Routledge
Taylor & Francis Group

Metadata for Special Collections in CONTENTdm: How to Improve Interoperability of Unique Fields Through OAI-PMH

MYUNG-JA HAN and CHRISTINE CHO

Content and Access Management, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA

TIMOTHY W. COLE

Mathematics Library, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA

AMY S. JACKSON

Digital Initiatives, University of New Mexico, Albuquerque, New Mexico, USA

Collection curators develop locally defined unique fields to support local requirements. As per the guidelines of the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH), Simple Dublin Core is the minimum requirement for exposing metadata to aggregators. Oftentimes the level of specificity of unique local fields is not translated well to Dublin Core, which may hinder the interoperability of the item metadata record. This paper researched 21 digital collections that were hosted in CONTENTdm. The objective was to explore the use and issues of unique fields in local context and recommend best practices that will increase the interoperability of metadata for special collections.

KEYWORDS   Dublin Core, metadata, mapping, CONTENTdm, OAI-PMH, interoperability, special collections, metadata standards

INTRODUCTION

Libraries wanting to digitize and provide online access to their special collections content must balance the tension between the requirements of local portals providing custom access to such content and the needs of aggregators wanting to provide integrated search and discovery services across special

Address correspondence to Myung-Ja Han, 1408 W. Gregory Dr. (MC-522), University of Illinois at Urbana-Champaign, Champaign, IL 61801, USA. E-mail: mhan3@illinois.edu

collection digital resources made available by multiple disparate institutions. (NB: There is no clear definition of the term "special collection" in library domain [Hewitt & Panitch, 2003]; we define special collection for this paper as materials that need special care and arrangement, or collections of materials that have been assembled for specific themes.) To deal with unique characteristics and context of special collection items, customized metadata schemes that include uniquely named descriptive attributes are often created for use in local implementations. However external aggregators need to collect and search against standardized or at least normalized descriptive metadata records. This paper examines the use of locally defined unique fields in item-level metadata descriptions of digital surrogates held in 21 collections from 11 CONTENTdm repositories and then examines the issues encountered when trying to map such locally customized metadata records to Simple Dublin Core or Qualified Dublin Core in ways intended to optimize interoperability. CONTENTdm implementations are especially well suited for this sort of examination because the application gives implementers license to add locally defined metadata fields and the freedom to then map or not map these local fields to Dublin Core however the implementer sees fit for purposes of metadata record export.

This paper stems in part from a finding described in the article *Dublin Core Metadata Harvested through OAI-PMH* (Jackson, Han, Groetsch, Mustafoff, & Cole, 2008). As described in this earlier paper, we found that "native metadata records are rich in meaning in their own environment, but lose richness in the aggregated environment due to the mapping errors and misunderstanding and misuse of Dublin Core elements" (Jackson et al., 2008, p. 18). This finding is closely related to a sharable metadata issue identified in a 2006 article by Shreeves, Riley, and Milewicz, that is, that "metadata may be of high quality within its local context, but may be compromised when taken out of this context for various reasons." Without context, useful local information may be lost, become insignificant, or become ambiguous and cause confusion to users in aggregators' environments. Since no single metadata standard works for every digital collection, it is inevitable for collection curators to develop and use locally defined unique fields for collections in their local environments. The challenge then is to support metadata aggregation and other forms of interoperability by maintaining context to the maximum degree possible, even while normalizing metadata records for sharing with others.

Shareable Metadata

Research on the importance of creating and providing interoperable metadata (i.e., sharable metadata) has evolved considerably in recent years, spurred on by the release and growing use of the Open Archives Initiatives Protocol for Metadata Harvesting (OAI-PMH),[1] the publication of the IMLS/NISO *A*

Framework of Guidance for Building Good Digital Collections (first edition 2001),[2] the online release of the *Cataloging Cultural Objects* guidelines (2006),[3] and disseminations from multiple recent and ongoing research projects in this domain (Bruce & Hillmann, 2004; Stvilia, Gasser, Twidale, Shreeves, & Cole, 2004; Shreeves et al., 2006; Elings & Weibel, 2007; Jackson et al., 2008). In addition, several projects built around metadata sharing have released best practices for creating interoperable metadata, for example, *Best Practices for OAI Data Provider Implementations and Shareable Metadata* (2005),[4] *DLF MODS Implementation Guidelines for Cultural Heritage Materials* (2006),[5] and Collaborative Digitization Program's *Dublin Core Metadata Best Practices* (2006).[6] This has encouraged collection curators to be more aware of common issues in creating interoperable metadata, such as consistency, encodings, and use of controlled vocabulary.

Additionally, in part because of the widespread adoption of OAI-PMH within the library community (Brogan, 2006), which mandates Simple Dublin Core as the minimum requirement when exposing metadata to aggregators, the quality of harvested Dublin Core metadata has garnered much attention in several venues (Lagoze, 2004; Tennant, 2004; Ward, 2004; Hutt & Riley, 2005; Shreeves, Knutson, Stvilia, Palmer, Twidale, & Cole, 2005; Jackson et al., 2008). The issue of improving metadata, while using Simple Dublin Core as a minimum requirement in the OAI-PMH environment, has created a host of problems because Simple Dublin Core is viewed as semantically limited and has a flat structure. In addition, problems entailing semantic ambiguity/vagueness and semantic overlaps in certain elements of Dublin Core, for example, <type> and <format>, <source>, and <relation> (Park, 2006; Jackson et al., 2008; Park & Childress, 2009), also contribute to metadata quality issues when records are aggregated in service provider environments. Yet researchers also have found that even in the context of Simple Dublin Core, aggregated metadata quality can be improved by creating, from the outset, metadata intended to support interoperability (Bruce & Hillmann, 2004; Shreeves et al., 2006) or by the consistent and correct mapping of custom local metadata to Dublin Core and/or other widely adopted standards (Park, 2005; Chan & Zeng, 2006; Zeng & Chan, 2006; Jackson et al., 2008). In this paper we will look at a range of metadata record mapping issues in the CONTENTdm environment.

Use of Unique Fields and Mapping Issues

While there has been good recent progress on understanding the character and importance of shareable metadata, comparatively less effort has been expended to date in examining the degree to which collection curators still locally develop and customize metadata standards for use in their local environments and how these customized metadata standards are mapped to standard metadata schemes, such as Dublin Core, for dissemination to

metadata aggregators. We hypothesize that locally defined unique fields, which potentially have substantial contextual information, could impede the interoperability of metadata since the contextual information in such fields as implemented could not be or was not being mapped to Dublin Core elements in ways to facilitate interoperability.

As an initial test for this hypothesis, we systematically examined current metadata and metadata mapping practices together. For this paper we looked at the locally defined unique fields created for a selection of digital special collections hosted in CONTENTdm to see how these fields were used and mapped to Dublin Core elements shareable through OAI-PMH. CONTENTdm is widely deployed by museums, archives and libraries to deliver local content digitized from special collections. While CONTENTdm derives its default metadata scheme directly from Dublin Core, the application allows implementers to add locally defined unique metadata fields on a collection-by-collection basis. The 21 collections examined, each with its own custom, "native" item-level metadata scheme, span a wide range of subjects and resource content types, and were created between January 1, 2001, and December 31, 2006, during which time understanding and sensitivity to metadata interoperability has evolved.

As described below, findings and results from this examination of current practice were informative. We identify and describe below several common shortcomings, which in turn suggest possible ways to improve the usefulness for interoperability of locally defined unique fields for special collection content.

CONTENTDM, DUBLIN CORE, AND INTEROPERABILITY

According to *SPEC Kit 298 Metadata*, published in 2007 by the Association of Research Libraries, CONTENTdm is one of the most used digital resource management tools in the library domain (Ma, 2007, p. 26). In the context of this discussion, there are two notable advantages of using CONTENTdm as a digital resource management tool.

First, metadata in CONTENTdm can be harvested through OAI-PMH. Being harvested by metadata aggregators is one way to expand user base and potentially increase use of a collection (Shreeves et al., 2006). In her 2006 report for the Digital Library Federation, *Contexts and Contributions: Building the Distributed Library*, Martha Brogan enumerates the now extensive use of OAI-PMH in the library community, noting, "As the digital content in repositories proliferates, efficient and consistent interoperability specifications are essential for effective downstream applications across a full spectrum of scholarly information arenas extending from e-research and e-learning to Web publishing and administrative computing," (Brogan, 2006, p. 2). Ultimately such applications benefit the end-user who gains knowledge

of and access to relevant information resources from a wider variety of sources. The inclusion of OAI-PMH functionality in CONTENTdm has been a boon to digital special collection curators and metadata aggregators alike.

Second, CONTENTdm allows collection curators to extend the CONTENTdm base metadata scheme (Dublin Core) with locally defined fields that can be used for more fully describing the resources in a collection—the field names do not need to conform to any standard. However, although CONTENTdm has an option that allows exporting metadata in its native format in the local environment, the raw, locally customized CONTENTdm metadata format cannot be disseminated through OAI-PMH, that is, metadata are harvestable from CONTENTdm in a Dublin Core format only, either Simple or Qualified.[7] (In our survey of 21 CONTENTdm collections examined for this paper, we found 14 collections use elements from Qualified Dublin Core, while 7 collections use elements from Simple Dublin Core only to make their collections available via OAI-PMH.)

However, note that the two advantageous features of CONTENTdm mentioned here are to a degree in tension. Implementers have a ready means of making metadata records available to external metadata aggregators as a way to enhance visibility of their digitized special collections content. They also are provided the means to customize metadata descriptions to local needs. The challenge is to do the latter in a manner that does not reduce the utility of metadata made available via OAI-PMH. Based on the observations discussed below, this is not consistently being done at present.

METHODOLOGY

Since 2001, the University of Illinois at Urbana-Champaign has been actively developing search and discovery services reliant on metadata harvested through OAI-PMH, for example, the Mellon Cultural Heritage Repository,[8] the Committee on Institutional Cooperation (CIC) Metadata Portal,[9] the Institute of Museum and Library Services Digital Collection and Content (IMLS DCC),[10] Illinois Harvest,[11] and most recently Opening History.[12] Among these, IMLS-DCC, Illinois Harvest, and Opening History are still ongoing. All three make extensive use of metadata records harvested from CONTENTdm.

Metadata Records Analyzed

For our analysis, we randomly selected 21 collections from the 11 CONTENTdm repositories being harvested by Illinois Harvest and/or IMLS DCC to analyze. From each collection, 5 to 10 records were randomly selected and harvested via OAI-PMH in XML Dublin Core format (151 records in total). (Fewer records were analyzed from collections if it could be determined that a collection had highly homogeneous record and field structure.) From each record, we extracted the Dublin Core <identifier> element containing

Image title	Horatio, Laertes
Performance title	Hamlet
Author/Composer	Shakespeare, William, 1564-1616
Theater	New Theatre (London, England)
Opening performance date	1934
Object	Costume rendering
Material/Techniques	Gouache; watercolor; pencil; Metallic (with two pencil studies)
Support	Paper
Dimensions	12 1/2" x 17 7/8"
Associated people	Gielgud, John (director);
Style or Period	German Renaissance c. 1500-1545
Production notes	http://images.library.uiuc.edu/projects/motley/prodnote/prodham.htm
Inscriptions	"Hamlet, Horatio (Laertes); H" upper right corner. "Motley 1934; MA" lower right corner. All in manuscript hand.
Description	Costume rendering for Horatio (Laertes) in Hamlet: black hat and shirt, gray doublet, silver-colored neckwear, black belt with silver-colored buckle, brown and black striped trousers with black tassels, black shoes; with two pencil studies.
Subject I (AAT)	Hat, black; Shirt, black; Doublet, gray (color); Trousers, brown, black, stripes, tassel; Shoe (footwear), black; Belt (costume accessory), black, medium gray; Neckwear, medium gray

FIGURE 1 Metadata displayed in local CONTENTdm environment.

the URL of the described resource in each local collection's CONTENTdm environment.

Using these URLs, we were then able to view the native metadata record as displayed in the collection's local CONTENTdm environment. This allowed us to compare descriptive content available in the local CONTENTdm environment with metadata disseminated by collection curators via OAI-PMH (see Figure 1 and Figure 2).

(For this analysis we examined only the publicly accessible native metadata, that is, the CONTENTdm record that collection curators decided to display to users. This meant we could not see any fields that might have been marked as "Hide" in CONTENTdm. We assume that fields hidden from users in the local environment are unlikely to be of utility to external metadata aggregators building search and discovery services.)

With samples of metadata from each collection's native environment, we were able to infer the local, native, CONTENTdm metadata scheme for each collection, and from these inferred schemes we were able to build a dataset of field names used for each collection. As part of our workflow for this analysis we separated the unique fields,[13] that is, locally defined unique fields, from Dublin Core elements, and organized data sets by collection in a Microsoft Excel spreadsheet. The unique fields were then subdivided by function, that is, descriptive, administrative, and technical metadata based on how the fields were used in the local environment. Unique fields were also examined to identify whether they might be from other standards, such as Categories for the description of Works of Art (CDWA), Visual Resources Association (VRA) Core, and Encoded Archival Description (EAD), all of which are metadata standards that are used for describing resources from museums and archives. For this process, we matched only on field names; we did not try to vet matches by inferring whether the semantic meanings of local fields identified matched meaning of the same-named field in the standard non-Dublin Core metadata scheme. As a last step before comparing native descriptions to OAI-PMH harvested records, we made an initial assessment

Title	Horatio, Laertes
Identifier	341114-003
	http://images.library.uiuc.edu:8081/u?/motley,1
Type	Image
Subject	Hat, black; Shirt, black; Doublet, gray (color); Trousers, brown, black, stripes, tassel; Shoe (footwear), black; Belt (costume accessory), black, medium gray; Neckwear, medium gray
Description	New Theatre (London, England)
	Paper
	German Renaissance; c. 1500-1545
	http://images.library.uiuc.edu/projects/motley/prodnote/prodham.htm
	"Hamlet, Horatio (Laertes); H," upper right corner. "Motley 1934; MA," lower right corner. All in manuscript hand.
	Costume rendering for Horatio (Laertes) in Hamlet: black hat and shirt, gray doublet, silver-colored neckwear, black belt with silver-colored buckle, brown and black striped trousers with black tassels, black shoes; with two pencil studies.
Extent	12 1/2" x 17 7/8"
Contributor	Gielgud, John (director)
Alternative	Hamlet
Contributor	Shakespeare, William, 1564-1616
Date	1934
Medium	Costume rendering
	Gouache; watercolor; pencil; Metallic (with two pencil studies)

FIGURE 2 Metadata displayed in aggregator's environment.

of likely mappings from a non-Dublin Core locally defined field to Dublin Core. We wanted in particular to see whether unique field names could be replaced with elements in Dublin Core and, if not, to determine the reason.

Finally we then compared the harvested OAI Dublin Core records with metadata gathered manually from the local CONTENTdm environment as a way to examine the as-implemented mapping practices for the unique fields, for example, how these fields were mapped to Dublin Core, and what fields were and were not mapped. The result revealed the possible impact of using locally defined unique field names on potential metadata interoperability.

Collections Represented in Sample

The sample of CONTENTdm collections analyzed in this paper was in many ways heterogeneous. As a way to evaluate the scope of collections analyzed, we characterized each collection according to item types included

in the collection and according to characteristics of the hosting institution. While this breakdown gives a qualitative sense of the universe of collections examined, we did not uncover any obvious dependencies between these collection/institution attributes and mapping quality or practice.

INSTITUTION TYPES

Twenty-one collections, selected from IMLSDCC and Illinois Harvest, were created by 11 institutions. Of those 11 institutions, 4 were large academic libraries; 1 was a government library; and there were 2 each of public libraries, museums, and independent research libraries.

ITEM TYPES CONTAINED IN EACH COLLECTION

We broke down collections by the item type(s) each contained. A collection can have more than one type of item, depending on what the physical collection is composed of and on local collection development policies, meaning some collections are represented more than once in the following breakdown. A majority of collections analyzed contain images (12)—there are four collections containing text, four containing still image items, three sound (audio) items, and two collections that have digital surrogates for three-dimensional physical objects (i.e., other than photographs or paintings). (See Figure 3). The institutions and collection analysis show that CONTENTdm is used in many different domains and for a variety of collections, with image type being the most popular.
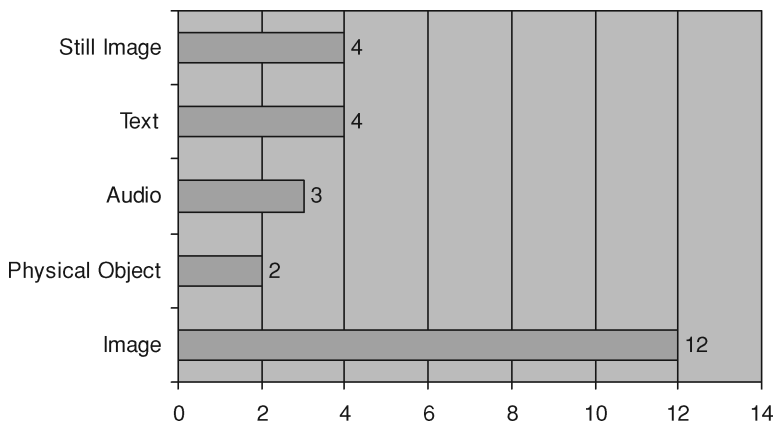

FIGURE 3 Collection by item type (n = 25).

FINDINGS

Collection curators use unique fields in their local environments to increase the use of their digital resources. The main issue is not about creating a universal metadata standard but, rather, about using unique fields in a more interoperable way.

Field Usage

The smallest number of fields used in a collection metadata scheme was 8 and the most was 28. The average number of fields used in a collection metadata scheme was 23.4. In total, and not accounting for the use of the same field by more than one collection, the 21 collections used 491 fields, including both unique fields and Dublin Core elements.

Only one collection examined relied exclusively on Dublin Core elements. All other collections used Simple and/or Qualified Dublin Core elements in combination with locally created fields or elements apparently taken from other metadata standards and reused in their local environments. For example, unique fields like <theatre> or <composition> were used alongside the Simple Dublin Core elements <subject> and <description>.

Unique Fields

Of the 491 fields used in the 21 collections examined, one-third of the fields, 171 (34.8%), are locally defined, unique fields that are not in Simple or Qualified Dublin Core. Of these 171 fields there were 135 unique field names (i.e., some non-Dublin Core field names were used by more than one collection, usually with seemingly equivalent or similar meaning). The number of unique fields used in each collection is quite different from collection to collection. While one collection did not use any unique field, another collection used 18 unique fields, and 6 collections used more than 10 unique fields. On average, 1 collection uses 8.1 unique fields. This is about one third of the total fields used in a collection. In other words, our inspection of CONTENTdm records showed that on average about one third of the descriptive properties associated with a digital resource were described using non-Dublin Core semantics.

From our inspection, we suspect that some collection curators employed specialized metadata standards to convey the specificity of the item. Metadata standards such as VRA Core 4.0, CDWA, ObjectID, and EAD are apparently being used frequently in local CONTENTdm metadata schemes, but since we did not talk directly to collection curators about this issue, this finding needs to be confirmed by subsequent investigations. When we matched unique field names with element names in these standards (again without confirming semantic meanings with collection curators), we found 42 out of

135 (33.1%) unique field names across all collections matched with elements from these standards. There were 6 from VRA Core 4.0 (e.g., <set>, <work>, and <collection>), 19 from CDWA/ObjectID (e.g., <material/technique> and <style or Period>), and 17 from EAD (e.g., <box number>, <shelf number>, and <folder number>). If correct, this verifies the notion that metadata standards are decided by characteristics of collection resources as well as domains (Elings & Weibel, 2007). Instead of creating local fields that are not conformant to any standards, using elements that are from broadly used standards in museum and archive domains may be an indicator that collection curators are aware of the importance of conforming to domain standards (presumably in part at least out of a desire to support interoperability).

Types of Unique Fields

The total of 171 unique fields alternatively can be broken down by the type of metadata the values in each field convey, though any such breakdown is of course subjective. In our judgment, 127 fields out of 171 (74.3%) non-Dublin Core elements were intended to convey descriptive metadata, 34 (19.9%) administrative, and 10 (5.8%) technical (see Figure 4). The breakdown shows that most unique, non-Dublin Core elements are created to better describe individual resources.

For the 127 unique fields used for descriptive metadata, we wanted to see if these field names could be replaced with suggested Dublin Core labels in their local environments. For this part of the research, we relied both on an inspection of field names used and on the values these fields contained. We did this before looking at how collection curators chose to map their non-Dublin Core elements to Dublin Core for purposes of dissemination through OAI-PMH. This helped us later in judging the consistency and correctness of mapping actually implemented by collection curators. Figure 5 shows how
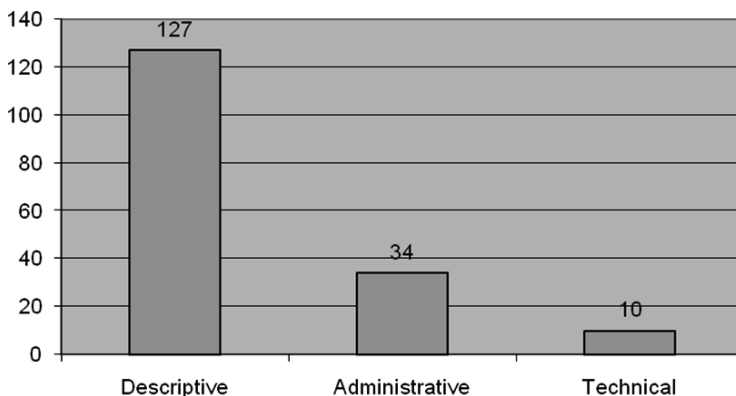

FIGURE 4 Types of unique fields (n = 171).

DC Element	No.	DC Element	No.
Contributor	2	Medium	1
Created	2	Publisher	5
Creator	6	Source	10
Date	13	Spatial	13
Description	12	Subject	6
Extent	9	Temporal	3
Format	6	Title	3
Identifier	3	Type	5
Is Part Of	8	Not Available	20

FIGURE 5 Unique descriptive metadata after normalization (n = 127).

many locally defined unique fields mapped in our best judgment to each Simple or Qualified Dublin Core element. As Figure 5 summarizes, we were able to determine with good confidence mappings for 107 out of 127 (84.3%) unique fields. The remaining 20 fields (15.7%) are so unique that we could not confidently map them to any Dublin Core element.

As shown in Figure 5 the element <spatial> can be used for 13 unique fields. Seven of the fields that can be replaced with <spatial> use the term "geographical" in their name instead of "spatial," as recommended by Dublin Core Metadata Initiative,[14] because the term "geographical" is more intuitive in meaning and familiar to users than the term "spatial." (Note that CONTENTdm uses metadata field names as labels for display purposes, thus encouraging collection curators to substitute more intuitive, locally contextual field names for generalized, often less intuitive Dublin Core element names.) Fields like <state>, <city>, and <country> could be also mapped to <spatial>. However, these more narrow and specific meanings are lost when performing this mapping.

<is Part Of> can replace eight unique fields that have "collection" in their names. A reason collection curators did not use <is Part Of> as a display label could be due to the fact that the element name itself was not as familiar to users as <collection>, <collection name>, or <collection title>.

<Format> and its refinement <extent> can also replace many unique fields (15 in all), for example, <dimensions>, <file size>, or <physical description>. Again, specificity is lost in such mappings. Also, we speculate that the field name <physical description> may have been popular because it comes from and is well defined in Anglo-American Cataloguing Rules Second Edition (AACR2). The information about <physical description> is added in data field 300 in Machine Readable Cataloging (MARC), and it contains more than one type of information, that is, it usually includes information of Dublin Core <extent> (subfields a and c), <type> (subfield a or e), and <description> (subfield b) in one field. When a MARC format record is converted for CONTENTdm collections, the structural differences between MARC and Dublin Core are easily overlooked even though interoperability is

affected. In addition <temporal>, <created>, and <medium> can replace six unique fields.

The elements <spatial>, <temporal>, <is Part Of>, <extent>, <medium>, and <created> are all refined elements of simple Dublin Core; <spatial> and <temporal> refine <coverage>, <is Part Of> refines <relation>, <extent> and <medium> refine <format>; and <created> refines <date>. Collectively these refined elements can be mapping targets for 36 unique fields (28.3%) used in local environments. Although these refined elements, that is, Qualified Dublin Core elements, were developed to express information in more narrow, specific meanings, the extensive use of non-Dublin Core elements to represent these concepts in native CONTENTdm environments suggests that even the Qualified Dublin Core properties are still not specific enough in terms of granularity (e.g., <spatial>) or intuitiveness (e.g., <is Part Of>). In the case of the Dublin Core element <spatial>, we found seven collections using <geographical coverage> instead of <spatial> as a display label. Note however, that the term "geographic" is also used in other metadata standards such as CDWA and VRA; so this observation may in some instances simply signal a preference on the collection curator's part for one set of semantics over another.

Dublin Core <date> can be used for 13 unique fields that include the date information for different manifestations, for example, <date original> and <date digital>. Seven out of 21 collections include <date> information for both physical and digital resources in their native metadata. Note also that the practice of describing both resources in the same record can be found in other elements, especially in <type> and <format>. This is in direct conflict with the Dublin Core one-to-one principle, which recommends that "Dublin Core metadata [describe] one manifestation or version of a resource, rather than assuming that manifestations stand in for one another" (Hillmann, 2005). However, previous research has shown that it is unavoidable to include information about both resources in practice (Shreeves et al., 2005). In order to provide information about two different manifestations, collection curators sometimes add qualifiers to Dublin Core elements, such as <date-original> and <date-digital>, which cannot be transferred to service providers.

Twelve locally defined unique fields can be mapped to <description>, including <production note>, <inscription>, and <comments>; however, some of the fields that can be mapped to <description> have unique, contextual information relating to the original resources, such as <English translation>, <poster text>, or <caption>, that is, collection curators would like to use more specific field names than a generic term <description>. In such instances there is a clear interest for purposes of the local portal in labeling these fields in a more distinct way than <description>.

Collection curators also use locally defined unique fields for the Dublin Core properties <creator>, <contributor>, and <identifier>. For <creator> and <contributor> fields, the locally created field usually describes the

role of creator or contributors in more specificity, such as <author>, <photographer>, <artist>, <interviewer>, <actor>, <producer>, or <participant>, rather than the more generic <creator> or <contributor>. As for <identifier> there are instances that collection curators used more specific terms for types of identifiers such as <image file name> or <local call number>. Since the definitions of Dublin Core elements are vague, it appears, quite reasonably, that collection curators want to use more specific names for their local users.

Another observation from our unique field analysis is that collection curators envision more use for their local metadata than just search and discovery. Clearly some collection curators try to capture administrative and technical metadata in addition to descriptive metadata in their local CONTENTdm environments. As mentioned above, of the 171 unique fields, 44 were not primarily descriptive—specifically 34 fields were used to record administrative metadata and 10 technical metadata. Administrative metadata provides information that manages a resource (such as location of physical resources and right statements) and digitized object (such as access and archival image files). Among 34 administrative metadata, 7 fields are for location of the physical resources, for example, <box number> and <shelf number>, and 4 fields are about rights statements, for example, <use> and <restrictions>. Twenty unique fields, such as <acquisition date> and <arrangement/identification note>, are used for resource management and three elements, such as <item order> and <order information>, are used for reproduction services. Out of 34 fields, 14 fields appear to have been taken directly from EAD.

Technical metadata is information that concerns the production process of the digitized item. Values in the 10 technical metadata fields range from the <color profile> to <digitization specification> and <digital production note>. Keeping administrative and technical metadata in addition to descriptive metadata is an indicator that collection curators are aware of the importance of preservation of digital resources. However, this begs the question of whether these metadata should be exposed to aggregators and users, and if so, then how this information could be transferred to Dublin Core since, by and large, Dublin Core lacks many elements useful for administrative purposes (Dublin Core <rights> being a noticeable exception).

The foregoing analysis shows that developing locally created metadata standards that include locally defined unique fields is a common practice in special collections since the items in collections require special needs in management and description. When it is possible, collection curators for the most part appear willing to borrow elements from established metadata standards including Simple and Qualified Dublin Core. However, it was also discovered that collection curators prefer using more granular, specific, and intuitive field names over using generic Dublin Core elements in local environments. In addition, collection curators need to create unique fields

not only for describing, but also for managing digital resources, something for which Dublin Core is not all that well suited.

Mapping

Mapping makes metadata in one standard available and useful in another standard. However, it should be noted that information loss is unavoidable in mapping because semantic meanings tend to be at least subtly different in each metadata standard as well as in each digital library (Attig, Copeland & Pelikan, 2004).

OAI PROVIDER

As has been mentioned, all the metadata created in CONTENTdm can be made harvestable by aggregators through OAI-PMH in Dublin Core format if collection curators check "Yes" for "Enable OAI" in their server setting. When collection curators build collections in their local CONTENTdm server, one of the first tasks is to create "Metadata Fields" for the collection. As part of this process, collection curators map each locally defined unique field to a corresponding Dublin Core element. Collection curators also have a choice not to map a specific field to any Dublin Core element if they want to keep those fields visible only in their local environments or for administrative purposes. This mapping process also decides what kind of metadata will be harvested through OAI-PMH, either in Simple Dublin Core or in Qualified Dublin Core. Our analysis showed that among 21 collections examined, 14 collections made their collection metadata available in Qualified Dublin Core (and Simple Dublin Core) and 7 collections in Simple Dublin Core only. When the choice to use Qualified Dublin Core was broken down by item type contained in each collection, collections containing at least some image items provided Qualified Dublin Core in 7 out of 12 instances. Collections that included still image items (4) and/or physical object items (2) mapped all of their collection metadata to Qualified Dublin Core (see Figure 6). The result is quite an encouraging trend that proves collection curators are aware of the fact that exposing metadata in Qualified Dublin Core format can improve the interoperability over providing metadata in Simple Dublin Core format.

MAPPING UNIQUE FIELDS

Among 171 unique fields, 128 (74.9%) fields were mapped by collection curators to Dublin Core and 43 (25.1%) fields were not mapped (see Figure 7). Of the 128 fields that are mapped, 93 fields (72.7%) were mapped

Type of collection	Qualified DC	Simple DC
Image	7	5
Physical Object	2	0
Audio	2	1
Text	3	1
Still Image	4	0
Total = 25	18 (72.0%)	7 (28.0%)

FIGURE 6 OAI metadata format by item type (n = 25).

consistently with our own analysis, and 35 fields (27.3%) were mapped at odds with our judgments.

SPECIFICITY AND GRANULARITY

Our initial analysis of mappings done by collection curators showed that the majority of unique fields were mapped to six Dublin Core elements. Among the total of 19 elements that are used for mapping by collection curators, <description> (23), <format> (15), <identifier> (14), <spatial> (13), <date> (13), and <creator> (9) are the 6 most heavily used (see Figure 8). The combined numbers of unique fields that are mapped to these 6 Dublin Core elements are 87 (68.0%) among 128 fields. The unique fields that were mapped to these 6 Dublin Core elements usually have more granular and specific terms in the local context than can be accommodated by Dublin Core elements. For this reason, although the mapping seemed right in a broad sense, such as <country> to <spatial> and <local call number> to <identifier>, information displayed in the aggregators' environments can seem confusing and ambiguous to users. Unless values in the more specific locally defined field are carefully crafted, mapping for a granular local field to a more generalized Dublin Core element cannot convey the same precise, unambiguous meaning as it does in locally customized applications (Shreeves et al., 2006). We recommend that fields useful only in the local context (e.g., local call number) are not exported to service providers.

Another example of local field names having more granular and specific meaning than Dublin Core elements is <description>. As shown in Figure 9, 23 unique fields were mapped to <description>. These unique

Type of field	Mapped	Not mapped		Mapped %
Administrative	20	14	34	58.8
Descriptive	102	25	127	80.3
Technical	6	4	10	60.0
Total	128	43	171	

FIGURE 7 Mapped fields by type.

Dublin Core mapping (n = 171)	Mapped by collection Curators (No.)	Mapping inconsistent with ours (No.)
Alternative	1	0
Contributor	1	0
Creator	9	2
Date	13	1
Description	23	11
Extent	6	1
Format	15	4
Identifier	14	7
Medium	1	0
Publisher	7	2
Relation	4	1
Rights	1	0
Source	7	2
Spatial	13	0
Subject	4	0
Temporal	4	1
Title	3	1
Type	2	0
Not Mapped	43	27
Total	171	60

FIGURE 8 Mappings of unique fields (n = 171).

fields lost context and specificity when mapped to the generic Dublin Core <description>.

Some unique field names mapped to <description> (for example, <comments>, <inscription(s)>, and <caption>—all commonly used terms for "Notes" as recognized by AACR2) are straightforward and seem unlikely to cause confusion. However, the information contained in other unique field names, such as <Location>, <Theatre>, <Theme>, and <Sub theme> can

Unique field name	Unique field name
Abstract date	Occupation/interviewee
Birth date/interviewee	Poster text
Caption	Production notes
Car Plan Number	Style or Period
Comments	Sub Theme
Condition	Support
Digitized Material	Technique
English translation	Theatre
Inscription(s)	Theme
Length of interview	Tracing Number
Location	Work dimensions-measurements
Notes	

FIGURE 9 Unique field names mapped to <description>.

Description	Conference Photographs
	F. W. Faxon Collections
	19 x 17 cm
	Photographic paper
	Documentary photography
	Picture #52:
	From Left: Walter Biscoe / A. H. Hopkins / J. I. Wyer, Jr. / J. C. M. Hanson / T. F. Currier
	Negative in UA 11,671
	Black and white photographs

FIGURE 10 A <description> field displayed in aggregator's environment.

be confusing to users in an aggregator's environment without any qualifier (see Figures 10 and 11).

Mappings that create undifferentiated repeated fields can also raise issues. In one case, a collection had five different unique fields: <Group Record>, <Subgroup Record>, <Series Record>, <Box Record>, and <Folder Record>. All of these individually named unique fields were mapped to <source>. (CONTENTdm does not currently provide a facility to combine fields as part of the mapping process.) In aggregators' sites, users are therefore confronted with five different numbers that are displayed in

Digitized Material	Black and white photographs
Year-Coverage	1906
Geographic-Coverage	Narragansett Pier
	Narragansett, Rhode Island
Names-Subject	Biscoe, Walter Stanley, 1853-1933
	Hanson, James Christian Meinich, 1864-1943
	Hopkins, Anderson H.
	Wyer, James Ingersoll, 1869-1955
	Currie, T. Franklin
Subject	Group Picture
Description	Picture #52:
	From Left: Walter Biscoe / A. H. Hopkins / J. I. Wyer, Jr. / J. C. M. Hanson / T. F. Currier p Negative in UA 11,671
Creator Corporate	F. W. Faxon
Creator Role	Photographers
Creator Contacts	83 Francis St., Backbay, Boston, Mass.
Publisher	
Theme	Conference Photographs
Sub Theme	F. W. Faxon Collections
Inclusive Dates	1894-1932
Language	English
WorkDimensions-Measurements	19 x 17 cm
Medium	Photographic paper
Technique	Documentary photography

FIGURE 11 Locally defined unique fields names that were mapped to <description>.

Source	99
	1
	14
	5
	10-19 [1898]

FIGURE 12 An aggregator's record without qualifiers.

<source> without any additional information regarding what each number means (see Figure 12).

ONE TO ONE PRINCIPLE

Mapping also can become problematic when one record includes information for both the physical and digital resources. In the native environment, the field names clearly identify what the information means, e.g., with prefixes such as <original> or <digital> added to descriptors. But when fields are mapped for export, confusion between manifestations of a resource can become an issue, since there are no corresponding elements in Dublin Core to represent different manifestations. In particular, <type>, <date>, and <format> are elements that may contain information about more than one manifestation. No element in Dublin Core can distinguish between different manifestations, for example, physical and digital, and usually conflicting fields are mapped to the same Dublin Core element, for example, <date>, <type>, and <format> (see Figure 13). Our analysis found seven collections that include <date> information about two different manifestations in their native records. In consequence, users in an aggregator's environment find conflicting descriptive values in one Dublin Core element name without any explanation.

<Native record>		<Aggregated record>	
Date.Original	unknown	**Date**	unknown
Date.Digital	2007-07-15		2007-07-15
Format.Original	photograph	**Format**	photograph
Format.Digital	jpeg		jpeg
Type.Original	Image	**Type**	Image
Type.Digital	Imag		Image

FIGURE 13 Two records with information for different manifestations.

QUESTIONABLE MAPPINGS

Among 128 mapped unique fields, 33 fields were mapped by collection curators in a way other than we anticipated based on our own examination of corresponding CONTENTdm descriptive records. A closer look at these 33 cases revealed two situations in which different mappings are likely: (1) when the locally defined unique field is so specific and/or poorly aligned to the Dublin Core abstract model that it is difficult to identify with confidence the closest corresponding Dublin Core element and (2) when there is a misunderstanding about the definition of a Dublin Core element.

As an illustration of the first situation, <description> is often used as a default Dublin Core element for mapping by collection curators when they have a difficult time finding any corresponding element but would like to make the information available to aggregators. Among 23 unique fields that are mapped to <description>, 11 were arguably mapped incorrectly, including <tracing number>, <style or period >, <theme>, <sub-theme>, and <work dimensions measurement>. These fields could have been mapped to other Dublin Core elements, for example, <work dimensions measurement> to <extent> and <style or period>, <theme> and <sub-theme> to <subject>, based on Getty's *Metadata Standards Crosswalk*,[15] and <tracing number> to <identifier>.

As an illustration of the second situation, we found that collection curators often misunderstand the definition of certain Dublin Core elements, notably <type>, <format>, <source>, <relation>, and <identifier>. The semantic overlaps in Dublin Core elements <source> and <relation> have been discussed elsewhere as has been the semantic ambiguity between <type> and <format> (Park, 2006; Jackson et al., 2008; Park & Childress, 2009). In addition to these elements, we found <identifier> was used haphazardly as a container for any field names in the local schema that include words such as <number> or <url>. While this is sometimes intuitively sensible, among 12 unique fields that were mapped to <identifier>, 5 should not have been mapped to any Dublin Core field or should have been mapped to an element other than <identifier>. This included a local, idiosyncratic <catalog number> value that should not have been mapped to any Dublin Core field for export, URLs for related content (e.g., parent collection) that should not have been mapped to <identifier> in the item-level record, mappings of <record group> to <identifier> (instead of <source>), and mappings of <repository> and <place kept> to <identifier> (rather than <publisher>, <relation>, or unmapped).

NOT-MAPPED FIELDS

Information is also lost when the collection curators simply do not map locally defined unique fields to Dublin Core. CONTENTdm provides the option to not map local fields to aggregators by mapping to <None>. It

Type of metadata	No.	Could be mapped
Administrative	14	3
Descriptive	25	23
Technical	4	1
Total	43	27

FIGURE 14 Types of Not-mapped fields.

is up to the collection curators whether the information should be made available for aggregators or kept in local environments. When mapping, collection curators can choose the CONTENTdm "Hide" option as well. So, collection curators have four different options for their metadata, that is, (1) make available to users in both local and aggregators environments, (2) hide to local users but make available to aggregators, (3) display in local environment but not to aggregators, and (4) keep only for administrators.

Of the 21 collections studied, 10 have not mapped publicly accessible fields in the local CONTENTdm display. Collectively across these 10 collections, 43 fields are not mapped. Of the 43 not-mapped fields, 25 are descriptive, 14 are administrative, and 4 are technical metadata (see Figure 14).

Out of 43 fields, our analysis suggested 27 could have been mapped to one of the Simple or Qualified Dublin Core elements (see Figure 15). For instance, <file size> could have been mapped to <extent> and <participant> to <contributor>. We also noticed that collection curators did not map the <collection> field to the corresponding Dublin Core <is Part Of> element, which can provide information regarding the collection as a whole. Seven unique fields have the term "Collection" in their names, such as <collection>, <digital collection>, <collection ID>, and <collection name>, and all of them are not mapped. Among not-mapped fields, we found nine fields that could be mapped to <source> (2) and <is Part Of> (7).

A complete list of locally defined unique field names that are not mapped (at least by some collection curators) to Dublin Core is provided in Figure 16. However, it should be noted that our analysis could not take into account collection curators' intentions. For instance, a collection curator

Correct mapping	No.	Correct mapping	No. (n = 27)
Contributor	1	Is Part Of	7
Date	1	Publisher	1
Date Created	1	Rights	1
Description	4	Source	2
Extent	2	Spatial	1
Identifier	1	Subject	2
Format	1	Type	2

FIGURE 15 Mapping not-mapped fields.

Field name	No.	Field name	No.
Accession number	1	General notes	1
Arrangement/identification note	1	Holding institution	1
Call number	1	Image file name	1
Campaign	1	Item number	1
Catalog number	1	Item order	1
Cite as	1	Location—collection	1
Collection	3	Location—folder	1
Collection name	1	Notes	2
Collections ID	1	Order number	1
Compound Doc	1	Ordering information	1
Contact information	1	Participants	1
Date (Original ISO)	1	People/Places	1
Date Digital	1	Production site	1
Date Original	1	PSMHS category	1
Digital collection	1	Resolution-measurements	1
Digital dimensions-measurements	1	Shelf number	1
Digitization specifications	1	State	1
File size-measurements	1	Type Digital	1
Format Original	1	Type Original	1
Full resolution	1	Use	1

FIGURE 16 Not-mapped unique fields (n = 43).

may reasonably have decided not to map locally created fields describing certain manifestations, <date digital>, <type digital>, or <format original>, in order to prevent confusion for users of the aggregators' sites. Not mapping the administrative metadata, such as fields related to reproduction services or contact information, seemed a right practice in terms of providing interoperable metadata.

CONCLUSION AND RECOMMENDATIONS

Based on the sample of CONTENTdm collections examined, collection curators make frequent use of locally defined unique field names in their local environment to describe their resources. In addition, these additional fields used provided descriptive information about the resources that was more specific and nuanced than could be achieved using standard Dublin Core elements. However, consistent with findings from Jackson et al. (2008) and Shreeves et al. (2006), we found that the frequent use of unique fields that make native metadata rich in context can impede interoperability if the field names are too unique and specific. This is because the contextual information associated with fields in the local environment, including more fulsome labeling, will be lost when the metadata is harvested by aggregators in Dublin Core format. Our data analysis as detailed here suggests six ways that metadata creators and collection curators can improve metadata interoperability.

First, keep a balance between specificity and generality in defining local fields. Although collection curators wish to describe their resources as specifically as possible, if the fields are too specific, some context will surely be lost in mapping to support harvesting. Our data analysis suggests that when the fields are too specific, collection curators choose either to not map the fields to Dublin Core at all or they choose to map to Dublin Core <description> as the default catchall. In both cases information is lost.

Second, decide at the outset which locally defined unique fields are intended only for the local environment and which should be made available to aggregators. Collection curators develop unique field names for descriptive, administrative, and technical purposes. Since not all the unique fields are used for resource discovery, collection curators can reasonably decide to keep some for use only in the local environment. Such decisions can be facilitated by examination of "metadata use cases," as opposed to basing such decisions on "there is no immediately obvious corresponding element in Dublin Core."

Third, be cognizant of how values will be created in the local environment. Don't segment a human-readable value into labeled component parts in your local system (e.g., don't segment the value string, "Record Series 4, Box 5, Folder 1" into three separate values "4," "5" and "1") unless absolutely necessary to support specific local search and discovery functions, since by doing so you render the values of much less use in an aggregated environment. Similarly, if you will describe both physical and digital manifestation properties in your local system using unique field names, consider whether you intend to follow the Dublin Core one-to-one principle, in which case only metadata about one manifestation will be mapped and made available to aggregators. Alternatively, if you wish to provide metadata about both manifestations in a single OAI-PMH Dublin Core record then consider strategies that you can deploy so that value strings, rather than just local field-naming conventions, make clear which value goes with which manifestation.

Fourth, maximize use of Qualified Dublin Core elements for labeling in the local environment. This obviates the need for subsequent mapping and avoids mapping errors, since semantics of Qualified Dublin Core are more refined and clear than Simple Dublin Core. We found that two thirds of the collections we examined provided their metadata in Qualified Dublin Core. However, collection curators seem surprisingly wary about using Dublin Core elements as a display label. Presumably this is because they would like to provide more intuitive and specific labeling for their local users. This is a reasonable and sometimes valid concern, but also consider using conventions in generating your value strings that would allow you to stay with Dublin Core field-naming conventions. The decision to forgo standardized metadata semantics should be based on the uniqueness and characteristics of digital resources as well as practices in that specific community.

Fifth, consider taking field names and definitions, if possible, directly from other metadata standards such as EAD, VRA Core, and CDWA when creating locally developed application profiles. This could improve interoperability, since the crosswalks between many established metadata standards and Dublin Core are reasonably well vetted and readily available. When collection curators do make a good effort to adapt and use fields from other metadata standards, the potential for interoperability is enhanced.

Sixth, share the logic of mapping decisions with aggregators. While we are in good agreement with a majority of mapping decisions made over our sample of collections, there is a sizable fraction of mapping decisions examined with which we disagree. Quite possibly the number of disagreements would be reduced if we had more information and context about how these decisions were reached. Unfortunately, there is as yet no clear communication channel where collection curators and aggregators can share mapping information. The commonly suggested way to share metadata information is to provide an application profile that includes data, content, and encoding standards in a project page, but the community may want to look at more robust ways to open lines of communication between provider and aggregator when it comes to sharing information about mappings used. Especially when the fields are from other than already established metadata standards, providing more information about mappings used will help aggregators understand and use the harvested metadata better.

For special collections, creating and developing local metadata standards is not a new practice and is still perceived as a necessary procedure because there is no single metadata standard that works for describing unique digital resources in all contexts. However, collection curators using CONTENTdm need to be more cognizant that this application offers the intrinsic facility for making their metadata available to aggregators via OAI-PMH. While a large and active user community has grown up around CONTENTdm, the importance of sharing metadata through CONTENTdm, and especially the issues to do with metadata interoperability, has not gotten the attention it warrants. To date, CONTENTdm User Groups have tended to focus more on local implementation issues and local metadata scheme design, rather than on the use of CONTENTdm as a way to push information out to the broader community beyond the local institution.

As described above there are inherent difficulties in mapping local metadata fields to standard schema like Dublin Core; however, these difficulties are not entirely intractable. Recently some success has been had with educational initiatives such as the *Metadata For You and Me*[16] collaborative work done at the University of Illinois and Indiana University (Shreeves and Riley, 2008). In addition, recent CONTENTdm User Group meetings have begun to feature presentations and discussions on these issues (Han & Yarasavage, 2009).

The usefulness of OAI-PMH as a method of metadata dissemination is strongly influenced by how well implementers consider interoperability when defining and populating locally defined unique metadata fields and, then, how well they map their local metadata scheme to Simple and Qualified Dublin Core. The challenge for CONTENTdm collection developers creating metadata schemes for their collections is to find the right balance between local design imperatives and interoperability.

NOTES

1. OAI-PMH: http://www.openarchives.org/pmh/
2. Third edition published in 2007 is available at http://www.niso.org/publications/rp/framework3.pdf
3. Cataloging Cultural Objects Guidelines: http://www.vrafoundation.org/ccoweb/index.htm
4. Best Practices for OAI Data Provider Implementations and Shareable Metadata: http://webservices.itcs.umich.edu/mediawiki/oaibp/?PublicTOC
5. DLF MODS Implementation Guidelines for Cultural Heritage Metadata: http://www.diglib.org/aquifer/DLF_MODS_ImpGuidelines_ver4.pdf
6. Dublin Core Metadata Best Practices: http://www.bcr.org/cdp/best/dublin-core-bp.pdf
7. Qualified Dublin Core was introduced in 2000 by the Dublin Core Usage Committee. CONTENTdm adapted Qualified Dublin Core as its mapping option since its first commercial release of version 3.1 in May 2001 and OAI-PMH function in 2003.
8. Mellon Cultural Heritage Repository: http://oai.grainger.uiuc.edu/
9. CIC Portal: http://cicharvest.grainger.uiuc.edu/
10. IMLS DCC: http://imlsdcc.grainger.uiuc.edu/
11. Illinois Harvest: http://illinoisharvest.grainger.uiuc.edu/
12. Open History: http://imlsdcc.grainger.uiuc.edu/history
13. For this study, we treated a field that is not from Dublin Core as a unique field.
14. http://dublincore.org/documents/dces/
15. Metadata Standards Crosswalk: http://www.getty.edu/research/conducting_research/standards/intrometadata/crosswalks.html
16. Metadata for You and Me: http://images.library.uiuc.edu/projects/mym/

REFERENCES

Attig, J., Copeland, A., & Pelikan, M. (2004). Context and meaning: The challenges of metadata for a digital image library within the university. *College & Research Libraries*, *65*(3), 251–261.

Brogan, M. (2006). Contexts and contributions: Building the distributed library. Retrieved February 3, 2009 from http://works.bepress.com/martha_brogan/6

Bruce, T. R., & Hillmann, D. I. (2004). The continuum of metadata quality: Defining, expressing, exploiting. In D. I. Hillmann & E. L. Westbrooks (Eds.) *Metadata in Practice* (238–256). Chicago: American Library Association.

Chan, L. M., & Zeng, M. L. (2006). Metadata interoperability and standardization—a study of methodology part I. *D-Lib Magazine*, *12*(6). Retrieved February 3, 2009 from http://www.dlib.org/dlib/june06/chan/06chan.html

Elings, M. W., & Weibel, G. (2007). Metadata for all: Descriptive standards and metadata sharing across libraries, archives and museums. *First Monday*, *12*(3). Retrieved from http://firstmonday.org/issues/issue12_3/elings/index.html

Han, M. J., & Yarasavage, N. (2009, July 13). *Creating a video collection using CONTENTdm* [PowerPoint slides]. Paper presented at the CONTENTdm User Group Meeting, American Library Association Annual Conference. Chicago, IL. Retrieved from http://hdl.handle.net/2142/13623

Hillmann, D. I. (2005). Using Dublin Core. Retrieved January 12, 2009 from http://dublincore.org/documents/usageguide/

Hewitt, J. A., & Panitch, J. M. (2003). The ARL special collections initiative. *Library Trends*, *52*(1), 157–171.

Hutt, A., & Riley, J. (2005). Semantics and syntax of Dublin Core usage in Open Archives Initiative data providers of cultural heritage materials. *Fifth ACM/IEEE-CS Joint Conference on Digital Libraries 2005* (pp. 262–270). New York: ACM Press.

Jackson, A. S., Han, M. J., Groetsch, K., Mustafoff, M., & Cole, T. W. (2008). Dublin Core metadata harvested through OAI-PMH. *Journal of Library Metadata*, *8*(1), 5–21.

Lagoze, C. (2004). The relationship between OAI-PMH and Dublin Core: Required, recommended or other? *Proceedings CERN Workshop on Innovations in Scholarly Communication: Implementing the Benefits of OAI* (OAI3), CERN (Geneva, Switzerland), February 12–14, 2004.

Ma, J. (2007) *Metadata, Spec Kit 298*. Washington, DC: Association of Research Libraries.

Park, J. R. (2005). Semantic interoperability across digital image collections: A pilot study on metadata mapping. In L. Vaughan (Ed.), *CAIS/ACSI 2005 Data, Information, and Knowledge in a Networked World, Proceedings of the 2005 annual conference of the Canadian Association for Information Science*, University of Western Ontario, London, Ontario, June 2–4, 2005. Retrieved January 12, 2009 from http://www.caisacsi.ca/proceedings/2005/park_J_2005.pdf

Park, J. R. (2006). Semantic interoperability and metadata quality: An analysis of metadata item records of digital image collections. *Knowledge Organization*, *33*(1), 20–34.

Park, J. R., & Childress, E. (2009). Dublin Core metadata semantics: An analysis of the perspectives of information professionals. *Journal of Information Science*. Retrieved May 29, 2009 from http://jis.sagepub.com/cgi/rapidpdf/0165551509337871v1–

Shreeves, S. L., Knutson, E. M., Stvilia, B., Palmer, C. L., Twidale, M. B., & Cole, T. W. (2005). Is "quality" metadata "shareable" metadata? The implications of local metadata practice on federated collections. In H. A. Thompson (Ed.) *Proceedings of the Twelfth National Conference of the Association of College and Research Libraries*, Minneapolis, MN, April 7–10 2005 (pp. 223–237). Chicago: Association of College and Research Libraries.

Shreeves, S. L., Riley, J., & Milewicz, L. (2006). Moving towards shareable metadata? *First Monday*, *11*(8). Retrieved January 30, 2009 from http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/1386/1304

Shreeves, S. L. & Riley, J. (2008) *Metadata for you and me: Current and emerging trends in metadata and content sharing* [PowerPoint slides]. Paper presented at the Digital Library Federation Fall Forum 2008, November 13, 2008, Providence, RI. Retrieved from http://hdl.handle.net/2142/9158

Stvilia, B., Gasser, L., Twidale, M., Shreeves, S. L., & Cole, T. W. (2004). Metadata quality for federated collections. *Proceedings of ICIQ04—9th International Conference on Information Quality* (pp. 111–125). Cambridge, MA: Massachusetts Institute of Technology.

Tennant, R. (2004). Bitter harvest: Problems & suggested solutions for OAI-PMH data and service providers. Retrieved February 28, 2009 from http://www.cdlib.org/inside/projects/harvesting/bitter_harvest.html

Ward, J. (2004). Unqualified Dublin Core usage in OAI-PMH data providers. *OCLC Systems and Services: International Digital Library Perspectives*, *20*(1), 40–47.

Zeng, M. L., & Chan, L. M. (2006). Metadata interoperability and standardization—a study of methodology part II. *D-Lib Magazine*, *12*(6). Retrieved February 3, 2009 from http://www.dlib.org/dlib/june06/chan/06chan.html