WORLD LIBRARY AND INFORMATION CONGRESS:
74ᵗʰ IFLA GENERAL CONFERENCE AND COUNCIL • 10 - 14 AUGUST 2008

Québec 2008

IFLA

CONGRÈS MONDIAL DES BIBLIOTHÈQUES ET DE L'INFORMATION :
74ᵉ CONGRÈS ET ASSEMBLÉE GÉNÉRALE DE L'IFLA • 10 - 14 AOÛT 2008

Date : 18/06/2008

Sharing Standards and Expertise in the Early 21st Century: Moving toward a Collaborative, "Cross-community" Model for Metadata Creation

Murtha Baca, Getty Research Institute
Elizabeth O'Keefe, Morgan Library & Museum

Meeting:	**156. Cataloguing**
Simultaneous Interpretation:	English, Arabic, Chinese, French, German, Russian and Spanish

Abstract

This paper provides a brief overview of the evolving descriptive metadata landscape, one phenomenon of which can be characterized as "cross-community" metadata as manifested in records that are the result of a combination of carefully considered data value and data content standards. he online catalog of the Morgan Library & Museum provides a real-life illustration of how diverse data content standards and vocabulary tools can be integrated within the classic data structure/technical interchange format of MARC21 to better describe unique, museum-type objects, and to provide better end-user access and understanding. The Morgan experience also shows the value of developing a collaborative model for metadata creation that combines the subject expertise of curators and scholars with the cataloging expertise and knowledge of standards possessed by librarians.

Keywords
cataloging codes, cataloguing, metadata, controlled vocabularies, thesauri, data content standards, FRBR, CCO, MARC, AACR, bibliographic records, art objects

Introduction
Almost every library owns some works of art or cultural objects that fall outside the parameters of traditional library collections, and that cannot be adequately described using the standard library cataloging rules and vocabularies. By the same token, many museum collections include some printed works, for which guidelines developed for describing unique objects may be inappropriate. Creating documentation for unique cultural items requires close collaboration between librarians and non-librarians; it also

requires seeking guidance from standards and vocabularies developed by and for other metadata communities.

At a time when libraries, archives, and museums are struggling to create the kind of immediate access and instant gratification that Google seems to offer, the area of metadata standards is experiencing a period of profound evolution. Emerging trends in metadata creation and maintenance include "schema-agnostic" metadata, "cross-cultural" metadata, and a judicious combination of, contribution to, and mapping of thesauri and other types of controlled vocabularies. All of these approaches have great potential for the future. For the present, most librarians must still work within existing frameworks: traditional integrated library production systems; the MARC format; and AACR and its putative successor, RDA. The Morgan Library & Museum's experience with the use of multiple standards to provide access to collections that consist of a mixture of library and museum materials illustrates some of the benefits and challenges of such an approach.

The Evolving Metadata Landscape

"Metadata"—which in many ways can be seen as a late 20th-early 21st-century synonym for "cataloging"—is seen as an increasingly important (albeit frequently sloppy, and often confounding) aspect of the explosion of information available in electronic form, and of individuals' and institutions' attempts to provide online access to their collections.[1] There is a considerable degree of validity in Michael Gorman's contemptuous assessment of "metadata,"[2] at least in some of its current forms (Gorman also decries "the wretched Dublin Core"[3] and we tend to agree, at least as a container for descriptive cataloging), but it is our contention that the librarians of the 21st century can make metadata creation into a viable, effective tool for enhancing access to the myriad resources that are now available in electronic form. The judicious, carefully considered combination of various standards can facilitate this.

A recent trend in metadata creation is "schema-agnostic" metadata. Although because of its origins in the library world RDA presupposes MARC as a vehicle for cataloging records, this emergent cataloging code could also be used with MODS, Dublin Core, or other metadata schemes. *Cataloging Cultural Objects* (CCO), though linked to schemas that have been developed by the art museum and visual resources communities such as CDWA Lite and VRA Core, can also be effectively used with more library-oriented standards such as MODS or MARC, as in the example of the Morgan Library & Museum. We shall also see how a variety of data value standards, from LCSH to the *Art & Architecture Thesaurus* (AAT) and others can be used together within a single record to enhance description and access.

Another recent trend in metadata creation is the notion of description (a.k.a. cataloging) as a collaborative, incremental process, rather than an activity that takes place exclusively in a single department within an institution (in libraries, this has traditionally been the technical services department). Metadata creation in the age of digital resources can and indeed should in many cases be a collaborative effort in which a variety of metadata—technical, descriptive, administrative, rights-related, and so on) is added incrementally by trained staff in a variety of departments, including but not limited to the registrar's office, digital imaging and digital asset management units, processing and cataloging units, and conservation and curatorial departments.[4] As the Morgan experience described below demonstrates, the participation of curators can be a critical

factor in the description of unique, museum-type objects. This kind of contribution from curators and other subject experts can enhance the intellectual value of records, while helping to cut time and costs for creating high-quality descriptive metadata. The incorporation of input from curators, scholars, and other subject experts is an area that institutions should actively pursue, if they want to provide rich, accurate descriptions of the non-bibliographic works in their collections. Information from non-cataloger subject experts could be routinely captured if there are effective methods for communication and collaboration between catalogers and curators. Expert social tagging—that is, the inclusion of keywords, names, and subject designators by experts who are not part of an institution's official cataloging unit—may also prove to be an effective method of enhancing descriptive metadata records.[5] But before that can happen, both the technical infrastructure (appropriate tagging software that can enable the layering of user-created metadata on top of structured metadata records) and the organizational and human behavioral changes (the notion that many people can contribute to the process of cataloging) will have to be in place.

What Are We Cataloging?
For many decades, in traditional library cataloging, what was being described was an expression or manifestation or item (to use the FRBR terminology) that was, in a sense, "static"—e.g., a particular edition of a novel or monograph, or an issue of a journal, or a musical composition or theatrical play published in a particular year. These types of published materials could be described in a MARC record that reflected something that was essentially unchanging. We might characterize these as largely static (albeit updatable) records for static resources.

More recently, catalogers have been faced with the challenges of describing works that are continuously evolving—resources such as Web sites, which change constantly, or online publications that are periodically updated. These types of publications are handled as "integrating resources," and the metadata records for them, which are occasionally updated to reflect the most recent versions, reflect their dynamic nature. Thus we have dynamic records for dynamic resources.

We might describe a third type of situation in which there are dynamic records for essentially static resources. The scenario here would be that of collaboratively, incrementally produced metadata records in a museum's collection management system, a visual resource collection's digital asset management system, or an integrated library system. In this case, works of art, architecture, or visual or material culture can be described in a process whereby the records are enriched through a workflow that involves staff from various units of the institution. For example, at the moment of digital capture, technical metadata is created (or machine-generated) and associated with the digital asset by staff in the imaging unit; metadata librarians (a new job title that has emerged since the turn of the 21st century) might add core collection-level or item-level descriptive metadata for the work that has been digitally rendered, or this data might be automatically "inherited" from the integrated library system or collection management system. Subsequently, the descriptive metadata might be enriched by trained staff in the cataloging units who have specific subject expertise, and additional metadata might be added or overlaid by researchers or scholars who work with the collections (in physical or

digital form) and attach indexing terms or access points to pre-existing records via some kind of social software application.

Cataloging of unique, museum-type materials— whether in a library production system or elsewhere—requires different approaches, different standards, different skill sets and subject expertises. For most of these materials, the "item in hand" is not the source of core information, as it is for published materials. Works of art, architecture, and material culture are not self-describing in the way that a published book or musical composition is.[6] Most of the works held in museums and similar collecting institutions do not have the equivalent of a title page that indicates the name of the work, who created it, and where it was created. In addition, for visual collections such as photo archives and what used to be called slide libraries—there is no "item in hand" at all—catalogers in these types of institutions are typically describing images of works of architecture in another country, or of tapestries that have been destroyed or stolen, or of drawings that are held in different collections, to give just a few examples. The sources of the descriptive information for these works and their images must for the most part be sought elsewhere than on or in the items themselves. There are also the issues necessarily raised by what amount to "hybrid" metadata records, in which some of the fields or elements describe the original work, and others describe the visual or digital surrogate of that work. *Cataloging Cultural Objects* deals explicitly with questions such as "What are we cataloging?" (work and/or image, groups, collections, items, related works) and how to create descriptive metadata for elements such as title, object type, date, style, and so on, that will enable users to find what they are seeking and to understand what they have retrieved.

The Morgan Library & Museum Experience: a Case Study

Background on the Morgan and Its Collections

The Morgan Library & Museum began as the private library of the financier and collector Pierpont Morgan (1837–1913). In the 1890s, Morgan began acquiring medieval and Renaissance manuscripts, literary and historical manuscripts, early printed books, master drawings and prints, and ancient near Eastern seals, as well as notable collections of art objects. Although most of the art objects were sold or given to art and cultural institutions after his death, the other collections remained in the possession of Morgan's son, J.P. Morgan, Jr. (1867–1943). In 1924, J.P. Morgan, Jr. established the library as a public institution, transferring ownership of the collections to a board of trustees and providing an endowment of $1.5 million for the use of the collections for research and the establishment of an art gallery. Over the years, the Morgan has continued to build on its original strengths, and to add important collections of music manuscripts, early children's books, Americana, and twentieth-century materials.

As the name indicates, the Morgan Library & Museum has a dual mission: it is both an independent research library and a museum. It is one of the few institutions in the United States that collects, exhibits, and sponsors research in the areas of illuminated manuscripts, master drawings, rare books, fine bindings, and literary, historical, and music manuscripts. As a research library, the Morgan is commited to making all its

collections available for study by qualified researchers in its Reading Room and Drawing Study Center. It is frequently the venue for classes, symposia, and lectures, and maintains a vigorous publications program. In fulfillment of its mission as a museum, the Morgan mounts several major exhibitions each year featuring material from its own collections and from museums and libraries in the United States and abroad, and lends material to other institutions for exhibition. It sponsors lectures, gallery talks, performances, and interpretative programming for the general public, students, collectors, and scholars.

The Morgan's holdings consist of a mixture of materials traditionally found in libraries and materials traditionally found in art museums. Textual material preponderates: the Morgan owns approximately 80,000 rare books, 80,000 reference books acquired to support research on the curatorial collections, and 45,000 manuscripts, while its art collections comprise about 35,000 items, including drawings, prints, ancient Near Eastern seals and seal impressions, and art and cultural objects. But many of the textual items are also art works. The collection of medieval and Renaissance manuscripts was formed to reflect the history of manuscript illumination and includes significant masterpieces from the ninth to sixteenth centuries and from all the major schools. The collection of literary and historical manuscripts contains illustrated works such as the Drake Manuscript, the original manuscripts of *Le petit prince* and Perrault's *Contes de ma Mère L'Oye*, and many artists' letters with sketches.

As might be expected, the dichotomy between library and museum is reflected in the education and training of the staff. Those who work with the printed book, music, and literary and historical manuscript collections are mainly professional librarians, or persons who have many years of library experience. The staff in charge of medieval and Renaissance manuscripts, drawings, prints, and art objects, and ancient near Eastern seals are art historians, whose experience is chiefly within museums. Most of the curators are responsible for documenting their own collections. Since those collections consist of unique items, expertise in art history, paleography, literature, or musicology is considered more important than training as a cataloger. Full-time professional catalogers are employed only by the departments in charge of reference and rare books.

History of Collection Information at the Morgan

Prior to the introduction of computer-based information technology, descriptions of the Morgan's holdings existed solely in paper form. Unlike most museums, where collection documentation is centralized within the office of a registrar or collection manager, at the Morgan collection information was created and maintained at the departmental level. Each curatorial department as well as the reference collection maintained its own handwritten accession books, recording brief descriptive and acquisitions data. A central card catalog located in the Reading Room contained fuller records for printed books and manuscripts. The Department of Drawings and Prints maintained a separate card file for its drawings, while the Department of Seals and Tablets relied on a comprehensive printed catalog issued in 1948 and a subsequently published report describing a small gift collection. Additional finding tools included periodically issued reports on new acquisitions, collection and exhibition catalogs, and typed and handwritten lists.

With the introduction of personal computers in the 1980s, electronic record-keeping became feasible for the first time. In 1980, the Morgan became a member of the Research Libraries Group (RLG), a bibliographic utility, and began contributing machine-readable records for printed books to the RLG union catalog, RLIN. The records were created by book catalogers, who adhered to the data standards used by American libraries: for data format, they used MARC (Machine Readable Cataloging); for data content, they followed the Anglo-American Cataloging Rules, supplemented by guidelines for rare book cataloging developed by the Office for Descriptive Cataloging Policy of the Library of Congress and the Bibliographic Standards Committee of the Rare Book and Manuscripts Section of the American Library Association. The RLIN records were used to generate cards that were filed in the Reading Room card catalog, which continued to be the authoritative source for the book collections.

The cataloging of other collections was less centralized and less standardized. The curators who cataloged manuscripts followed informal guidelines and best practices developed over the years by manuscript curators and librarians. They did not enter records for manuscripts in RLIN; this was due partly to a shortage of computers and dedicated lines, and partly to the fact that use of the MARC format required more time and training than curators who were also expected to plan exhibitions, acquire new material, and deal with research enquiries could afford. They elected instead to document new acquisitions in tagged word-processing files. These allowed rudimentary searching and sorting, and could generate printed lists or labels; but the end result was further fragmentation of collection information, since cards for manuscripts were no longer added to the main card catalog. Drawings were cataloged according to descriptive conventions used in published catalogs for drawings; the curators created these records in word-processing files. The level of description varied considerably: drawings for which entries were being written for published catalogs or exhibition labels rated lengthy entries, but the bulk of the collection was documented only in very brief records based on a subset of information from their card file.

By the mid-nineties, the entire reference collection and about 15% of the rare book holdings were listed in RLIN. Some curatorial departments had brief or full in-house records for most of the items in their collections, while other departments had only a fraction of their holdings in machine-readable form. The fragmentation of collection information, and the inconsistencies apparent when information was brought together for comparison, made it increasingly desirable to create a single, comprehensive database. By this time, larger libraries and museums had begun to acquire online systems, and the Morgan's librarians and curators looked with envy on these systems. Consequently, a committee made up of curators and librarians began to explore the feasability of computerizing collection information and centralizing it in an online database.

Two of the most important issues the committee had to address were whether a single system could or should be used, and whether that system should be a library system or a museum system. A library system was clearly more appropriate for printed books, which were already being cataloged into a MARC-based system, and for modern manuscripts, since many libraries were routinely entering records for modern manuscripts into RLIN or their local systems. To test whether the MARC format would work for other collections, the librarians in charge of cataloging reference books began to experiment with creating MARC records for medieval and Renaissance manuscripts and for ancient

Near Eastern seals. These collections were chosen for two reasons: their curators were willing to authorize the experiment, and abundant descriptive data was available for these collections.

There was some precedent for the application of the MARC format to the description of medieval and Renaissance manuscripts. The Library of Congress and several other libraries had entered records for medieval and Renaissance manuscripts in RLIN, and an article on the topic had been published by Hope Mayo, a curator in the Dept. of Printed Books at the Morgan.[7] After several sample records were entered in RLIN, the curators were asked to review them. The records were judged to provide an acceptable level of description, more than appeared in manuscript censuses such as those compiled by DeRicci or Faye and Bond, and they were far easier to update and to disseminate than any published inventory could ever be.

There was no precedent for using the MARC format to document the seals collection, and no formal standards within the field for documenting this type of object. The absence of standards simplified the task. The librarians analysed the data elements that appeared in the entries in the published catalog, discussed their analysis with the curator, and then proceeded to map the information to MARC, and create in-house MARC-format records for the entire collection using Minaret, a MARC-based data capture tool used by the Library of Congress. No effort was made to apply AACR, since the descriptions so clearly came from a completely different metadata universe: for example, seals never have titles (they are identified by inventory number) or known creators, and there is no agreement in the field of ancient Near Eastern studies about chronology.[8]

But the curators in charge of drawings, prints, and art objects remained unconvinced that library standards were appropriate for documenting their collections. They felt that records created in accordance with library cataloging rules would not capture the complexities of art historical information or conform to its descriptive conventions, and that a collection management system like those used in museums would be preferable for the Morgan's art works.

In 1996, the decision was made to acquire a single system to document all the collections, and to choose an integrated library system for that purpose. There were obvious financial benefits to acquiring one system rather than two; moreover, it was felt that a single system would make it easier to achieve the goal of integrated access across collections. The choice of a library system rather than museum system was dictated by a number of considerations, including the preponderance of textual materials, the need to support the library data format standard (MARC) in order to be able to share records for books with other libraries, and the fact that library systems were (and continue to be) more robust and more in step with current technology than museum systems.

One of the authors of this article, Elizabeth O'Keefe, the former Head of Cataloging at the Morgan, and currently Director of Collection Information Systems, was given responsibility for coordinating the acquisition and implementation of the new system, and for overseeing the creation or conversion of collection information and its transfer into the system. She worked closely with the new Head of Cataloging and Database Maintenance, Maria Oldal. In the narrative that follows, it should be understood that these two staff members were the librarians principally involved in collaboration

with the curators on system implementation, data conversion, and application of cataloging rules.

It was understood from the start that the various data projects would involve close collaboration between the librarians, the curators, and the various permanent and project catalogers. Within this framework, the curators would always have the final say on the substance of the descriptive data, i.e. attributions, dating and localization, medium, support, provenance, etc., while the librarians would cast the deciding vote on issues of data mapping, system implementation, selection of appropriate controlled vocabularies, and choice and application of descriptive standards. This powersharing arrangement worked very well for the conversion of the book and manuscript collections, since much of the existing documentation conformed to standard library cataloging practices, or could be made to fit these standards without too much effort. The seals collection had already been converted, and posed no problems.

The drawings, prints, and three-dimensional art and cultural objects were the last collections to be converted. By this time, the librarians felt confident that the information for these works could be easily accommodated in the library catalog. After all, cylinder seals—far more exotic objects than drawings—had been integrated without great difficulty. But when the librarians sat down with the drawings curators to discuss the conversion and mapping of information for their collections, both sides discovered major differences in their expectations about what the system could or should do, and in their views of the nature of collection information and of the cataloging process, the uses to which collection information should be put, and the data elements required to describe items.

The librarians had no experience with the cataloging tradition and environment within which the drawings curators worked. Descriptive conventions that are taken for granted within the world of art historians struck the librarians as aberrant, counterintuitive, and just plain wrong. Data elements considered crucial by the art historians were considered minor by the librarians. The librarians were also unfamiliar with the dynamic nature of art information. They assumed that once a work was cataloged, the record for it would be revisited only if the form of an access point changed (if, for example, a subject heading was replaced by a newer term, or if the name of an author changed as result of a change in marital status). The unanticipated fluidity of collection information also had implications for the budget, since frequent record revision would require additional software licenses and ongoing training.

For their part, the art historians had little experience with the behavior of data within a file or aggregation. Although most of them had prior experience elsewhere with museum collection management systems, which use a database structure, their primary tool at the Morgan for recording collection information was word-processing software, and the primary output for this information was published catalogs. They tended to create one word-processing document per catalog entry, and work on one entry at a time; within the entry, data was arranged in much the same way that it would appear in the publication. Although entries in published catalogs of drawings tend to be quite structured, the structure is generally conveyed by formatting (paragraphs, fonts, and some use of labels) rather than by explicit fielding in the database sense. This made automated conversion very difficult, so briefer, less accurate descriptions entered in tagged word-processing files had to be used for the first phase of the conversion.

The drawings curators were less accustomed than the curators of textual collections to sharing data across institutions or even across departments (the printed book and manuscript departments shared a common card catalog, and the drawings department maintained its own card catalog). They were more reluctant to surrender preferences for, say, a particular form of an artist's name in order to facilitate data sharing. Although conceding that it would be a fine thing in theory if all repositories were to pool their records for art objects, they dismissed this as an unrealistic goal: art historians could not even agree upon the meaning of the underlying concepts, much less reach consensus on the use of the same terminology.

The drawings curators were also reluctant to provide public access to information on their collections until it had been reviewed and revised with the same rigor as information destined for publication. It was difficult for them to believe that records in an online catalog would not be judged by the same standards as entries in a published catalogue; they feared the censure of their peers if incomplete records containing unverified information were exposed to the world. At one point, the librarians were asked to investigate whether the library system could be configured to display preliminary records in a different color, or with a warning in red at the top of the screen advising users that the records had not been checked and verified. (The answer was no.) The curators' reluctance to make what they considered to be substandard information publicly available did, however, have one positive effect: they enthusiastically supported a request for additional funding to upgrade their records.

The preliminary discussions made it clear that one of the first things that had to be done was to identify which differences of opinion involved significant issues and which involved merely cosmetic or trivial matters. For example, title case (i.e., capitalization of every significant word, *Portrait of a Man and Woman with a Dog*) is used for art works, while AACR prescribes sentence case (*Portrait of a man and woman with a dog*). Another example is the art historians' preference for expressing measurements in millimeters rather than centimeters, and adding measurements in inches as well. Early in the proceedings, the librarians agreed that the curators could follow the stylistic conventions they preferred within their own records for art works.On the other hand, the curators had to sacrifice italics when transcribing inscriptions (italics cannot be used in MARC records, or, for that matter, within fields in any database).

Some variations in descriptive practice differed more in degree than in kind, and could be accommodated fairly easily within a library system. For example, description of the physical characteristics of library materials is generally brief. The book-centric nature of library cataloging rules means that the medium and the support most often occurring in library materials—ink and paper—are never specified. Measurements for books and manuscripts generally include only the height, not the length or the width. But AACR and MARC do make provision for more detailed description, if desired, so it was possible to include appropriately detailed descriptions of support and medium and dimensions. Here is a description of the medium and support of a drawing, entered in the 340 (Physical Medium) field:

340 \\$a Paper $a Pen and black ink, with gray wash**,** over traces of pencil**;** framing line in black ink

And here is a description of the dimensions of the Stavelot Triptych, a 12th-century reliquary:

> 340 $b Wings open: height: 19 in. (484 mm.), width: 26 in. (660 mm.), Left wing: height: 19 in. (484 mm.), width: 6 3/8 in. (161 mm.), depth: 1 1/8 in. (28 mm.), Center: height: 18 7/8 in. (480 mm.), width: 12 1/2 in. (318 mm.), depth: 1 3/4 in. (45 mm.), Right wing: height: 19 in. (484 mm.), width: 6 1/4 in. (158 mm.), depth: 1 1/8 in. (28 mm.), Medallions: diameter: 4 1/4 in. (108 mm.), with beading

Other disparities could not be resolved so easily. Some of the major areas of difference between cataloging traditions and how they were resolved are described below.

Self-describing versus Non-self-describing Resources

Most items in library collections are self-describing. They come with prepackaged descriptions that take the form of a title page or title page equivalent; the title page formally and authoritatively presents information about the item. It is the cataloger's job to transcribe what appears on the item into the appropriate fields in a MARC record. Information supplied from sources other than the title page and other approved locations must be bracketed.

At the Morgan, the librarians' first impulse was to apply the same rules to art works. For example, if the description of an Audubon drawing included a note that read: "Inscribed in graphite at lower right, "Siurus inereus. - / Cat Squirrel / Dec 9th 1841 / J.J.A.","," they assumed that the information should be parsed to the appropriate MARC fields. The creator's name as it appeared on the object would be recorded in the 245$c (Statement of Responsibility) subfield, the title-like descriptive phrase would be recorded in the 245$a (Title) subfield, and the date would be recorded in the 260$c (Publication, Distribution, etc. (Imprint) subfield. But as the librarians talked more with the curators about the nature of the information appearing on the object, it became clear that this approach was misguided.

Art and cultural works do not have formal title pages. Information appearing on or affixed to the object is only one of many sources, and it is not regarded as authoritative. It may represent conjecture or wishful thinking by some previous owner (not every work inscribed "Rembrandt" is by Rembrandt). It may be difficult to interpret (is a four-digit number a year, or a lot number from a sale?); or it may be insufficiently descriptive (art historians prefer a descriptive title, in the language of the cataloger, over a word or phrase appearing on the object that may or may not represent the artist's preference). The object cataloger takes pains to record every marking or inscription that appears on the item; this helps to identify it, and may provide clues that someone else can interpret (for example, the notation may be recognized as that used by a particular collector or dealer). But the cataloger is free to prefer information from other sources, such as accompanying documentation, sales catalogs, inventories, catalogues raisonés, or published research, or to make judgments based on his or her own expert knowledge.

As a result of their discussions, the librarians agreed to diverge in several particulars from the rules used for book cataloging. When describing art works, brackets are not used in the Morgan catalog for titles taken from other sources. In the world of self-describing objects, brackets signal a departure from the norm; in a world where the opposite is true, their use is pointless. All inscriptions and markings are recorded in the 562 (Copy and Versions Identification Note) field. As is customary in object cataloging, not just the content but the location of the inscription and the instrument and medium in which it is inscribed are recorded. Information appearing on the object may be used as evidence supporting attributions, titles, etc.; if so, it is also recorded in the appropriate MARC fields. Recording the information in both places clarifies which information appears on the item and which information represents the cataloger's judgment of what is actually true. However, the 245$c (Statement of Responsibility) field is never used for object records. There is a fundamental difference between the formal statement of responsibility appearing on a published item and the occurrence of an artist's name on an object. Art historians are not necessarily handwriting experts; artists may write their name on works they did not create; and it seemed wiser not to use this field to record unofficial statements of responsibility. The revised record for the Audubon drawing would read:

> 100 1\ $a Audubon, John James, $d 1785-1851.
> 245 10 $a Eastern Gray Squirrel.
> 260 \\ $c 1841.
> 562 \\ $a Inscribed in graphite at lower right, "Siurus inereus. - / Cat Squirrel / Dec 9th 1841 / J.J.A."

Another consequence of the non-self-describing nature of art works is the fluidity of descriptive information. Attributions and dates are constantly changing; so are the titles assigned to the works by custom or by art historians. Changes in title may result from language variations (if a painting owned by a Swedish collector is acquired by an English museum), or from reinterpretation of the content of a representational work (a drawing in the Morgan's collection formerly titled "The First Passover" is now called "Romans at a Feast"), or from a change in institutional or even departmental style (the "Female with Child" of one generation of curators may become the "Woman with a Child" of the next).

The Morgan curators pointed out that researchers often seek an art work under a former attribution or title, so that it was important to record this information in addition to the "preferred" title or attribution. It was agreed that former attributions would be traced as added name entries, qualified by $e "formerly attributed to" ($e is the MARC subfield used for relator terms). Variant titles were more difficult to handle within the AACR and MARC framework. Books often bear variant titles, such as cover titles, spine titles, and half-page titles; these are recorded in the bibliographic record, because they pertain to the manifestation as defined in FRBR, while variant titles for a textual work are entered in the authority record. Since the titles of art works change so frequently, following this rule would have required establishing authority records for most of the items in the collection. It was decided that variant titles for art works would be entered in the bibliographic record instead.

Different Concepts of Authorship

Personal authorship is the norm within the library world. Most works in library collections are created by one or more known persons. Corporate authorship is recognized by AACR, but restricted to a narrow range of possible relationships between a work and a corporate body—for example, official publications recording the activities of the organization, or liturgical works used by specific denominations. AACR does not regard families as capable of authorship; however, the successor code to AACR, RDA, will recognize families as authors.

Works created by known artists were easily accommodated within the Morgan's library system. The only contentious issue was the form of name used. The curators frequently preferred a name form different from the form authorized by AACR. Since artist names also appear in records for rare and reference books and for modern manuscripts, and since these records are coded as standard when contributed to bibliographic utilities, the librarians insisted on using the AACR form for all names. If the form of name preferred by the curators differs from the form used in the heading, the variant is recorded in the 545 (Biographical or Historical Data) field. The 545 field, which also includes very brief biographical data (what is known as "tombstone text" in museum circles), displays in the Morgan's online catalog immediately below the main entry:

> **Author/Artist**: Parmigianino, 1503-1540.
> **Biographical Data**: Girolamo Francesco Maria Mazzola, called Il Parmigianino. Parma 1503-1540 Casalmaggiore

Art catalogers take a much more liberal view of corporate and family authorship than book catalogers do. Works are routinely attributed to families of artists, such as the Bassano family, a family of painters active in Venice and the Veneto from the early 16th century to the 17th century, or to corporate bodies that produce objects, such as the Wedgewood or Tiffany companies. The librarians agreed to the use of corporate or family names as main entries for the handful of objects owned by the Morgan that were created by families or by corporate bodies.

Works that cannot be attributed with certainty to a known creator were far more problematic. Librarians are of course familiar with textual works of unknown authorship such as the *Song of Roland* and *Beowulf*. Book catalogers deal with these works by leaving the author field blank, and use the title as the entry element for citations. The Morgan librarians' assumption that a similar approach would be used for art works was met by a firestorm of protest from the curators.

Works that cannot be attributed with certainty to any known person are much more frequent in the realm of art and material culture. Since titles are both unstable and non-distinctive ("Female Nude" may change to "Nude Woman,"and then to "Study for Susanna and the Elders"), it is impossible to rely on them as a stable citation. As a result, a nuanced terminology has been developed to express degrees of uncertainty, and to relate works by unknown artists either to known artists or to schools or cultures:

Attributed to Holbein
Formerly attributed to Holbein
School of Holbein
After Holbein [this may refer to a work by a known artist modeled on Holbein's work, or to a work by an unknown artist modeled on Holbein]
German, 16th century.

When used as access points, the known artist's name is inverted:

Holbein, Hans, 1497-1543, School of.

The curators pointed out that linking the object to the known artist's name minus the qualifier was simply inaccurate, while omitting any artist's name from the record would make the objects unretrievable. Since attributions of this type appear in publications and on exhibition labels, there was no justification for eliminating them from the online catalog. The curators' arguments were irrefutable; the librarians not only agreed to the use of these attributions in the online catalog, but persuaded the Cataloging Advisory Committee of the Art Libraries Society of North America (ARLIS/NA) to submit a proposal to the MARC Advisory Committee to define a new MARC subfield for this information. The proposal was accepted, and a new subfield $j (Attribution Qualifier) was added to the MARC format—evidence that library data standards can be hospitable to metadata from other communities.

Different Importance of Object Type

Perhaps the greatest challenge was how to treat object type. Librarians attach little importance to object type, focusing instead on the information content of their collections. Object type information is distributed throughout the bibliographic record; the field varies depending on the type of object. The object type most often found in libraries, "Book," is recorded nowhere in the bibliographic record; information on object type for non-books may appear in the 007 (Physical Description Fixed Field) field, the 245 $a (Title) subfield, the 245 $h (Medium) subfield, which is used to record what AACR refers to as the General Material Designation, the 300 $a (Extent) subfield, or the 655 (Index Term-Genre/Form) field. During the initial mapping of collection information for drawings, prints, and art objects, the Morgan librarians chose the 300$a field for the object type (i.e. 1 drawing, 1 painting, 1 sculpture), and the 655 field for object type index terms.

But a little experimentation with how this would work in the online catalog showed how unsatisfactory this mapping was, at least within the context of the way library systems currently present data. The 300$a field does not display in initial result sets, so users would have to drill down to the individual record display to discover whether a work titled "Virgin and Child" was a drawing, a painting, a print, or a sculpture. This was unacceptable, given the centrality of this data element for object cataloging. Object type is the single most important piece of information about an art and cultural work. It is the logical focus of the record, and a major organizing principle for

museum collections, just as subject is for library collections. Museum collection information systems define a single, dedicated field for object type which is prominently displayed in search results. Users can browse or search by this element, and input and display screens can be tailored to the type of object. For example, the labels for a drawing would read:

Artist:
Title:
Date of creation:

while the labels for a book would read:

Author:
Title:
Date of publication:

After reviewing the options, the Morgan librarians decided to record the object type in the 245 $h subfield as well as the 300 $a subfield, and to expand the list of terms that could be used in the 245 $h subfield. Of the brief list of General Material Designation (GMD) data values approved by AACR for the 245 $h subfield, only "realia," "graphic," and "art original" apply to the Morgan's art works, and the curators found "graphic" too broad, "realia" meaningless, and "art original" risible (at least in the context of the Morgan's collections). The same term is used in both the 245 $h and the 300$a: e.g.

245 10 $a Virgin and Child $h [drawing]
300 \\ $a 1 drawing

The 245$h field is separately indexed, allowing users to combine a search for object type with some other parameter, such as creator name, title, or provenance. Although the Morgan's production system does not allow customization of labels based on object type, more inclusive labels were devised: for example, "Author/Creator" does duty for both textual and art works.

Creation Information versus Publication Information

In library collections, the place and date of publication of a work are the focus. Both usually appear on the item, and are transcribed along with the name of the publisher in the imprint area of the record. Places, dates, and names associated with the printing, manufacture, and distribution of the item may also be recorded in this field. In museum collections, creation rather than publication is the most important event associated with an object, and the place and date of creation must usually be supplied by the cataloger.

Date of creation was the easier of the two elements to fit into the AACR/MARC environment. Both of these standards define the date area within the imprint area to include date of creation as well as dates associated with publication. Determining the date

of creation is often difficult. Dating is based on internal and external evidence, including the life dates of known creators, or the period associated with the style of the work. For example, a work in the style of the New Kingdom (the last of three great periods of ancient Egyptian civilization), 18th Dynasty, would be associated with the dates of that dynasty. Fortunately, responsibility for dating fell on the curators rather than the librarians.

Art and cultural works are studied in the context of the originating culture, which is usually associated with a geographical area, rather than in the context of the exact place of creation, which can seldom be identified with certainty. If the work is by a known creator, it is studied in the context of the nationality or culture with which the creator is predominantly identified, rather than the place where the work was created. Thus works by Gauguin painted in Tahiti are regarded as French rather than Tahitian. For drawings and objects, the originating culture is recorded in the 655 $x (General Subdivision) subfield of the Genre/Form indexing field:

> 655 \7 $a Drawings $x French $y 18th century. $2 aat
> 655 \7 $a Sculpture $x German $y 16th century. $2 aat

The place of creation is generally unknown for drawings; but many medieval and Renaissance manuscripts can be localized by evidence within the manuscript, such as a dedication to the abbot of a monastery, or liturgies or devotions associated with a specific dioceses. Initially, the known or conjectural place of creation for these manuscripts was recorded in a 500 (General Note) field. This field is invisible in brief search results, and can not be indexed separately. Subsequently, it was decided to use the 260$a (Place of publication) field for this information, although neither AACR nor MARC defines this subfield for creation location. As records for medieval manuscripts or objects with known places of creation are updated for other purposes, the place of creation information is transferred to the 260 $a field.

Importance of Subject Access

Most libraries acquire material for its information content. At the Morgan, only reference material is acquired for its information content; items in the curatorial collections often have informational value, but the primary focus is on the objects as artifacts. Curatorial interest in providing subject access varies considerably, depending on the object type, the perceived needs of researchers, and the resources available for providing subject analysis.

There were no special problems with mapping subject information in existing collection documentation to MARC. Free-text descriptions of content were mapped to the 520 (Summary, etc.) field, and the 6XX fields were used for subject indexing. The 520 field is used extensively in records for three curatorial collections. Most records for literary and historical manuscripts contain summaries, some quite lengthy, of the content of letters and documents. Records for seals and seal impressions and for individual illustrated pages in medieval and Renaissance manuscripts include detailed descriptions of their pictorial content.[9]

Records for all collections except music, drawings, and art objects contain subject indexing terms. Subject analysis is irrelevant to music; musical genres are indexed instead. Subject access is not currently provided for drawings and art works, because the curators feel that researchers are interested primarily in artist, school, and period, not pictorial content. This is an area where the descriptive metadata could be enriched by trained catalogers; although catalogers do not necessarily have the expertise to recognize what is represented in an art work, they could supply standardized access points for real and imaginary persons, places, events, built works, and flora and fauna identified in the titles. Prints fall somewhere in between: original artist prints are viewed as important primarily as artifacts, but the Morgan also owns sizable collections of published prints, most of which are political caricatures or portraits. Subject indexing is considered desirable for the published prints, but resources for the task are not available, so subject indexing is sketchy and sporadic.

Authorized Headings and Controlled Vocabularies

The formulation of artists' names was the main area of conflict when establishing guidelines for standardized access points for names and titles. Applying AACR to the names of persons with other roles, such as collectors, donors, and patrons, which also appear in controlled form on records, did not provoke the same strong feelings. Generally speaking, the curators acquiesced in the use of AACR as a standard for the formulation of name and title headings; their recommendations to the catalogers about authoritative sources for names proved very helpful in constructing authority records for submission to the Library of Congress Name Authority File.

The Library of Congress Subject Headings (LCSH) list is used for reference and and rare books, literary and historical manuscripts, and for the records created to describe medieval and Renaissance manuscripts as whole entities. Records for individual pages within medieval and Renaissance manuscripts contain subject indexing terms drawn from a vocabulary of more than 28,000 terms developed by the Index of Christian Art for indexing medieval art works. A local list of subject terms derived from the published catalog for the seals collection is used for indexing the pictorial content of seals and seal impressions. These domain-specific vocabularies provide a level of specificity that LCSH could never offer: for example, Index of Christian Art subject headings that appear in Morgan records include "Christ—Among Doctors" (19 occurrences), "Christ—Anointed by Woman" (15 occurrences), and "Christ Child—Bathing" (9 occurrences), while records for seals include terms such as "Nude bearded heros," "Nude goddesses with open veils," and "Pigtailed figures."

Controlled vocabularies are also used to provide access by object type or genre. The *Art &Architecture Thesaurus* (AAT) is the main source of terminology for object types for art works; the vocabularies developed by the Rare Book and Manuscripts Section (RBMS) of the American Library Association for physical characteristics and formats associated with books and manuscripts are another source. Local terms are used as needed—for example, "dated manuscripts" and "carpet pages." The curators welcomed the introduction of controlled access points for object type: not only do they permit browsing, but they make it very easy to supply quantitative information about the

number of items in their collections ("how many?" questions used to be dreaded, because they were so difficult to answer in a paper environment).

Conclusion

Both curators and librarians benefited from the implementation of an integrated online catalog at the Morgan. The quality of information for curatorial collections was significantly improved by the clean-up and normalization that took place during data conversion. The ability to locate with a single search information about one curatorial collection or many has transformed every aspect of collection management, as well as the services provided to researchers. And the decision to use the MARC format (in addition to the very practical considerations of the significant additional work and money that would have been involved in implementing more than one information system) has made information on the curatorial collections accessible outside the confines of the Morgan's own catalog: records for all the Morgan's holdings will soon be searchable in Open WorldCat, the public interface to the OCLC database.

Implementing an integrated online system was an achievement in and of itself for the librarians. In the process, they also acquired a greater understanding of and respect for other cataloging traditions. Librarians sometimes have a rather parochial and proprietary attitude towards cataloging, a conviction that book cataloging and library cataloging codes are the only "real" cataloging, and anything else is amateurish dabbling. Learning more about the descriptive practices of other metadata communities and their underlying logic helped immeasurably in making decisions about mapping and system implementation.

"Pure" collections—that is, collections consisting entirely of books, or entirely of art objects—are becoming increasingly rare. This is doubly true of virtual collections: even a small library can acquire an impressive collection of visual surrogates of books, art works, objects of material culture, or natural history specimens. One size will not fit all when it comes to providing access to these resources: not only does each metadata community craft data standards that apply to its own core areas of interest, but it lacks the inclination and the expertise to create standards suitable for use outside its own sphere. In order to satisfy user expectations for one-stop information shopping, cultural institutions must be prepared to judiciously use a variety of metadata standards, and to find ways to blend the various data elements, descriptive conventions, and vocabularies into a harmonious whole. The Morgan Library & Museum's experience in building an integrated catalog for library and museum materials though close collaboration between librarians and curators and a carefully thought-out metadata strategy might serve as a roadmap for other institutions faced with similar challenges.

[1] For an overview of metadata and its applications on the Web, see M. Baca *Introduction to Metadata*, revised edition (Los Angeles: The Getty Research Institute, 2008). Also available on line at http://www.getty.edu/research/conducting_research/standards/intrometadata/.

[2] See M. Gorman, "RDA: The Coming Cataloguing Debacle," available at http://www.slc.bc.ca/rda1007.pdf, where the father of AACR asserts, among other things, that "It is hard to believe that the world's libraries have taken metadata seriously."

3 M. Gorman, "Authority Control in the Context of Bibliographic Control in the Electronic Environment," in A. Taylor and B. Tillett, eds., *Authority Control in Organizing and Accessing Informtation: Definition and International Experience* (Philadelphia: Haworth Press, 2004). Available on line at http://www.haworthpress.com/store/E-Text/View_EText.asp?a=3&fn=J104v38n03_03&i=3%2F4&s=J104&v=38.

4 See "Practical Principles for Metadata Creation and Maintenance," in M. Baca, ed., *Introduction to Metadata* revised edition (Los Angeles: The Getty Research Institute, 2008), pp. 71-72.

[5] Social tagging, also referred to as collaborative tagging, social classification, and social indexing, is the method by which individuals and groups can associate terms, names, etc. (called "tags") with digital resources in an online "social" environment. The results of social tagging in online environments such as Flickr, where users can simply add tags to digital objects, are often referred to as "folksonomies"—a misnomer, since a taxonomy is an orderly classification, and a folksonomy is simply a group of uncontrolled keywords that may or may not be ranked according to how often they are used.

[6] For more on what a "work" means in the museum world, and its implications for cataloging, see M. Baca and S. Clarke, "FRBR and Works of Art, Architecture, and Material Culture"in Arlene G. Taylor, ed., *Understanding FRBR* (Westport, CT: Libraries Unlimited, 2007).

[7] H. Mayo, "MARC Cataloguing for Medieval Manuscripts: an Evaluation," in Wesley M. Stevens, ed., *Bibliographic Access to Medieval and Renaissance Manuscripts: a Survey of Computerized Data Bases and Information Services* (New York : Haworth Press, 1992), pp. 93-152.

[8] The project is described in E. O'Keefe, "From Cuneiform to MARC: a Database for Ancient Near Eastern Cylinder Seals Owned by the Pierpont Morgan Library," in L. McRae and L. White, eds., *ArtMARC Sourcebook: Cataloging Art, Architecture, and their Visual Images* (Chicago: American Library Association, 1998), pp. 180-195.

[9]These records are created by catalogers from the Index of Christian Art at Princeton University as part of a collaboration between the Index and the Morgan to photograph and describe all significant illustrations in the Morgan's collection of medieval and Renaissance manuscripts.