Archival Science (2024) 24:897–922
https://doi.org/10.1007/s10502-024-09465-7

ORIGINAL PAPER

Locating yourself in the historical record: challenges of provenance and metadata schemas in the library of congress's digital materials

Kathryn Manis[1] · **Patricia Wilde**[2]

Accepted: 13 August 2024 / Published online: 11 September 2024

Abstract

The Library of Congress (LOC) is an inherently political institution with immense reach. With 151.6 million visits and 520.3 million page views in 2022, its digital collections put the LOC's repository of materials in the hands of users around the world, informing the kinds of narratives we tell about our past for purposes of the present. While more accessible, these collections are not always appropriately or transparently contextualized, creating significant barriers to access and often perpetuating biased or offensive language and attitudes. This matter stems from principles of provenance and metadata schemas, standards that govern how context is preserved and made available. As scholars working with digital information and literacy argue, the ubiquity of attributing authority to web-based information makes nuanced, accurate, and accessible context for digital collections increasingly necessary. Shortcomings in contemporary provenance and metadata practice are even sharper in the case of image and graphic narrative collections since prevailing descriptive standards were not designed with visual content in mind. These intersecting and at times contradictory concerns demonstrate both the complicated tension between provenance's failures and its apparent necessity, and the ways it continues to affect applications of metadata. Exemplifying these complexities, we discuss two LOC case studies: the Webcomics Web Archive and Liljenquist Family Collection of Civil War Photographs. Illustrating the constraints of provenance and its circulation in metadata, these collections highlight the accessibility and equity issues that particularly impact visual materials.

Keywords Provenance · Photographs · Webcomics · Metadata · Library of congress subject headings (LCSH)

Extended author information available on the last page of the article

Introduction

An inherently political institution, the Library of Congress (LOC) houses an expansive array of archival materials that plays an essential role in the preservation and maintenance of American identity and culture. Founded in 1800, it began as a legislative library, though as its charge expanded to include the general public, it became, in effect, the national library of the United States. To make its holdings more accessible to a larger swath of people, the LOC started experimenting with digital offerings in the 1990s. Though it began by digitizing documents from about forty collections, ranging from George Washington's Papers, 1741–1799 to An American Ballroom Companion: Dance Instruction Manuals, ca. 1600–1920, it has since grown to include hundreds of digital curations containing both facsimile and born-digital artifacts. In offering these materials online, the LOC puts them into the hands of users from around the world. Illustrative of the reach and influence of these archives, the LOC (n.d.-c) recorded 151.6 million visits and 520.3 million page views in 2022 alone. But while technology makes these sources more immediately available, digital archives also "pos[e] enormous challenges for both researchers and memory institutions," as Lise Jaillant et al. explore in a recently published special issue of *Archival Science* (Jaillant et al. 2022 p 285). When inappropriately contextualized or otherwise obscured, digital archives can perpetuate biased or offensive language and attitudes or become generally inaccessible. These matters stem from principles of provenance and metadata schemas, standards that govern how context is preserved and made transparent to users.

As scholars working with digital information and literacy argue (McPherson 2018; Noble 2018), the ubiquity of attributing authority to web-based information and the embedded inequities of many of our most widely-used digital tools make nuanced, accurate, and accessible context for digital collections increasingly necessary. This is further punctuated by questions of authorship, ownership, and origin. While a substantive body of LIS scholarship takes to task the idea of single and/or narrow creatorship (Jimerson 2009; Wurl 2005) when contextualizing archival collections, fewer practical guidelines exist to implement what Nathan Sowry (2014) and others refer to as "pluralist provenance." This alone seems counterintuitive as even our most conventional XML finding aids offer far more possibility for description than is utilized in most cases (Light and Hyry 2002). In addition, surveys of existing archival finding aids, such as Katherine Wisser and Jackie Dean's 2013 study, reveal that individual institutions vary wildly on how and to what extent they utilize available EAD elements, suggesting that existing recommendations fall far short of serving diverse, contemporary descriptive needs (Wisser and Dean 2013).

Shortcomings in contemporary provenance and metadata practice are even sharper in the case of image and graphic narrative collections, since prevailing descriptive standards were not designed with visual content in mind. Indeed, image and text cataloging were discrete activities in the US until 1982, when Elizabeth Parker's *Graphic Materials: Rules for Describing Original Items and*

Historical Collections was published and widely adopted (Parker 1982). These intersecting and, at times, contradictory concerns demonstrate both the complicated tension between provenance's failures and its apparent necessity and the ways it continues to affect applications of metadata, specifically through LOC Subject Headings (LCSH) and information structuring. Exemplifying the complex relationship between availability and discoverability, we discuss two LOC case studies in this article: the Liljenquist Family Collection of Civil War Photographs and Webcomics Web Archive. Though thematically divergent, they are representative of growing efforts to offer digital archives not only for the benefit of researchers, but for the general public. With this capacious audience in mind, these community-driven archives make digital materials widely available and accessible to the masses. But like many archives, the Liljenquist Family Collection and the Webcomics Web Archive have shambolic origins, which are further complicated by the visual nature of the artifacts that they house. Illustrating the influence of provenance and its circulation in metadata, these collections specifically highlight the ways descriptive and cataloging choices inform how visual materials living in digital spaces are equitably located and interpreted.

Legacy approaches to archival description: the challenges of visual collections

Provenance, broadly defined as the origin or history of an object, is designed to protect the context of records and is associated with the more precise principles of original order and *respect des fonds* (Millar 2002; Moore 2008). It is also used to verify the "legitimacy" of collections and the actions archivists undertake to make them available. As Shelley Sweeney describes it, provenance often serves as "the basis not only for establishing authenticity and sometimes for dating records but also for their acquisition, appraisal, arrangement and description, and retrieval" (Sweeney 2008, p 194). However, as Peter Horsman observes, such a "grouping of records… in reality never existed at any one point in time, outside of the archives" (Horsman 2002, p 22), emphasizing the imposition of order on primary source materials as an inherently interpretive act. Though the profession generally agrees that the context of archives as preserved through guideposts like provenance is both important and necessary, many scholars raise concerns about the ways in which it is ultimately applied. Highlighting issues like the homogenous demographics of the LIS workforce, intentional, biased absences in the historical record that continue to impact researchers and collecting policies, shortcomings in a (albeit, understandable) preoccupation with "efficiency," and the systemic privileging of donor and institutional priorities, among other challenges (Bastian 2006; Christen & Anderson 2019; Millar 2002), archivists are highly aware of the limitations of this legacy principle.

In addition, the difficulties of archival labor—staffing shortages, processing backlogs, shrinking budgets, and otherwise under-resourced repositories—are ever present in the choices archivists make from acquisition to description to access. Arising from these tensions are opposing philosophies about how much time, attention, and intervention the archivist should undertake in preparing a collection for research. An

enduring approach to managing these contrasting needs is Mark Greene and Dennis Meissner's 2005 "More product, less process" strategy, which advocated for limited descriptive effort in favor of greater discoverability (Greene and Meissner 2005). Further underlying such perspectives and their application, though, is the belief that archivists ideally remain neutral and leave the research to researchers. As one stop along the neutrality debate spectrum, this emphasis on staying out of research remains a subject of scrutiny, particularly in light of social justice and decolonizing perspectives on libraries and archives. Such variable approaches leave us to ponder questions Horsman raised in 2002 and which remain relevant twenty years later:

> What then is the provenance of the record and to what extent does the content of the record play a role in suggesting a wider context? How far should archivists go in establishing a context that will enable the full interpretation of the record? Without expanding the context of the record beyond its immediate provenance, the record and its meaning have not fully come together (Horsman 2002, p. 283).

The problems of provenance Horsman raises are particularly germane to visual materials, a term used to describe pictorial artifacts. While there are differing theories on when and where the concept of provenance originated (Sweeney 2008), it has most traditionally been employed to contextualize textual records, the conventional priority of the archive. Joan explains it this way: "whereas photographs, prints, and watercolors are generally acknowledged to be historical, they are not normally understood as archival; while their value is readily acknowledged to be informational, it is seldom viewed as evidential" (Schwartz 2002, p 146). Consequently, the principle of provenance often does not adequately address the unique dimensions of images. "[U]nlike textual records," maintains Note (2019), "visual materials are often removed from their original locations and filed in subject files without further description. The context and purpose of an image is often not conveyed to those who were not present at the time of the event being captured." Such contextual occlusion significantly constrains how visual materials are understood and subsequently put to use.

While provenance is essential to the interpretation of images, such descriptive context also helps make them accessible through the generation of metadata. Defined by the Society of American Archivists (n.d.) as "information about data that *promotes discovery*, *structures data objects*, and supports the administration and preservation of records" (emphasis added), metadata schemas were adopted in the 1980s as LIS institutions shifted from card catalogs to databases and attention to discoverability became more central. This "data about data" includes fields such as item title, creation date, content creator, the LCSH, etc. But traditionally untethered from provenantial concerns, descriptions of visual materials often focus on the image's physical attributes (Schwartz 2002). Expounding on the problems of this "denotative fiction," Cara Finnegan asserts that image description is an inevitably subjective process that will never lead to a definitive answer to the "deceptively simple question 'What's this a picture of?'" (Finnegan 2006, p 118). She illustrates the repercussions of these discursive choices through her own experiences in the Prints and Photographs Room at the LOC. Studying visual depictions of poverty,

Finnegan examines photographs in the Farm Security Administration-Office of War Information file. As she explains, she had difficulty locating one particular photograph because of LCSH. The photograph she was trying to track down depicted a man, whom she assumed be "a farmer or sharecropper," but the image was classified under the subject heading "shack" (Finnegan 2006, p. 117). Illustrating how different ways of seeing can inform description practices, Finnegan's example makes clear how metadata affects what visual materials emerge (or not) in searches. Such naming and categorization disparities create further barriers with born-digital collections, where sheer file and collection sizes preclude manual searching (Talhoom and Bell 2022). Add to that, issues associated with web archiving and the minimal metadata that arises from it (Jaillant 2022, p 433) and it becomes clear just how much visual archival material in particular can become essentially unfindable for even the most seasoned archivists and researchers.

While numerous metadata frameworks exist, from internationally adopted standards to disciplinary taxonomies to local solutions, LCSH are both particularly relevant to the collections examined and a subject of extensive criticism. We acknowledge that there is much more to metadata than subject headings. However, given our focus on the LOC and the role it plays in how information circulates and is legitimized in the USA and our investigation of the imbrication of metadata practice with principles of provenance, we have zeroed in on this schema and the ways it informs metadata organization, generally. From Adler's (2017) work on the discriminatory histories of headings for Queer subjects, to the categorization of Native and Indigenous subjects as historical, to the decade-long grassroots movement to change the "illegal alien" subject heading, it is clear that the language libraries use has embodied consequences. Indeed, such standardized arrangement and content fields demonstrate what leading institutions deem important; namely, collectors and donors of archives, institutional cohesion and connection, quantity over quality, and, in many ways, the needs of researchers who already utilize libraries and archives. As Adler succinctly explains, "Of great consequence are the authorities under which taxonomies are produced and the cultural and political contexts that play into the question of what counts as knowledge" (Adler 2017, p. 8). Our case studies, while made available by the LOC, have significant implications for community-based research, U.S. genealogy, equity in the historical record, and the complicated ways that difference is understood, documented, and obscured. As such, both the collections themselves *and* the protocols developed and implemented by the LOC to make them accessible bear close examination as the ultimate signposts of current disciplinary standards and values regarding community access.

In sum, metadata, particularly as rendered by monolithic mechanisms like LCSH, play a significant role in the discoverability of digitized and born-digital collections and the ways in which researchers understand the materials they encounter. Far from a mathematical or purely systemic approach, metadata must strike a careful balance between under and over description, though the profession isn't always attentive to the impact of this tenuous equilibrium. As Heather MacNeil (1995) puts it:

> The tendency to describe metadata in metaphorical terms, e.g., in relation to archival inventories, has distracted attention from consideration of what.meta-

> data are in substantial, concrete terms. They are, in fact, *records created and used in the conduct of affairs of which they form a part*. (MacNeil 1995, p 26, emphasis added)

Further complicating provenance's limitations, metadata schemas must also make visible both the *content* and *origins* of library materials (Deelman et. al. 2010; Hartig 2009), which is particularly fraught when image and text coexist significantly in archival collections. And, like the principles of provenance, metadata librarians diverge in how they understand and utilize metadata standards for both physical and digital collections. Examining the ways in which our adherence to legacy approaches to description and arrangement fall short of our necessarily expanding discoverability mandates, the case studies that follow engage the unique qualities of visual archives and the communities they invite and represent.

Case study one: the Liljenquist family collection of civil war photographs

Acquired by the LOC in 2010, the Liljenquist Family Collection of Civil War Photographs offers site users a noteworthy visual experience of the American Civil War. Featuring over 7,000 photographs of Union and Confederate soldiers, it is one of the largest compilations of Civil War images available online. Though substantial, this collection offers a limited perspective of the war. Because "seeing comes before words," in the language of John Berger (1972, p 7), how this nation-defining event is visually depicted shapes how it is remembered. While the Liljenquist Family Collection "include[s] representation of African Americans" (LOC, n.d.-b), photos of the lone white soldier are, by far, most prevalent. Though new additions continue to be made, fewer than 150 images in total feature Black Americans. An effect of provenance that "valorize[s] and venerate[s] white western masculinity" (Drake 2016), this archival scarcity is further amplified by the metadata used to describe and make accessible these images. While conventional metadata practices have historically sought to manage, classify, and render discoverable text-based artifacts, such approaches do not always adequately attend to digital archives that house visual materials, particularly those depicting marginalized and underrepresented groups.

While the LOC houses a range of photographs documenting the Civil War, mirroring traditional archival standards that keep separate the *fonds,* the Liljenquist Family Collection is assigned its own landing page (Fig. 1), offering contextualizing information that demonstrates how provenance influences archival structure. Reinforcing Howard Zinn's observation that "the most powerful, the richest elements in society have the greatest capacity to find documents, preserve them, and decide what is or is not available to the public" (Zinn 1977, p 20), the collection was created by a business man named Tom Liljenquist and his three young sons, Jason, Brandon, and Christian. After finding bullets from the war near their home in Virginia, they became interested in collecting Civil War photographs, focusing specifically on images "of enlisted men, the young men who fought on the front lines, rather than the famous generals of the war" (LOC, n.d.-b). Indicative of the many

Fig. 1 Screenshot of the landing page for the Liljenquist family collection of civil war photographs. https://www.loc.gov/pictures/collection/lilj/ Accessed 20 Jan 2024

layers of provenance that inform such a corpus, they purchased images from antique shops, eBay, Civil War auctions, estate sales, and photography shows. The Liljenquists' financial investment in the Civil War photographs is significant. According to an article in *Arlington Magazine*, the photographs cost "anywhere from $250 for a cracked ambrotype of a young soldier to $19,000 for a well-known portrait of an African-American soldier with his family" (Rosenberg 2012). A rare and valuable collection, Liljenquist donated it to the LOC to share "this important history…with the public" (Rosenberg 2012).

Though the photographs have passed through many hands over the years, the LOC centers the Liljenquists' history of ownership, foregrounding their vision of the collection. Evidence of this influence is perhaps most readily observed in the donor perspective essay published on the site. Written by then-teenaged Brandon Liljenquist (n.d.), "The Last Full Measure: The Liljenquist Family Collection" is suggestive of how provenance can shape public interpretation. Specifically speaking to how the collection might be used, by whom, and to what effect, Brandon begins

by recounting how the photograph of a young white Union soldier named George W. Weeks helped complicate his view of the war. When he and his brother read a letter that Weeks wrote to his mother, they "laughed" at his belief that he served "the greatest army that was ever known." Hinting at the ongoing legacy of Lost Cause ideology, Brandon "knew which was the greater army. The bravery and fighting spirit of the Confederate army was legendary. When equally equipped, the Confederate army always outmatched the Union army." While Lost Cause visions generally maintain that the South engaged in a heroic fight over freedom, culture, and states' rights, this romanticization includes the belief that the Confederacy lost the war, not because their army was inferior, but because the Union had more resources.

Though Brandon never concedes otherwise, he notes that he felt greater empathy for Weeks after learning that the soldier contracted malaria and died. Acknowledging this change in perspective, Brandon writes "We came to learn the ideals an army embraces are what make it great, not its military prowess. Weeks and his fellow soldiers were the emancipators of a race." Exemplifying the perspective the photographs can help cultivate, Brandon's anecdote is suggestive of the influence the collection may have on how the war is remembered. Further driving home this point, Brandon recounts how he shared the photographs with his classmates, demonstrating how the images can be used in educational settings. As he reports, "Our classmates, familiar only with Civil War generals pictured in textbooks, were amazed to see how many of the images depicted soldiers their age and younger." But, he continues, "the biggest surprise for everyone was seeing images of African American soldiers. Our classmates were unaware of the significant contribution these soldiers made to the Union victory." Though the war ended over 150 years ago, the white androcentric memorializing patterns that defined the conflict post-Reconciliation continue to haunt contemporary perspectives of this nation-defining event, as Brandon's own experiences and those of his classmates suggest. While the various roles that Black Americans played in the war are essential to understanding the full scope of the conflict, only a modicum of artifacts in this collection document their martial experiences.

This archival scarcity is, in part, a reflection of the systemic inequities that Black Americans faced during the war and in the years after. "Social circumstances shape what information may be known, what may be recorded, and what may not, and how it may be recorded," as Tom Nesmith rightly observes in his argument for societal provenance, an approach that makes explicit the various forces that enable the formation of records (Nesmith 2006, p 352). While developments in photography and print technology allowed the Civil War to be recorded in novel ways, there is a general "lack of images of black soldiers" in archives, as confirmed by Deborah Willis (2021, p viii). Though Black men and women—both enslaved and liberated—contributed in myriad ways to the Union, until the Emancipation Proclamation was issued in 1863, Black men were prohibited from enlisting in the Union Army. Given these constraints, roughly 180,000 Black men in total volunteered as soldiers, comprising about 10% of the Union forces. As more Black men enlisted, they were photographed with greater frequency, suggesting that "the subjects were aware of the significance of the moment and sought to preserve it," Willis explains (2021, pp 9–10). But photographs were

also expensive, costing "$0.25 for the smallest images to $2.50 for the largest—approximately $6 to $60 in today's dollars," with even higher prices in the South where opportunities for pictures were less readily available (Willis 2021, p. 10). Given these circumstances, the Liljenquists' choice to prioritize portraits of soldiers thus limits the number of photographs of Black Americans in their collection, exemplifying how provenance is built into the DNA of a given archive.

While archival scarcity has significant repercussions for how the public remembers the war, this exiguity is intensified by the metadata that make these materials discoverable. Amalia S. Levi (2019) describes it this way:

> What gets into the archives gets to be preserved 'forever' and is enhanced with access tools through archival description and metadata that usually reinforce Western assumptions. These records are then retrieved and used in scholarship, and eventually become part of historical narrative in a process that shapes what we remember and what we forget. (Levi 2019, p 131)

While Levi is discussing text-based archives, her point is broadly applicable to visual collections, including those assembled by the Liljenquist family: although there are a limited number of photographs of African Americans in this archive, how they are classified, described, and shared by the LOC can impact their discoverability and how they are interpreted. These effects are a consequence of how visual materials have historically been cataloged. As Schwartz (2002) asserts:

> Traditional item-level descriptions of photographs, indexed by subject and credited to the photographer, but without adequate contextual information about their functional origins of provenance, or clear links to such contextual information, transforms archives into stock photo libraries, reducing photographs to their visible elements, and conflating photographic content and photographic meaning. (Schwartz 2002 p 157)

But compounding the problems arising from how visual materials are described and made discoverable, standard approaches to metadata have historically failed to attend to the circumstances of archiving marginalized populations. Though characterized as a neutral undertaking, traditional descriptive and bibliographic practices tend to make more accessible sources that reflect dominant white visions of the past while ignoring, minimizing, or otherwise obfuscating extant traces of the "othered" in the archive.

The consequences of these descriptive enterprises can be observed by examining the metadata used to catalog the visual materials comprising the Liljenquist Family Collection. Each photo is accompanied by a digitized record that includes, among standardized metadata fields, the Library of Congress Subject Headings. Adding to the litany of criticisms directed at the LCSH, all photographs of Black men and women in this collection are tagged as "African Americans," though no equivalent term exists to describe white Americans depicted therein. Though subtle, this approach to LCSH racializes people of color while normalizing whiteness. But further reflecting and perpetuating socially constructed racial differences, many images of Black Americans contained within this collection have yet

to be identified by name. While many white soldiers, too, go unnamed, this omission affects soldiers of color to much greater extent. Only about a dozen African Americans in this collection have been identified. "Names are important," as Alexis Antracoli and Rawdon (2019) rightly assert:

> Those who are named are empowered, and those who remain nameless are at best marginalized, and at worst erased. The exclusion of people's names from descriptions of historical records is not merely a barrier to scholarly research. Rather, it is both a symptom and cause of the violence of whiteness in our society (Antracoli and Rawdon 2019, p 323).

Without proper names or other identifying features, photographs are often assigned vague and repetitive titles and headings, rendering them discursively indistinguishable. Several photographs of Black soldiers, for example, are titled "Unidentified African American soldier in Union uniform." Typical LCSH for portraits of Black soldiers include tags like "United States—Army—People—1860–1870," "African Americans—Military service—1860–1870," and "Soldiers—Union—1860–1870." When relevant, some titles and LCSH account for physical elements of the photo, including "Backdrop," "Rifle," "Flag," and "Tent," which can help differentiate one photograph from another, but outside of the setting, the LCSH tends

Fig. 2 Unidentified African American soldier in union uniform, AMB/TIN no. 2782

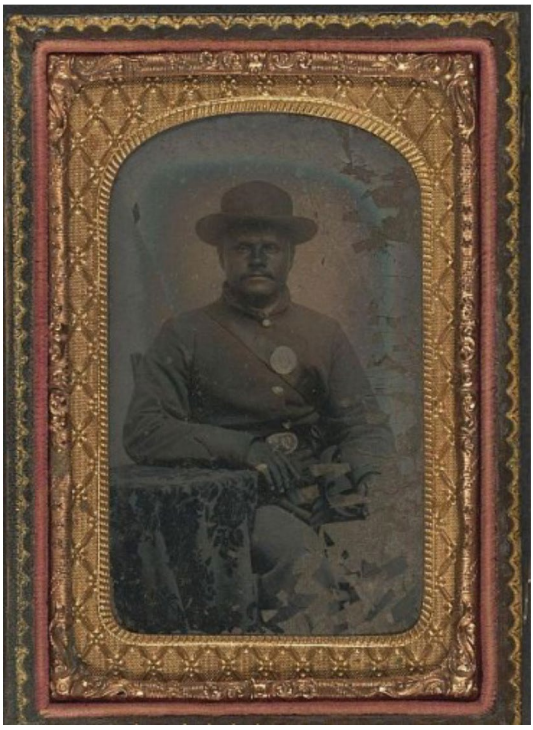


Fig. 3 Unidentified African American soldier in union uniform, AMB/TIN no. 2549

to focus on categories of race, gender, and age, eclipsing other unique qualities and contexts. To this end, a portrait can share the same title and the same LCSH and still look rather different, demonstrating how textual practices can conflate visually divergent images as Figs. 2 and 3 illustrate.

Images capture racialized difference, which is reinforced by metadata when they are described, sorted, and made discoverable. But metadata can also mask the inequities captured in the image, and more broadly, the social and political epoch. "[T]he camera is never neutral," writes John Tagg. "The representations it produces are highly coded, and the power it wields is never its own" (Tagg 1993, pp 63–64). Exemplifying this possibility, the photograph titled "Camp Brightwood, D.C.–Contrabands in 2nd R.I. Camp" (Fig. 4) depicts a group of Union soldiers and Black "contrabands," the name given to the thousands of enslaved who fled to Union camps when the war erupted. These war refugees played a significant part in the war and its politics, though only a few images documenting this particular context are included in the Liljenquist Family Collection, foreclosing possibilities of a more nuanced vision of the war. Taken between 1861 and 1863, this particular photograph features Union Captain Beriah Brown, Captain John P. Shaw, and Lt. James B. Fry, all of whom are named in the metadata, while the Black individuals are described as "African American men and boy" (Camp Brightwood). As previously noted, the inability to name the archive's most disenfranchised both mirrors and advances asymmetrical power dynamics, but photographs, too, can function similarly when not adequately contextualized. This image, which captures racial inequities, exemplifies

Fig. 4 "Camp Brightwood, D.C.–Contrabands in 2nd R.I. Camp"

such a phenomenon. Although all in the photo wear uniforms, the white officers' attire is more ornate; their swords are also in full view in contrast to the contrabands who have no weapons. But the differing positions, postures, and facial expressions also speak to the white supremacy that informed even Union regiments. The white men all appear at ease and in control, standing or lounging comfortably on a chair, while the war refugees are positioned below them. Sitting on the ground are the two youngest contrabands; their brows are furrowed, as they gaze in the direction of the photographer. Despite the racialized disparities made visible in this photograph, there is no discussion of contrabands in the image summary. And at the time of writing this article, the African Americans in the photo were described as "Fugitive slaves" in the LCSH, characterizing them as the problem, not the product, of an unjust system.

The inequities made visible in the "Camp Brightwood" photograph come more clearly into focus when juxtaposed with photographs of a similar ilk. By itself, "Camp Brightwood" might not immediately signal the racial prejudice that war refugees endured in these camps, but taken collectively, this smattering of images reveals patterned ways that white soldiers asserted their dominance over the African Americans. In many of these group photos (Figs. 5 and 6), the Black male, often a boy, is seated on the ground, sometimes holding a pitcher or pan, while the white soldiers, in contrast, stand erect usually with their guns in hand. Though a notable number of these group shots exist, the LOC does not offer metadata that help render this visual feature legible. While scholars have mixed thoughts on

Fig. 5 Brady's national photographic portrait galleries, "Union soldiers with rifles stand in front of tent "10" at Camp Cameron, Washington, D.C., as an African American boy sits in front shining boots"

how best to "resist and transform normative archival description practices without creating newer but still equally damaging or silencing practices," as K.J. Rawson asserts in his discussion of queer archiving (Rawson 2017, p 18), to ignore how these images mirrored and contributed to systemic discrimination is to replicate this vision. Because the LCSH, in particular, have a taxonomizing function, grouping artifacts with similar qualities and attributes determines how images are connected (or not) and how users read a given photograph, as demonstrated by the prescribed themes offered on the collection landing page, which topics, such as "United States Colored Troops" and "African Americans in Uniforms." Attending to the commonalities and divergences that specifically affect marginalized figures in archival collections is one possible way to mitigate past injustices, particularly in the context of the World Wide Web.

Fig. 6 Brady's National Photographic Portrait Galleries (1862), "Twelve generals in uniforms, some identified, and a young African American boy at Cumberland Landing, Virginia, May 14, 1862"

As this case study demonstrates, the provenance of the Liljenquist Family Collection reflects white-centered, patriarchal logics, which are amplified by LOC metadata applications. Primarily featuring portraits of Civil War soldiers, the vision of the past that it evokes largely centers the experiences of white men, though the war's impact was far more capacious. While some photographs of African Americans are present in the collection, the LOC's metadata does not always make these images accessible in meaningful ways, contributing to archival scarcity and sanitized views of the past. Such examples renew the urgency of attending to questions about how visual materials in the archive should be articulated to best "educat[e] the public, suppor[t] social justice, and facilitat[e] scholarly research" (Holterhoff 2017, para. 31), particularly in online spaces.

Case study two: the Webcomics web archive

Launched in June 2017, the Webcomics Web Archive is focused on "comics created specifically for the web and supplement[ing] the Library of Congress' extensive holdings in comic books, graphic novels, and original comic art" (Library of Congress, n.d.-d). While collecting is ongoing, sixty-two webcomics are preserved and made available through this unique archive. One of over 500 born-digital collections at LOC and one of three devoted to comics works and cultures, the description notes that many of the preserved items feature authors and subjects "not traditionally represented in mainstream comics." At the cutting edge of early-Internet culture, webcomics provide documentation of the places, topics, and art forms that united people searching for community in the 1990s and beyond. Like self-published zines and

underground comix, webcomics build connections and circulate information among folks not represented or welcomed by the mainstream. Yet, the metadata for the Webcomics Web Archive collection focuses on academic theories and ontological distinctions used by scholars to classify and compare comics types. Further, provenance information is reduced to a statement about web capture and potential content warning and supplemented by sparse and generic metadata fields. Finally, the archive struggles, as many catalogers of graphic narrative do, to account for both the image and text content of these works. Such disparities provide an opportunity to acknowledge, as Anna Dahlgren and Karin Hansson have, that "different parties have different needs for descriptive metadata" (Hansson and Dahlgren 2022, p 22), and to reimagine prevailing standards to better address searchability and equity with image-based collections. In particular, they invite consideration of what might be included in context for image-text materials like comics and who the audiences for these materials really are.

Megan Halsband and Stephanie Grimm note that both early and contemporary webcomics are available via shared hosting locations or "hubs" (Halsband and Grimm 2019, p 123). Historical examples include Keenspot and Webcomics Nation, and modern authors utilize sites like Webtoon and Webcomics Hub. Creators also leverage expanding social media resources like Tumblr, Instagram, Facebook, and individual blog sites to connect with and grow readerships. These platforms allow artists to serialize their work and host multiple projects in one place and allow readers to find new comics based on those they already love. Most significantly, they provide the basis for enduring community forums through comment sections, live blogs, and extensive hyperlinking. Many webcomics creators interact directly with readers this way and even make immediate or eventual changes according to feedback. And webcomics fans are themselves incredibly active. They debate and discuss comics in these public spaces, often raising alarms about discriminatory attitudes among creators and fandoms and sometimes having a significant impact on the fate of once-beloved projects (illustrative examples include *Sinfest* (Webcomics Review 2024) and its creator's anti-trans attitudes and the recent doxxing of Stonetoss, the artist behind the neo-Nazi webcomic, *Red Panels* (Wiggins 2024). In these ways, the community interactions surrounding webcomics are fundamental to understanding and contextualizing them. Unfortunately, current web archiving tools do not consistently provide full captures of interactive web content and many of these hosting sites have broken components or are entirely defunct.

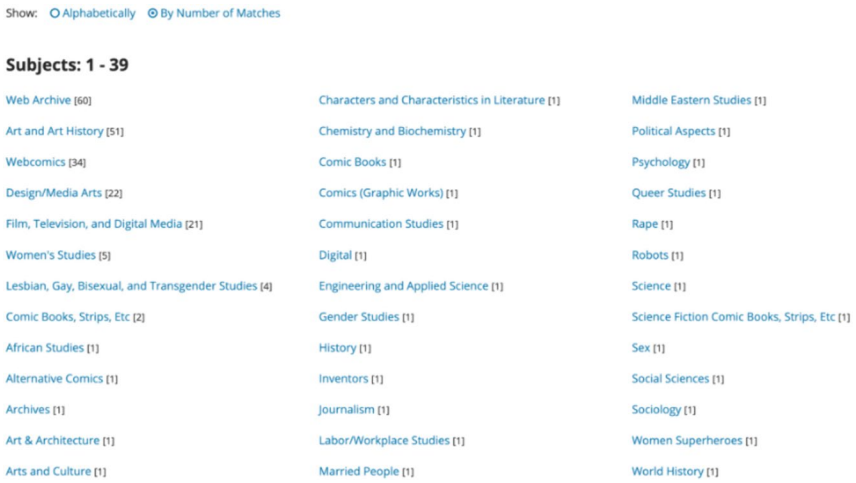
The particular complexities of webcomics are not limited to their participatory nature, ephemeral hosting, or counter-cultural impulses; they also include their multi-modality which creates challenges to standardization for repositories generally and for catalogers and metadata librarians in particular. How and why to reimagine image-based metadata and description is complicated for graphic narrative formats, most obviously because their structural elements—gutters, speech bubbles, panels etc.—cannot be easily accounted for. And though comics-specific mark-up languages—like John Walsh's Comic Book Markup Language (CBML) (Walsh 2012)—are available, they are infrequently taken up beyond the institutions that originate them. On a philosophical level, however, attitudes about whether these items should be classified as children's literature, periodicals, adult fantasy/fiction,

or something else, still undergird decisions to collect and preserve them. The roots of this disconnect reside with the Comics Code Authority's (CCA) assembly in 1954, following a censorship case against DC Comics and mid-century publishers with imprints in "horror" and "crime." Alleging that comics are solely for children (and assuming that those children are always white) the CCA accused publishers of exposing innocent readers to "unseemly" and "inappropriate" content and heavily sanctioned them. Implicit was the belief that comics are a "low brow" and juvenile medium and that being so makes them less worthy. While this article isn't exactly about biases against comics, it is relevant to point to the inaccuracy in these latent attitudes. Scholars like Howard and Jackson (2013), Kuruhara (2023), Aldama (2010), and others have highlighted the legacy of Black, Latino/Latinx/Latine, Indigenous, and Queer comics in the USA, as well as the diverse comics frameworks in other countries. Very few graphic narrative traditions are authored for readerships of exclusively young children, and there is more to be said about the idea that children only consume "lesser" media. We foreground these enduring biases because they continue to impact the webcomics archive and others like it, despite the admirable goal of highlighting underrepresented authors and characters.

Like the LOCs other web-based archives, the Webcomics Web Archive is organized into thematic "collection items," which either include the full run of a single title or the multiple works of a creator. This means that an "item" could include a few web pages, or hundreds, depending. The interface is comparable to the search engines and databases at public libraries, schools, and research repositories that users may have already encountered. On the homepage, visitors are greeted with image thumbnails for the most frequently searched-for collections. The bulk of the page features the collection items, while banners at the top allow users to filter results by categories, and a navigation bar on the left allows you to further refine according to common metadata fields. Similarly ubiquitous, each discrete item is tagged with LCSH that enable and structure their searchability. Subject categories provide essential opportunities for researchers to discover works using more terms than are possible via metadata details alone. And related controlled vocabularies contribute to shared lexicons allowing patrons to be successful across libraries, in spite of small variations in language employed. At the time of writing, the most common subjects attributed to the collection are those focusing on genre/medium or on scholarly approaches to comics. This is true of the expanded subject list as well, which includes thirty-nine LCSH (Fig. 7).

While the overall focus is clearly on academic disciplines or types of art production, the detail and specificity as applied to individual works is inexplicably variable. For example, if a user refines a search based on the "Women's Studies" heading and clicks on *Priya's Shakti*, one of five results, the subjects attributed to the record are both slightly more specific and broader, perhaps anticipating general subject searches not just academic ones (Fig. 8).

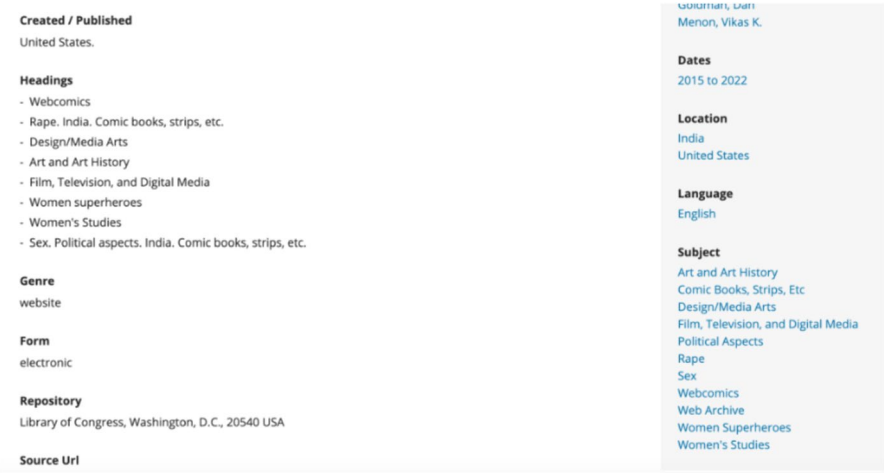
These additional possibilities include subjects related to location and culture, identifying that the comic is set in India, as well as thematic ones like "Women Superheroes," "Rape," and "Political Aspects." Such descriptive practices provide a clearer picture of the webcomics' subject, focus, and purpose, and anticipates the strategies and interests of a far greater range of users. But this does not appear to be


Show: Alphabetically By Number of Matches

Subjects: 1 - 39

Web Archive [60]
Art and Art History [51]
Webcomics [34]
Design/Media Arts [22]
Film, Television, and Digital Media [21]
Women's Studies [5]
Lesbian, Gay, Bisexual, and Transgender Studies [4]
Comic Books, Strips, Etc [2]
African Studies [1]
Alternative Comics [1]
Archives [1]
Art & Architecture [1]
Arts and Culture [1]
Characters and Characteristics in Literature [1]
Chemistry and Biochemistry [1]
Comic Books [1]
Comics (Graphic Works) [1]
Communication Studies [1]
Digital [1]
Engineering and Applied Science [1]
Gender Studies [1]
History [1]
Inventors [1]
Journalism [1]
Labor/Workplace Studies [1]
Married People [1]
Middle Eastern Studies [1]
Political Aspects [1]
Psychology [1]
Queer Studies [1]
Rape [1]
Robots [1]
Science [1]
Science Fiction Comic Books, Strips, Etc [1]
Sex [1]
Social Sciences [1]
Sociology [1]
Women Superheroes [1]
World History [1]

Fig. 7 LCSH list for library of congress's Webcomics web archive. https://www.loc.gov/collections/webcomics-web-archive/ Accessed May 21, 2023


Created / Published
United States.

Headings
- Webcomics
- Rape. India. Comic books, strips, etc.
- Design/Media Arts
- Art and Art History
- Film, Television, and Digital Media
- Women superheroes
- Women's Studies
- Sex. Political aspects. India. Comic books, strips, etc.

Genre
website

Form
electronic

Repository
Library of Congress, Washington, D.C., 20540 USA

Source Url

Menon, Vikas K.

Dates
2015 to 2022

Location
India
United States

Language
English

Subject
Art and Art History
Comic Books, Strips, Etc
Design/Media Arts
Film, Television, and Digital Media
Political Aspects
Rape
Sex
Webcomics
Web Archive
Women Superheroes
Women's Studies

Fig. 8 Landing page for Priya's Shakti in LOC Webcomics web archive. https://www.loc.gov/item/lcwaN0014023/ Accessed May 25, 2023

consistently utilized or considered throughout. Other examples, like *Darwin Carmichael is Going to Hell* and *Fried Rice,* both award-winning and fan-favorite comics tackling serious themes, have only the most generic subjects applied. While they are wildly different texts—a darkly-humorous fantasy comic and an intimate autobiographical work based in Malaysia—three of the four subject tags attributed to them are the same.

Each of the included works is far more complex than is marginally captured by accompanying descriptors, which is further compounded by their equally robust

textual and image-based content. And this is significant because very little additional metadata outside of subject terms is actually provided. Beyond fields that are redundant for searchers like "genre," "repository," and "format," only occasional short summaries and direct links to still-active hosting sites provide added context. Readers familiar with comics or creators attempting to locate them using subject searching—or plain text searching, as many instinctually do—would likely be unsuccessful. Simultaneously, patrons exploring the archive with relevant subject interest who *aren't* familiar with individual works would also struggle, except in rare exceptions.

Curiously, the total schema of subject headings and their seemingly haphazard application is both functionally and ideologically antithetical to the stated focus on underrepresented creators and topics. Put simply, such focus is nowhere evident in the records or metadata for the webcomics collection. Worryingly, this might suggest to both new and experienced researchers that academic categories and oblique genre distinctions are the only subjects represented in the archive. And many unanswerable questions arise from the disconnect. Why, for instance, are two items ascribed "comic books, strips, etc." while one is relegated to "comic books" and yet another to "comics (graphic works)"? Why are only thirty-four of the sixty-two items categorized as "webcomics"? What is the significance of tagging fifty-one of the items as "art history" and only one as "art and culture"? Are those comics about art history as a discipline or are they evidence of the history of art? And, for all of the subject headings that have been applied, are they referring to the text, the images, or both?

It is reasonable to assume that users without academic backgrounds comprising the majority of topical attributions will connect items described as "Women's Studies" to stories and themes relevant to the experiences of those identifying as women, possible women creators, and the likelihood of women characters. And most controlled vocabularies would connect "women" and "women's studies," making those items discoverable to searchers utilizing general words and phrases. Alternatively, it's not reasonable to assume that users will know their success is highly contingent on these mechanisms and devise their search strategies accordingly. And not all of the field-based categories provide such easy, visible connections. Indeed, currently, a wide variety of the content that has been painstakingly preserved in this urgent archive is not meaningfully searchable. With a stated purpose of serving researchers who are both "new and experienced" (Library of Congress, n.d.-a), these factors call into question whether LOC descriptive practices, and by extension, the many thousands of libraries that follow their model, are actually serving such an expansive range of users.

In an interview in November of the launch year, LOC information technologist Abigail Grotke said of this groundbreaking collection that, "While we've been archiving a variety of types of websites since 2000, comics were added to the Library's web-archiving program in 2014. […] This presented new challenges as these were more visually focused than sites included in prior collections'' (Hartsell 2017). In spite of this recognition, the collection does not include an explanation of or key to how Grotke's team handled those challenges, leaving questions about how, why, and by whom these metadata fields were selected above numerous other possibilities, unanswered. Obviously, decisions must be made. But transparency is

significant when such decisions influence what is findable and ascribe value to the included webcomics, making an argument about why, when, and even if one should interact with them. Though manual searching is currently possible given its relatively finite size, this collection is presumably growing and many born-digital collections are far too large for such a work around.

Beyond the mechanics of searchability, materials in the web archive are described at collection level, rather than individual instance. As Librarian With Issues (2017), whose blog focuses on the challenges of graphic narrative cataloging notes, this complicates description if title, numbering, or creator names shift, or if there are long periods where new issues aren't generated. And it can severely limit the efficacy of subject headings. As they explain it, "because you are including several trades on the same record, you can't have specific subject headings, they have to be broad enough to describe the series as a whole" (Librarian With Issues 2017, par. 6). While this is referring to print comics, the difficulty of serialization certainly applies. If one instance of *Hyperbole and a Half* is about anxiety and the next one is about bad roommates and you're addressing the collection level, how could you possibly make both of those optimally searchable with any subject schema? And how can you account for that variety with more expansive metadata without going too far in the other direction? While librarians working with sequential art formats note that balance between labor and peak accessibility is difficult and hard to come by (Bailund et.al. 2023), one might dare to look to the LOC for transparency about their choices in these complex cases, if not for more innovative solutions.

Finally, because it is comprised of website-captures, the surrounding community aspects are, at best, sporadically included. Interactions surrounding webcomics are central to their context and origins, can help track their evolution over time, and can verify a capture's veracity, which is to say that community is the most effective source of provenance for these materials. In the example below (Figs. 9 and 10), hyperlinks to other issues or areas within the webcomics' original website, Twitter links, and "shout out" sections are largely preserved, whereas advertisements and other image-based content are not. Exemplifying the challenges of web archiving, however, the interactive elements have only been variously active thanks to newly-broken links and the manual workflow of the utilized ArchiveIt tool. This is not a straightforward premise, as scholars, librarians, and creators of electronic art already know. With the end of Adobe Flash on December 31, 2020, for example, numerous websites and works of digital art and literature became nonfunctional. Many of those early and innovative works only exist—sadly, in screen capture format, which doesn't preserve the functionality of the original—thanks to the efforts of organizations like the Internet Archive and the Electronic Literature Organization (ELO).

Ultimately, due to LOC's descriptive choices and policies related to web archives, (Halsband & Grimm 2019, p 133), neither the archive's supposed focus on BIPOC, Queer, and women creators, nor the important discursive purpose of many of the selected works are clear or findable. Catalogers and archivists continue to wrestle with effectively accounting for the unique characteristics of works of graphic narrative utilizing current standards. Moreover, ideological debates about what materials are worthy of inclusion in repositories are ongoing and falling behind the larger zeitgeist. And webcomics are even more complicated. Their insular and communal

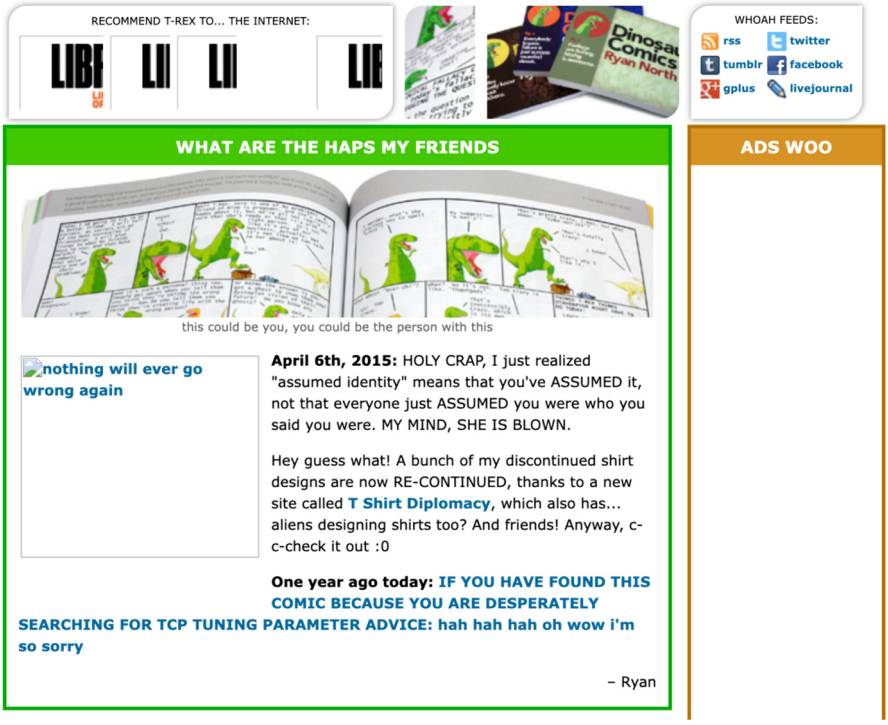


Fig. 9 Example of a webcomics page with some active and inactive elements. "Dinosaur Comics," library of congress, Webcomics Web Archive Collection. Retrieved Jan 30, 2024

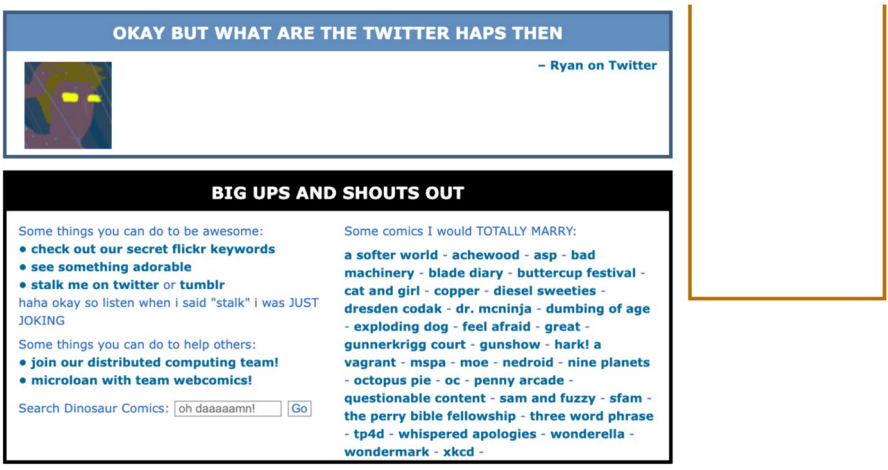


Fig. 10 Example of a webcomics page with some active and inactive elements. "Dinosaur Comics," library of congress, Webcomics web archive collection. Retrieved Jan 30, 2024

circulation, intimate depictions of personal experience, and mixed media format make them particularly important for understanding not just early (and increasingly disappearing) web culture, but the ways in which individuals have created and located spaces where they are accurately represented, when such representations are not available in mainstream culture.

It is important to recognize the foresight and efforts of LOC librarians to collect and preserve webcomics in any capacity given their paucity from most institutions. But it's also worth examining why the intent of the collection's curators and the descriptive and organizational schemas applied to it are at odds. These tensions bring to mind Jaillant's important assertion that, "*Not all users of digital collections have the same needs*. … Instead of one-way guidelines designed by archivists and targeted to users, it is essential to co-design strategies to make digital archives easier to access" (Jaillant 2022, p 421, emphasis added). They likewise recall Drake's observation that:

> The descriptive challenges facing born-digital archival records is not technical or social, but both technical and social. Considering one without the other obscures how humans use information technology, which is to achieve both technical and social ends. In many ways, provenance as a theory or practice is poorly positioned to address either challenge (Drake 2016).

Making community-centered, complicated visual collections truly accessible first requires a careful analysis of the full range of users and then necessitates a partnership to generate search and design structures that are more meaningfully and widely useful to *all* of those readers.

Addressing scarcity and discoverability: visual representation and inequity in LOC collections

By analyzing provenance and metadata in visual collections, our case studies engage the intersections of descriptive principles and the stakes involved in leaving cornerstone concepts like provenance unexamined or unrevised over time. Though it bears the imprint of the many direct and imperceptible hands that give rise to the archive, "Provenance," as we (2022) describe it elsewhere "… is often characterized as an impartial, practical approach to archival organization, yet is anything but. These principles of arrangement are necessarily human, informed by the prejudices and proclivities of particular subject positions—whether that be the collection's creator, the processing archivist, or as is often the case, both" (Manis and Wilde 2023, p 69–70). In creating and arranging the archive, these originating sources invariably leave their mark, reflective of both personal and institutional priorities. But while this influence is inescapable, how and to what extent it is explicitly acknowledged in finding aids and archival records varies. Providing context about the collection and its creator, the choices archivists make in establishing and following the principles of provenance during processing affects how researchers understand the origins, purpose, and meaning of the archive.

The collections we analyze have noteworthy differences: they document distinct time periods, vary in scope and volume, feature a spectrum of visual material (i.e., digitized photographs and born-digital multimedia narratives), attract different audiences, and offer researchers a range of contextual and provenancial information. Perhaps most significantly, they diverge in how they approach (or don't) the legacy of archival exclusion and reparative efforts to amplify underrepresented voices in the historical record. Despite these differences, however, both demonstrate the challenges and complexities of attending to provenance and accessible metadata in the case of visual materials and highlight how pervasive these challenges are in the USA.; even at the most impactful and arguably most provisioned library (despite its own limitations) in the nation. More pithily put, they demonstrate how across time periods, types of digital materials, and topical foci and in spite of the visibility and relative accessibility of digital collections at LOC, complicated origins mean messy provenance, which leads to messy metadata. Such elision of boundaries, strategies, and categories renders many important materials effectively invisible utilizing current standards of description and arrangement. While the Liljenquist Family Collection includes considerable information about its origin, collectors, and intended use, the Webcomics Web Archive offers only minimal contextual detail. Following standard metadata practice, both examples rely on the inherited LCSH schema, notably impacting the discoverability of historically marginalized experiences and much of the visual content in each archive.

Of particular relevance are tensions between the sociopolitical impact of LCSH and the sheer difficulty of enacting changes to them. Specifically, we are concerned with the role this system plays in inevitably limiting and defining what language is available to describe the many individuals and communities comprising the historical record. "The choice of standard terms requires careful consideration for, in our use of language, in our systems of ordering, in our fields of classification, and in our rules for description, we privilege and we marginalize," Schwartz explains. "The words we choose to describe what we do reflect our view of the world, the values we hold, the things of this world that we value" (Schwartz 2002, p 147). There are, on average, 4,000 new headings and subheadings added to the now-extensive list of possible LCSH every year, along with a mechanism for suggestions from LIS professionals and community members (Library of Congress 2017). However, changes to the broader organizational structure are non-existent, and updates based on principles of equity and inclusion are subject to a much more convoluted process. Referring to Congressional resistance to the "illegal alien" update, for example, Melissa Adler insightfully highlighted that "members of Congress view subject headings to matter and were compelled to intervene" (Adler 2017, p 11). Even the LOC acknowledges that the patchwork approach to updates and restructuring has created inconsistencies and barriers to use. The introduction to the thirty-ninth edition of the LCSH schema notes that, "Because the list has expanded over time, it reflects the varied philosophies of the hundreds of catalogers who have contributed headings. Inconsistencies in formulation of headings can usually be explained by the policies in force at the varying dates of their creation" (Library of Congress 2017, p. xiii).

As Itza Carbajal asserts, "the practice of metadata creation has been tackled without much forethought, concern, or vision for its liberatory or restorative potential,

likely a result of metadata serving solely as a management tool for the growing amount of data" (Carbajal 2021, p 93–4). Attending to the ways metadata highlights or occludes marginalized populations can significantly impact epistemic equity. Because archives, particularly those held by highly revered and influential institutions like the LOC, "have long been considered a source for writing the history of the modern nation and…for the forging of a national identity and consciousness…," as JJ Ghaddar and Michelle Caswell observe, "[t]he story of a nation's origin, its history and myths, serve as a vital script for citizenship and guide citizens in understanding who does and does not belong to the nation, and their place in the world" (Ghaddar and Caswell 2019, p 75). Creating and reinforcing structures of power, images possess a unique constituting force. Though "[t]he way we see things is affected by what we know or what we believe" (Berger 1972, p 8), what we see concomitantly contributes to what we *think* we know and what we *think* we believe. With a wide reach, digital archives that house visual collections stand to be particularly influential in shaping ocular narratives of the past that inform the present, providing meaningful opportunities for citizens to locate themselves in the historical record.

Declarations

Conflict of interest Kathryn Manis and Patricia Wilde declares that they have no conflicts of interest.

References

Adler M (2017) Cruising the library: perversities in the organization of knowledge. Fordham University Press, NY

Aldama FL (2010) Multicultural comics: from zap to blue beetle. University of Texas Press, Austin

Antracoli AA, Rawdon K (2019) What's in a name? Archives for black lives in Philadelphia and the impact of names and name authorities in archival description. In: Sandberg J (ed) Ethical questions in name authority control. Library Juice Press, Sacramento, pp 307–336

Bailund A, Holloway SW, Kuni K, Tomaras D (2023) Reprints, reboots and retcons: standardizing comics cataloging with the best practices guide from the GNCRT. Art Libr J 48(3):62–68. https://doi.org/10.1017/alj.2023.11

Bastian JA (2006) Reading colonial records through an archival lens: the provenance of place, space and creation. Arch Sci 6:267–284

Berger J (1972) Ways of seeing. Penguin Books, London

Brady's National Photographic Portrait Galleries [1861] Union soldiers with rifles stand in front of tent "10" at Camp Cameron, Washington, D.C., as an African American boy sits in front shining boots [Between 1861 and 1863] Library of Congress https://www.loc.gov/item/2021637276/. Accessed 20 Jan 2024

Brady's National Photographic Portrait Galleries (1862) Twelve generals in uniforms, some identified, and a young African American boy at Cumberland Landing, Virginia [published by E. Anthony, 501 Broadway, New York from photographic negative in Brady's national portrait gallery, 14 May] https://www.loc.gov/item/2023631086/. Accessed 20 Jan 2024

Camp Brightwood, D.C.—Contrabands in 2nd R.I. Camp [Between 1861 and 1865] library of congress https://www.loc.gov/pictures/item/98500959/. Accessed 20 Jan 2024

Carbajal IA (2021) Historical metadata debt: confronting colonial and racist legacies through a post-custodial metadata praxis. Across Discipl 18(1/2):91–107

Christen K, Anderson J (2019) Toward slow archives. Arch Sci 19:87–116

Deelman E, Berriman B, Chervenak A, Corcho O, Groth P, Moreau L (2010) Metadata and provenance management. In: Shoshani A, Rotem D (eds) Scientific data management: challenges, existing technology, and deployment. CRC Press, Boca Raton

Drake JM (2016) Radtech meets radarch: towards a new principle for archives and archival description. Medium https://medium.com/on-archivy/radtech-meets-radarch-towards-a-new-principle-for-archives-and-archival-description-568f133e4325.

Finnegan CA (2006) What is this a picture of? Some thoughts on images and archives. Rhetor Public Aff 9(1):116–123

Ghaddar JJ, Caswell M (2019) 'To go beyond': towards a decolonial archival praxis. Arch Sci 19:71–85

Greene M, Meissner D (2005) More product, less process: revamping traditional archival processing. Am Arch 68:205–263

Halsband M, Grimm S (2019) Panel problems: issues and opportunities for webcomics archives. Art Doc: J Art Libr Soc North America 37(2):119–140

Hansson K, Dahlgren A (2022) Open research data repositories: Practices, norms, and metadata for sharing images. J Am Soc Inf Sci 73(2):303–316. https://doi.org/10.1002/asi.24571

Hartig O (2009) Provenance information in the web of data. LDOW 538

Hartsell M (2017) New online: collecting web comics and culture, 31 Oct. 2017. Library of Congress Blogs https://blogs.loc.gov/loc/2017/10/new-online-collecting-web-comics-and-culture/

Holterhoff K (2017) From disclaimer to critique: Race and the digital image archivist. Digital Hum Q 11(3):137–147

Horsman P (2002) The last dance of the Phoenix, or the de-discovery of the archival fonds. Archivaria 54:1–23

Howard SC, Jackson RL II (2013) Black comics: politics of race and representation. Bloomsbury publishing, London

Jaillant L (2022) How can we make born-digital and digitized archives more accessible? Identifying obstacles and solutions. Arch Sci 22:417–436

Jaillant L, Aske K, Goudarouli E, Kitcher N (2022) Introduction: challenges and prospects of born-digital and digitized archives in the digital humanities. Arch Sci 22:285–291

Kuruhara RK (2023) He inoa 'ala: Scent, memory, and identity in Indigenous comics. J Asian Am Stud 26(3):357–378. https://doi.org/10.1353/jaas.2023.a913084

Levi AS (2019) Intersectionality in digital archives: the case study of the Barbados synagogue restoration project collection. In: Bordalejo B, Risam R (eds) Intersectionality in digital humanities. Amsterdam University Press, Amsterdam, pp 127–147

Librarian with Issues (2017) Three ways to catalog graphic novels series. https://librarianwithissues.wordpress.com/2017/07/19/three-ways-to-catalog-graphic-novel-series. Accessed 1 Jan. 2024

Library of Congress (2017) Introduction. Thirty-ninth edition of Library of Congress subject headings https://www.loc.gov/aba/publications/Archived-LCSH39/LCSH39%20Main%20intro.pdf. Accessed 1 Jan. 2024

Library of Congress (n.d.-c) General information. https://www.loc.gov/about/general-information/. Accessed 1 Jan. 2024

Library of Congress (n.d.-d) Web comics web archive: About this collection. https://www.loc.gov/collections/webcomics-web-archive/about-this-collection/. Accessed 1 Jan. 2024

Library of Congress (n.d.-a) About the library. https://www.loc.gov/about/. Accessed 1 Jan. 2024

Library of Congress (n.d.-b) Background and scope. Liljenquist family collection of civil war photographs https://www.loc.gov/pictures/collection/lilj/background.html. Accessed 1 Jan 2024

Liljenquist B (n.d.) The last full measure: The Liljenquist family collection. Library of Congress, Prints and Photography Reading Room https://www.loc.gov/rr/print/coll/633_lilj_measure.html. Accessed 1 Jan. 2024

MacNeil H (1995) Metadata strategies and archival description: comparing apples to oranges. Archivaria 39:22–32

Manis K, Wilde P (2023) A matter of order: the power of provenance in the mercury collection of Marion Lamm. In: Kirsh G et al (eds) (2023) Unsettling archival research: engaging critical, communal, and digital archives. Southern Illinois University Press, Carbondale, pp 67–80

McPherson T (2018) Feminist in a software lab: difference and design. Harvard University Press, Cambridge

Millar L (2002) The death of the fonds and the resurrection of provenance: archival context in space and time. Archivaria 53:1–15

Moore LJ (2008) Restoring order: the Ecole des Chartes and the organization of archives and libraries in France, 1820–1870. Litwin Books, Sacramento
Nesmith T (2006) The concept of societal provenance and records of nineteenth-century Aboriginal-European relations in Western Canada: implications for archival theory and practice. Arch Sci 6:351–360
Noble S (2018) Algorithms of oppression: how search engines reinforce racism. New York University Press, NY
Note M (2019) The 'margins of archivy': archival descriptions of visual materials https://www.margotnote.com/blog/2019/09/02/archival-description. Accessed 1 Jan 2024
Parker EB (1982) Graphic materials: rules for describing original items and historical collections. Library of Congress, Washington
Rawson KJ (2017) The rhetorical power of archival description: classifying images of gender transcription. Rhetor Soc Q 48(4):1–25
Rosenberg M (2012) Common valor. Arlington Magazine, https://www.arlingtonmagazine.com/common-valor/. Accessed 21 Jan 2024
Schwartz JM (2002) Coming to terms with photographs: descriptive standards, linguistic 'othering', and the margins of archivy. Archivaria 54:142–171
Society of American Archivists (n.d.) Metadata. Dictionary of Terms https://dictionary.archivists.org/entry/metadata.html. Accessed 11 Jan 2024
Sweeney S (2008) The ambiguous origins of the archival principle of 'provenance.' Libr Cult Rec 43(2):193–213
Tagg J (1993) The burden of representation: essays on photographs and histories. University of Minnesota Press, Minneapolis
Unidentified African American soldier in Union uniform. [between 1863 and 1865] AMB/TIN no. 2549 Library of Congress https://www.loc.gov/pictures/collection/lilj/item/2010650484/. Accessed 20 Jan 2024
Unidentified African American soldier in union uniform. [between 1863 and 1865] Call number AMB/TIN no. 2782 Library of Congress https://www.loc.gov/pictures/item/98500959/. Accessed 20 Jan 2024
Walsh JA (2012) CBML: comic book markup language. https://dcl.luddy.indiana.edu/cbml/ Accessed 1 Aug 2024
Webcomics Review (2024) Sinfest. https://thewebcomicsreview.tumblr.com/archive/tagged/sinfest Accessed 1 Aug 2024
Wiggins C (2024) After alleged neo-nazi named, X updates doxxing policy; now many Libs of TikTok posts violate it Advocate https://www.advocate.com/technology/x-doxxing-policy-update. Accessed Aug 2024
Willis D (2021) The black civil war soldier: a visual history of conflict and citizenship. New York University Press, NY
Wisser K, Dean J (2013) EAD tag usage: community analysis of the use of encoded archival description elements. Am Arch 76(2):542–566
Zinn H (1977) Secrecy, archives, and the public interest. Midwest Arch 2(2):14–26

Kathryn Manis (she/they) is an Instruction and Community Engagement Librarian in the University of Georgia (UGA) Special Collections Libraries. Their research and publications are interdisciplinary, engaging comp/rhet arguments, Queer/LGBTQIA + histories and scholarship, Library and Information Science (LIS) research, and decolonial praxis. She recently defended her dissertation for a Ph.D. in rhetoric and composition, entitled "Unexamined Standards and Complex Histories: Case Studies on Rhetoricity, Transgression, and Decoloniality in Archives," with distinction.

Patricia Wilde (she/her) is an Associate Professor of English and Director of Composition at Washington State University Tri-Cities. Her scholarly pursuits focus on feminist rhetorics, specifically as they concern history, historiography, and public memory; rhetorics of the American Civil War; histories of rhetoric and composition; archival research; and composition studies, with a focus on access-oriented teaching and dual enrollment programs. Her in-progress book manuscript, *Winning Our Wonder: Rhetorical Re/Constructions of Women in the American Civil War*; interrogates the powerful relationships between rhetoric, materiality, public memory, and the web.

Authors and Affiliations

Kathryn Manis[1] · **Patricia Wilde**[2]

✉ Kathryn Manis
maniskathryn@gmail.com

Patricia Wilde
patty.wilde@wsu.edu

[1] Special Collections Libraries, University of Georgia, Athens, USA

[2] Department of English, Washington State University Tri-Cities, Richland, USA