



The Cataloguing Cultural Objects Experience: Codifying Practice for the Cultural Heritage Community

Erin Coburn

The J. Paul Getty Museum, Los Angeles, United States

Elisa Lanzi

Smith College, Northampton, MA, United States

Elizabeth O'Keefe

The Morgan Library & Museum, New York, United States

Regine Stein

Bildarchiv Foto Marburg, Phillips Universität, Marburg, Germany

Ann Whiteside

Massachusetts Institute of Technology, Cambridge, United States

Meeting:

107. Cataloguing

WORLD LIBRARY AND INFORMATION CONGRESS: 75TH IFLA GENERAL CONFERENCE AND COUNCIL

23-27 August 2009, Milan, Italy

<http://www.ifla.org/annual-conference/ifla75/index.htm>

Abstract

A body of cataloguing practice has coalesced around Cataloguing Cultural Objects: a Guide to Describing Cultural Works and Their Images (CCO) since its publication in 2003. CCO is a manual for describing, documenting, and cataloguing cultural works and their visual surrogates. The primary focus of CCO is art and architecture, including but not limited to paintings, sculpture, prints, manuscripts, photographs, built works, installations, and other visual media. CCO also covers many other types of cultural works, including archaeological sites, artifacts, and functional objects from the realm of material culture.

This paper examines the influence of CCO and its implementation in cataloguing settings for the museum and image library community. In drawing together three diverse scenarios, the authors have identified common strategies for particular challenges in each one. The showcase projects

include: (1) the development of a *de facto* standard for contributing cultural and natural history collections to union catalogues and digital repositories by harmonizing the CDWA Lite and museumdat XML schemas; (2) the use of CCO in the Society of Architectural Historians Architecture Resources Archive (SAHARA) project, a shared online archive of photographs that document architecture and cultural landscapes worldwide. The SAHARA project developed a cataloguing template for use by scholars and librarians; and (3) the application of CCO alongside other guidelines in records for cultural works in library settings.

Emerging CCO cataloguing practice has resulted in a significant body of records from the museum and image library communities headed for LAM (library/archive/museum) integrated access environments. The authors comment on how cataloguing decision-making (e.g., differing concepts about a "work") may impact the convergence of records in these environments.

Introduction

Libraries have a long-standing tradition and mission-critical responsibility to shelter, document, protect, preserve, and ensure public access to knowledge and resources. The application of well-established and adopted standards is one way to guarantee that these activities are achievable and sustainable. Yet the library community recognizes that the documentation of cultural heritage collections, and the means for facilitating access to them, needs to be extensible or adaptable to reflect the growing demands presented by technology and the possibilities inherent in the networked environment. Cataloguing is an area that is affected by increased expectations from diverse audiences in how collections are discovered, and in how resources are used, and is adapting accordingly, notably with RDA: *Resource Description and Access*¹ as a replacement for AACR2. Significantly, RDA has not sought to reinvent a data content standard for bibliographic access, but rather builds on the foundation and success of AACR2.

In the last decade, there have been many other standards-based initiatives that have not only built upon existing standards, but have also crossed the library, archive, and museum divide in how collections are documented and discovered through the networked environment. One such emerging practice that was created to address the absence of a data content standard for describing, documenting, and cataloguing cultural works and their visual surrogates, and which thoughtfully recognized the limitations presented by AACR2 to fill this void, is *Cataloguing Cultural Objects: a Guide to Describing Cultural Works and Their Images* (CCO).²

The CCO initiative (a project of the Visual Resources Association Foundation, begun in 2001) provides guidelines for describing works of art and is based on the VRA Core 4.0³ and *Categories for the Description of Works of Art* (CDWA)⁴ element sets. Unlike those two schemas, however, CCO employs generic concepts that can be used with other metadata element sets (e.g., Dublin Core, MODS, MARC⁵). The cultural heritage community had never published guidelines (like AACR for the bibliographic community) that met the unique and often idiosyncratic descriptive requirements of one-of-a-kind cultural objects. The decisions that cataloguers make when describing cultural works are framed by the cataloguer's perception of how a work of art is defined. CCO is a data content standard intended to inform the decision-making processes of cataloguers and builders of cultural heritage systems. As community-specific metadata standards proliferated, there was a growing awareness that CCO could bridge disparate communities by prescribing common practice for describing cultural works.

Since its inception CCO has been a welcome addition to the corpus of cataloguing codes. Its impact has traversed geographical and organizational boundaries. For example, CCO has been recommended for use in large aggregated databases, such as the California Digital Library Shared Image Service⁶ and the Mellon ARTstor⁷ digital library hosting program. CCO is listed as one of the data content standards in NISO's *A Framework of Guidance for Building Good Digital Collections*.⁸ International projects have adopted CCO as well, for example, the *Electronic Catalogue of Bulgarian Cultural Historical Heritage*,⁹ the Centro de Documentacion de Bienes Patrimoniales (Chile),¹⁰ and the State Museums of Berlin/Institute for Museum Research¹¹. And finally, CCO influenced RDA as it was being developed with an awareness of standards for resource description from outside the library world.

Emerging CCO cataloguing practice has resulted in a significant body of records from these museum and image library communities headed for LAM (library/archive/museum) integrated access environments. While these new records will improve such catalogues because CCO practice facilitates sharable metadata, legacy records may provide some interesting dilemmas in the same system. For example: older “flat” records versus hierarchical records for complex works; dilemmas concerning “of-ness” and “about-ness”; and differences in the concept of a “work.” This last issue may be compounded by improper use of FRBR as well.¹²

A new research project, the “Museum Data Exchange”¹³ is using CCO to help analyze large bodies of data harvested from museum databases. The project is funded by the Andrew W. Mellon Foundation and operated by the RLG Programs of OCLC. Project director Günter Waibel (OCLC) observed,

“While it [the project] uses the same data structure (CDWA Lite XML), all participants are aware that rules to populate that data structure with data content may vary considerably from institution to institution. *Cataloguing Cultural Objects* is becoming a household name, but a good bit of the data shared probably predates the emergence of this data content standard, let alone its local implementation.”¹⁴

Let us now look at three diverse implementation settings that demonstrate how CCO can provide a common ground for cultural heritage cataloguing.

The Role of CCO in Harmonizing Cultural Metadata: CDWA Lite and museumdat

CCO is unique in that it is poised to address cataloguing cultural works and their visual surrogates independent of the data structures that manage these collections, and regardless of the community that houses these works. CCO recognizes that museums, libraries, archives, and image library collections all contain unique cultural works or representations of them, and have an increasing responsibility to create access to these works in the online environment. Another initiative that recently emerged to address an absence, specifically an appropriate technical solution for facilitating access to unique cultural works in the networked environment, is *CDWA Lite*.¹⁵

CDWA Lite is an XML schema for encoding core records for works of art and material culture. It relies upon existing standards to achieve its objective of a low-barrier way to enable

institutions to contribute their collections information to union catalogues. CDWA Lite is based on the data elements and guidelines in *Categories for the Description of Works of Art* (CDWA), a framework for documenting and organizing information on cultural works and images. The specification recommends using guidelines from CCO to assist with selecting, ordering, and formatting data used to populate its elements. The schema and guidelines encourage use of controlled vocabularies and authorities and the delivery and sharing of metadata records follows the *Open Archives Initiatives Protocol for Metadata Harvesting* (OAI/PMH).¹⁶

This standards-based initiative grew out of an identifiable need in the museum community for a more efficient and sustainable model to contribute collections to union catalogues and digital repositories. The team that developed CDWA Lite recognized the absence of a data structure for unique cultural works with a technical format for expressing that data in machine-readable format. Furthermore, the CDWA Lite team members realized that a solution was needed to eliminate the overhead that is commonly associated with contributing to union catalogues; ensure a method for being able to provide updated, accurate information about works accessible in the online environment; promote the idea that data integrity and accuracy should occur at the primary source or repository of the collection; and create a mechanism for bringing users back to a resource in its native environment, where learning more about a work of art can take place in the context of its larger collection.

The incorporation of existing data structure, content, value, and technical format standards is intentionally “lightweight,” to encourage and facilitate use even by small institutions in cataloguing, online publishing, and exposing metadata. The result is a packaged solution that makes using standards simplistic. Therefore the potential for the distribution of collections in the networked environment is all the more attainable.

CDWA Lite is made up of 22 elements, of which 19 are for descriptive metadata and 3 for administrative metadata; only 9 elements are required. The elements reflect the core descriptive documentation traditionally captured about works in cultural heritage collections.

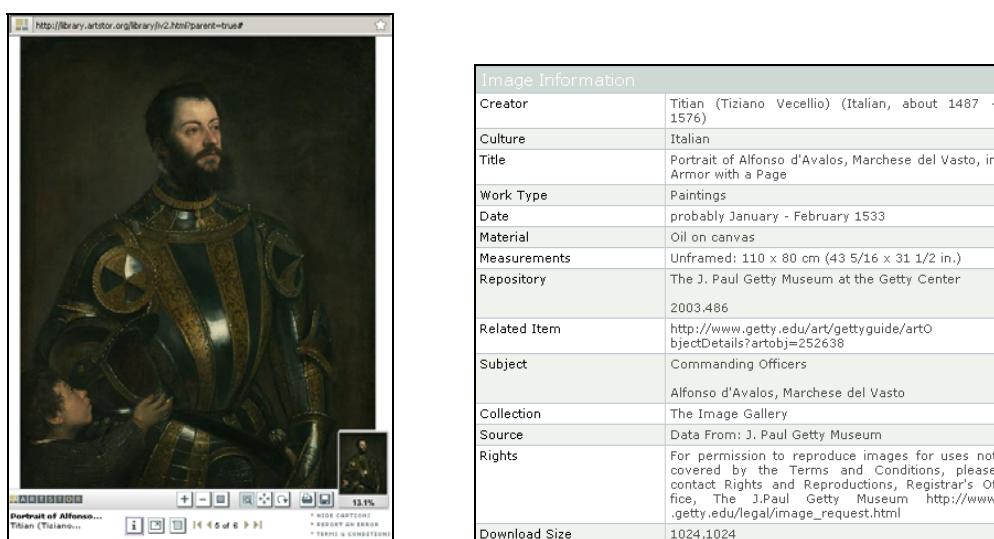
1. Object/Work Type	12. Display Creation Date
2. Title	13. Indexing Dates
3. Display Creator	14. Location / Repository
4. Indexing Creator	15. Indexing Subject
5. Display Measurements	16. Classification
6. Indexing Measurements	17. Description / Descriptive Note
7. Display Materials/Techniques	18. Inscriptions
8. Indexing Materials/Technique	19. Related Works
9. Display State/Edition	20. Rights for Work
10. Style	21. Record
11. Culture	22. Resources

CDWA Lite Element Set

A unique characteristic of CDWA Lite is that it creates a division between display and indexing elements, which is consistent with the recommendations of *Cataloguing Cultural Objects*. CCO recommends that certain display data be encapsulated for presentation with the end user in mind. This might involve concatenating values from various fields, or removing certain sensitive

information or administrative data for local purposes, in order to achieve meaningful descriptive information for the end user. Furthermore, display fields allow for the expression of uncertainty or ambiguity, which is common with art information. Indexing elements, on the other hand, provide values traditionally from controlled vocabularies or authorities, which ensures consistency and accuracy along with more effective retrieval. With indexing elements, CDWA Lite allows attributes to have a respective URI (termsource and termsource ID), which creates the opportunity for the identification of a term in the larger context of a controlled vocabulary.

An example of the application of CDWA Lite can be described in looking at a painting from the J. Paul Getty Museum's collection, which was harvested according to CDWA Lite by the digital library ARTstor, as illustrated below.



The screenshot shows a painting of Alfonso d'Avalos in armor, with a small figure in the foreground. To the right is a detailed table of metadata:

Image Information	
Creator	Titian (Tiziano Vecellio) (Italian, about 1487 - 1576)
Culture	Italian
Title	Portrait of Alfonso d'Avalos, Marchese del Vasto, in Armor with a Page
Work Type	Paintings
Date	probably January - February 1533
Material	Oil on canvas
Measurements	Unframed: 110 x 80 cm (43 5/16 x 31 1/2 in.)
Repository	The J. Paul Getty Museum at the Getty Center 2003.486
Related Item	http://www.getty.edu/art/gettyguide/artObjectDetails?artobj=252638
Subject	Commanding Officers Alfonso d'Avalos, Marchese del Vasto
Collection	The Image Gallery
Source	Data From: J. Paul Getty Museum
Rights	For permission to reproduce images for uses not covered by the Terms and Conditions, please contact Rights and Reproductions, Registrar's Office, The J. Paul Getty Museum http://www.getty.edu/legal/image_request.html
Download Size	1024,1024

This painting contains exhaustive descriptive information in the J. Paul Getty Museum's collection information management system. However, the CDWA Lite schema is not intended to re-create all the descriptive elements for a work, but rather to serve as a minimal set of information needed to facilitate access to a resource in a "union" environment. In this example, a limited amount of descriptive and administrative information about this painting by Titian was made available to the aggregator (ARTstor). Focusing on the creator information, ARTstor presents the CDWA Lite element, "Display Creator" in this way:

Titian (Tiziano Vecellio) (Italian, about 1487-1576)

The record that the Getty Museum contributed to ARTstor also contained indexing elements for the Creator, in addition to the display element. These indexing elements are encoded in the schema as follows:

```
<cdwalite:displayCreator> Titian (Tiziano Vecellio) (Italian, about 1487 -  
1576)</cdwalite:displayCreator>  
<cdwalite:indexingCreatorSet>  
<cdwalite:nameCreatorSet> <cdwalite:nameCreator type="personalName" termsource="ULAN"  
termsourceID="ulan500031075"> Titian </cdwalite:nameCreator>  
<cdwalite:nameCreatorSet>  
<cdwalite:nameCreatorSet> <cdwalite:nameCreator type="personalName" termsource="ULAN"  
termsourceID="ulan500031075"> Vecellio, Tiziano </cdwalite:nameCreator>  
<cdwalite:nameCreatorSet>  
<cdwalite:nationalityCreator>Italian</cdwalite:nationalityCreator>  
<cdwalite:vitalDatesCreator birthdate="1487" deathdate="1576"> about 1487 - 1576  
</cdwalite:vitalDatesCreator>  
<cdwalite:genderCreator> male </cdwalite:genderCreator>  
<cdwalite:roleCreator termsource="AAT" termsourceID="aat300025136">painter  
</cdwalite:roleCreator>  
</cdwalite:indexingCreatorSet>  
</cdwalite:indexingCreatorWrap>
```

Indexing elements contain information that facilitates search and retrieval, in addition to assisting aggregators with filtering and sorting search results. At the J. Paul Getty Museum, the creator information largely comes from an artist authority file, which is then mapped appropriately to CDWA Lite indexing and display elements.

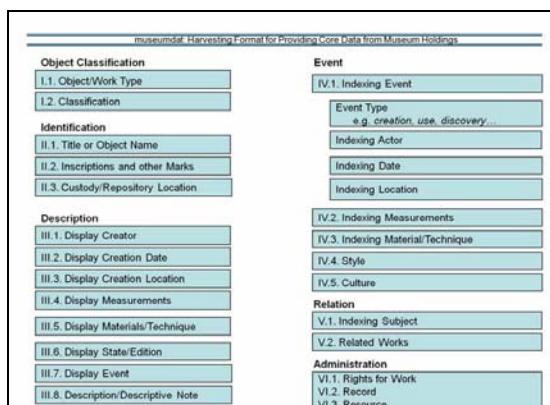
Furthermore, it is worth pointing out again that the specifications for CDWA Lite include guidelines for how best to populate elements, which are derived from *Cataloguing Cultural Objects*. For example, for the element “Display Creator,” CDWA Lite advises:

Formulated according to data content rules for creator display in CCO and CDWA; may be concatenated from the Indexing Creator elements, if necessary. The name should be in natural order, if possible, although inverted order is acceptable. Include nationality and life dates. For unknown creators, use one of the conventions illustrated in the following examples: "unknown," "unknown Chinese," "Chinese," or "unknown 15th-century Chinese."¹⁷

CDWA Lite has proven to be successful as a low-barrier way to contribute collection metadata to union resources. Its implementation of *Cataloguing Cultural Objects* guidelines has allowed for the concept of a “work” to be properly accommodated in its framework, and furthermore properly positioned for integrated access opportunities. As a result, material that addresses the description of unique objects with shared practices are beginning to converge in the online environment, and especially through resources that are aggregating from museums, libraries, archives, and the image library sectors (LAM).

In the relatively short amount of time that the CDWA Lite schema has been available, a great deal has happened with respect to its use, implementation, analysis, and widespread adoption. New software called OAICatMuseum, based on the Online Computer Library Center (OCLC) open source software OAICat, was developed to allow for CDWA Lite XML records to be harvestable according to the OAI PMH model, which requires the Dublin Core XML schema as the “lowest common denominator” for harvesting metadata records.¹⁸ Collection management vendors have begun to create mechanisms for exporting CDWA Lite records from their systems and to make them available for harvesting according to OAI PMH.¹⁹ Digital repositories and portals have begun to harvest CDWA Lite records, or to allow for records to be contributed using the CDWA Lite format.²⁰ And various communities across the world have begun to evaluate the relevancy and ease of use of CDWA Lite—for example the MuseFusion²¹ project in Taiwan, and the “Museum Data Exchange Project.” This latter initiative involved collaboration from RLG museum partners to create a suite of tools based on CDWA Lite that help facilitate its ease of use and implementation.

One of the most significant developments with CDWA Lite has come through the Documentation Committee of the German Museums Association in the creation of an XML schema called *museumdat*, which expands upon the CDWA Lite schema in order to be more inclusive of natural history and cultural history collections, and brings the elements in alignment with the event-oriented approach of the CIDOC Conceptual Reference Model (ISO 21127:2006).²² Museumdat provides a semantic framework and treatment for events surrounding an object by adding an additional element for events to CDWA Lite, bringing the total number of elements to 23, and then reconfiguring the elements to best represent the events-based approach. However, this schema also reduces the number of required elements to only three.²³



museumdat Element Set

Museumdat is structured into five primary categories, which is in accordance with the CIDOC CRM Core Metadata Element Set. It also brings administrative elements into a category, adds attributes that introduce multilinguality into the format, and provides a mechanism for data conversion control. Aside from these changes, museumdat very much maintains the focus and intent of CDWA Lite. For example, the specification document for museumdat keeps the guidelines from *Cataloguing Cultural Objects* for populating elements intact, where appropriate, and also keeps the possibility for both display and indexing elements.

The museumdat XML schema was published in 2007 and in a short amount of time has seen a level of adoption and enthusiasm similar to that of CDWA Lite. In addition to use within the German museum community, museumdat has become the standard delivery format for museum object data to the Bibliotheken Archive und Museen (BAM) portal, which allows for searching across library, archive and museum collections in Germany.²⁴ It is also being used by the regional museum network Digicult Schleswig-Holstein and the image archive Bildarchiv Foto Marburg, and is being considered as an accepted metadata format for Europeana, a European Commission funded project that allows for searching across cultural collections in Europe.²⁵

The CDWA Lite and museumdat XML schemas have been successful in providing a model that identifies a minimal set of information needed to facilitate resource discovery, and a solution that serves as a low-barrier way for institutions to participate in contributing collections to union catalogues and digital repositories. Both of these initiatives recognized that the growing expectation for universal discovery means there need to be ways for everyone to be able to participate in making cultural heritage available online. Existing standards can be successfully used to achieve this. CDWA Lite has led the way in this effort, but museumdat has broadened the scope to be more inclusive of diverse collections, and more flexible in allowing for describing concepts and relationships in cultural heritage documentation.

The creators of CDWA Lite quickly recognized the achievements of museumdat in combining the best aspects of both the event-oriented, multi-disciplinary approach of CIDOC CRM, and the relative simplicity and core elements offered by CDWA Lite. Purpose specific implementations of schemas can be useful and sustainable. However the appeal and potential of combining the efforts of CDWA Lite and museumdat into a common schema was determined to be an ideal opportunity. A CDWA Lite/museumdat Working Group has been established with key stakeholders from both initiatives, and tasked to create a new schema that builds off the foundation of CDWA Lite, and incorporates the enhancements of museumdat. The end result will be a schema that allows information from library, archive, museum and image library collections to be made available in a standardized format for contribution to the networked environment, and for facilitating resource discovery.

The harmonization of CDWA Lite and museumdat intends to create a de facto standard for contributing cultural and natural history collections to union catalogues and digital repositories. The support that these two initiatives have received from the community, whether from vendors, aggregators or collecting institutions, and the commitments from the CDWA Lite/museumdat Working Group to combine forces and create a common schema, is a testament to the role existing standards and new principles can have in providing a solution that crosses multiple sectors to achieve a shared goal: resource discovery for all.

Scholars and Cataloguing: the SAHARA Image Archive

In March 2008 the Society of Architectural Historians (SAH) received a grant from The Andrew W. Mellon Foundation to develop the SAH Architecture Resources Archive (SAHARA), a dynamic online library of images of architecture and landscape for research and teaching.²⁶ The need for such a resource was articulated in a Scholarly Communication Institute (SCI 4), also sponsored by The Andrew W. Mellon Foundation and held at the University of Virginia in Charlottesville, Virginia, summer 2006.²⁷ The Institute's goal was to provide an opportunity for leaders of SAH, architectural historians, librarians, publishers, technologists, and higher

education administrators to study, develop, and implement institutional and discipline based strategies to advance scholarly communication in the context of the ongoing digital revolution for the field of architectural history.

The Scholarly Communication Institute in Virginia resulted in two grants awarded to SAH. The first was to develop a prototype electronic version of the *Journal of the Society of Architectural Historians*. The second was for the SAHARA, with the expectation that scholars, librarians, and institutional leadership would join together to create a shared online resource that would both enrich the field of architectural history and create a new collaborative model for visual resources and art and architecture libraries. For the first time, instead of creating repetitive digital archives at each individual university, SAHARA will enable collaboration resulting in the creation of a highly authoritative resource with global coverage that supports new research and scholarly publications as well as enhancing university-level teaching.

The expectation is that SAHARA will change the way visual resources and art and architecture librarians at those institutions conduct their work. Instead of developing separate, independent collections of architectural images for each institution, scholars and librarians will contribute images and metadata to SAHARA, a shared resource that will be widely available. Redundancy in collection building will be minimized, which will lead to a reduction in redundant original cataloguing as well. This has been a goal of the visual resources community for a very long time, and SAHARA provides a model for testing this concept.

SAHARA is a partnership encompassing a wide range of individuals and institutions. Members of the SAH leadership are one set of partners—Executive Director Pauline Saliga, President Dietrich Neumann, and First Vice-President Dianne Harris. An initial set of library partners—Brown University (Providence, United States), Massachusetts Institute of Technology (Cambridge, United States), and the University of Virginia (Charlottesville, United States)—were identified based on knowledge of the schools' architectural history programs and the keen interest of the visual resources and architecture librarians to become involved. The technology partner is ARTstor, a major image repository with the technological infrastructure and organizational expertise in place, along with staff who can contribute to a project such as SAHARA.

SAHARA will be a peer-reviewed resource, much like a learned society journal. One of the goals is to help foster the idea that image content is as valuable a research tool as textual content, and that the particular viewpoints represented in scholars' images are directly related to their thinking about sites in the built environment. The SAHARA collection will be comprised of two overlapping groups of images: the "Members' Collection" and the "Editor's Choice Collection." Any SAH member is able to upload his/her images to SAHARA, assuming the images meet the baseline criteria for technical quality. Those images will reside in the "Members' Collection". Images chosen for the "Editor's Choice Collection" will be selected from the "Members' Collection" based on a higher technical quality as well as factors that include uniqueness of the content and the need for content in particular areas to develop the collection. This review and selection will be conducted by editorial teams composed of SAH scholars and librarian partners who will be "area editors" based on geography and time period (e.g., Renaissance Italian architecture).

During the first year, the partners concentrated on the development of an ingest tool that will allow individual scholars to personally upload images, create metadata about them, and finally, add the content to the SAHARA collections. This model of scholars working to help build collections and make them accessible is a new model that requires us to think differently about the roles of scholars and librarians in terms of collection building and cataloguing. Scholars are specialists in particular areas of their domain; we therefore assume that the expertise of the scholars who contribute to SAHARA will inform the cataloguing and accessibility of the images and will result in high-quality metadata. In addition, the area editors also bring scholarly expertise that will further enhance the quality and authoritativeness of SAHARA as a peer-reviewed collection.

Although CCO is the standard for cataloguing the built environment, our SAHARA scholar colleagues were not aware of the cataloguing standards that exist. There was an assumption that the standards (as developed by librarians) would not be domain-specific enough, and would not allow for the specificity of terms that architectural historians might use in search and retrieval (e.g., “cancello” a type of Early Christian screen used in Rome). We therefore had considerable discussion about appropriate metadata for architectural images amongst the SAHARA scholars and librarians. The initial schema concept was developed in the planning phase by a team of scholars and librarians who discussed the needs of scholars for finding image content related to the built environment, and the metadata that required to make that content discoverable. The goal for the metadata schema is a challenging one. We needed to strike a balance between encouraging scholars to become engaged in the process of developing quality digital resources and making the process of cataloguing images less burdensome. While the goal is to try to transform scholarly work habits, SAHARA cannot just turn scholars into full-scale cataloguers. The schema also needs to meet the requirements of cataloguers and computer systems staff, with clean data that can be used for effective search and retrieval and shared across resources. Finding that balance is not necessarily an easy task and we suspect that our first iteration is likely to undergo some changes after we test it.

The SAHARA project currently offers two “views” of the schema. One brief “view” for scholar input that does not necessarily include authority work as they go. The other “view” of the schema can be used by cataloguers or by scholars interested in providing full metadata records. The workflow model we are considering is one wherein librarians at participating institutions share the load of doing authority work for SAHARA and clean-up of data input by scholars. There will be controlled vocabulary lists for particular fields to assist people in their cataloguing. Other terms, such as names of creators or geographic place names, will be derived from the ARTstor Name Authority and Country files. Linking to available and existing authority files and controlled terminology will both aid in consistency of data input and make it possible to help train people who are not used to cataloguing.

The schema has undergone considerable revision, but has consistently maintained a strong relationship to the VRA Core 4.0, and embodies the principles found in *Cataloguing Cultural Objects*. ARTstor staff have contributed their metadata expertise and technical knowledge in helping us to refine the schema, ensuring that it will work within the ingest tool. The ARTstor technologists, with programmers at the three initial partner institutions, have also developed an XML schema to facilitate the movement of data from the local cataloguing tools to ARTstor. In

the future, the hope is that institutional contributors will also use the ingest tool in the full record mode, rather than exporting data from numerous cataloguing systems to ARTstor, a method that requires the data to be massaged before it can be published.

The SAHARA schema uses many of the elements found in CCO. They include Creator, Title of the building or site, Work Type, Date, Location, Style, Source, View Type, View Description, Materials and Techniques, Measurements, Description, and Subject. Within the Creator information, one can include the Creator's nationality, the extent of the creator's role in the building or site, and attribution information. With Title information, a contributor can provide alternate titles, as well as the names of complexes of which a building might be a part. Location information includes Street address, City, State, and Country. There is growing interest in the use of geo-referencing in the field of architecture, and thus SAHARA includes the option of including a point reference for the building or site, using longitude and latitude.

One of the most important aspects of creating a resource like this (i.e., distributed with multiple contributors) is the ability to take advantage of authority files and controlled terminology. Working with our colleagues at ARTstor, we have linked the Creator field to the ARTstor Name Authority File (derived from the Getty ULAN²⁸). The Title field, while not connected to an authority file, does use an auto-fill feature so that if a contributor enters the name of a building or site that already exists in SAHARA, the name will appear as a selection. This will facilitate a measure of consistency in data entry. As the content in SAHARA grows over time, the auto-fill feature will become more useful, as more titles of built works will be found in the collection.

Other fields require a controlled terminology. These fields include View Type, Narrow Classification (which in this case is used to describe Work Type), and Country (taken from ARTstor's country list). The terms for these fields were chosen by a collaborative group of SAHARA scholars and librarians using the Getty's thesauri. The development of the Work Type list, called Narrow Classification in SAHARA, was a long process. The number of terms used as Work Types in most image cataloguing databases results in a very long list. A pull-down pick-list of hundreds of terms is unwieldy and could be a disincentive for contributors. SAHARA created a list that is extremely short, and therefore far broader than originally conceived. The list, rather than serving to distinguish buildings or sites by specific functions, serves as a way to classify buildings and sites in broad categories. The concept of Work Type, in this instance, has been diluted, to meet the pragmatic needs of those who are not professional cataloguers.

Not specific to CCO itself, but critical to SAHARA, is the inclusion of administrative fields to delineate the photographer,

The screenshot shows the 'DESCRIBE THE IMAGE' interface with the following fields:

- Title/Name**: Fields for Title/Name of Work (e.g., Santa Croce, Pazzi Chapel) and View Type (e.g., Exterior, Facade).
- Creator**: Fields for Creator Name (e.g., Brunelleschi, Filippo) and Role (e.g., architect).
- Location**: Fields for City/Country (e.g., Florence, Italy) and Repository (e.g., J. Paul Getty Museum).
- Chronology**: Fields for Date (commission date: 1429; creation date: 1442-ca. 1465) and Chronology (earliest date: 1429, latest date: 1465).
- Physical Description**: Fields for Description and Commentary.
- Style**: Field for Style (e.g., Renaissance).
- Source**: Fields for Photographer (e.g., Friedman, David), Contributor (e.g., Kee, Jeffrey E.), Image Date (e.g., Winter 2007/2008), Image Earliest Date (e.g., 2007), and Image Latest Date (e.g., 2008).

the contributor of the images, the copyright, and usage rights of the contributed images. In order to try to find the balance between asking for some metadata, but not too much, we have delineated required fields. These fields are close to CCO recommended required fields and include: Title, View Type, Broad Classification (which aids in search and retrieval in large digital repositories), Location, Date, Style, and Source information relating to contributor, copyright, and usage rights.

In discussions about metadata for the SAHARA project, several scholars noted that there is a pressing, urgent need for a built work name authority file. For example, such an authority file would help disambiguate the various Palazzo Corner in Venice. Cultural heritage cataloguers have also voiced this wish for many years. In response, the Getty Vocabulary Program is developing a new thesaurus, the Cultural Objects Name Authority (CONA).²⁹ SAH is particularly interested in working collaboratively on the development of CONA. Again, the expertise of the scholarly community would be of great benefit to such an authority file.

As scholars began to share their images and catalog them in SAHARA, one of the first questions we received concerned the issue of “what am I cataloguing?” If one is cataloguing a building, the location is the geographical place in which the building sits. But, a scholar attempting to catalogue a map asked us, “How can one catalogue something like the Nolli map when the “location” field is still required?” She observed that clearer definitions for what goes into SAHARA and what belongs in ARTSTOR may be needed. Or, on the other hand, that that two repositories need to become more interrelated and fluid in terms of content and cataloguing. This scholar has hit upon an issue that cataloguers think about daily – how to bring together diverse cataloguing viewpoints into shared systems in ways that make sense to end users. We can address the basic question in our cataloguing guidelines, but the issue of how content converges in repositories has to be a collective response among cataloguers and repository providers.

SAHARA was launched on April 1, 2009, with a seed collection of approximately 9,500 images that contributors can add to using the ingest tool. As scholars contribute and use the metadata template, they are also providing feedback about our metadata schema, the ease of use of the template, and specific metadata fields. For example, scholars have told us that the View Type term list is too short. We have had the same feedback on the Narrow Classification list. These comments bring us back to our original discussions about how to create term lists that are short enough to pick from versus lists that are long enough to be valuable. In our deliberations about metadata fields, a decision was made that the Style field should be required. But, we have had many people comment that requiring Style is not helpful because not all buildings, sites, or landscapes can be pinned to a definitive style. Other comments reflect a lack of understanding about how fields are to be used, the values intended for specific fields, and why some kinds of information are needed for access to a collection that is envisioned to be in the hundreds of thousands over time.

The SAHARA project team has created a feedback log so that we can make informed decisions about suggested changes to the metadata schema and the use of fields resulting in agreed upon changes during the next year. We are also developing a set of cataloguing guidelines that will be shared broadly in the Society of Architectural Historians community, which we hope will both educate scholars and assist in the cataloguing within SAHARA. As part of our outreach efforts,

we also are engaging librarians who work with SAHARA scholar contributors to help them understand and use the metadata fields.

SAHARA is in the process of not only building a collection, but of educating scholars, to think critically about metadata as a practice and to select metadata that will provide the best access. As SAH members begin to contribute to the collection, and as architecture and visual resources librarians become involved with the cataloguing and editorial processes, the goal is to build a collaborative community focused on creating a new model of scholarship in architectural history. SAHARA is one possible model in which librarians can engage with scholars to define these new roles and CCO is providing vital guidance in this effort.

Applying CCO in a MARC/AACR world

The first reaction of many librarians to hearing that a library uses CCO in its online public access catalog (OPAC) is likely to be, “Why would you want to?” Isn’t CCO intended for use in visual resource collections and art museums, rather than in library collections, which have their own set of data standards? Why mix standards from two different worlds?

It is true that library collections consist chiefly of printed publications, and that the data standards³⁰ and information systems used by libraries were developed for, and work best when applied to, traditional library materials. But there are very few libraries that do not own at least a handful of art and cultural objects. These may include: portraits of founders or donors; artwork gifted for decorative purposes; and art or cultural objects that come to the library with someone’s papers, or because they have some association with existing textual collections. In most cases, there are too few objects to justify the cost of setting up a separate database. Moreover, the existence of a separate database complicates collection management activities such as inventory and circulation, makes it difficult to provide integrated access to the entire collection, and precludes contributing records for the objects to larger aggregations of library metadata such as OCLC’s WorldCat.³¹

The most effective way to establish internal control over objects and to provide access to them is to document them in the OPAC. The records need not be elaborate; accompanying documentation may be minimal, and library staff will probably lack the expertise to supplement or correct it with a description that would pass muster with an art historian. Even a brief description, linked if possible to an image of the object, will provide basic identifying information which, when disseminated through the OPAC, may elicit additional information from users. Librarians who possess more substantive documentation, or who can tap the expertise of art historians, conservators, dealers, or collectors can create more detailed descriptions. Whatever the length of the record, librarians will find CCO an invaluable guide for the selection and formulation of information appropriate for the description of art and cultural works.

What follows is a discussion of key issues encountered when attempting to apply CCO to the cataloguing of art and cultural works in the collections of The Morgan Library and Museum (New York, United States). The Morgan’s collections consist preponderantly of printed books and periodicals, manuscripts, and music, but also contain important collections of art and cultural works, such as drawings, prints, ancient Near Eastern cylinder seals, paintings, sculpture,

decorative objects, and cultural artifacts as diverse as Voltaire's briefcase, John Ruskin's lead soldiers, and a lock of John Keats' hair. The institution's decision to acquire the Voyager library system and to provide access to all materials through the system's OPAC coincided with the Visual Resources Association's decision to develop the CCO guidelines. Several Morgan librarians were able to participate in the development process as part of the editorial board and to apply the emerging standard to the description of their own collections.

Although it is possible to create "pure" CCO records in MARC, there is little advantage to doing so. Differences in stylistic conventions between the *Anglo-American Cataloguing Rules* (AACR) and CCO (for example, the latter does not use ISBD punctuation) can be jarring, and differences in the way data elements are parsed raise problems for display, indexing, and retrieval. Within a library context, CCO is best applied as a supplement to library cataloguing standards, to bring out characteristics of objects that are not covered by rules formulated for textual works and published items.

The object or work type is the single most important piece of information about an object; in the words of CCO, "The Work Type establishes the logical focus of the catalog record."³² In the world of library cataloguing, object type is considered "carrier" information, as distinct from content information; it characterizes the delivery medium for a particular manifestation of a work, rather than the essential nature of the work. The object type most commonly found in library collections is recorded nowhere in the bibliographic record, the assumption being that unless otherwise indicated, the item described is a book. The object type for non-books is recorded in various fixed and variable fields within the MARC record; none is entirely satisfactory for object cataloguing.

The MARC 300\$a (Extent) subfield, part of the physical description area, does not display in initial result sets and is unlikely to be indexed in many library systems, since it is used mainly to record pagination. The MARC 245\$h (Medium) subfield, which is used for recording the General Material Designation (GMD), is preferable for display and indexing purposes, but the only object-related terms defined by AACR for the GMD are *graphic*, *art original*, and *realia*, all of which are far too general to be useful. Morgan librarians chose to record the object type in the 245\$h subfield, but to substitute more appropriate terms. Specific object types such as *drawing*, *painting*, or *sculpture* are used for items with titles that describe what the work depicts, as in this stage design by the 19th century Italian artist Pelagio Palagi:

245 10 \$a Interior of a Vast Roman Fortress \$h [drawing]

For items lacking pictorial content, where the title conveys the object type, the more general "object" is used:

245 10 \$a Embroidered 18th-Century Italian Waistcoat Made for Count Gasparo Gozzi \$h [object].

(Note that the title includes the name of the person for whom the coat was made. Many cultural objects derive their meaning and value from their association with famous persons, rather than

from their innate value as artifacts. Conveying this relationship in the title makes the objects more accessible to users.)

Genre terms (MARC field 655) for work type (from the *Art and Architecture Thesaurus*³³), subdivided by culture and date, are used to enhance retrieval and to provide browsable lists:

655 _7 \$a Drawings \$x Italian \$y 18th century. \$2 aat
655 _7 \$a Drawings \$x Italian \$y 19th century. \$2 aat

Library cataloguing rules were designed for the description of published items. They assume that items come pre-packaged with a title page containing a formal description; information not appearing in the prescribed sources on the item is bracketed. Unpublished objects do not have title pages, and much of the description must be supplied, based on a variety of different sources, including the cataloguer's judgment. Since supplied information is the norm, brackets are not used.³⁴ Even more surprising to librarians, information appearing on the object itself, such as inscriptions and markings, is not privileged over other sources. CCO reflects art cataloguing practice in preferring a supplied title that fully describes the pictorial content or function of the object to a description appearing on the object, even one in the hand of the artist. At the same time, it recommends recording variant and former titles and carefully transcribing all inscriptions and markings. Here is how these recommendations are translated into a MARC record in our catalog:

100 1_ \$a Zuccari, Federico, \$d 1542 or 3-1609.
245 10 \$a Allegory of Sin \$h [drawing]
246 33 \$a Pianto, Peccato, Spavento \$h [drawing]
246 33 \$a Allegoria del Peccato \$h [drawing]
562 __ \$a Inscribed in black chalk, by the artist, "PIANTO / PECAT[O] / SPAVENTO"; at lower edge at center, in pen and brown ink, "Zuccaro"; on mount, in lower left corner, in pen and brown ink, "Zuccaro"; on verso of mount, in pen and brown ink, "Pa Auctionkost P. 1-9".

This gives users the best of both worlds: a meaningful title in the language of the catalog as well as access to everything written on the object, which may provide clues about attribution or provenance. Note that in this example, information relating to the artist is recorded in the MARC 562 (Copy and Version Identification Note) field, rather than in the 245\$c (Statement of Responsibility) field. Creator information written on a unique object by someone who may or may not be the artist does not carry the same weight as a formal statement on a printed title page.

The authors of CCO wisely chose not to reinvent the wheel by drafting rules for the formulation of name headings. Cataloguers are free to select their own sources for name authorities, so librarians need not worry about conflicts in the OPAC between the headings that provide access to their objects and those that provide access to secondary material. CCO does deviate from AACR in its view of the creative role of corporate bodies, families, and unknown artists. Unlike AACR, CCO considers corporate bodies such as the *Worcester Royal Porcelain Company* or *Tiffany and Co.* to have primary responsibility for the objects produced in their manufactures; families such as the Bibiena family of Italian artists and theatrical designers are regarded as the

primary creators of art and cultural works attributed to the family in the 18th century.³⁵ The Morgan's records for art and cultural works conform to CCO in treating corporate bodies and families as main entries for the objects they produce or create.

It is more difficult for librarians to accept the need for anonymous creator headings. Library cataloguers deal with a textual work of uncertain or unknown authorship by omitting the author field from the record and making the title the primary access point. But this situation is relatively uncommon for textual works; works of art that cannot be attributed to a known artist or even to an artist identified by a distinctive phrase such as the *Achilles Painter* are so prevalent that art cataloguers have developed a range of different ways to provide users with some context for the work's creation:

Attributed to Francesco Salviati.
Formerly attributed to Francesco Salviati.
Workshop of Francesco Salviati.
Follower of Francesco Salviati.
After Francesco Salviati.
Italian, 16th century.

Access points based on these attributions can be integrated into OPAC heading browses with surprising ease. In the Morgan's catalog, the AACR name form for Salviati, which is used to index both art works attributed with certainty to Salviati and printed secondary material reproducing his work, appears first in the name browse, followed by the non-AACR headings with qualifiers:

Salviati, Francesco, 1510-1563.
Salviati, Francesco, 1510-1563, after
Salviati, Francesco, 1510-1563, attributed to.
Salviati, Francesco, 1510-1563, formerly attributed to.
Salviati, Francesco, 1510-1563, workshop of.

Headings for attributions to cultures are also used (although the Morgan prefers "Anonymous" to the CCO-recommended "Unknown"):

Anonymous, Italian, 16th century.

Library cataloguing rules for physical description are limited in scope and in the amount of detail required; they focus chiefly on the extent of an item (i.e., the number of pages, leaves, volumes, fiches, reels, etc.), the presence of illustrations, and the measurements (in one dimension only, height, for books). Because each art work is a unique physical object, precise and detailed description is needed for purposes of identification. The task is complicated by the fact that the number of different object types is virtually unlimited; someone accustomed to describing graphics may be stymied by the challenges presented by a coin or a banjo clock. CCO's chapter on Physical Characteristics will provide a lifeline to librarians struggling with object description. The chapter, which is twice as long as any other data element chapter, offers guidance on recording information concerning the measurements, materials and techniques, editions and

states, inscriptions and markings, and nature of a wide variety of different object types. The section on measurement is particularly helpful to non-specialists, who might not think of including information on shape (for an oval miniature), weight (for a carved gem or a megalithic stone), or size (for an article of clothing).

Both MARC and AACR, especially the AACR-compliant codes developed for use in cataloguing special collections, make provision for more detailed physical description when desired; It is possible to fit CCO-style physical descriptions into OPAC records without too much difficulty. The MARC 340 (Physical Medium) field is defined for “physical description information for an item that requires technical equipment for its use or an item that has special conservation or storage needs”³⁶; it is more granular than the 300 (Physical Description) field, so that medium and support can be recorded in separate subfields. Here is an example of our use of the 340 field for a 15th-century Italian sculpture:

340 __ \$b Sculpture - height: 12 3/4 in. (315 mm), width: 11 3/8 in. (290 mm), depth: 6 1/2 in. (165 mm); Base - height: 3 1/2 in. (90 mm), width: 14 1/4 in. (363 mm) depth: 9 in. (227 mm) \$c Terra cotta with polychrome decoration.

Here is a much simpler description, for a 19th-century Venetian crystal locket:

340 __ \$b 7 x 4 cm \$c Crystal and silver.

In summary, a little information goes a long way in providing access to objects in library collections. Used as a supplement to library data standards, CCO provides librarians with the basic tools for creating records for art and cultural works that can live in harmony with the records for their mainstream collections. In another arena, the question of images in a MARC environment is going to loom larger as image collections (e.g., art photograph archives) are integrated into library collections and eventually, LAMS environments. Visual resources cataloguers routinely deal with issues arising from deciding whether to catalog a group of objects as a series or a set or as discrete objects and how to structure the object/image relationship. Librarians have experience in dealing with multi-work series, single issues or broken sets of periodicals, and archival collections. CCO could be a means of helping both communities deal with these complexities in the realm of objects and images.

Conclusion

Fortunately, the last few decades have seen intensive development of data standards for describing cultural works, resulting in a theoretical foundation encompassing a range of viewpoints. Driven by the rapid growth of technology and the educational mandates of cultural institutions to provide access to information about works of art, *Cataloguing Cultural Objects* provides a common framework in this effort. Today, a cross-section of museums, library special collections, and pictorial collections use CCO with a range of descriptive metadata element sets and specialized controlled vocabularies to catalogue and share information about cultural works.

¹ RDA: *Resource Description and Access*. American Library Association, Canadian Library Association, and the Chartered Institute of Library and Information Professionals <http://www.rdaonline.org/>

² Baca, M.; Harpring, P.; Lanzi, E.; McRae, L.; Whiteside, A., Eds. *Cataloging Cultural Objects: A Guide to Describing Cultural Works and Their Images*; American Library Association: Chicago, 2006.

³ VRA Core Categories 4.0, Visual Resources Association
<http://www.vraweb.org/resources/datastandards/intro.html>

⁴ Categories for the Description of Works of Art (CDWA). J. Paul Getty Trust, College Art Association
http://www.getty.edu/research/conducting_research/standards/cdwa/

⁵ Dublin Core Metadata Initiative <http://dublincore.org/> ; Metadata Object Description Schema (MODS). Library of Congress <http://www.loc.gov/standards/mods/>; MARC21 Formats. Library of Congress <http://www.loc.gov/marc/>

⁶ California Digital Libraries, Shared Metadata Working Group, *Metadata Submission Guidelines for UC Shared Images*. 26 March 2009 http://www.cdlib.org/inside/projects/image/msg_ucs1.pdf

⁷ ARTstor Images for Education and Scholarship <http://www.artstor.org/index.shtml>

⁸ NISO, *A Framework of Guidance for Building Good Digital Collections* 16 April 2008
<http://framework.niso.org/node/5>

⁹ Trendafilova, Malina. "Development Environment for Building Common Catalogue for Representation of the Culture-Historical Heritage of Bulgaria." *Bulgarian Academy of Sciences Cybernetics and Information Technologies*. Volume 7. no. 1 (2007), 95 – 105.
http://74.125.95.132/search?q=cache:16UMdXEgH2AJ:www.cit.iit.bas.bg/CIT_07/v7-1/95-105.pdf+Electronic+Catalogue+of+Bulgarian+Cultural+Historical+Heritage&cd=2&hl=en&ct=clnk&gl=us&client=firefox-a

¹⁰ The documentation system for the state museums of Chile, SUR™, is based on CDWA and the manual developed for museum professionals uses CCO as its model. Nagel, Lina, ed.: *Manual de registro y documentación de bienes culturales*. Santiago de Chile: Andros Impresores, 2008.

¹¹ State Museums of Berlin, Institute for Museum Research /Staatliche Museen zu Berlin, Institut für Museumsforschung <http://www.smb.museum/ifm/index.php?ls=12&topic=Home&lang=en&te=ja&tf=ja>

¹² "FRBR and Works of Art, Architecture, and Material Culture," by Murtha Baca and Sherman Clarke, in Arlene Taylor, *Understanding FRBR: What It Is and How It Will Affect Our Retrieval Tools* (Westport, Connecticut and London: Libraries Unlimited, 2007).

¹³ Museum Data Exchange Project. OCLC/RLG Programs.
<http://www.oclc.org/programs/ourwork/collectivecoll/sharecoll/museumdata.htm>

¹⁴ Waibel, Günter. "Museum Date Exchange: Asking the right questions " 20 March 2009, OCLC/RLG Programs, Hanging Together blog. <http://hangingtogether.org/?p=644>

¹⁵ CDWA Lite http://www.getty.edu/research/conducting_research/standards/cdwa/cdwalite.html

¹⁶ Open Archives Initiatives Protocol for Metadata Harvesting <http://www.openarchives.org/pmh/>

¹⁷ CDWA Lite Specification: A list of Elements, Tags, Description, and Examples, p. 4-5
http://www.getty.edu/research/conducting_research/standards/cdwa/cdwalite.pdf

¹⁸ OAICatMuseum 1.0 <http://www.oclc.org/research/software/oai/oaicatmuseum.htm>

¹⁹ COBOAT metadata publishing tool <http://www.oclc.org/research/software/coboot/default.htm>

²⁰ ARTstor Images for Education and Scholarship, Personal Collection Guidelines
http://help.artstor.org/wiki/index.php/Personal_Collections

²¹ Museum Computer Network, Taiwan Chapter. *MuseFusion* <http://mcntw-stds.teldap.tw:8080/musefusion/>

²² CIDOC Conceptual Reference Model. <http://cidoc.ics.forth.gr/>

²³ museumdat <http://www.museumdat.org/index.php?ln=en&t=home>

²⁴ BAM: Portal zu Bibliotheken Archiven und Museen (BAM) <http://www.bam-portal.de/>

²⁵ DigiCULT Museen Schleswig-Holstein <http://www.digicult-sh.de/>
Bildarchiv Foto Marburg <http://www.fotomarburg.de/>
EDL Foundation. Europeana <http://www.europeana.eu/portal/>

²⁶ Society of Architectural Historians Architecture Resources Archive (SAHARA)
<http://www.sah.org/index.php?src=gendocs&ref=HOME&category=Sahara%20HOME>

²⁷ University of Virginia Library, *Scholarly Communication Institute SCI 4 Architectural History 2006*,
<http://www.uvasci.org/archive/architectural-history-2006/>

²⁸ The Getty Research Institute. *Union List of Artist Names Online* (ULAN)
http://www.getty.edu/research/conducting_research/vocabularies/ulan/

²⁹ The Getty Research Institute, *Cultural Objects Name Authority* (CONA)
http://www.getty.edu/research/conducting_research/vocabularies/contribute.html#cona

³⁰ The Morgan Library and Museum, in common with almost all American libraries, uses AACR as a data content standard and MARC21 as a data format and communications standard. Institutions that use other library-specific data standards may encounter different problems when attempting to integrate object descriptions in their catalogues. Author O'Keefe would be very interested to hear about their experiences.

³¹ OCLC Online Computer Library Center, Inc. *WorldCat* <http://www.worldcat.org/>

³² *Cataloguing Cultural Objects*, p. 48

³³ The Getty Research Institute. *Art & Architecture Thesaurus Online*. (AAT)
http://www.getty.edu/research/conducting_research/vocabularies/aat/

³⁴ The fact that *RDA Resource Description and Access*, the cataloguing code that is intended to supersede AACR, dispenses with brackets for items that lack a formal title-page, may be an indication that the different metadata communities are beginning to learn from one another.

³⁵ *RDA*, reflecting the influence of archival cataloguing, will treat families as capable of authorship

³⁶ *MARC 21 Format for Bibliographic Data. 1999 Edition*. <http://www.loc.gov/marc/bibliographic/bd340.html>