing, showed encouraging results as a viable method to crowdsourced metadata. Olivieri, Shabani, Liu and Sokhn⁵ discuss a mobile application that employs the use of aggregated cultural heritage metadata as well as the evaluation of user profiles leading to the identification of prospective participants well suited to perform metadata crowdsourcing tasks.

Other relevant and notable publications exist that also inform on the benefits of crowdsourcing. For example, Cairns⁶ shares how social tagging, specifically folksonomies, has increased access to online museum collections. Cairns explains the potential for co-creation of knowledge between experts and non-experts to aid in recognising cultural influences on language and developing interests among communities. In their book, Alemu and Stevens⁷ make a case for institutions to work toward creating mixed metadata for description, which is essentially combining standards-based metadata with user-constructed metadata.

The crowdsourcing work done at USU differs from common practices outlined in recent publications in that much of the focus is on direct and personal interaction with stakeholders. While USU utilises online tools to successfully crowdsourc metadata for public collections, the output is less significant when compared with the results of other, more in-depth methods. As new collections are created, concentrated efforts are made to collect as much information as possible from collaborators and experts. The following work offers an overview of the outsourcing, community events, metadata interviews and ethnographic description the present authors have implemented for efficacious crowdsourcing and improved equitable representation within descriptive metadata.

**Outsourcing**
Outsourcing involves the production of descriptive metadata by stakeholders outside the CMS unit. Stakeholders consist of Special Collections and Archives curators, students conducting fieldwork projects, collection donors, community partners and digital collection users. These groups have been identified as those who can best help achieve the goal to
improve collection descriptions. Special Collections and Archives curators often possess additional familiarity and understanding of items they oversee in a collection, making them potential sources of knowledge. Student groups work to contribute to the creation of specific, themed physical and digital collections and can gather metadata for materials in real time as they are made. Collection donors and community partners, often called community scholars, are uniquely qualified to give thorough descriptions of items because they are often the creators of content or are closely associated with the subject matter or area of research.

USU’s outsourcing efforts are coordinated through instruction and one-on-one meetings. When instruction is given, or in-person meetings are conducted with donors, student groups and others, those who will be creating descriptive content are taught the value of metadata and how such data should be structured, as well as the significance of fairly representing stakeholders involved with a collection. Of course, being fairly represented is also vital to community partners and donors, and their expertise is taken seriously. Guidance is given on the use of a metadata template spreadsheet, which is the key data-gathering tool used for outsourcing metadata description. This simple tool asks the following questions regarding the creation and subject matter of a resource: who, what, where, when and how. Participants create primary metadata for items (all properly named using a USU naming protocol) as they enter information into this spreadsheet, which also aids in giving the data preliminary structure and added context. The spreadsheet is submitted with associated digital object files for processing and incorporation as a new digital collection.

The spreadsheet is also used to repurpose the metadata to create Encoded Archival Description (EAD) guides for this material. All metadata produced outside the CMS unit undergo a review for quality control, followed by editing and augmentation to meet the best practices, application profiles, and applicable regional, national and international standards.

**Community events**

USU also organises community events with specific stakeholders as an effective crowdsourcing technique. These events focus on a specific topic or collection in SCA and involve collecting new materials as well as descriptive metadata. Importantly, such occasions allow CMS and SCA staff to interact with the community of researchers, donors and area experts where they live, work and recreate.

When planning crowdsourcing events, USU library staff follow basic guidelines. The following information can be adjusted and reordered to meet the partner’s needs — for instance, with some events, objectives are determined first, for others it is important to work with partners to evaluate the needs of each stakeholder before figuring out the ultimate goal:

- determine objective;
- research relevant community organisations/events;
- partner with community liaison(s);
- submit and receive institutional review board approval if needed; and
- establish strategy to reach goal:
  - where (location of event),
  - what (resources needed),
    - staff
    - equipment
    - forms
  - who and how (publicity).
Event logistics

For the crowdsourcing events to be successful, the project team first needs to decide what the institution wants to get out of the experience. For instance, are there existing items that need identification and description, or is the planned activity focused on collecting new donations and gathering descriptive metadata? Once an objective is determined, it is necessary to research relevant community organisations or events that are similar in scope to the established goal.

One USU event involved describing previously unidentified photographs, focusing on Mendon, Utah, a small town ten miles from the university. After researching community organisations, the team decided to reach out to the town’s historical society for help with the project. Finding them very open and receptive, the team endeavoured to find a member of the historical society or another knowledgeable community member with the right contacts and influence who could act as community liaison for the project. Community liaisons play a vital part in a successful project. They can help determine the best venue for an event, and the type of resources needed. Most importantly, they contact and encourage community members to attend.

When arranging such a large-scale event, it is important to find a location that is well known and easily accessible for the targeted audience, and that can accommodate the equipment, materials, library staff and community members needed to accomplish the objective. Moreover, it needs to be a place that inspires stakeholder ownership of the project. Other logistical considerations include figuring out equipment, staffing and transportation needs. When Utah State University’s Special Collections and Archives hosted the Wetlands History Project community event at the Bear River Migratory Bird Refuge, located 40 minutes from the university, it required vans to move equipment, supplies and staff to the site. This took extra planning but was needed to achieve the goal.

Figure 1: Metadata Specialist (Becky Skeen) and SCA Photograph Curator (Dan Davis) gathering metadata from informant (M. Craig Dangerfield) at the Wetlands History Project Event, May 6, 2017

When it comes to staffing an event, there is a fine line between too many and too few. Too few staff might leave people or donors waiting; too many will result in some having nothing to do. Everyone needs to be busy and engaged, but not feel overwhelmed. Furthermore, it is especially important that staff assigned to collect metadata have sufficient time to talk with each person in order to get the rich, diverse and effective information that is desired. This not only
helps with metadata creation, but also promotes inclusivity and project ownership within the designated community. Special Collections and Cataloging and Metadata staff often collaborate with community partners long after the collecting day, which builds trust and aids ongoing digital and physical donations and collection management. Along with staffing, it is necessary to create the forms and templates needed to collect the required information as quickly as possible. For USU sponsored events, these forms are used for tracking purposes and to record associated metadata.

Figure 2: Example of possible staffing for a crowdsourcing event, including roles and responsibilities
The next step is to publicise the details of the event. Outreach and publicity for such an occasion can be difficult; however, the best way to alleviate these concerns is to use the expertise and community knowledge of the project liaisons. USU finds the most successful promotional materials are posters, fliers, community newsletters and word-of-mouth (e.g., event announcements at community meetings, host site staff telling their patrons about the event, etc). Additionally, the project team actively shares USU’s crowdsourcing methodology at regional and national conferences and in publications.

Figure 3: Publicity (flyer) for Mendon, Utah Digital History Collection
**Long-term considerations**

While preparation is the key to staging a successful event, it is also important to think past the event itself. One of USU’s biggest challenges was how to gather metadata from community members after the experience was over. For the Wetlands History Project, community members brought in more materials than could be processed during the designated timeframe, which required donors to lend items to the library for digitisation. However, without the donors being present at the time of digitisation, the problem of efficiently gathering the descriptive metadata emerged. Two primary solutions to this problem were formed, each depending on the donor’s computer literacy and usage. If the donor was not comfortable using computers, a physical metadata collecting booklet was created and sent out to the contributor, who then completed the booklet and mailed it back to the library in a self-addressed self-stamped envelope. The same process was also used for those with higher computer literacy skills, but, instead of a physical booklet, the donor was sent a basic metadata spreadsheet to fill out and then return to the library.

USU’s crowdsourcing events produced numerous positive outcomes. Relationships with community members were strengthened as participants became excited to share their stories and expertise. Furthermore, the turnout was excellent, yielding long-reaching benefits for all involved. USU gathered robust metadata, while community members donated to an enduring public project. Finally, as USU planned and implemented events, and dealt successfully with the aftermath, it provided an opportunity to take a second look at existing workflows and identify ways to streamline and better facilitate metadata gathering.

**Metadata interviews**

Crowdsourcing for collecting and obtaining metadata for digital (or physical) projects can also include personal metadata interviews, utilising an informant with first-hand knowledge of the collection items or the subject matter. Crowdsourcing using metadata interviews offers an opportunity for a primary expert and metadata specialist to hybridise the process of creating descriptive content. The following segment highlights two Utah State University collections for which metadata were created using this in-depth crowdsourcing strategy.

**Barre Toelken’s Fieldwork Image Collection**

The Barre Toelken Fieldwork Image Collection\(^\text{1}\) represents half a century of folklore fieldwork. It was created by internationally renowned folklorist, Barre Toelken,\(^\text{2}\) and serves as an example of when a personal interview with the creator is most beneficial for producing metadata. Toelken’s collection encompasses a wide range of genres, such as customs (cemeteries), folk communities (Native Americans, mostly Navajo), folk art and customs (Austria, Germany, Japan and the USA). In developing this collection, both SCA and CMS team members worked with Toelken in order to provide an intimate view and understanding of the multi-themed collection, which resulted in detailed and accurate data with which to create the best possible metadata.

Over a two-month period, Randy Williams, Fife Folklore Archives Curator, and Heidi Williams, SCA student employee, worked with Toelken and his wife Miiko Kubota Toelken at their home to gather information about each of his fieldwork images, adding these data to a metadata spreadsheet. From these efforts, Anna-Maria Arnljots, USU Metadata Specialist and folklorist, created the metadata for this collection. Arnljots checked the metadata item by item to ensure quality control. She also worked directly with Toelken and his family (making numerous visits to their home and later the care facility), in order to gather as much data as possible to bolster the metadata and descriptive narrative for the nearly 1,300 items. In this case, the metadata specialist served as a conduit between the researcher seeking data and the informant with first-hand knowledge of items in the collection.
Without these personal metadata interviews, it would have been impossible to correctly identify the important details necessary for rich metadata — not only the date and time, names, locations and cultural significance, but the level of detail that only Toelken would know. He and his family provided an intimate level of crowdsourcing, yielding granularity, accuracy and quality for the data. An example of this is the Navajo Wedding Basket. The descriptive data for Toelken’s image of a Navajo wedding basket illustrates the intricate work of creating metadata and the reason that it is beneficial to have direct access to the collector creator or community scholar who understands the value of the item, rather than relying on broad generalisations from library staff or unvetted community contributors for more information.

Figure 4: Navajo wedding basket with traditional motif – Source: http://digital.lib.usu.edu/cdm/singleitem/collection/barretoelke/id/900/rec/16

By conducting personal metadata interviews with the expert, one can learn not only what the item reveals, but the nuances that may otherwise be hidden. In this case, the Navajo basket is not an ordinary basket, but rather a ceremonial basket with a distinctive pattern that serves as a metaphoric representation of the Navajo creation story. The traditional utilitarian function has evolved through the years, but ceremonial baskets, such as the Navajo wedding basket, are still in use. An intimate view of an item through the eyes of the expert illuminates and brings the otherwise unknown or ordinary to the realm of extraordinary.

Student fieldwork collections
Utah State University’s Special Collections and Archives collaborates with the institution’s Folklore Program to preserve examples of folklore in the 50-plus-year USU Student Genre Collection, generated by students in folklore courses. A detailed guide for folklore collecting helps student ethnographers/metadata specialists. The assignments require students to conduct personal (or self) interviews with informants on ten folk topics, such as customs, beliefs, tales and jokes, and legends. One notable example is St. Anne’s Retreat — a local place about which supernatural stories, laced with ancient themes, are shared within the local majority group. The SCA’s Fife Folklore Archives Student Genre Collection: Supernatural, Non-religious Legends contains over 140 examples of stories relating to the legend, collected through interviews with local teens and college students. These stories depict various themes and
versions. One example, told by Diane S., a senior at USU was told to student ethnographer Natalie H., in February 1985, while travelling past the derelict location of St. Anne’s Retreat. Diane noted that she learned this legend while growing up from the ‘big kids who always told this to little kids to scare them’:

‘There is an old nunnery up Logan Canyon. Years and years ago the nuns were sent up there who got pregnant by the priests so no one else would find out about their indiscretions. There was a main house where everyone would meet and there were also four smaller houses where they would sleep. And there was also a big swimming pool. And as soon as the babies are born the babies were to be drowned in the pool. Now this is old and deserted, but if you go there at night you can still hear the babies cry.’

Student crowdsourcing (item collecting and metadata), using the informant data required as part of class assignments, is an example of metadata collection from students conducting folklore fieldwork by interviewing individuals participating in the folklore (in this case a legend). In this case, the informant presents their own version of how they perceived their individual experience (here, a visit to a haunted place), or by re-telling a story they may have heard from a friend.

The metadata elements gathered by the student fieldwork assignment are fairly detailed, including title (name of story); genre (folklore genre the legend belongs to); informant (narrator of the story); context (where the story came from), text (experience is quoted verbatim), and texture (stylistic notation of how the informant presented the information). This in itself provides significant data about the collection and its content and offers important details for the metadata specialist. The major role of the metadata specialists involved with student fieldwork collections is as a trainer and a resource. They provide education on proper metadata practice, especially with the use of encoded archival protocols. This may also include the appropriate use of terminology taken from various thesauri and controlled vocabularies.

One commonality between the crowdsourcing examples is that the conduit, i.e. the metadata specialist, remains the expert on how to best translate and parse the content of the various accounts and thereby create more structured and useful metadata. The folklore curator and folklore-trained metadata specialist use a folk lens when describing these collections, recognising areas of cultural sensitivity, terminology and other nuances that might otherwise go unnoticed. Taking the time to conduct personal interviews with donors/creators and informants in student folklore collections, is to more inclusively and authentically describe these rich cultural collections, which is ultimately central to the discoverability of the collection.

Ethnographic efforts

Similar to the personal interview, SCA also engages in community-based ethnographic fieldwork efforts (oral history) to gather, describe and preserve the ‘voices’ of diverse communities in the collecting purview. Ethnographic efforts serve as hands-on teaching tools, like the student fieldwork collections, and provide a rich area for metadata crowdsourcing endeavours. These efforts help improve representation of diverse, often excluded communities, in an archive, and help bring equity and inclusivity to the historical record. Over the past 80 years, folklore fieldwork practice, bolstered by advancements in recording technologies, has enabled researchers to document the lived experiences of the everyday man and woman, from the bottom up. Additional progress has been made through the work of curators, archivists and cataloguers to help researchers make this rich ethnographic material accessible and discoverable. In fact, archivists have evolved their professional scope to include active participation of community scholars in all aspects of the ethnographic work. For the past 20 years, Randy Williams has laboured to move this work forward by collaborating with historically
excluded communities to bring the voices of often-overlooked communities into archival records. Williams and Duncan note:

‘To do this work ethically and responsibly, these projects must begin with community inclusion at all levels of the work, from concept through completion and presentation. USU curators plan their oral history work using ethnographic documentation best practices and obtain institutional review board approval for trained interviewers, including those from the community’.

Today, this work includes involving community scholars to describe archival products — audio, transcript, images, associated documents — from community-based ethnographic projects. This work builds on efforts often undertaken with donors and ultimately adds to the quality of metadata that accompanies objects in digital collections.

USU’s Voices from Drug Court: Cache Valley Drug Court Oral History Project provided a first step with this work. The Drug Court project was the brainchild of Andrew Dupree, then-Drug Court participant and now graduate, who suggested that USU’s ethnographic efforts be applied to Drug Court. Dupree noted in his interview with Williams and Duncan:

‘It was the last thing you [Williams] spoke about [at the Cache Valley Refugee Oral History Project presentation]. Right before the close of the presentation; and you got up and you said that it was very important to you to do this work because — I think (to paraphrase you), you said, I’d like to give a voice to those who don’t “have a voice,” or who have very little voice. And, yes that hit me really hard; it resonated really hard with me, as someone who had been currently on drug court at the time’.

Dupree worked with Williams and Duncan to create, execute and curate the Voices from Drug Court project. Language from each interview was used to create metadata — making the rich materials from the one-of-a-kind insider collection available to researchers. Using the community-approved questions and vetted transcripts, Williams and the CMS team created rich controlled vocabulary terms. The assistance of community scholars, however, was not employed to clarify or even construct that vocabulary list. When presenting on this collaborative work at a regional library conference, a librarian audience member offered additional drug-related metadata terms from her work with people with substance abuse disorder. At this point, Williams and Andrea Payant, USU Metadata Librarian, created a plan to train community scholars involved in Informing the National Narrative: Stories of Utah’s Opioid Crisis (an offshoot of Voices from Drug Court), on metadata techniques. This project is a joint effort between SCA, CMS, USU Extension faculty, and community scholars. In an effort to build on other crowdsourcing efforts, and to bring greater equity, diversity and inclusivity to new SCA collection descriptions, the project work plan included making the library responsible for training the interview team on metadata best practices in order to help with the description of interview products.

The training included sessions on naming protocols, coding and metadata (description) spreadsheet use. During the metadata/description session, Payant reviewed the importance of metadata with community scholars, while giving them latitude for description. This will not only help researchers and ‘insiders’ to search the collection but also gives the community that created the collection ownership of all aspects of the work. Thus, although the work entails more effort at the outset, the work aids in eliminating (often-unintentional) pejorative, sexist, racist, possible stigma-induced language, etc, from the metadata. Efforts to give voice to excluded and overlooked communities should first aim to do no harm, and a community-described material mindset makes this a reality, not just a desire.
Conclusion
This approach allows USU to reach its goal of partnering with excluded or under-represented communities to create and describe important archival collections that better tell the stories of such communities. Historically, cataloguing practice hindered the possibility for extensive metadata description due to limited descriptive space (eg cataloguing cards), technology (eg no world wide web), and professional practice (eg ‘this is how we have always done it’). Today, however, these limitations are a thing of the past, and as such, USU works conscientiously to create diverse, equitable, and inclusive collections. Of course, it is not enough merely to accumulate materials. Library teams must also describe the resources in the most complete and effective way possible so that these rich ethnographic assemblages give voice to local, state and national narrative that can be discovered and used by those seeking such knowledge. It is said, ‘if you build it they will come’, but for USU, the saying ‘if you describe it they will come’ is much more appropriate. The application of this philosophy ensures that people are able to find relevant materials, from a judge to a drug court graduate, from a community scholar to a distinguished professor. This philosophy is more equitable and more thoroughly represents not only the subject matter of collections but also the very people who contribute to these collections. After all, ‘metadata is a love note to the future’, giving voice to the lived experience.

References
12. Toelken B. Barre Toelken Fieldwork Figure Collection [Internet]. USU Digital History Collections. Utah State University, Special Collections & Archives; [cited 2019 Jul 25]. Available from: http://digital.lib.usu.edu/cdm/landingpage/collection/barretoelke


20. Scott, J. Metadata is a love note to the future. [blog on the Internet] [place unknown]: [Scott]; 2011 Sep 29 [cited 2019 Jun 6]; [1 screen]. Available from: