# Why Can't I Find You? A User-Centric Study and Recommendations for Locating Archival Information

Selena Aleman, Courtney Christner and Molly McGee
School of Information, University of Texas at Austin

**Abstract**
This literature review will explore the evolution of and recommendations for improving archival finding aids. The first sections will introduce finding aids and explore current studies on how users search for records and information, using case studies and user-centric research. The next section will explore ontology and EAD (Encoded Archival Description) as a user-friendly enhancement for finding aids, and discuss what archives are currently doing to implement their use. The final section will describe current experiments and ideas involving crowdsourcing as a method archives may use to deal with backlog or rapidly changing collections. This will lead into a recommendation for further research on adopting universal standards for finding aids and subjects in order to develop an archival network to improve the process of searching for resources in archives.
**Keywords**: finding aids; archives; ontology; EAD; Encoded Archival Description; crowdsourcing; archival commons;
**Citation**: N/A
**Copyright**: Selena Aleman, Courtney Christner, Molly McGee
**Contact:** mollymcgee316@gmail.com

## 1    Introduction

The transition of archives from the physical to the digital realm has brought with it a myriad of benefits and challenges, but the two most important functions of archival repositories have remained the same: 1) preservation of records and collections and 2) organizing information about collections and making that information easily accessible to visitors. Although archivists rely on a series of preservation standards to keep collections in prime condition to fulfill the first function, there is currently not a standard for repositories to organize information within a collection nor for how they make that information known to the public. Archivists by principle have remained fiercely independent as a way to avoid creating a bias in the historical record, thereby actively avoiding the creation and enforcement of an industry standard. Furthermore, while there have been many attempts in the past to create descriptive standards and subject authorities, the most commonly used collection of standard headings provided by the Library of Congress is not used universally. Additionally, the Library of Congress subject authority database is difficult to navigate and slow to update, thereby creating outdated subject headings that are difficult for archivists to change.

Over the years, archive users have grown from a small selection of experts with a thorough knowledge of collections, to a group of users ranging from academic researchers to students or casual visitors who may or may not understand how to navigate traditional finding tools (Milligan, 2006). Modern researchers are also demanding a wider diversity within collections, requesting to see materials not traditionally present in repositories. Previously, researchers could only locate collections and records by using on-site paper finding aids, documents prepared by archivists for the dual purpose of inventory organization and explaining collections while providing researchers with contextual information about a collection. Today, online finding aids and digital archives may widen the number of searching possibilities to include anyone with Internet access. Not only can repositories provide users with improved finding aids containing a network of subjects linking to other related subjects with specific ontologies, but many also offer users the ability to search their archives using keywords or subjects to discover records.

The goal of this literature review is to research current finding aid format standards and explain current efforts to improve finding aids by creating easily connected, searchable descriptive terms with ontologies by utilizing archival tools like Encoded Archival Description (EAD). This paper will attempt to project how repositories may further improve finding aids and eliminate cataloging backlogs by harnessing the power of their users. The paper will conclude by identifying future areas of research necessary to making archival finding processes and finding aids more accessible to the public and more useful to users. Specifically, this literature review will recommend further research about implementation of a universally used subject authority standard across repositories or research that seeks to answer questions such as the following:

- Are finding aids in their current formats an effective method of discovering information within and about collections for researchers and other users?

- Should finding aids be standardized for the use of all users of archives?
- How has technology affected the use, format, and flexibility of finding aid content?
- Are there any approaches that would allow archivists to maintain and consistently update finding aids and description as collections grow and change?
- If a standard subject authority does not help make a collection more accessible, what can be done to optimize the information regarding specific collections housed in repositories?

## 2       Methodology

Although this literature review discusses the implementation of ontology and linked data, the research featured here comes primarily from archival journals rather than fields with established ontologies because the focus is on optimizing archival information. While the research was focused on the specific goal of making records and collections easier to find in an archival repository, the authors examine existing descriptive standards and explain the need for a universal archival standard, a unified subject name authority, or ontology recommendations for the writers of finding aids. Since a secondary goal is to suggest actions repositories may take to improve finding resources through innovative cross-disciplinary techniques, this review also features articles on crowdsourcing in the context of archives and similar institutions.

In order to narrow our search area and limit our information to more recent publications, we decided to restrict our references to 21st century works written in English. Exceptions to this rule are several works discussing archival theory and postmodern archival criticism, which were still deemed timely and relevant because they are often cited in modern articles on similar subjects.

## 3       A Brief History of Archival Finding

In order to understand the tools of archival finding, it is necessary to understand the historical progression of archives by examining the time and culture-induced evolution of their purposes, audiences, and media. The democratization of information within an archive is a relatively modern concept, thus, influencing the types of information being sought and creating new demands for how this information is presented.

Many early archival repositories were established by government institutions for the purpose of retaining records relating to the functions of the government, legal transactions, or recording historical events. One of the first government archival institutions, the Archives Nationales, was founded in 1789 for the purpose of serving as a repository for laws, rights, and constitutional foundations of the French government. Soon this archive, and others like it, grew to include public records such as censuses, legal documents, and genealogical records that had previously been stored by local parishes (Milligan, 2006).

As time has progressed, the number of people who desire and are allowed access to archives has likewise increased. A common problem in early archives, especially those functioning under a monarchy or similarly imperialistic regime, was that only select scholars or those conducting official business were allowed access to most collections (Milligan, 2006). Governments derived power not only from carefully selecting which records to include in the archive, thus writing history to their liking, but also from reserving the right to deny access to those who did not propose to use the holdings as the authoritative power deemed the archive should be used. Many modern archivists feel they must be free of industry standards to maintain their independent non-biased position, fully cognizant of the complicated power they hold to potentially alter the historical record. However, the rise in literacy and the democratization of governments, archives, and knowledge has led to a desire for archives to provide equal and widespread access. Additionally, the recent popularity of personal genealogy research has led to a new body of researchers interested in a formerly unneeded way to search for records of individuals.

Early archival repositories only organized inventory information, known as the principle of provenance which focuses on maintaining original order and subject or donor context, rather than focusing on the broader context and meaning of a collection. Originally, finding aids were intended to serve only as collection management tools for archivists, but eventually gained a secondary use as research aids (Daines and Nimer, 2011). Finding aid organizational method was derived solely by archivists, who prefer maintaining provenance and was not meant to be seen by the public, who prefer more intuitive organization. The current archival workflow process as Purcell (2003) describes, is used by many archivists to better organize the information within a collection in a three-step process. The first step is to take an inventory of the physical collection and any other materials. Second, a representative sample of the collection is taken and used to group the collection into broad subject terms. The final step involves organizing the information collected in the collection and creating a subject based collection guide. Although this subject guide was specific to the organization studied, Purcell (2003) found the subject

guides helped researchers find materials they needed and rediscovered institutional collections that were overlooked without the subject guide.

Early paper finding aids allowed researchers to follow the same methods as archivists to discover relevant records due to the nature of early archival research, which primarily consisted of academics or historians. Traditionally, many archivists had historical research backgrounds, overlapping with researchers who were predominantly historians; therefore researchers could traditionally locate material due to similar methods and jargon used within both fields. However, with the modern diversity of archivists joining from various fields of study and the implementation of digital systems, new ways of finding for diverse researchers who are not always historical researchers must be implemented. These diverse figures entering the archival community are accustomed to a wide variety of terms, languages, and geographical backgrounds, which creates a non-uniform taxonomy. While properly designed archival systems allow researchers to find records based on controlled keywords or names, archival finding aids have also found new uses and applications in their digital format.

The evolution from analog to digital and new roles played by archivists and researchers brings the nature and purpose of finding aids into question once more. Questions as to what should be done to finding aids arise from their history and original purposes. Should finding aids hold to their original dual-purpose format, as a collection management tool for archivists and a finding aid for researchers? Or, should the finding aid system be redesigned—perhaps split into two separate systems—in order to be most useful for both parties concerned? More importantly, how are finding aids used, as opposed to perceived usage by archivists, and how can repositories adapt their practices to meet the needs of both internal and external users?

## 4        User Interaction with Finding Aids

Although archivists have frequently discussed efficient methods to organize and represent the overall information within a collection, few have studied the information-seeking behavior of users accessing the information within a repository. Some archivists have suggested researchers should have access only to the inventory while others have argued that a note on the collection and an index of subject terms should be included for ease of search. In the 1960's, archivists began a collective discourse regarding the observation of many researchers using the subject indexes included with an inventory list as access points to identify potential collections to research (Dearstyne, 1993). Nearly ten years later, these observations led to the development of the Machine-Readable Catalog (MARC) record becoming an industry standard adopted by the guiding professional organization, the Society of American Archivists, to organize the subject terms used to describe a collection (Hunter, 2003). Many saw the development of MARC records as an information breakthrough to help standardize the organization of information within a collection, yet as previously stated, few have studied the efficiency of this breakthrough. Additionally, these new methods employed by archivists are adapted from traditional cataloging methods developed by librarians. As repositories collect a larger diversity of materials, these traditional methods must be updated and specialized to fit with the needs of a repository, not a book in a library.

Moreover, there is not an industry standard for MARC record creation or subject headings; some repositories add these to all collections while others may only add these to select collections and some repositories do not use subject headings at all. Archivists can select a universal taxonomy found through two databases maintained by the Library of Congress, yet many archivists fail to use this standard. Once again, the databases used by archivists for subject headings are maintained by librarians, these terms are created to catalog books and may not necessarily reflect the items within an archival collection.

Adding to the standards set by libraries for repositories, the Library of Congress, has suggested collections should contain a maximum of six subject titles, however most finding aids have been shown to contain well over the recommended number of subject headings (Cox and Czechowski, 2008). This lack of standardization for finding aid subject use has made archival information seeking behavior difficult to study. To further complicate the research of finding aid effectiveness, there is not a standard for finding aids; several repositories have their own style guides, some post full finding aids online while others only post an inventory and others only use paper finding aids, because of individual archivist preferences, institutional practices, or available resources.

The existing research is unclear regarding the effectiveness of subject guides, and lacks clear recommendations and calls for action. One study found novice researchers struggle with the jargon used within an archive and finding aid, therefore many researchers do not even rely on subject indexes to ascertain a collection (Daniels and Yakel, 2010). Many novice researchers have even suggested repositories should provide access to comments written about a collection from previous researchers, not an archivist, who many found to be unhelpful to those unfamiliar with archival jargon (Yakel and Torres,

2003). Daines and Nimer (2011) further noticed the distinction for finding aid organization between archivists and novices; archivists preferred an organization like the provenance standard used with most finding aids, while novices preferred an ontologically-based system with a more limited context.

The disconnect between researchers and archivists muddles the open access mission of many repositories to provide access to the myriad of materials found within their collections. Many find this disconnect as a reason to avoid archives, rather than explore and contribute to the information within. Finding aids are usually created by one person who may not fully understand the context of a collection, and some finding aids have been found to contain non-factual information, outdated offensive language, or not representative of the people or communities of the collection (Caswell, 2017). Some archivists are aware of complications with representation, however, the information entered in an archive will serve as an official historical document. Adding non-traditional voices and viewpoints to the stringent historical record, gives credibility to groups previously unrecognized, a complicated issue currently being discussed by archivists and other interested parties. On the other end of this muddled situation of adding new representations to the archive, many archivists feel they must be free of industry standards to maintain their independent non-biased position. Archivists are fully cognizant of the complicated power they hold to potentially alter the historical record.

Thus, the focus shifts from how archivists can improve their own finding aids to how can archivists use the suggestions of researchers to allow greater access to users. Archivists can benefit from information seeking behavior research to create more efficient searching methods for researchers while maintaining traditional archivists job duties. Archivists can maintain traditional processing methods including arrangement of a collection and creating basic descriptions passed on to crowdsourcers for approval or suggestions. Additionally, can crowdsourcing the information provided to researchers regarding a collection introduce greater organization as opposed to archivists using a universal taxonomy?

## 5    Implementation of Ontology with Archival Tools

The traditional method of finding information within an archive is the finding aid. Its goal is to articulate for the researcher: the creator, dates of creation, extent, types of material used and topics referenced as well as the contextual knowledge of the collection (Gracy, 2015). Finding aids are vital for researchers to help them search for relevant information located in collections; however, they have not fully adopted new ontology implementation through various standards such as Machine- Readable Cataloging (MARC), Encoded Archival Description (EAD), CIDOC Conceptual Reference Model (CIDOC CRM), and Dublin Core (DC). The benefits of using such standards with access points as descriptions allows keywords associated with collections to be linked to data either internally (located in the archives catalog) or externally (to internet sources). Linked data can provide better contextual information throughout one archive catalog and could possibly link related material located across local, state, or national archival institutions, thus creating a more optimized way for searching for material within the archive or within a network of archives.

Gracy (2011) describes archival description and linked data in a recent preliminary study of some of the challenges and possibilities of linked data in EAD finding aids as well as MARC records.  If the potential access points in MARC and EAD cataloging practices are semantically tagged to other material either within the archival catalog or to other material located in linked institutions, it can provide more significant information to the researcher as well as assisting the archival staff in delving deeper into their own research on much broader topics (Gracy, 2015). While the primary goal for semantically linking data should be to better assist researchers, it also assists archivists in their own research for their collections. According to Eito-Brun (2014), "...with EAD-encoded finding aids, it would be possible to access detailed descriptions of records and information resources available at archives located worldwide; fonds and collections related by provenance but geographically or administratively dispersed could be virtually integrated on the internet." The possibilities for bridging provenance locally, statewide, nationally, or even global is no longer an impossible feat, but more of an under-utilized reality. The benefits to having records linked semantically gives researchers in the institution greater access to material, grants archivists the ability to perform archival tasks with ease, and provides greater contextual provenance as records can be linked beyond physical boundaries.

Another tool used to implement ontology is CIDOC CRM. CIDOC CRM's main goals are to promote a shared understanding by providing a semantic framework in order to map information, facilitate information integration between cultural heritage institutions and to serve as a guide for conceptual modeling (Bountouri & Gergatsoulis, 2011). This tool is highly useful for mapping ontology between collections housed within one institution while also being implemented in conjunction with EAD. As many

archival institutions are not networked to one another yet, building an ontology internally for one institution would help researchers visiting that particular archive. The problem with CIDOC CRM is that a domain expert must operate the tool (Bountouri & Gergatsoulis, 2011). Archivists do not always have the technical skills needed to operate with CIDOC CRM and therefore the tool is often not used.

Lack of technical skills are not the only limits to utilizing linked data in archives. EAD's current structure reinforces the archival practice of placing emphasis on structural elements of the document rather than on the various data points within the document (Gracy, 2015). This stems from finding aids' original usage as an inventory organization method rather than a finding tool. While changing how EADs are structured seems like a simple fix, the fact that a large percentage of collections held within archives have no description at all while waiting in the archival backlog poses an even larger problem. These collections and the inventory information associated with it are hidden from researchers because basic processing has not occurred. Even if the institution begins to use linked data by using EAD, a clear majority of the collection will still have not undergone initial processing, not to mention the need for item level descriptions necessary to link items to keywords. Another large hurdle for archival institutions to overcome is standardization of description. Even the Society of American Archivists Working Group on EAD does not speculate on overall standards for descriptions as the international community uses different standards (Society of American Archivists, 1999).

The archival community needs to increase the interoperability (Bountouri & Gergatsoulis, 2009) of archival finding aids as an initial step towards increasing the implementation of ontology. Establishing a standard across archives is the first step. Each archive has its own way of describing their collection and while two institutions could both use archival description methods like EAD and MARC, their initial frameworks are different. Consistent standards in finding aid taxonomy choices are essential to building ontologies. For appropriate data to be linked, the initial taxonomy must exist to continue to build on related topics. While ontology attempts to build relationships with different words, two different frameworks cannot communicate together effectively. To remedy this issue, a crosswalk method (Bountouri & Gergatsoulis, 2009) can be proposed to build semantic mapping connections to different frameworks. To build a crosswalk, the individual must be a domain expert and this technical skill many archivists to do not have which makes mapping between frameworks difficult and costly. Defining a universal standard between archives or archival networks would greatly benefit ontologically linked data.

## 6      Crowdsourcing Archival Tasks

While technological advances have allowed for enhancements to finding systems in archives, the funding and staffing of archives has unfortunately not seen similar growth. The ever-increasing number of records and information archives receive has left archivists with insurmountable backlogs (Eveleigh, 2014). While most archival tasks, like appraisal and preservation, require specific and often complex archival skills and knowledge, others, like description, can often be handled by volunteers.

Fortunately, technology once again provides archivists with innovative solutions to dealing with inevitable backlogs. Several studies discuss the possibility of using crowdsourcing techniques in archives for tasks like cataloging, description, and transcription, and others have tested the efficacy of gamifying such tasks to improve the engagement of volunteers (Paraschakis, 2013; Ridge, 2011). While the possibility of harnessing the power of volunteers to supplement basic metadata using something like a descriptive tagging system is appealing, additionally exciting is the idea that crowdsourcing could allow archives to sustainably continue to improve or update finding aids or cataloging as collections grow and change (Eveleigh, 2014). This has become increasingly important with the advent of postmodern criticism of archival theory, which posits the idea that description is an ever-evolving process, and that records are constantly in the process of being created or recreated (Gorzalski, 2014).

Another benefit of crowdsourcing is a greater diversity of perspectives on archival records within finding aids and descriptions, rather than just the perspective of the archivist who typically creates them (Light, 2002; Newman, 2012). While provenance of a record traditionally described the relationship between the record and its originator or original owners, postmodern archival theory suggests that everyone who appraises, interprets, describes, or otherwise touches a record has become a "creator" of the record and is also included in the record's provenance (Nesmith, 2005). If researchers or even readers who form interpretations of a record are indeed part of the provenance, these researchers suggest that documentation of their contributions should be added to the record somehow. A full list of possible crowdsourcing tasks is suggested by Mia Ridge (2011), including "Tagging; Debunking (i.e. correcting/reviewing content); Recording a personal story; Linking; Stating preferences; Categorizing; and Creative responses." Personal responses and annotations from researchers would be useful to other researchers and even archivists themselves, but could add a layer of complexity into quality assurance

for archivists. The result would be a content strategy like that of other commons projects like Wikipedia, featuring a reliability (and potentially comments or annotations that are completely searchable) that can only be the result of a peer-produced and peer-edited online project (Evans, 2007). In addition, the greater diversity of perspectives would enhance the context of the record and allow for more useful networks to exist between records, creators, and users (Anderson, 2009).

Some archives and other organizations have already attempted crowdsourcing activities with favorable results. One distinction organizations like the United States National Archives and Records Administration (NARA) make is establishing that rather than enlisting the help of a crowd, they are offering a chance for amateurs ("citizen archivists", in NARA's case) to contribute to the value of the collection by adding descriptive tags to photographs or records (Owens, 2013). A task requiring less expertise but that adds infinite value to collections is transcription, famously used by ReCaptcha, but also implemented by NARA, the Smithsonian, and multiple university libraries and archives. Not only does crowdsourcing enlist the help of the public to improve collections, it has also increased engagement between these volunteers or users and the contents of the archive, in some cases even drawing in donors who might not have learned about the archive's projects otherwise (Owens, 2013).

Aside from the quality assurance question, why are archives so slow to involve users in the creation of documentation and description? Although the technology is readily available and the need for assistance is undeniable, archivists have been hesitant to incorporate user input or commentary into description or cataloguing primarily, some researchers posit, because it involves sacrificing a certain amount of archival power associated with description and arrangement (Gorzalski 2014; Noordegraaf et. al., 2014). In order to preserve that power yet still make use of solutions such as crowdsourcing, archivists may see their roles shift to becoming content moderators, organizers and facilitators of capable volunteers (Evans, 2007) rather than creators in the traditional sense.

While archivists have successfully leveraged crowdsourcing to achieve goals within their particular repositories, turning users to the task of linking data or creating networks between repositories seems to be a logical next step. However, before an archival network is possible, repositories must agree to implement a description subject authority standard. The need for subject authority standards has been a common theme since crowdsourcing was first suggested in Huvila's (2008) article which discussed participatory repositories as needing to include "deeper involvement and more complex semantics" rather than simply the size of the users' contributions or the simple kinds of information they could provide. Without a set of standards to guide archivists who plan to harness the power of users to improve archival resources, all attempts to do so will result in more confusion and unhelpful descriptive terms. Therefore, before crowdsourcing or user-driven efforts to improve archival networks can be attempted, an attempt at uniformity must occur, where archives must agree nationally or even globally to define the vocabulary they will use to describe their collections.

## 7      Defining a Universal Descriptive Standard

Before the shift to digital archives, description was typically contained in the form of a paragraph detailing the provenance and details of a record, such as might be found in a finding aid (Pitti 2004). However, the implementation of ontology resulted in the search ability of individual search terms and categorization by these subjects. This opened a new frontier in terms of user access in archives. However, a universal descriptive standard has yet to emerge and has left archival catalogs unable to communicate effectively with one another. Adoption of the finding aid for researcher's use led to decisions for usability but have not led to a universal descriptive method.

Quickly realizing that splitting the description into categories would enable easier search, archivists have taken steps to begin the process of standardizing names and vocabularies using resources like the following, addressed in a 2015 SAA report *Describing Archives: A Content Standard* (DACS) which initially began as a joint effort between Canadian and U.S. archives to build a content standard between the two countries (Society of American Archivists & Hensen, 2004). DACS is related to other standards like MARC 21, EAD, APPM, Anglo-American Cataloguing Rules (AACR2) and two conventions written by the International Council on Archives: ISAD (G) and ISAAR (CPF). This content description standard was created as a means of replacing rules in AACR2 and allows more specific guidance in the description of modern archival materials (Society of American Archivists & Hensen, 2004). DACS failed to bridge standards used in Canada and the U.S., but have become a standard that can be used in the U.S.

Library of Congress Authorities contain rules for formatting standardized names and other descriptive elements, originally designed for cataloging books not archival descriptions.  These authorities are set by the Library of Congress as subject headings and are often used in archival description.

However, these subject headings can often be outdated and difficult to change. The Library of Congress suggests limits to subject titles, but archival finding aids often contain more subject headings than the limit. This restricts how many possible ontological matches could be made if only a limited number of subject headings can be used. The Library of Congress also maintains two databases for universal taxonomy, but it is rarely used for much of the same reason (Cox and Czechowski, 2008).

In addition, subject headings that are authorized by the Library of Congress do not have to be the only subject headings used in an archive. Archivists may use any heading within their subject headings, which can make searching for specific terms difficult if the heading is not an authorized subject heading. Another standard implemented in the Library of Congress is Anglo-American Cataloging Rules 2 (AACR2). AACR was first implemented in the Library of Congress in 1967. These rules used different descriptive terms for subject headings than what had previously been used and confused researchers (Taylor, 2012). AACR and AACR2 implemented much of the same rules, but still confused researchers. Past the confusion for researchers, most libraries and archives still used to card-catalog and were limited in space to how many subject headings could be used (Taylor, 2012). This poses a problem to linked data ontology as many subject headings can be related to one another and limiting subject headings does not build a large enough taxonomy to implement an ontology (Vitali, 2004). The goal of ontology is to relate search terms in a meaningful way for researchers to find information that is relevant to their questions and to help archivists to become more familiar with their own collections.

The last cataloging standard to be implemented at the Library of Congress is Records Description and Access (RDA). This standard could be used in a hybrid manner with MARC coding as the main standard, but RDA information embedded within it can be difficult to transfer into a completely new RDA standard. The biggest implementation factor for RDA is the lack of abbreviations needed in the standard. Abbreviation was needed in AACR2 as a means to match the card catalog which is human readable and not machine readable (Maurer & Panchyshyn, 2014). Since RDA is a machine-readable ontology, it is more likely to be used in a library and archival setting, thus, making it a better choice than AACR2.

The International Standard Archival Authority Record built ISAAR (G) and ISAAR (CPF) for the purposes of building a standard across international borders. ISAAR (G) built basic principles of archival description, implementing a statement of articulation through a hierarchy of levels from fonds to item (Vitali, 2004). The ISAAR (CPF) is built more specifically for corporate bodies, persons, and families. It allows archivists to describe information about the creator, and historical context of the record, an aspect that traditionally did not implement as its own category in the finding aid, but interwoven with other information in the finding aid (Vitali, 2004). To build a global archival network ISAAR (CDF) would be a standard to look at and is helpful in building related subjects to one another; but American and Canadian archivists have yet to implement these standards which creates problems with global ontological implementation.

However, while archival description practices have certainly improved, the goal of ontology is to relate search terms in a meaningful way for researchers to find information that is relevant to their questions and to help archivists to become more familiar with their own collections. Many standards mentioned above were implemented before ontology through technology could be of any use. The issue with having a multitude choices for archival institutions in terms of standards makes externally linked data between institutions is highly improbable. Also, there is still not a complete or universal standard for description of records that would allow the implementation of crowdsourcing techniques to create the network of archival records many researchers would like to see. Many collections may impose descriptive standards upon their own collection, or such standards may even grow to include entire states or regions, but until a universal subject standard can be established, a complete, searchable network of archival finding aids or collections will not be possible.

## 8    Conclusion
The most difficult aspect of improving the use and creation of finding aids is compromising the promise of technological advances with the restrictions and traditions of repositories and archivists. While archivists prefer a provenance based system and are worried that a digitally linked ontology will break down this standard of practice, even minimally linked data subject terms could bring users into to the archive to experience the information in its provenance context. While this literature review focuses on user-centric studies, it does not aim to remove provenance from archival practice. Instead, this literature review has discussed user-centric studies on finding aid use, current technical tools archivists use to improve finding and create additional resources for users, and possible techniques like crowdsourcing that can be used to create a sustainable model for finding aid improvement within archival collections.

Archives have developed the tools needed to implement linked data to form ontological relations between items held in their own collections and to repositories throughout the world. Undertaking such a task would be beneficial to both researchers and archivists with finding information by linking related subject terms together. However, with a lack of description standardization across the archival community, statewide, nationally, and internationally, building a relatable ontology seems to be unlikely as the amount of archival back log is extensive and in order to build an ontology connecting every record is a daunting task for archival professionals. Building crosswalks between different standard frameworks can be done in order to link subject terms, but this skill requires a domain expert. Many archivists to not have the skill associated with this kind of task and the unlinked data does not help researchers.

The difficulties of building linked data records may be eased with the use of crowdsourced archival description. While many archivists do not want to relinquish control over archival descriptions, crowdsourcing could allow for quicker descriptions which would result in easier use for researchers. Crowdsourcing also allows for more terminology to be utilized in subject headings as crowdsourcing allows marginalized communities the opportunity to describe the collections that they created. Irresponsible tagging could occur, but with oversight and limitations put in place crowdsourcing could be a viable option (Caswell, 2017).

While this paper discusses many of the benefits of such a system, some drawbacks and limitations to our research must be acknowledged. Because the idea of crowdsourced archives, or an "archival commons" is new, few experiments have been conducted, and thus it is hard to say what kind of opposition such efforts will meet from archivists. Crowdsourcing models and implementation of user-centric formats and organization prioritize the researcher (the outside user) over the archivist (the internal user). Archivists have traditionally resisted confinement to descriptive standards because it detracts from their power to describe as they want. However, further research may shed light on solutions to these problems, or enhance our understanding of how crowdsourcing or other methods could improve description and finding in archives. It may even be possible that crowdsourcing description might bypass the wariness archivist's have concerning a universal descriptive standard, if a wide cloud of search terms could all be linked back to a single idea, thus allowing freedom and diversity in description while maintaining a networked archival ontology.

Our suggestions for future research include studies on:
- Possibility of implementing a universal description standard for use in archives and libraries.
- Efficacy of crowdsourcing techniques in creating or improving description in archival finding aids.
- Further research into user interactions with finding aids and other archival search tools.
- Further study of how crowdsourcing could give voice to marginalized communities in archival settings.

A brief look at the history of archival finding reveals that technology has only enhanced the findability of records and collections. Although repositories have been slow to adapt to new technologies due to insufficient funds, a reluctance to abandon traditional methods, or the overwhelming amount of work such a transition would demand, it seems necessary that archivists remain on the lookout for new opportunities to improve their systems and collection organization. While some of the limitations imposed on repositories make it difficult to keep up with an "on-demand" culture, archivists can work to improve their finding systems, and although it seems an insurmountable task, perhaps a universal standard will one day enable all archives to join in a great network of information.

## 9    References

Anderson, S., & Allen, R. (2009). Envisioning the Archival Commons. *The American Archivist*, 72(2), 383-400. Retrieved from  http://www.jstor.org/stable/27802694

Bountouri, L., & Gergatsoulis, M. (2009). Interoperability Between Archival and Bibliographic Metadata: An EAD to MODS Crosswalk. *Journal of Library Metadata*, 9(1–2), 98–133. https://doi.org/10.1080/19386380903095107

Bountouri, L., & Gergatsoulis, M. (2011). The Semantic Mapping of Archival Metadata to the CIDOC CRM Ontology. *Journal of Archival Organization*, 9(3–4), 174–207. https://doi.org/10.1080/15332748.2011.650124

Caswell, M. (2017). *Identifying and Dismantling White Supremacy in Library and Information Science*. Presentation, School of Information, University of Texas at Austin.

Cox, E., & Czechowski, L. (2008). Subject Access Points in the MARC Record and Archival Finding Aid: Enough or Too Many? *Journal of Archival Organization*, 5(4), 51-59. DOI: http://dx.doi.org.ezproxy.lib.utexas.edu/10.1080/15332740802153476

Daines, J. G., & Nimer, C. L. (2011). Re-imagining archival display: Creating user-friendly finding aids. *Journal of Archival Organization*, 9(1), 4-31. DOI: 10.1080/15332748.2011.574019

Daniels, M. & Yakel, E. (2010). Seek and You May Find: Successful Search in Online Finding Aid Systems. *The American Archivist*, 73(2), 535-568. Retrieved from http://www.jstor.org.ezproxy.lib.utexas.edu/stable/23290758

Dearstyne, B. (1993).    *The Archival Enterprise: Modern Archival Principles, Practices, and Management Techniques*. Chicago: American Library Association.

*Describing Archives: a content standard, second edition (DACS)* (2015). SAA Standards Committee. Retrieved from http://files.archivists.org/pubs/DACS2E-2013_v0315.pdf

Eito-Brun, R. (2014). Context-based aggregation of archival data: the role of authority records in the semantic landscape. *Archival Science,* 15(3), 217-238. Doi: 10.1007/s10502-014-9215-3

Evans, M. (2007) Archives of the people, by the people, for the people. *The American Archivist*, 70(2), 387-400. Doi: http://dx.doi.org/10.17723/aarc.70.2.d157t6667g54536g

Eveleigh, A. (2014). Crowding out the archivist? Locating crowdsourcing within the broader landscape of participatory archives In M. Ridge (Ed.) *Crowdsourcing our cultural heritage*, (pp. 211-226). New York City, New York: Routledge.

Freund, L. and Toms, E. G. (2016), Interacting with archival finding aids. *J Assn Inf Sci Tec*, 67, 994–1008. doi:10.1002/asi.23436

Gorzalski, M. (2014). Examining user-created description in the archival profession. *Journal of Archival Organization,* 11(1-2), 1-22. DOI: http://dx.doi.org/10.1080/15332748.2013.866858

Gracy, K. F. (2015). Archival description and linked data: a preliminary study of opportunities and implementation challenges. *Archival Science*, 15(3), 239–294. https://link-springer-com.ezproxy.lib.utexas.edu/article/10.1007%2Fs10502-014-9216-2

Hunter, G. (2003). *Developing and Maintaining Practical Archives: A How-To-Do-It Manual*. New York: Neal-Schuman Publishers.

Huvila, I. (2008) Participatory Archive: toward decentralized curation, radical user orientation, and broader contextualization of records management. *Archival Science*, 8(3), 15-36. DOI:10.1007/s10502-008-9071-0

Light, M., Hyry, T. (2002). Colophons and Annotations: new directions for the finding aid. *The American Archivist,* 65(2), 216-230. DOI: http://dx.doi.org/10.17723/aarc.65.2.l3h27j5x8716586q

Maurer, M. B., & Panchyshyn, R. S. (2014). Understanding the Why: A Case Study in Managing the RDA Implementation. *Cataloging & Classification Quarterly*, 52(3), 259–284. https://doi.org/10.1080/01639374.2013.859644

Milligan, J. (2006) "'What is an Archive' in the History of Modern France." In A Burton (Ed.), *Archive Stories*. Durham: Duke University Press. Retrieved from http://ebookcentral.proquest.com.ezproxy.lib.utexas.edu/lib/utxa/detail.action?docID=1168379

Newman, J. (2012). Revisiting archive collections: developing models for participatory cataloguing. *Journal of the Society of Archivists*, 33(1) pp. 57-73. Doi: 10.1080/00379816.2012.666404

Nesmith, T. (2005), "Reopening Archives: Bringing New Contextualities into Archival Theory and Practice," *Archivaria,* 60, 259-274. Retrieved from http://archivaria.ca/index.php/archivaria/article/view/12523/13660

Nimer, C. & Daines III, J. G. (2008). What do you mean it doesn't make sense? Redesigning finding aids from the user's perspective. *Journal of Archival Organization*, 6(4), 216-232. DOI: 10.1080/15332740802533214

Noordegraaf, J., Bartholomew, A., & Eveleigh, A. (2014). Modeling Crowdsourcing for Cultural Heritage. In N. Proctor, & R. Cherry (Eds.), *Museums and the Web 2014: selected papers from an international conference* (pp. 25-37). Silver Spring, MD: Museums and the Web LLC.

Owens, T. (2013). Digital Cultural Heritage and the Crowd. *Curator*, 56(1), 121-130. DOI: 10.1111/cura.12012

Paraschakis, D. (2013). *Crowdsourcing Cultural Heritage Metadata Through Social Media Gaming*. (Master's Thesis), Available through Google Scholar: http://hdl.handle.net/2043/16114

Pitti, D. (2004) Creator Description: Encoded Archival Context. *Cataloging & Classification Quarterly*, 38(3-4), 201-226, DOI: 10.1300/J104v38n03_16

Ridge, M. (2011). Playing with Difficult Objects: Game Designs to Improve Museum Collections. In J. Trant and D. Bearman (Eds.), *Museums and the Web 2011: Proceedings*. Toronto: Archives & Museum Informatics. Retrieved from http://www.miaridge.com/wp-content/uploads/2013/01/MiaRidge_MSc_Dissertation_OpenAccess.pdf

Society of American Archivists (Ed.). (1999). Encoded Archival Description application guidelines: version 1.0. Chicago, Ill: Society of American Archivists.

*Describing archives: a content standard.* (2004). Society of American Archivists, & Hensen, S. L. (Eds.). Chicago: Society of American Archivists.

Taylor, A. G. (2012). Implementing AACR and AACR2: a personal perspective and lessons learned. *Library Resources & Technical Services*, 56(3), 122–126.

Vitali, S. (2004). Authority Control of Creators and the Second Edition of ISAAR(CPF), International Standard Archival Authority Record for Corporate Bodies, *Persons, and Families. Cataloging & Classification Quarterly*, 38(3–4), 185–199. https://doi.org/10.1300/J104v38n03_15

Yakel, E. and Torres, D. (2003) AI: Archival Intelligence and User Expertise. *The American Archivist,* 66(1), 51-78. DOI: http://dx.doi.org/10.17723/aarc.66.1.q022h85pn51n5800

Zastrow, J. (2014). Crowdsourcing Cultural Heritage: 'Citizen Archivists' for the Future. *Computers in Libraries*, 34(8), 21-23.

## 10  Appendix

### 10.1  Group Member Contributions to Draft

Selena
- Overview
- Introduction
- Evaluation
- User Interaction with Finding Aids
- Editing
- Scripting, filming, and editing video

Courtney
- Risks
- Implementation of Ontology
- Defining a Universal Descriptive Standard
- Editing
- Formatting
- Scripting, filming, and editing video

Molly
- Abstract
- Introduction
- Methodology
- Crowdsourcing Archival Tasks
- Defining a Universal Descriptive Standard
- Conclusion
- Milestones
- Editing
- Scripting, filming, and editing video