Metadata Obscura: Refocusing digital collections through the lens of art history

Stefanie Hilles
Wertz Art and Architecture Library at Miami University, Ohio, United States
hillessa@miamioh.edu

Alia Levar Wegner
Walter Havighurst Special Collections and University Archives at Miami University, Ohio, United States
wegnera3@miamioh.edu

Copyright © 2019 by Stefanie Hilles and Alia Levar Wegner. This work is made available under the terms of the Creative Commons Attribution 4.0 International License: http://creativecommons.org/licenses/by/4.0

Abstract:

Art librarians in academic libraries often rely on generalists in metadata and digital collections departments to accurately describe visual collections. When these partnerships are successful, students and researchers in art disciplines can use their subject training to discover and contextualize visual resources. However, art researchers may experience disruption and disconnection in their research when they encounter visual collections that were digitized without proper attention to disciplinary expectations. This is especially important in academic libraries, whose primary mission is to support students and faculty, and where university users are inherently subject specific in their research. This paper discusses the decisions and workflows implemented to revise a popular digital collection of 19th and 20th-century trade cards held at the Walter Havighurst Special Collections at Miami University. By aligning subject metadata with the methodologies of art history, librarians improved the accessibility and discoverability of these visual materials for art researchers.

Keywords: digital collections, metadata.

Collection History

The Charles and Laura Dohm Shields Trade Card Collection is currently housed in the Walter Havighurst Special Collections and University Archives at Miami University. Donated by Charles Shields in 1987, the collection contains approximately 2200 trade cards and advertisement postcards from the late 19th and the early 20th centuries. The earliest trade cards were simple monochromes that advertised a tradesman’s commercial services. By the
19th century, however, trade cards had become an increasingly common means of advertising a variety of products and services including medicines, beverages, clothing, and farm equipment. Trade cards were especially popular in the United States where they were distributed in stores, included in product packaging, and even sold as collector sets. The popularization of trade cards coincided with the spread of chromolithographic printing in the 1860s. As a planographic printing process, chromolithography enabled printers to combine words and images in seamless, creative ways at an affordable cost. Advertisers capitalized on the popularity of rich chromolithographic prints to attract new customers.

Trade cards are challenging to digitize because they defy easy categorization. Are they art objects or media artifacts? As a predominantly visual medium employing traditional iconography, trade cards can be considered artistic works in a popular culture context. Unlike traditional artworks, however, trade cards were produced to persuade customers to buy specific products and almost always rely on text to convey commercial information and, as such, fall into the history of graphic design. The difficulty categorizing trade cards also impacts how librarians digitize these materials. Subject analysis must consider the trade card’s artistic content as well as its commercial purpose. Miami University Libraries’ long history digitizing the Shields Trade Card Collection illustrates how difficult it is to maintain a dual subject focus amid changing metadata standards and academic trends. The many iterations of the digital collection over the years prompted the recent initiative to refocus attention on the specific disciplinary methodologies and needs of art researchers for a new generation of users.

**Project Background**

The Shields Trade Card Collection was first digitized in 1999, making it Miami University Libraries’ oldest digital collection. The collection was digitized in three iterations between 1999 and 2019. The project originated as a proof of concept for an in-house image management database five years before the library adopted CONTENTdm as their primary digital asset management system in 2004. During the first iteration of the project, from 1999-2001, the physical collection was imaged for the first time and metadata was created for each trade card. The original project team included the former head of the Art & Architecture Library, who added a much needed art historical perspective to the digital collection. The earliest collection metadata contained fields for physical measurements of the materials and administrative condition notes, which are important metadata elements in art cataloging. Unfortunately, this art subject focus was not sustained in subsequent iterations of the digital collection.

During the second phase of the Shields Trade Card Collection, between 2004-2009, a new project team reimaged the trade cards and revised the metadata to reflect an advertising and marketing focus. Librarians consulted business school faculty about their user needs and added specialized controlled vocabularies to describe commercial subjects, such as the Standard Industrial Classification system. During this phase, the project team hid the original metadata fields that described the cards’ physical medium and measurements from public view. While these changes may have opened up the digital collection to new disciplinary audiences in the social sciences, the lack of art subject terms and physical measurements made the collection less relevant to art users. The third and most recent phase of the digitization project sought to recover and enhance the digital collections’ art focus. In this iteration, from 2018-2019, the current Digital Collections Librarian and the Arts and Humanities Librarian teamed up to redesign the digital collection to emphasize art methodologies and improve collection access for art users. The project team reviewed contemporary scholarship on the information needs of
art researchers to ensure that the revised subject metadata would address the current information seeking practices of academic art users.

**The Information Needs of Art Researchers**

Much of the literature describing the needs of art researchers notes the lack of attention paid to their research needs when compared to other disciplines (Cobbledick, 1996; Cowan, 2004; Beaudoin, 2005; Hemmig, 2008; Hemmig, 2009; Mason & Robinson, 2011; Lo & Chu, 2015; Lee & Haddow, 2018). In general, studies separate art researchers into three categories: practicing visual artists, studio art students at universities, and art historians. The information needs of visual artists and studio art students are similar as both groups demonstrate a strong preference for serendipitous browsing (Toyne, 1977; Pacey, 1982; Day & McDowell, 1985; Stam, 1995; Cobbledick, 1996; Frank, 1999; Hemmig, 2009; Lo & Chu, 2015; Münster, Kamosiöri, Friedrichs, & Kröber, 2018). In his literature review of the information needs of visual artists, Hemmig (2009) identifies five purposes that motivate visual artists to conduct research, “inspiration, specific visual elements, knowledge of materials and techniques, market and career guidance, and knowledge of current trends in the artworld” (p. 683). Three of these motivations are also highlighted in the literature detailing studio art student’s information needs: inspiration, specific visual elements, and materials and techniques (Münster et al., 2018; Lo & Chu, 2015). Art historians, more traditional in their approach to research, frequently rely on searching to find information (Beaudoin & Brady, 2011). Their information seeking strategies are akin to the methods used across the humanities, although art historians rely on visual images and objects more than other humanities disciplines (Beaudoin & Brady, 2011).

While the literature on how art researchers use image database collections is more robust than those that deal with their broader information needs, the results are comparable. Browsing is the preferred method of information access for studio art students, especially regarding their user needs for inspiration, specific visual elements, and materials and techniques (Brett, 2013; Lo & Chu, 2015; Münster et al., 2018).

Unfortunately, the digital image needs of visual artists remain under studied in the literature. However, given that artists seek visual information more than textual information (Toyne, 1977; Pacey, 1982; Day & McDowell, 1985; Beaudoin & Brady, 2011), their well-documented use of browsing as a information seeking strategy, and the fact that their general information needs are similar to those of art students, it is reasonable to assume that visual artists would also browse digital image collections for inspiration, specific visual elements, and materials and techniques.

Art historians are more adept than visual artists and art students at using the search functions within image databases. Frost et al. (2000) observe that art historians prefer the search function over browsing five to one when compared to generalists. However, art historians often browse digital image collections as well (Frost et al., 2000; Münster et al., 2018). Graham and Bailey (2006) find that art historians use digital image collections to get ideas at the beginning of their research and that, “The serendipitous discovery of visual material not previously known about was considered a great advantage of digital image collections” (p. 22). Münster et al.’s (2018) more recent article draws similar conclusions stating that, “images can often inspire the initiation of a research project through generating ideas for subjects to pursue or triggering further information-seeking activity” (p. 370).
When using digital image collections, art historians value both browsing and searching functions. Frost et al. (2000) note that, “Preference for searching vs. browsing was tied more to the task at hand than to relative expert status. Since experts, by definition, are likely to be more familiar with artist names, types of artistic media, and titles of specific works, they would be more likely to be able to effectively use a search option than would generalists” (p. 294). They conclude that both information searching and browsing should be supported in a digital image database to best meet user needs.

**Refocusing the Collection for Art Researchers**

With a better understanding of the information seeking behavior of art researchers, the project team assessed the Shields Trade Card Collection from an art user perspective. Special attention was paid to the collection’s coverage of art subjects and the collection’s browsing and searching functionality. The informal assessment revealed usability problems for art researchers. While trade cards can be considered art objects, especially within the history of graphic design, the Shields Trade Card metadata proved insufficient for art research needs. Searching ability was particularly affected by the absence of art-related metadata. For example, art users were unable to search the earlier digital collection for traditional art information such as measurements, medium, technique, object type, style/period, and culture. Moreover, the Getty’s *Art & Architecture Thesaurus* (AAT) was not used as the controlled vocabulary for description and access, creating additional problems for discovery by art researchers. Art historians, in particular, were impacted by the omission of the AAT, as they are the group of art users most likely to use the controlled vocabulary for searching.

Ambiguous metadata and the lack of art subject indexing terms also impacted the browsing ability of art researchers. Basic art information taught in art history courses such as artist, title, and date were inconsistently applied across the collection or omitted. The absence of trade card titles in the metadata confused students who needed to accurately cite the materials for research. Dates were also listed in an idiosyncratic manner. For instance, an undated trade card was typically described as “circa 1900” in the date field. Although controlled vocabularies are typically associated with information retrieval, the absence of the Getty’s *Art & Architecture Thesaurus* also affected browsing ability. When shared across a collection, controlled vocabularies can be displayed as browsable facets that users can use to explore related items. The lack of structured art metadata in the collection provided an insufficient number of access points for art users.

Given the importance of both browsing and searching strategies for art researchers and the limited art subject coverage in the collection, the decision was made to update the Shields Trade Card Collection with robust metadata relevant to art researchers. Special attention was paid to increasing the art subject coverage of the collection as this would also improve browsing and search results for art users. Enhanced subject classification included both the addition of controlled vocabularies for art terms as well as textual transcripts and free-text subject descriptions of the illustrated content. As Hepburn and Fiscella (2006) conclude, “the research suggests that extensive descriptions have a positive influence on retrieval of the records. The authors find correlation simply between the number of unique instances of words used in the image record and the number of retrievals. The correlation was stronger in cases where a creator was unknown or the image was not attributed to a named person” (p. 351), as is the case with the Shields Trade Card Collection.
The new project team identified three ways to revise the collection to bring it more in line with art research needs: enhance art subject access, standardize titles, and add salient art metadata fields such as technique and measurements. Creating additional art subject access points proved to be the most impactful and time-intensive part of the repurposing project and will be discussed in greater detail below.

New Template

After the preliminary assessment of the digital trade card collection revealed usability problems for art researchers, the project team decided to create a new metadata template for the collection that was better suited to visual resource description. The digital collection originally used a Qualified Dublin Core metadata schema. Although Dublin Core is a versatile and popular metadata schema, it proved inadequate to capture the wide artistic and subject context of 19th and 20th-century trade cards. In the revised collection, VRA Core was selected as the most appropriate metadata schema given its focus on visual resource description. Adding the VRA Core fields for measurements, work type, and technique allowed for more art description of the materials. These fields also created additional art access points through the use of specialized controlled vocabularies for art collections.

The Digital Collections Librarian also changed the content standard used to describe the trade cards. While the previous iterations of the digital collection followed the Library of Congress’s AACR data standard for most fields, the revised collection adopted the Cataloging Cultural Objects (CCO) data standard for material description with one significant exception. Titles were created according to the Descriptive Cataloging of Rare Materials (Graphics), the descriptive cataloging standard widely used by the rare book and manuscript library community. Unlike CCO, which is primarily used to describe visual objects that lack formal titles, DCRM(G) provides guidelines for cataloging graphic materials “typically found in rare book, manuscript, and special collection repositories,” such as prints, posters, advertisements, photographs, and other materials that usually include text, such as trade cards (ACRL, 2013, p. 7). Following DCRM(G) guidelines, the trade card title was taken from the advertising headline, caption, or the prescribed illustration title, when available. Adding titles helped differentiate the trade cards and provided more context for users browsing the collection.

Improving subject access for art users was a central goal of the Shields Trade Card revision project. Adopting a VRA Core metadata template allowed the project team to create new art history-related fields such as measurements, technique, style/period, cultural context, and work type. With the exception of measurements, these new fields all used the Getty’s AAT controlled vocabulary. Another Getty vocabulary, the Thesaurus of Geographic Names (TGN), was used for the subject location field. Adding AAT as a controlled vocabulary improved the discoverability of the collection for art researchers by facilitating the textual analysis of visual media. “Without verbal/textual interpretation, the content of the image is locked within itself - - inaccessible to others and individualized by each viewer” (Molholt & Peterson, 1993, p.31). Since their inception in 1980, the Getty vocabularies have been not only a classification system, but a system of retrieval, “…used as search assistants in database retrieval systems and on the Web. They are knowledge bases that include semantic networks showing links and paths between concepts; these relationships can make retrieval more successful” (Harpring, 2010, 67). These changes also address art researchers’ needs in terms of inspiration and specific visual elements by facilitating browsing and searching. The addition of the technique and object type fields provide access points for art users seeking information about materials and...
printing techniques. Taken together, the new art focused metadata realizes the three main motivations that underlie visual artists’ and art students’ research needs.

Despite the aim of the revision project to enhance the subject focus for art users, the AAT was not adopted as the controlled vocabulary for the general subject field. Instead, the project team continued using the Library of Congress’s *Thesaurus for Graphic Materials (TGM)* as the preferred controlled vocabulary for general subject indexing. This decision was made for two reasons: *TGM*’s suitability to index the wide range of artistic and commercial subject matter depicted in trade cards, and the project team’s aim to repurpose as much of the original metadata as possible to save time and labor. With its emphasis on general subject classification, the *TGM* is an appropriate controlled vocabulary to index the broad range of subjects depicted in American trade cards. Unlike 19th-century European trade cards, which were produced for a small number of industries and manufacturers, American trade cards were issued by manufacturers, retailers, and distributors in virtually every industry (Twyman, 2013). The content depicted in these trade cards consequently covers a wide range of human activity and objects more appropriately covered by the generalist *TGM* subject classification than the *AAT* (Harpring, 2013, p. 79). To improve art subject coverage of the materials, the Digital Collections Librarian developed new subject classification guidelines to index the artistic content depicted in the trade card illustrations as well as the card’s commercial function. Existing subject terms were reviewed for exhaustivity and specificity of art subject coverage and new general subject terms were added to index relevant art historical content.

Art subject access was also enhanced with the addition of subject description notes in the new metadata template. Following the recommendations of the *CCO*, free-text notes were added to the metadata records to describe the content of trade card illustrations in clear and simple language (Baca, Harpring, Lanzi, McRae & Whiteside, 2006). In addition to summarizing the aboutness of the illustration, the description notes aided subject classification by describing the relationships between indexed terms. As Hourihane (2002) discusses in his chapter on image subject access, free-text description notes complement indexed terms. He recommends that catalogers of visual resources use both controlled vocabularies and free-text descriptions to classify subject content. The description notes also provide a crucial subject access point for users with accessibility concerns. By describing the illustration in simple language, the description note also provides a narrative to help users “visualize” an image when interfacing with the digital collection through screen readers and other assistive technologies.

In addition to improving subject classification, free-text description notes also assist in image retrieval and discoverability. The trade cards in the Shields Trade Card Collection often lack known artists and formal titles, thus requiring description for retrieval. The additional text supplied by free-text description notes provide more metadata for keyword searching. Description notes are particularly valuable for art users who are conducting searches for inspiration or specific visual elements. Münster et al. (2018) note:

> Usually, users access image repositories through a keyword-based search. This calls for the translation of visual needs into text. In art history, images function as a surrogate of an art object. The focus is on the depicted item, which is typically a known item like a painting or sculpture and specifically searched for using, e. g., title or artist. But, photographs of a cityscape or building, which are of value to architectural history specifically need a description for the retrieval process (p. 375).
Keyword searching was also improved by the inclusion of full-text transcripts of the trade cards using optical character recognition technology (OCR). Transcripts of the printed text provide a similar benefit to art users as description notes. While the description notes were used to describe the illustrated content of a trade card, textual transcripts captured the advertising content of the cards, enabling this information to be searched as well. Users of the collection can now employ multiple information seeking strategies to explore the artistic and commercial subjects in the Shield Trade Card Collection.

The trade card example, *Birdsall & King, Dealers in clothing, hats, caps*, demonstrates the improved usability of the new template for art researchers. Valuable information including title, measurements, cultural context, style period, object type, and technique have been added, enabling successful browsing and searching by creating more access points. This key change is supported by the findings of Graham and Bailey (2006) when they interviewed art historians about their needs when using digital image collections. They discovered that art researchers “...wanted standardized and enhanced descriptive metadata to accompany all digital images, including thumbnail images, full details about individual artists, techniques and dimensions of works, genre, time periods, and so on” (p. 23). Technique is of particular importance, as it fulfills visual artists’ and art students’ needs for information on materials and techniques, in this case chromolithography and die cutting. Browsing options are expanded as a direct result of linked fields. The differences between the two templates can be seen in figure 1 and figure 2. Although a formal usability assessment is still in the planning stages, the new art-focused template has produced improved search results for art users by providing more art subject coverage and facilitating browsing and searching strategies.
Next Steps: Fostering cross-discoverable visual resource collections

Improving the Shields Trade Card Collection metadata for art historical access also better prepares the collection for cross discoverability. To ensure that the metadata can be widely shared by aggregation platforms, the Digital Collections Librarian mapped the VRA Core metadata template to Qualified Dublin Core, the recommended schema for harvesting metadata with OCLC’s WorldCat Digital Collection Gateway. Mapping the metadata template to a commonly used interoperable metadata schema also paves the way to later include the Shield Trade Card Collection in the Digital Public Library of America (DPLA). While aggregators like OCLC and DPLA increase the discoverability of the materials for a more general audience,
the adoption of the Getty controlled vocabularies helps make the Shields Trade Card Collection cross-discoverable with other visual resource collections.

Linked open data (LOD) provides another means of preparing the collection for future cross-discoverability. In 2014, the Getty released AAT and TGN as LOD, allowing collections that use the vocabularies to be integrated into the semantic web and leading to enhanced discoverability. Cole, Han, and Szylowicz (2015) argue that LOD will break down silos and make new connections between disparate collections. This is exactly what art researchers want as Green and Courtney (2015), Palmer, Zavalina, and Fenlon (2010), and Zorich (2012) note. “The challenge before scholars is now to make connections among and within huge sets of digitized data and create new knowledge from them.” (Henry & Smith, 2010, p. 108). Increased discoverability also leads to increased collection use. While the potential of LOD in digital special collections has yet to be fully realized, using controlled vocabularies with linked open data ensures that the Shields Trade Card Collection is better prepared for future integration into the semantic web.

Conclusion

The revision of the Charles and Laura Dohm Shields Trade Card Collection illustrates the importance of balancing art subject access with the information seeking practices of art researchers. When the Digital Collections Librarian and Arts and Humanities Librarian at Miami University teamed up to improve the Shields Trade Card for art users they also considered the beneficial impact that additional art subject access would have on browsing and searching capabilities. Subject specific metadata was added, in part, to facilitate subject specific information seeking practices. The new VRA Core template adds more robust art metadata tailored to art researchers, including: title, object type, technique, cultural context, and measurements. The project team also adopted the Getty’s Art and Architecture Thesaurus for key fields to enable art researchers to browse or search disciplinary terms. With these metadata adjustments, art researchers are able to work with the collection more efficiently using their disciplinary search strategies, obtain results in categories relevant to their research, and compare these results across common themes and types. Moreover, by making the Shields Trade Card Collection more searchable for an academic art audience, the collection become more cross-discoverable, particularly with other art museum databases, where metadata is already focused on the needs of art researchers as a user group.

References


