Factors Affecting Controlled Vocabulary Usage in Art Museum Information Systems

by Alison Gilchrist, Metropolitan Museum of Art

[The following article was the winner of the 2001 Gerd Muehsam Award. The award recognizes excellence in a paper written by a graduate student on a topic relevant to art librarianship or visual resources curatorialship.]

Introduction

The last decade has witnessed a significant change in the way museums manage their collections information. Making the leap from card-based or homegrown systems to 'off the shelf' collection management databases is now a familiar task for the largest and most prominent collections all the way down to those of small local historical societies. Consequently, the shift toward robust digital collections management presents an opportunity for institutions to reconsider the depth, accuracy, and potential uses of their data, and thus implement systems that will best help them achieve optimal standardization for improved retrieval.

In addition to comprehensive data structures and sophisticated pre-defined functionalities for common museum tasks, commercial museum database systems have also begun to integrate vocabulary capabilities for indexing and retrieval of collection records, a trend that corresponds to the development of improved vocabularies for cultural heritage information and increased consumer demand for access. Standardized terminology applied to key data fields is a demonstrated way to ensure improved retrieval performance in most database environments, whether it means selecting terms from an established vocabulary tool or agreeing on a set of authority terms to be used and enforced locally. The most common controlled vocabularies in use for most museum collections include the Getty's Art and Architecture Thesaurus® (AAT), the Union List of Artist Names® (ULAN), and the Thesaurus of Geographic Names™ (TGN). Non-Getty terminology sources include the Library of Congress’s Thesaurus of Graphic Materials (LCTGM I and II), the Revised Nomenclature for Museum Cataloging, and ICONCLASS. Each source varies widely in scope, content, and purpose, and no one source is sufficient by itself to address the wide variety of data required to adequately catalog museum collections.

Vocabularies and corresponding browser tools (most commonly the AAT) bundled with commercial management software provide users the capability of linking appropriate descriptive terms to a record in the database. Not only does linking from an established source help encourage consistent use of terminology, but in the case of faceted or hierarchical thesauri, it enhances searching capabilities as well. Once linkages are established, queries can be broadened or narrowed based upon the position of the term in the hierarchy, synonyms, and/or related concepts. For example, a query on the term portrait may be much too broad to satisfy an information need, but it can be further narrowed to terms such as self-portrait, group-portrait, etc. Conversely, a query on self-portraits can be broadened to include all portraits indexed within the database. Moreover, a query on the term portrait will return items indexed under the term portraiture and vice versa.

Controlled vocabularies for museum use continue to gain attention in the literature as they improve in scope and accuracy and as more institutions migrate their data into electronic collection management systems. Several successful demonstration projects involving networked image and data resources have also raised the level of awareness among museum professionals. The Getty Research Institute (GRI) has been at the forefront of promoting controlled vocabulary usage and educating users on how to incorporate vocabularies into cataloging practices in order to add value to their intellectual property. In spite of efforts by GRI and others to disseminate the benefits of terminology control, the perception among most in the museum community is that vocabulary control is not being widely adopted.

To that end, this study was designed to assess the accuracy of that perception— to discover where museums stand in the divide between theory and practice when it comes to using controlled vocabularies to catalog their collections. Using a representative sample of thirty fine arts museums throughout the United States, each with an electronic collections management system, this study was designed to answer the research question: What factors affect the adoption and usage of vocabularies in art museum collection management databases?

Methodology

An interview protocol was designed for the basis of all inquiry (see Appendix). The nature of questions ranged from simple reporting of known attribute information to those that solicited opinion and personal experience concerning terminology control. By incorporating a wide range of question types (attribute, beliefs, attitudes, and behavior) the instrument design intentionally aimed for both breadth and depth on the topics of vocabulary usage, museum resources, and involvement in networked initiatives.

Participants were selected according to a rigorous set of criteria to ensure that they were likely to be representative of current practice and behaviors, based upon the premise that
those museums that have vocabulary tools available are more likely to consider using them. Thus, all participants had to be currently using an electronic collections management system from a commercial vendor whose product includes the option to purchase and install vocabulary tools. Candidate database systems were identified using the widely recognized collection management database reviews from both the Canadian Heritage Information Network (CHIN)7 and the Museum Documentation Association (MDA).8 Eligibility was based upon the presence of a vocabulary browser or utility incorporated with the system (latest release) for both data entry and query construction. From these candidates, the three receiving the highest review scores, and those with the most comprehensive client lists, were chosen. In order to achieve the desired diversity in the sample, it was important that the client museums represented a wide range of collection size, staff size, and geographic location. Additionally, only fine art collections in the United States were considered, in order to maintain both consistency among the sample and to better conform to the body of previous and ongoing research in this area.

Next, within each vendor's client list, candidates were further limited to institutions that are accredited by the American Association of Museums (AAM). Recognition by this professional body ensures that these museums have undergone rigorous peer review and adhere to the standards and best practices of the professional community.9 AAM accreditation was utilized as both a criterion to further limit the sample and to identify those most likely to have high interest in and awareness of issues involving data standards. Finally, ten institutions from each of the three vendors' client lists were selected based upon their relative collection size, staff size, annual attendance, and geographic location.10

The study questions were administered to the thirty individuals self-identified as the in-house contact responsible for, or most familiar with, the institution's database system. Most often these individuals were within the museum registrar's office, but in a few instances they were curators, catalogers, information technology professionals, or database administrators. Participants were assured that neither their identity nor their institutional affiliation would be compromised if they consented to take part. Twenty-two of the thirty agreed to be interviewed by telephone while eight opted to respond to the questions via e-mail. Not every participant provided answers to every question.

Results

I. Vocabulary Usage

Are American fine art museums in fact utilizing controlled vocabularies? Just under half of the sample (46.7%) said that the availability of a controlled vocabulary and browser utility was a factor in the decision to purchase their software. Interestingly, only 50% of those institutions for which it was a factor actually have it installed. For the entire sample, 19 out of 30, or 63.3%, have a vocabulary and browser utility installed as part of their collections management system, although only 11 of the 19 (57.9%; 36.7% of the total sample) are using it when entering data into the system. A total of 18 institutions among those who responded (60%) are using at least one controlled vocabulary reference when entering data into their systems.

The number of institutions that reported using a controlled vocabulary for data entry was spread comparably among users of the 3 collection management software packages. AAT and ULAN users each comprised half the clients for 2 of the systems and two-thirds for the third. Several participants, however, remarked that even when given the option of installing AAT or ULAN locally with their database system, they opted not to. Reasons included lack of ready access to the most recent term list updates without having to wait for system upgrades, reserving valuable disk space for their own data storage, and ease of use. For some, it remains easier to have a Web browser open in the background than to have to travel through a large number of screens within the database. The consistent distribution of vocabulary use within each vendor group along with the 57.9% use rate among those who have a vocabulary installed suggests that average use of controlled vocabularies, whether local or via the Web, is in the 50-60% range.

During interviews, most participants commented on the circumstances that have led to their museum's current state of terminology control. Among the most representative responses on both sides of the argument were:

- Our curators are still emotionally attached to the old terminology, and as long as they stay consistent we don't mind.
- The newer curators are a lot more agreeable to adhering to some strict sense of how to catalog something whereas the ones who have been here forever don't buy in as quickly.
- We got into the habit of making up terms all over the place. Nobody sat down to say 'before you create your own terms why not look it up and see if it's already there?'
- Vocabulary controls were not added to the database until one of the upgrades that were available sometime after it was purchased. By then, a system or pattern of entering data and new records was already in place.
- Using them [vocabularies] is tough because they're still in the process of building them, plus they are proprietary and people don't want to give up their terms to use somebody else's.
- Standardized terminology is a must.
- Vocabulary control is very important to us.

Fourteen respondents indicate they regularly use the AAT, 14 use ULAN, and 7 use TGN. Nobody within the sample reported using ICONCLASS or LCTGM 1 or II, and only 2 respondents mentioned they occasionally consult The Revised Nomenclature. Eight museums use both AAT and ULAN, 7 use AAT and TGN, and 7 use both ULAN and TGN.

Comments garnered from interviews also reveal some of the most significant barriers to usage of individual vocabularies. AAT, the most widely known and integrated vocabulary, prompted the most feedback. The following statements represent the pertinent comments taken from interview responses in regard to factors affecting adoption behavior of specific vocabularies.
Regarding the AAT:

- The AAT is too complicated. It’s over-structured. Most people feel it’s very complicated... I envy the simplicity of the Library of Congress subject headings. Cataloging a work of art is far more complex than cataloging a book.

- I find it annoying. [The AAT] goes too far, it’s too clunky. Also, our collection is encyclopedic—it covers all times and all parts of the world. I just don’t know that using it is that helpful for our collection.

- Intellectually, it’s still hard to grasp. People don’t know how deep or broad to go into the hierarchy.

- We don’t think of terms within a hierarchy—the curators don’t see or use the terms in the same way.

- Yes it’s installed, but we just don’t have the staff to deal with it.

- ...steep learning curve...

- It’s a good functionality and a great feature, we’re just not using it. Part of the reason is that the terms for the types of objects we have are not highly developed. But the hierarchy is a good concept.

- It’s useful once you get the hang of it.

Regarding ULAN:

- We tried to use ULAN as a name source, but not only is it geared toward Western artists, but it’s odd with names in [other languages].

- For non-Western art, I don’t even bother to look at those tools. An old French dictionary has more Japanese artists listed than ULAN.

- ULAN is very limited for Asian collections—well, all non-Western, and we have a broad collection. We’d like to only have to use one or two sources.

- There are very clear cases when ULAN is wrong, so our local information is preferred.

- ULAN comes in handy sometimes, but sometimes we don’t trust the dates. But probably of all of them ULAN is the easiest to use.

Regarding TGN:

- TGN—doesn’t have all places we need and preferred term isn’t our preferred term—they use language of country rather than English. Sometimes we want to use historical terms for place rather than current term.

- I would use TGN also although I get too caught up in the information because it’s too interesting, so I try not to go in there.

Statistical analysis of the data revealed strong correlations between two sets of variables: those who use the AAT and TGN and those who use ULAN and TGN. Unfortunately, due to the very small sample size, it is problematic to conclude with any level of confidence that use of one Getty vocabulary significantly correlates to use of another. It is more plausible to assume that institutions that are familiar with how to use one or more of the Getty vocabularies are likely to consider using others as time and individual data requirements permit.

Aside from an external vocabulary, 26 of the 30 institutions (86.7%) reported that they are using customized lists of terms (referred to throughout this study as authority control). Most often these term lists are used in fields that describe object types, materials, and geographic origin, but also include artist/creator names, styles, or periods. Preferred sources for terms mentioned by interviewees include Getty tools, the Revised Nomenclature, Hall’s Dictionary of Subjects and Symbols In Art (London: J. Murray, 1979) and the Bibliography of the History of Art ([Electronic Resource] J. Paul Getty Trust and Institut de l’Information Scientifique et Technique, ©2001)(BHA). Other common sources indicated were curators, educators, scholarly publications, and institutional memory ("the way it’s always been done"). One-third of the sample uses only their own authorities when cataloging museum objects. Only 2 institutions reported utilizing neither vocabulary nor authority control in their systems.

II. Museum Resources

Just under half of all participant institutions reported 3 or fewer staff members responsible for entering new data into the database. The maximum number of cataloging staff in an institution using vocabulary control is 12, and the average is 4.1. The average number of cataloging staff among institutions not using vocabulary control is 8.9. These numbers suggest that fewer staff with data maintenance responsibility may be a factor in successful adoption of controlled vocabulary practices.

The average level of academic training among catalogers was difficult to ascertain with any accuracy, but typical responses were either bachelor’s or master’s degrees in studio art, art history, or museum studies. In 5 instances, respondents indicated that they or someone on the cataloging staff held a library science degree. All 5 of these institutions use both authority control and vocabulary control.

This strong association comes as no surprise, as many museums are beginning to absorb the advances that libraries have made in terminology control and data standards in cataloging systems. A recent article published by members of the Solomon R. Guggenheim Museum’s documentation program points to this very trend: "Automated libraries are a generation ahead of collection management systems and librarians can share with their colleagues their hard-won knowledge gained from their experiences..."12 In that vein, one interviewee observed, "libraries have developed more of a best practice system, but museums have systems that just automate museum work—which is just practice, not best practice." Expressing doubt that the two disciplines can ever fully integrate, another participant contributed, "library systems are rigid in categorization, but museums resist any hard and fast rules. We don’t like to pigeonhole curators into trying to name things the way they should be. Instead, we use ‘guidelines,’ and I don’t know if there is ever going to be a merging."

Twenty-six of the 30 museums have at least 90% of their permanent collection cataloged in their present database system.
The total number of electronic catalog records created by institutions is often much greater than the number of objects in the collection as many museums create records for objects brought in as loans or possible acquisitions. In relation to a 1996 MDA study\(^\text{13}\) that found museums with the largest number of records more likely to use or develop terminology sources, the results of this present study concur somewhat. Out of the 5 museums in the sample with the largest number of records (100,000 and greater), only one uses controlled vocabulary, but all 5 use authority control. Within the next tier of 5—those with electronic holdings in the range between 40,000 and 99,999 records—4 out of 5 are using vocabulary control, and all 5 are using authority control. In the smallest 5 collections in the sample (0-5999 records), 2 out of 5 use vocabulary control and 3 out of 5 use authority control.

### III. Attitudes and Behaviors Toward Collections Management

In order to gauge whether users' attitudes toward, and experiences with, their collection management software influence vocabulary adoption behavior, participants were asked to rate their satisfaction with several aspects of the system based upon their personal experiences. Responses were ratings on a Likert scale\(^\text{14}\) of 1 to 5, with the following designations: Highly Satisfied (5), Satisfied (4), Somewhat Satisfied (3), Somewhat Dissatisfied (2), Dissatisfied (1). The tabulated results of these rankings appear below.

<table>
<thead>
<tr>
<th>Satisfaction Indicator</th>
<th># responses</th>
<th>Mean</th>
<th>Median</th>
<th>Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>Satisfaction with the vocabulary browser tool included with the database (if applicable/installed)</td>
<td>17</td>
<td>3.65</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Satisfaction with the quality of data contained in the database (in regard to consistency, thoroughness, accuracy, etc.)</td>
<td>29</td>
<td>3.55</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Satisfaction with the data contained in query results (in regard to consistency, thoroughness, accuracy, etc.)</td>
<td>27</td>
<td>3.86</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Overall satisfaction with the database for daily use to accomplish job objectives</td>
<td>27</td>
<td>4.25</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>

Satisfaction with the browser tool revealed a strong statistical correlation among institutions that use the AAT when cataloging records (R = .509, p < .001). As previously stated, the primary vocabulary being incorporated into commercial collection management software is the AAT; thus, users at institutions that are using the AAT have more experience using the tools included with the database. The fact that users have an overall positive view of the tools is promising news to both vendors and the museum community. Again, as the sample size is quite small, commenting on the significance of any correlations from the dataset is problematic.

Interestingly, satisfaction with data quality negatively correlates with the number of items in the collection (R = -.422, p < .05), the total number of employees (R = -.436, p < .05), and the museum's annual attendance (R = -.499, p < .01). One might conclude that the larger an institution's operations are, the more difficult it becomes to maintain a consistent standard of thorough and accurate data. These numbers align closely to the previous observations that the more staff members a museum has cataloging the data, the less likely it is to use vocabulary control. As noted above, only one of the institutions with 100,000 electronic records or more is using vocabulary control. The average satisfaction with data quality among those five largest institutions was below the sample average at 3.2, and among those responses was one of only 2 'dissatisfied' (response = '1') ratings received for any of the questions assessing satisfaction.

Since the median length of current database ownership for the sample at the time of the study (excluding previous systems and older versions of the same software) was just 3 years and 4 months,\(^\text{15}\) many institutions were and are still continuing to go through the time-intensive and tedious process of cleaning up records that have been converted out of older systems while also maintaining day-to-day operations. During interviews, many participants spoke in detail about how their institution's priorities concerning collections cataloging affected data. For the most part, implementing new data quality standards has taken a lower priority behind entering new object records and cleaning up records already in the system. Representative comments from those interviewed include:

- A big chunk of my time right now is spent going back and cleaning up the data from previous conversions.
- Our museum has decided to go for breadth rather than depth in the system, so we're getting a lot of data in, just not as clean as one would like.
- The thought of going back to clean up is just overwhelming. Sitting at the computer and having to enter these terms is just so tedious and our staff have so many other things to do. We had resistance even to computing data in the first place.

A number of museums indicated that they now have or are forming committees to address data standards including vocabulary control in the collections database, but these cooperative efforts among curators, registrars, educators, and IT staff are slow to develop and slow to reach consensus. In general, vocabulary control in collections management systems has yet to register as a significant priority for many institutions, but the forming of committees demonstrates that many are starting to think about data quality.

Emphasis on speed over quality was a trend noted by many. Efforts to populate databases have been spurred on by the move to get collections data linked to the Web at a rapid pace. Remarks to that effect include, "We've been trying to finish the basic core fields in order to get something on the Web" and "most major museums have a Web site now, and it's competitive to get collections up there." One registrar from a museum whose collection is completely Web-accessible noted, 'I'm grateful for
our administration that they’ve been so forward thinking. They’ve made the commitment.” Another interviewee in the opposite position commented, "Getting our data in shape needs to be an institutional mandate of how to spend time, energy, and money, and it needs to come from the director."

The perception that most museums have their collections online, however, may not be entirely accurate. From this sample only 9 museums of 29 responding (31%) say that at least some of the information in their database is available through their Web site (whether directly or through data export to an external system). One-third of these museums are affiliated with universities that maintain some form of electronic union catalog of museum, library, and archival holdings. Even when museums do make collections data available, it is typically only a percentage of the total collection or records that represent the institution’s signature works. One database administrator surveyed argues, "what good is having it on the Web unless the whole database is there so you get a complete search?"

As the third indicator of satisfaction, participants were asked whether or not they use any vocabulary tools when constructing queries within the collection management system. Eleven out of 30 (36.7%) responded that they did, though most did not make a clear distinction between using a utility that allows manipulation of the query based on term linkages and choosing a pre-defined term from a drop-down list of authority terms. The results of perceived user satisfaction with query results did not correlate to any other factors, including whether or not an institution uses controlled vocabularies. The recurring comment concerning query satisfaction was typically a variation of "what you get out is only as good as what you put in."

Some users expressed frustration at the limitations of the query tools provided with the software and constraints in the database design itself that prevent searching on some fields. Many noted that they frequently had to construct multiple queries to get a complete set of results. However, in spite of these frustrations, satisfaction with query results received a slightly higher score than satisfaction with the quality of the data itself, most likely because participants recognize the limitations of the system and have learned ways to work around them.

The high score recorded for the final statement regarding overall satisfaction with the database system for daily use reflects a positive trend. Museum personnel are growing increasingly more comfortable with using technology, and the systems and tools they are using continue to improve. One participant, relatively new to her system, remarked, "I’m happier with it every day." Institutions that were the least satisfied typically had older installations of their systems and were either awaiting or contemplating upgrades. In one instance, an interviewee felt that the system his museum had purchased was simply not a good fit for the nature of the collection and the types of functions they needed the system to fulfill.

Although it is not reflected in the satisfaction rating, a clear division in attitudes emerged between those museums that see their database purpose primarily as a registrar’s tool, and those that consider it more of a global information system. One participant observed: "collections management systems for internal use were not geared for making data pretty for the public. [They are] a workhorse tool for large numbers of objects with a large amount of information which can work for you and leave a trail." Those in the latter category tended to express doubts and frustration with the tools currently available on the market for linking data from the database directly to the Web site and integrating large numbers of related images. While a few felt they had been promised something better than they purchased, most were optimistic that the tools will continue to improve with time and experience. All acknowledged that it is better to have something than nothing, and for overall daily use, the systems gained a high satisfaction rating.

IV. Networked Initiatives

Half of the institutions in the study indicated that they share their collections data in some form of networked access. Though not a factor in selection for participation in the study, AMICO members composed exactly one-third of the sample. Five out of the 10 AMICO members are among those using controlled vocabularies. Several non-AMICO member institutions expressed interest or plans to contribute to the AMICO library in the future. Costs and time involved to prepare data for export were frequently cited as drawbacks or barriers to AMICO membership, but overall it was highly desired among those that are not currently members. From an administrative point of view, however, the idea of contributing resources to a large-scale networked initiative is generally secondary to providing information via institutional Web sites. Once museums achieve an acceptable level of data quality, they may revisit the decision to contribute to data-sharing projects such as AMICO as a way of extending their internal efforts toward data quality and access.

Other types of partnerships mentioned by study participants include contributions to union library catalogs as previously stated, and established consortia such as the University of California at Berkeley’s Conceptual and Intermedia Arts Online (CIAO). Two museums indicated they are in the planning and pilot phases of data-sharing projects with other museums or regional institutions.

A representative from one museum indicated that his institution’s interest in networked access is shifting toward metatagging projects, standards initiatives, and search-engine tools for cross-database querying. "We’re moving away from wanting to contribute physical datasets to other databases.” Instead, the idea is for a sophisticated portal that links numerous distributed databases, searchable with one engine that has built-in vocabulary tools to handle differences in terminology as well as spelling variants and synonyms. The portal concept would diminish some of the costs and time required to export customized datasets, and it is, in theory, infinitely scalable. Developments in metadata crosswalks among disparate element sets will also help to increase the accuracy and precision of retrieved sets from distributed databases. Controlled vocabularies, while still of great importance locally, may prove to be not as necessary in the networked realm when it comes to guaranteeing high precision retrieval capabilities.

Conclusion

Assuming that just under half of the nation’s fine art museums have adopted vocabulary control as a regular cataloging practice, the professional museum community still has a long way to go to achieve objectives toward this particular aspect of data standardization, if in fact it is deemed universally desirable.
F-d-users can no longer claim that they are waiting for electronic versions of the vocabularies or that they are blindly following in the footsteps of their predecessors. The advent of the Internet has made both excuses obsolete, for now access to both vocabularies and best practices for museum data standards is ubiquitous.

Additionally, the role that collections management software plays within museums will continue to evolve. Personnel in education, user services, and exhibitions departments are finding ways to adapt the software for their unique purposes, and in doing so, challenge the traditional view that these databases are to be used exclusively as registrar’s tools. As Howard Besser observed in 1997, “... recent advances [re: technological limitations] are likely to promote the convergence between these two camps.” With increasing levels of use for more diverse purposes, museums may learn through experience how valuable a standardized vocabulary can be when it comes to capitalizing on their data assets.

Murtha Baca, Head of Vocabulary Standards for the Getty Research Institute, cites knowledge—or lack thereof—chief among the obstacles to using vocabularies. “Vocabularies have to be available and museums have to know about them and know that they are important.” She also points to the changing roles of museum personnel, noting that previously non-existent job titles such as "Data Standards Administrator" will start to become more prevalent, and with them, more comprehensive and enforceable data standards programs. Additionally, Baca indicated that work on a new Getty publication is currently underway, one that addresses subject access for image collections that may assist end users in learning how to incorporate vocabularies into everyday practice.

The gap between awareness of vocabularies and skill in using them perhaps could be further narrowed if institutions commit to continuing professional education for staff that extends beyond introductory training seminars. Dissemination of innovation also could be greatly enhanced if collections management personnel were to develop local, regional, or national networks among themselves to foster support and collaboration.

No single profile accurately characterizes the type of museum that is a good candidate for successful implementation of vocabulary control standards, although the results of this study strongly indicate that the presence of individuals with information and/or library science education is a factor in improving adoption behavior. While this study also shows that institutions with smaller numbers of records and fewer staff actively involved in cataloging are more likely to use vocabularies, actual implementation varies widely with regard to the nature of their collections, the priorities of their administrations, and individual staff workloads. Among those participants who are currently using vocabularies, many indicated that they are actively writing their own documentation and conducting small training seminars within their museums to supplement training received from vendors.

The results of this study show long-standing habits and traditional traditions emerging as the most significant barriers need to be overcome before a museum can objectively approach terminology control. The advent of networked resources and the drive to the Internet have sparked a great deal of change in the way museums view their intellectual capital, but the realities of change remain governed by availability of resources including staff time, training, and financial and administrative support. Close behind tradition are the tools themselves. Complaints of complexity, inaccuracy, steep learning curves and insufficient scope are serious concerns for many, and are factors that have a profound effect on an institution’s decision to adopt one or more vocabularies. Finally, the time involved presents a substantial obstacle to museum personnel who are barely managing to complete daily workloads without the added burden of a terminology control project.

The author wishes to thank Dr. Helen R. Tibbo and the faculty of the School of Information and Library Science at the University of North Carolina at Chapel Hill for their assistance with this paper.

APPENDIX: Interview Protocol

(For items marked with "*" please estimate if exact figures are not known or readily available)

1. What software package is your institution currently using for collections management?
   What version?

2. Is this the first commercial database system your institution has purchased?

3. How long have you had this system? ___yr(s) ___mo(s) *

4. Total number of objects owned by the museum? *

5. How many museum-owned objects are currently recorded in the database? *(For instances of multiple objects with one accession number, estimate the number of accession numbers.)

6. How many total object records are currently in the database? (accounting for instances where the museum records loans, potential acquisitions, etc.) *

7. How many staff enter new data into the database? *

   In general, what positions do these staff members hold? What is their level of academic/professional training?

8. Have you personally used any vocabulary tools in your own career or training?

   If yes, what tools, and describe your impressions or experiences.

9. Was the availability of vocabulary tools or a built-in vocabulary browser (such as the Getty’s Art and Architecture Thesaurus, Union List of Artist Names, Thesaurus of Geographic Names, ICONCLASS, or Library of Congress’s Thesaurus of Graphic Materials) a factor in your institution’s decision to purchase this system?

10. Are any vocabulary tools (such as those named above) installed on your system?

    If yes, which ones? Continue to question 11.

    If no, why not? Skip to question 13.
11. Do catalogers at your institution regularly use formal vocabulary control sources when entering new records?
   If yes, which ones?

12. Do database users at your institution regularly use vocabulary tools when querying the database (e.g., querying on a broader or narrower term in a hierarchy)?

13. Do catalogers at your institution regularly use authority control (i.e., an in-house or customized list of terms, not a formal thesaurus) when entering new records?

14. Does your institution offer staff training in how to use the database?
   If yes, describe the nature and availability of training. In what aspects (e.g., data entry, querying, report writing, World Wide Web access, etc.)?

15. Please rank the following based on your personal experience, and feel free to comment on any of the selections:
   Satisfaction with the vocabulary browser tool included with the database?
   (If applicable/installed)
   Highly Satisfied Somewhat Satisfied Somewhat Dissatisfied Dissatisfied

   Satisfaction with the quality of data contained in the database (in regard to consistency, thoroughness, accuracy, etc.)?
   Highly Satisfied Somewhat Satisfied Somewhat Dissatisfied Dissatisfied

   Satisfaction with the data contained in query results (in regard to consistency, thoroughness, accuracy, etc.)?
   Highly Satisfied Somewhat Satisfied Somewhat Dissatisfied Dissatisfied

   Overall satisfaction with the database for daily use to accomplish job objectives?
   Highly Satisfied Somewhat Satisfied Somewhat Dissatisfied Dissatisfied

16. Is the database, whether in part or whole, accessible to the public?
   If yes, how? (check all that apply):
   kiosks located in gallery space in conjunction with objects on view
   a designated computer area apart from the collection
   by appointment only
   via the museum's Web site
   via another Web site

17. Does your institution currently collaborate or have future plans to collaborate in any form of networked access to your collections data with any other institution (i.e., museum-library partnerships, AMICO membership, union catalogs or databases, etc.)?

18. Do you have any other comments or experiences relating to your institution’s use of controlled vocabularies or vocabulary tools that you would like to share?

Notes
10. Collection size was estimated from institution Web sites and confirmed by participants asked to self-report. Staff size and attendance was obtained from The Official Museum Directory (Washington, DC: The American Association of Museums, 2001). Geographic distribution was determined in accordance with the regional classifications used by the American Association of Museums located at http://www.aam-us.org/awiso/map.htm.
11. Statistical Analysis on Selected Vocabulary Variables

<table>
<thead>
<tr>
<th></th>
<th>AAT</th>
<th>ULAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pearson's R</td>
<td>.584</td>
<td>.584</td>
</tr>
<tr>
<td>P</td>
<td>.001</td>
<td>.001</td>
</tr>
<tr>
<td>Pearson's Chi-Square</td>
<td>.002</td>
<td>.002</td>
</tr>
<tr>
<td>N (sample size)</td>
<td>29</td>
<td>29</td>
</tr>
</tbody>
</table>

13. In 1996 the MDA conducted a survey of terminology in UK museums in which only 257 of 2000 questionnaires mailed to museums of all types of collections throughout the UK were completed. In spite of a low response percentage, the results indicate that "many of those institutions which reported the largest number of machine-readable records use—or are developing—thesauri." Survey of Terminology in [UK] Museums, 1996, MDA ©1996. http://www.mda.org.uk/survey/htm (accessed 27 June 2001).
14. A Likert scale is a standard statistical method used to measure attitude via the extent to which a person agrees with a set of statements. Each degree of agreement is given a numerical value, most commonly on a scale from one to five, and numerical values can then be calculated from all the responses.

15. Mean is calculated as the sum of all response values divided by the total number of responses; also known as arithmetic mean or average. Median is the middle value of all responses, and mode is the value of the response that occurred most frequently.

16. Short-term system ownership reflects the rapid changes in the software and technology industries along with a surge in system turnover just prior to the year 2000 due to Y2K incompatibility issues.

