Survey of Benchmarks in Metadata Quality: Initial Findings

White Paper

Prepared by the Metadata Quality Benchmarks Sub-Group

Digital Library Federation | Assessment Interest Group | Metadata Working Group

Steven Gentry   Project Archivist | University of Michigan
Meredith L. Hale   Metadata Librarian | University of Tennessee, Knoxville
Andrea Payant   Metadata Librarian | Utah State University[1]
Hannah Tarver   Department Head, Digital Projects Unit | University of North Texas
Rachel White   Head of Metadata and Cataloging Services | Colgate University
Rachel Wittmann   Digital Curation Librarian | University of Utah[2]

May 2020

---

[1] Co-facilitator.
[2] Co-facilitator.

Acknowledgements

Table of Contents

List of Figures and Tables

Introduction

The following report details the response data gathered from the Survey of Benchmarks in Metadata Quality, which was deployed under the Digital Library Federation's Assessment Interest Group, Metadata Working Group. The published survey questions[1] and anonymized response data[2] are included along with an overview of results and synthesis of the group's analysis.

Background

The Digital Library Federation's Assessment Interest Group (DLF AIG) was created in 2014 to address and solidify digital library standards, tools, and practices.[3] To address this, the DLF AIG developed numerous working groups, including the Metadata Working Group in 2016. This working group strives to "collaboratively build guidelines, best practices, tools, and workflows around the evaluation and assessment of metadata used by and for digital libraries and repositories."[4]

The Metadata Quality Benchmarks sub-group was formed within the DLF AIG Metadata Working Group to investigate how to formulate general guidelines for measuring the quality of metadata. As a first step, the Metadata Quality Benchmarks sub-group wanted to gather more information about methods already used by the digital library community for measuring characteristics of metadata quality. The group drafted and distributed a survey to determine:

1. What metadata requirements and standards are commonly implemented in libraries, archives, museums, and other cultural heritage organizations;
2. The methods and criteria used to evaluate metadata quality in these institutions; and
3. Gaps in knowledge and practice related to metadata quality.

Methods

The sub-group drafted a series of questions related to metadata quality assessment and created a Qualtrics survey (hosted at Utah State University) to be publicly circulated within the digital library community. After Institutional Review Boards at the University of Utah and Utah State University deemed the study non-human research and exempt from IRB oversight,[5] the sub-group promoted the survey through various domain listservs relevant to the metadata profession.[6]

---

[1] See Appendix A for the full list of questions.
[2] See Appendix B for a link to the complete anonymized dataset of responses.
[3] DLF AIG wiki page: https://wiki.diglib.org/Assessment
[4] DLF AIG Metadata Working Group's "About" page: http://dlfmetadataassessment.github.io/About
[5] University of Utah IRB# 00121527, Utah State University IRB# 10273.
[6] See Appendix C for the full distribution list.

The survey opened on May 23, 2019 and responses were collected until July 10, 2019. To minimize duplicate responses, survey instructions asked that only one metadata expert provide responses for each organization, including both individual institutions and entities functioning as aggregators. After the survey closed, members of the sub-group reviewed anonymized data to determine trends in responses. Free-text responses were reviewed and coded for recurring themes in order to categorize and analyze response data.

Survey Completion Rate

Respondents were required to answer only two questions in the survey:

1. Confirmation of their consent to take part in the survey.
2. How many repositories they wished to include in the repeatable portion of the survey (i.e., the number of times they wanted that section to repeat).[7]

Although 240 respondents consented to take the survey, only 107 respondents (45%) fully completed the survey. Another 44 respondents (18%) partially completed the survey, yielding a total or partial completion rate of 63%. The remaining 89 respondents (37%) consented but did not answer any questions. The repeatable portion, geared toward individual repositories, was initiated 142 times.

Results

Data analysis in this paper mirrors the structure of the survey, which comprised four sections:

- **Section 1** (questions 2-8) -- Information about respondents' institutions and professional experience.

- **Section 2** (questions 9-18) -- Metadata practices and technology within repositories.

- **Section 3** (questions 19-21) -- Metadata elements included (i.e., required, recommended, or optional) and elements that are evaluated within repositories.

- **Section 4** (questions 22-26) -- Implemented and aspirational metadata evaluation practices at respondents' institutions.

When reading this report, please note that data for responses in sections 2 and 3 represent repositories, while responses in sections 1 and 4 are institution-wide (i.e., may apply to multiple repositories); textual references may switch between "respondents" or "responses" and "repositories" depending on the appropriate data units for that section. Metadata element names are italicized (e.g., *title*, *creator*, *identifie*r) in order to differentiate these names within the text. In some cases, we reference specific sections or questions in footnotes using the notation §# for the section number, and append .# for a question number.

---

[7] §2.11 through §3.21

Section 1: Respondent Profile

This section asked general questions about respondents' organizations in order to provide context for the rest of the survey. These questions relate to the sizes and types of respondents' organizations, number and experience of staff members doing metadata work, and the kinds of work that respondents are doing with metadata.

Respondents either chose from a suggested list or entered additional institutional types to describe their place of employment. Of the 152 answers to this question, most respondents (120) identified a library setting:

- Academic library or department ----- 84 respondents (55%)[8]
- Special libraries or organizations --- 22 respondents (14%)[9]
- Public libraries ------------------------- 14 respondents (9%)
- Archives (e.g., within a library) ------ 16 respondents (11%)[10]
- Museums ---------------------------------- 6 respondents (4%)
- Consortia ---------------------------------- 3 respondents (2%)
- Aggregation projects -------------------- 2 respondents (1%)

The bulk of the write-in responses were grouped with pre-set categories, but there were also 2 "other" responses without further specification.

Perhaps reflective of the high number of respondents working in academic libraries, nearly half of the respondents noted that their institutions employed at least 101 employees. Other responses were somewhat evenly split:

- 1-10 employees ------- 28 organizations (19%)
- 11-50 employees ------ 30 organizations (21%)
- 51-100 employees ---- 21 organizations (14%)
- 101+ employees ------- 67 organizations (46%)

With regard to full-time employees (FTEs) who work with non-MARC metadata (see Figure 1), 81% had fewer than 6 FTEs working in this area. The most common response was between 1-2 employees (36%), followed by 3-5 employees (26%). The "less than 1 FTE" category accounted for 13% of responses, which included 5 respondents who said there are 0 FTEs working with non-MARC metadata at their organizations, and 12 who have employees working on metadata part-time (totalling less than 1 FTE).

---

[8] Includes 2 respondents who self-identified as working within an academic college.
[9] Includes 8 respondents who self-identified as corporate, state, and school libraries.
[10] Includes 3 respondents who self-identified as university library departments and a historical society.

Figure 1. Breakdown of Non-MARC Metadata Staffing
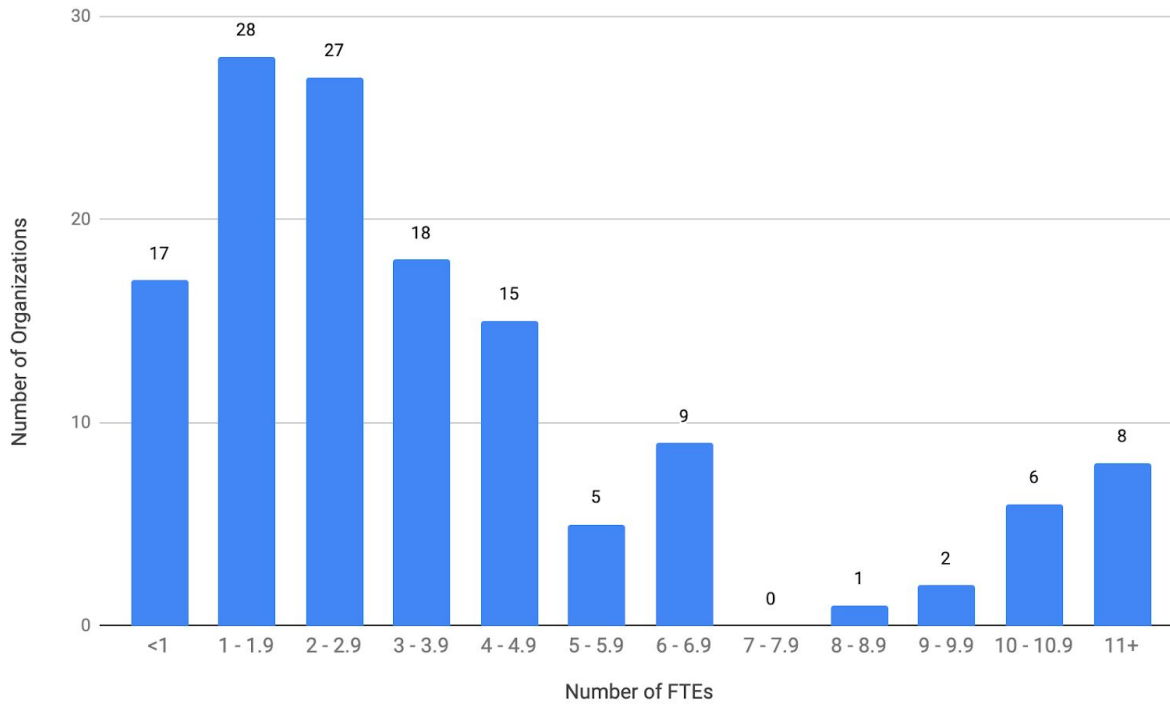
Although the number of employees who work with non-MARC metadata represents a small portion of their staff, respondents also noted a high amount of professional expertise--65% indicated they had been working with such metadata for more than 5 years, with 59% of this number working with non-MARC metadata for more than 10 years (see Table 1). Additionally, 76% respondents reported working with MARC metadata for some period of time, with the majority of respondents indicating that they have worked with MARC for 10 or more years.

Table 1. Respondents' Years of Experience Working with non-MARC and MARC Metadata

# Years	Experience with non-MARC metadata		Experience with MARC metadata	
	# Respondents	% Respondents	# Respondents	% Respondents
Never	7	5%	35	24%
0-4 years	44	30%	28	19%
5-9 years	39	27%	20	14%
10+ years	55	38%	62	43%

Respondents were asked to provide more information about the types of tasks they perform with metadata at their workplaces and to choose multiple options as relevant (see Table 2), resulting in 606 selections (or write-in answers) from 145 respondents. Most respondents reported managing existing metadata, including migrating, remediation, and enhancement; followed by setting guidelines and best practices. Responses from the "other" category included creating metadata or performing authority work; training, consulting, or administrative work; advising and developing best practices; and metadata outreach and assessment.

Table 2. Respondents' Metadata Responsibilities

Metadata Task	Number of Responses	Percentage of Respondents (n = 145)
Creating descriptive metadata	118	81%
Setting guidelines and best practices	128	88%
Supervising metadata creators	92	63%
Quality control checks	119	82%
Managing existing metadata (e.g., migrating, remediation, and enhancement)	132	91%
Other (write-in answers)[11]	16	11%
Management, training, advising	7	5%
Technical management	8	6%
Authority control	2	1%
Access	2	1%

---

[11] Highlighted rows are tasks mentioned multiple times in the 16 free-response answers.

Section 2: Metadata Basics

This section asked foundational questions about repository implementations, including various infrastructures or Digital Asset Management Systems (DAMS), Metadata Application Profiles (MAPs) and schema alignments, and use of controlled vocabularies. These questions are meant to establish general points of similarity across digital libraries, before moving on to more specific aspects of metadata elements and quality assessment.

The majority of respondents (67%) indicated their organizations manage 1-2 repositories, while the other third of the 134 respondents for the question[12] provided a wide range of responses (see Figure 2). The second most common category was 3-4 repositories (21%) and some organizations reported managing up to 20 repositories, though the number dropped off sharply after more than 5 repositories.

Figure 2. Number of Repositories Per Responding Organization
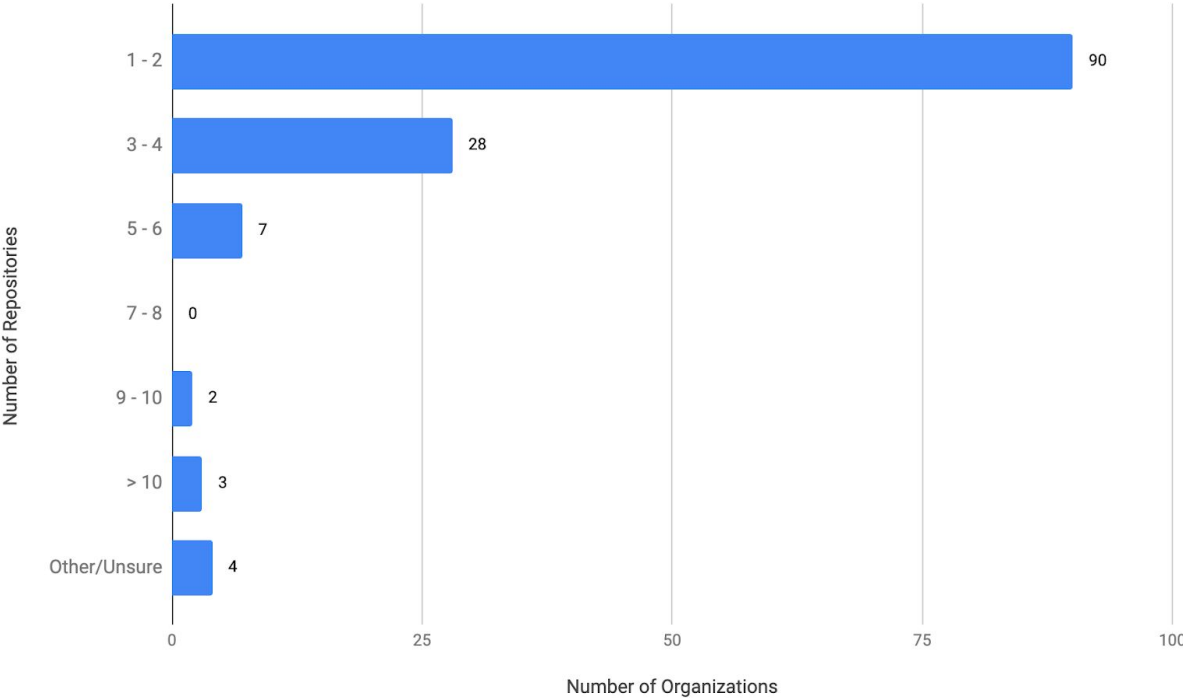
Portions of section 2--which addressed metadata practices--were repeatable, as respondents completed this section for each repository at their organization, for a total of 142 individual digital repositories; note that these results may not necessarily reflect the overall practice at an organization.

---

[12] §2.9

For each of the individual repositories in the repeatable section, respondents noted whether the content represented digital collections (primarily digitized or born-digital cultural heritage materials), institutional repositories (primarily resources produced by the organization and/or constituent members, such as scholarly works), or both kinds of content:

- Digital collections ----------- 75 repositories (53%)
- Institutional repositories --- 22 repositories (16%)
- Both -------------------------- 45 repositories (32%)

The types of repository infrastructures or Digital Asset Management Systems (DAMS) in use were widespread, including a relatively high number of homegrown systems (10%) and a large number (around 20%) of systems or combinations of systems in use by a single repository. In most cases, respondents selected a single type of DAMS for the repository, but there were some responses that marked multiple systems for a total of 147 selections (or write-in answers) across 141 responses. The most common systems were CONTENTdm, Islandora, and Dspace; however, the most-used system--CONTENTdm--reflects only 18% of total responses (see Figure 3).

Figure 3. Repositories' DAMS Usage
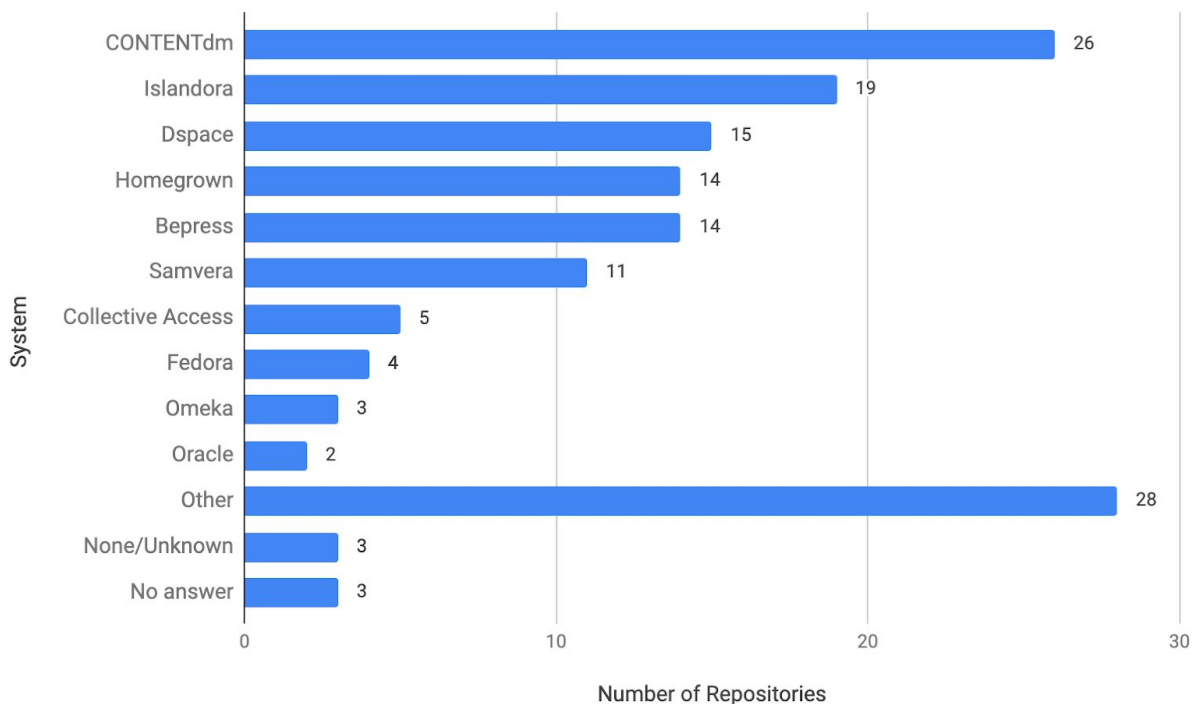
To get a general sense of the level of conformity expected within the repositories, the survey asked whether each of the repositories used a Metadata Application Profile (MAP) and also whether the MAP was locally written or if the repository aligns with consortial MAPs (e.g., DPLA, Mountain West Digital Library, etc.). Overall, MAP usage was a slightly higher-than-even split (see Figure 4), with 80 repositories (57%) implementing MAPs.

Figure 4. Implementation of MAPS by Repository
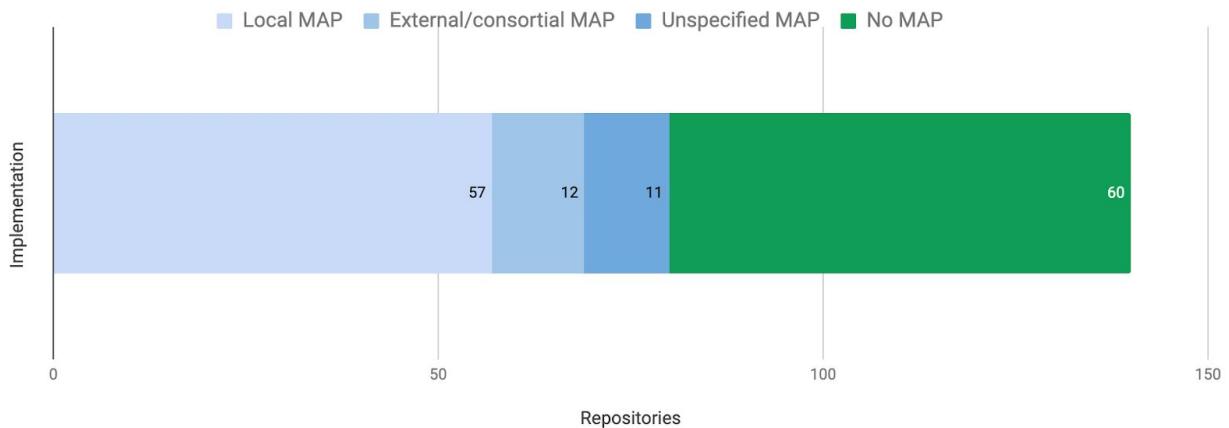
Respondents also noted the specific metadata schemas and controlled vocabularies used in each repository by choosing from a list of common options or writing in other values in a free-text field.[13] There were 114 responses regarding schemas (see Figure 5), including more than 51% reflecting Dublin Core-based schemas, and one-quarter using MODS. Beyond those frequently-used schemas, there were significantly fewer shared schema types. At least 5 repositories (4%) use a local schema[14] and 3 repositories (3%) are not using a schema.

---

[13] Write-in answers were folded into existing categories or broken into additional categories when there seemed to be overlap.

[14] Some numbers are less certain due to ambiguity in free-text responses.

Figure 5. Schema Usage by Repository
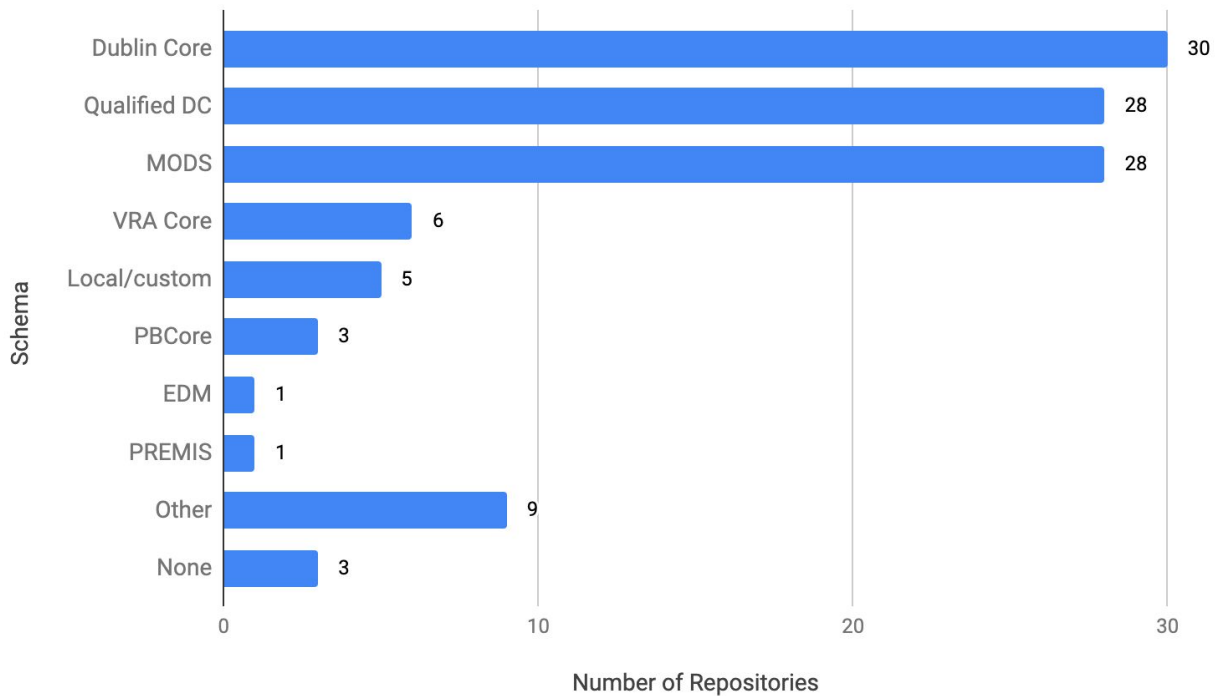
Comparing the reported schemas with the repositories being used, the results show that while institutions may choose to use the default schema for a particular platform (e.g., MODS for Islandora), there is some flexibility in schema decisions. It is also possible to use no schema at all. One respondent answered this question by stating, "Samvera is schema-less" and three others specifically selected "None" as their answer.

Since respondents could choose multiple options to identify the controlled vocabularies used in each repository, there were 107 individual responses with 352 total selections for this question, including 33 free-text answers (see Table 3). Clear front-runners included Library of Congress vocabularies--such as Library of Congress Subject Headings and Library of Congress Genre/Forms--and those curated by the Getty Museum, including the Art & Architecture Thesaurus. Several vocabularies were not included as an option in the survey question but were written in multiple times; some vocabularies (e.g., the Library of Congress Name Authority File) may seem to have fewer-than-expected responses since they were only counted from write-in responses.

Table 3: Breakdown of Responses Regarding Controlled Vocabulary Usage

Controlled Vocabulary	Number of Repositories	Percentage of Responses (n = 107)
FAST Subjects	16	15%
GeoNames.org	19	18%
Getty vocabulary databases		
Art & Architecture Thesaurus	65	61%
Thesaurus of Geographic Names	30	28%
Union List of Artist Names	17	16%
Library of Congress vocabularies		
Genre/Forms (LCGFT)	47	44%
Subject Headings (LCSH)	85	79%
Thesaurus for Graphic Materials (TGM)	35	33%
Medical Subject Headings (MeSH)	6	6%
Other	*33*	*31%*
LC Name Authority File	7	7%
Virtual Name Authority File	2	2%
Local/Custom	7	7%
PBCore	2	2%
DCMI Type	4	4%
None	4	4%

At least 7 repositories use some form of custom or locally-created vocabulary and only 4 repositories use no controlled vocabularies. However, based on the free-text answers, some respondents interpreted this question to mean "only subject vocabularies" (e.g., "...there are no subject headings") while others interpreted it more inclusively as "any controlled vocabulary" by writing in ISO 639-2 (language codes), rightsstatements.org, and various name or material type vocabularies. It is unclear if this interpretation affected any of the results, especially in terms of respondents leaving out additional non-subject vocabularies that they may use.

Responses were also split with regard to the use of local or regional controlled vocabularies, as 54 repositories (47%) do not use local or regional controlled vocabularies and 60 repositories (53%) use local vocabularies to describe names (persons, organizations, and academic departments), geographic locations, material types, genres, or specialized fields of study.

The survey also asked for the size of each repository, using the number of metadata records as a measurement. Of the 144 responses, most replied with rounded numbers--such as "thousands" or "hundreds of thousands"--although 4 respondents did not know the number of items, or marked this question as "not applicable." After breaking responses into groups based logarithmically around that representation, the size of the repositories creates a rough bell curve (see Figure 6). A 41% majority of repositories (58) were in the middle of the distribution in the "tens of thousands" category. Only 8% of repositories were extremely large--consisting of millions or tens of millions of records--while 7% of repositories had fewer than one thousand records.

Figure 6. Size of Respondents' Repositories
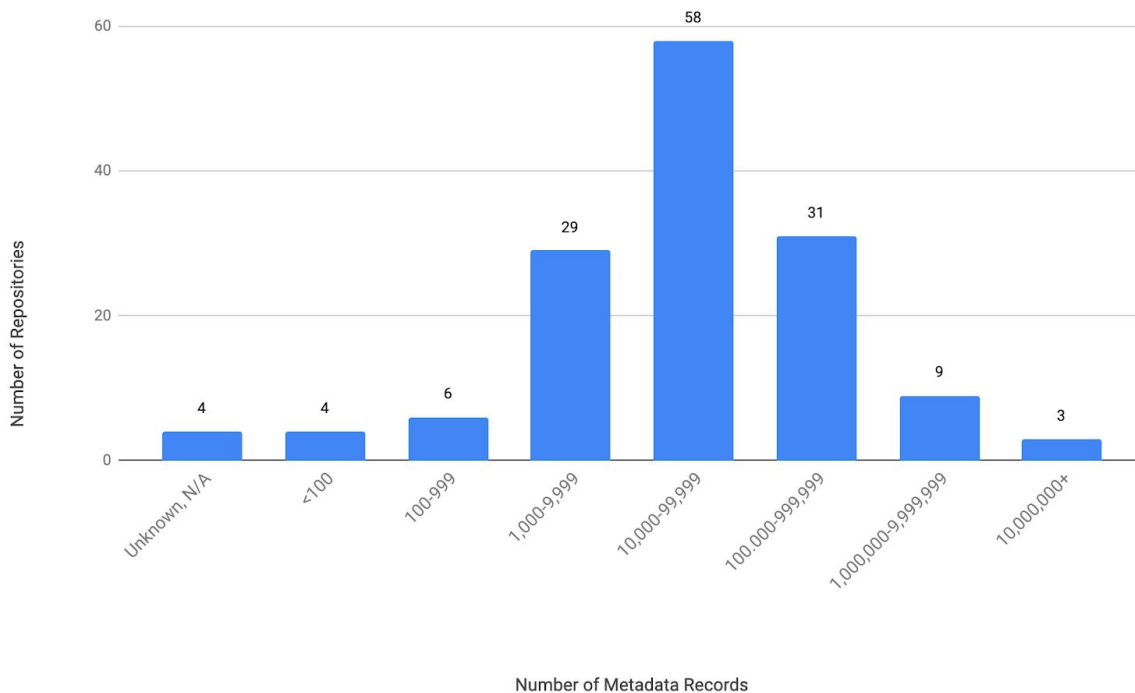
Section 3: Metadata Elements Grids

This portion of the survey included two "grids" listing 26 commonly-used metadata elements, answered for individual repositories. When completing the first grid,[15] respondents selected whether each specific element is required, recommended, or optional in the repository, with instructions to not choose an option if the element is not included (unavailable) in that repository.

The elements with the largest number of responses in this grid--occurring in 121 repositories--are *title* and *creator*, with responses of 115-120 for six additional elements (see Table 4). The least common elements occurring in only 84 repositories are *digitization specifications* and *table of contents*; however, most elements are widely shared, with only five other elements representing fewer than 100 repositories.

Table 4. Most- and Least-Frequent Elements, by Number of Repositories that Require It

Element	Req[16]	Rec	Opt	Total	Element[17]	Req	Rec	Opt	Total
title	115	3	3	121	*physicalLocation*	36	18	39	93
identifier	93	13	9	115	*genre*	26	42	31	99
rights	75	26	16	117	*isPartOf*	17	29	53	99
date	54	54	12	120	*transcription*	8	27	58	93
creator	46	64	11	121	*digitization specifications*	8	19	57	84
subject	36	56	28	120	*table of contents*	4	9	71	84
description	33	47	35	115	*relation*	4	24	70	98
contributor	16	52	50	118					

Overall, there is relatively little difference in total element distribution--i.e., all of the elements on the list are shared by 84-121 repositories, rather than certain elements being extremely common and others appearing in very few repositories. Not all elements are equally represented in terms of being required vs. recommended or optional (see Figure 7); however, every element from the list is required by at least one repository, and there is significant correlation between total frequency and required or recommended elements. The eight most frequent elements (Table 4) comprise all five of the most-often recommended elements (*date, creator, subject, description,* and *contributor*) and three of the most-often required (*title,*

---

[15] §3.19

[16] Columns denote elements identified as required (Req), recommended (Rec), or optional (Opt).

[17] Elements on the left side of the table are most frequent, highlighted elements on the right are least frequent.

identifier, and *rights*). There are two outliers that are required 68% of the time, but occur in fewer repositories:

- *type* --------------- required by 75 out of 111 repositories
- *collection title* --- required by 74 out of 109 repositories

Similarly, the least frequently-occurring elements also include four of the five elements most often designated "optional" (*relation, transcription, table of contents,* and *digitization specifications*), along with *alternative title*, which is optional 85% of the time (93 out of 110 repositories).

Figure 7. Element Frequency by Repository
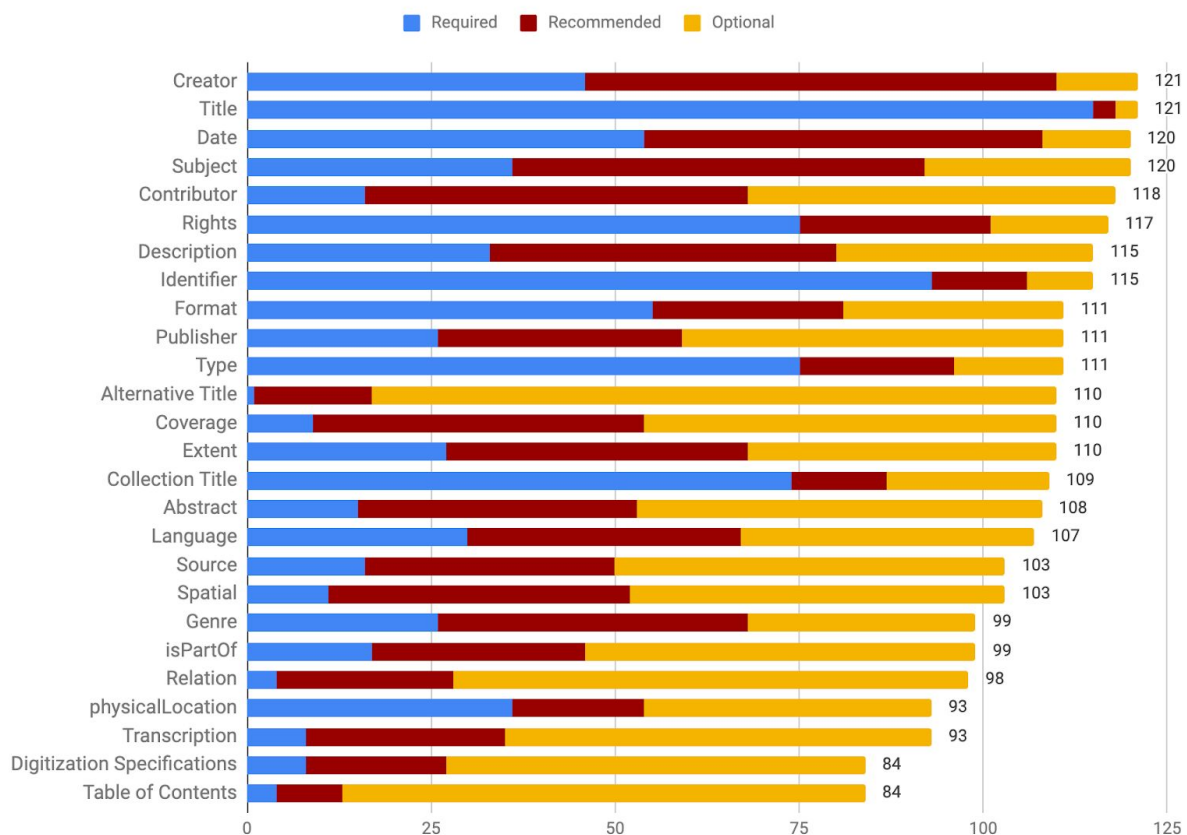
Each of the grids in this section was also followed by free-text input fields to allow respondents to account for additional metadata elements in local usage. These answers have been organized into general categories based on separate sections that asked about required, recommended, and optional elements (see Table 5). Given the nature of this section, both responses and interpretation are somewhat subjective and difficult to classify, especially in cases where elements are only used (or required) for certain material types in the repository (or where there was not enough context to determine the best fit).

The most frequently-occurring additional elements written in by respondents appear to be related to administrative or technical metadata, access and use, archival information (e.g., information about the physical materials), format-specific elements (especially related to video, photos, or theses and dissertations), and various miscellaneous elements (particularly for local information). There are also a number of extremely specific elements (e.g., *ISSN*, *time period*, *translated title*, etc.) for information that may be folded into the standard grid elements for repositories that use qualified schemas (e.g., *identifier*, *coverage*, *title*, etc.).

Table 5. Categorized Respondent-Submitted Metadata Elements[18]

Category	Example Elements	Req[19]	Rec	Opt
Access/Use	*usage or access restrictions, dcterms.rights, license, periodEmbargo*	7	2	4
Acknowledgement			1	
Administrative/ Technical Metadata	*date digitized, metadata creator or cataloger, filename, processing method*	21	6	2
Author/Creator	*affiliation, creator role, biographical note, creator ORCID*		4	7
Archival	*provenance, physical location, box, archival series*		14	4
Citations	*preferred citation, attribution*	3		
Collection	*digital collection title, collection description*	2		
Contact Information	*contributing partner or depositor information*	3		
Coordinates	*geographic coordinates, latitude/longitude*		2	
Coverage	*temporal, time period*			2
Date	*date created, century*	1		2
Deposit	*depositor name, depositor date*		2	
Disclaimer		2		
Donor/Institution	*department or college, partner, contributing institution*	6		2
Format/Medium	*document type, medium*		4	
Format-Specific Elements	*video captioning, thesis/dissertation department, cast/credits, genus/species*	15	6	9

---

[18] Numbers in this table are approximate, representing individual elements rather than responses; it was not always clear when respondents were referring to a single element containing multiple pieces of information or vice versa.

[19] Columns denote elements identified as required (Req), recommended (Rec), or optional (Opt).

Identifier	*call number, DOI, ISSN*	2	2	5
Keywords		2		
Note	*local note, general note, source of title*		4	3
Publication	*publication status, volume/issue, original publication*	3	5	7
Relationship	*is format of, has part*			3
Title	*translated title, journal title*			2
URLs/URIs	*PURL, link to digital version, isShownAt*	3		
Other	*available for reproduction, object name, content flag, sponsor, inscription*	8	7	4

The second metadata grid[20] listed the same elements but asked whether each is "evaluated" or "not evaluated" in the repository, with instructions to only respond for elements available in the repository (see Figure 8). The largest total response in this grid was 119 for *title*[21] (87 repositories evaluate the element) and the fewest responses--81--were for *digitization specifications* (19 repositories evaluate the element) and *transcription* (23 repositories evaluate the element). In the evaluation grid, the most frequently-evaluated elements are *title*, *creator*, *date*, and *subject*; these are also the most frequently-available elements based on the data in the first grid.

---

[20] §3.20

[21] In the first grid, some elements occur in up to 121 repositories, so the exact correlation between element occurrence and evaluation in the data is unclear.

15

Figure 8. Element Evaluation by Repository, in Order of Frequency
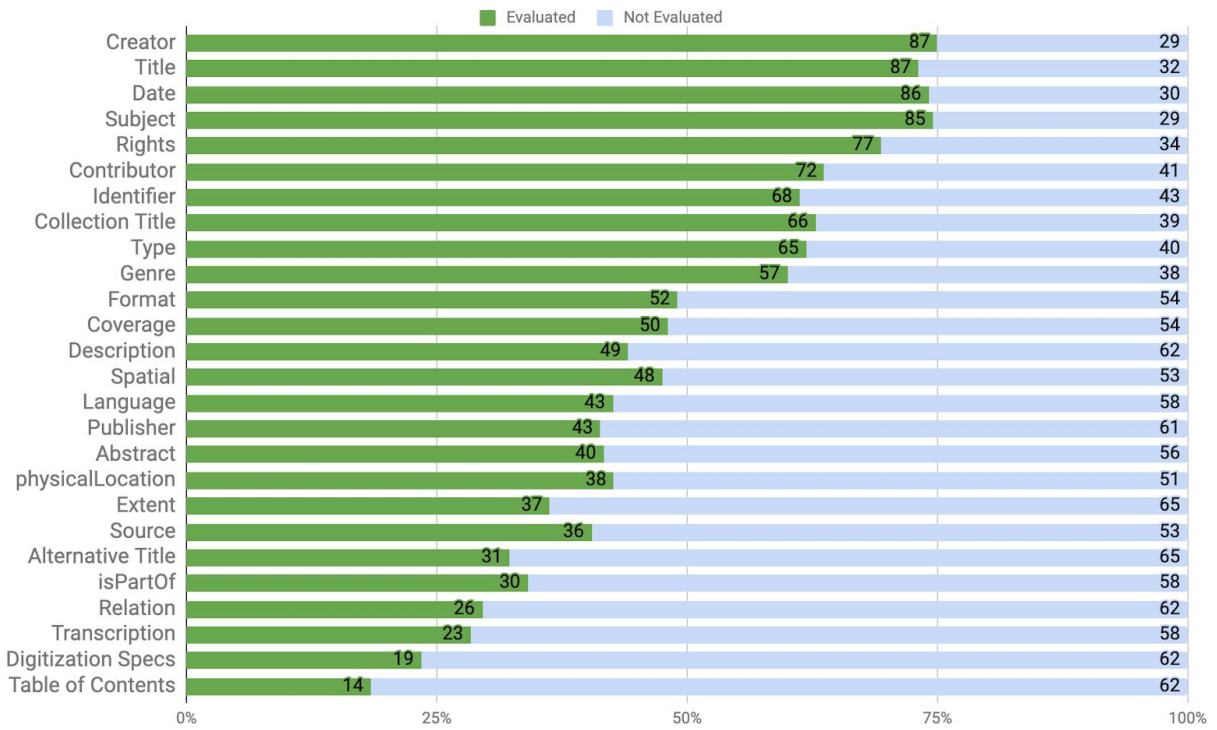
Some correlation of total numbers is expected, given that organizations would not be evaluating elements that they do not use (see Table 6). Elements in the middle of Figure 7 (available in 100-115 of the repositories) tend to be evaluated slightly less than those that are most frequent, aside from three elements that are evaluated proportionately more often: *identifier*, *type*, and *collection title.* Additionally, the *genre* element--which is among the less-frequently-available elements--is evaluated in 60% of the repositories where the element is available.

Table 6. Most- and Least-Frequently Evaluated Elements[22]

Element	# Repositories Evaluating	# Repositories Not Evaluating	Available in # Repositories[23]	# Repositories Requiring
title	87	32	121	115
creator	87	29	121	46
date	86	30	120	54
subject	85	29	120	36
rights	77	34	117	75
description	49	62	115	33
extent	37	65	110	27
alternative title	31	65	110	1
relation	26	65	98	4
digitization specifications	19	62	84	8
table of contents	14	62	84	4

In terms of the free-text option regarding evaluation of elements not on the list, there were only 14 entries and respondents often provided general answers rather than specific element names, such as:

- Every element in usage for a particular repository--e.g., "All the values described in the answer previous - if the data is available, the metadata is evaluated."

- Elements checked on a collection-by-collection basis--e.g., "collection specific fields."

- Organizations using holistic sampling methods that may not account for every element or evaluate specific elements reliably--e.g., "I look at some records as a sample."

The respondents who did list specific elements replied with a broad range of element types that had little overlap.

---

[22] The first rows are most-frequently evaluated and highlighted rows are the least-frequently evaluated.
[23] Data in the two right columns comes from the previous question, §3.19.

Section 4: Metadata Quality Assessment

In the fourth section, respondents answered several questions about metadata evaluation practices--both actual and aspirational--at their institution. Topics addressed tools used by respondents to evaluate metadata quality; the kinds of characteristics that respondents evaluate--or would like to evaluate--when judging a metadata record; and the methods that respondents employ to evaluate metadata. The survey also asked respondents to provide any final feedback.

With regard to the tools used to assess metadata quality,[24] respondents were able to choose multiple tools--for a total of 164 answers from 76 respondents--and selections demonstrated several kinds of tools as more heavily used than others (see Table 7). The most frequently-used tools included spreadsheets, OpenRefine, and MARCEdit. The free-text portion of this question included database tools and XML tools such as Oxygen XML Editor or Schematron. No respondents selected Gadget, LibreCat/Catmandu, or LODrefine from the list of options.

Table 7. Statistics for Metadata Evaluation Tool Usage

| Evaluation Tool	Number of Organizations	Percentage of Respondents (n = 76)
Spreadsheets	63	83%
OpenRefine	43	57%
MARCEdit	24	32%
DPLA OAI Aggregator Tools	11	14%
Metadata Breakers	3	4%
Python pandas	2	3%
Metadata Quality Control (MDCQ) by AVP	2	3%
Database tools (e.g., SQL)[25]	4	5%
XML tools (e.g., Oxygen XML Editor, Schematron)	7	9%

---

[24] §4.22
[25] Highlighted rows are tools mentioned multiple times in free-response answers rather than list options.

In addition, the survey asked respondents to rank seven aspects of quality[26] in the order that their institution views their importance. A high-level overview of responses to this question is provided in Figure 9. Examining the 89 responses to this question reveals several clear trends based on how often each was ranked as a top or second-highest priority:[27]

- **Accuracy** --------------------------------- 57 institutions (64%)
- **Accessibility** ---------------------------- 11 institutions (12%)
- **Completeness** ------------------------- 52 institutions (58%)
- **Conformance to expectations** ----- 19 institutions (21%)
- **Consistency** ---------------------------- 29 institutions (32%)
- **Provenance** ------------------------------ 4 institutions (4%)
- **Timeliness** --------------------------------- 2 institutions (2%)

Based on these numbers, the metadata quality aspects that respondents viewed as the most important were accuracy, completeness, and consistency. Respondents demonstrated clearly that some aspects were not high-level priorities--including conformance to expectations, accessibility, provenance, and timeliness.

Figure 9. Rankings of Metadata Quality Aspects by Institutional Importance


---
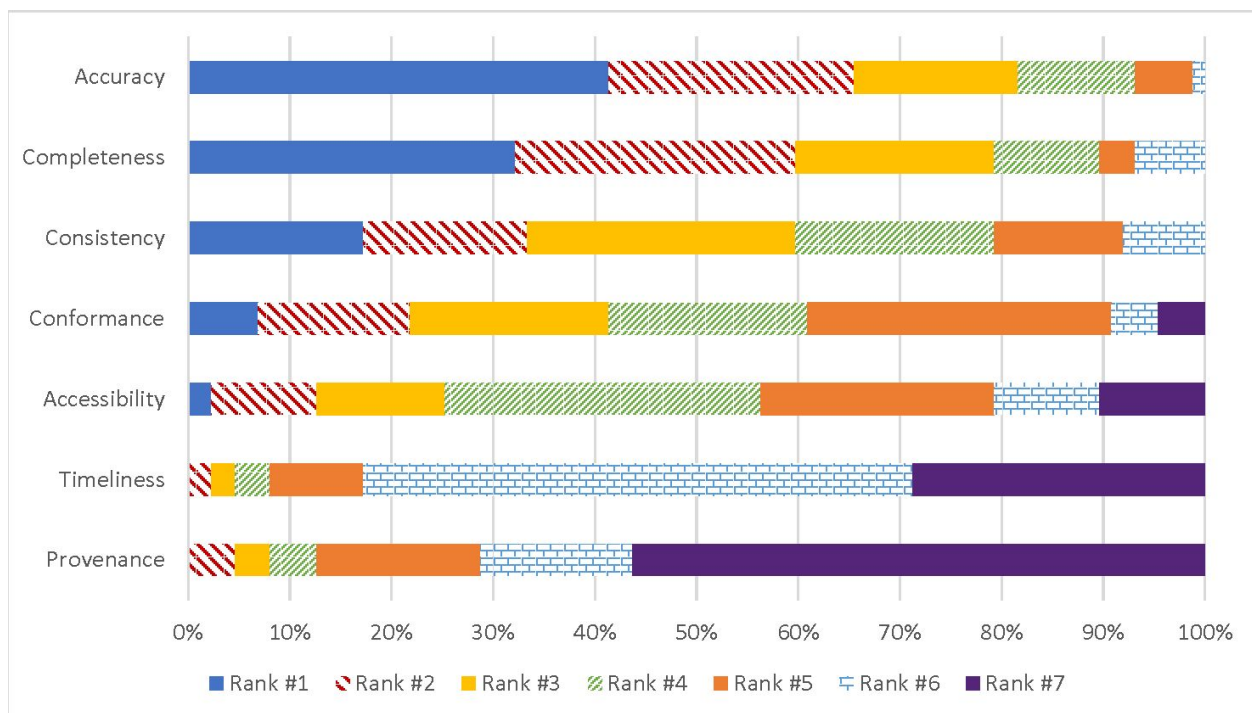
[26] A definition for each aspect is provided in Appendix A.
[27] Provenance and timeliness were not ranked first by any respondent, so numbers for those aspects only reflect the number of times they were ranked second.

Question 24 asked respondents how they measured the same seven qualities in their own metadata. Across the 48 free-text responses, several themes emerged as respondents described their more common evaluation techniques.[28] The responses demonstrated a reliance on manual methods (e.g., spot-checking entire records or comprehensive reviews); established requirements or standards, including controlled vocabulary implementations; and various tools and scripts (see Figure 10).

Figure 10. Methods of Metadata Quality Evaluation Based on Free-Text Answers
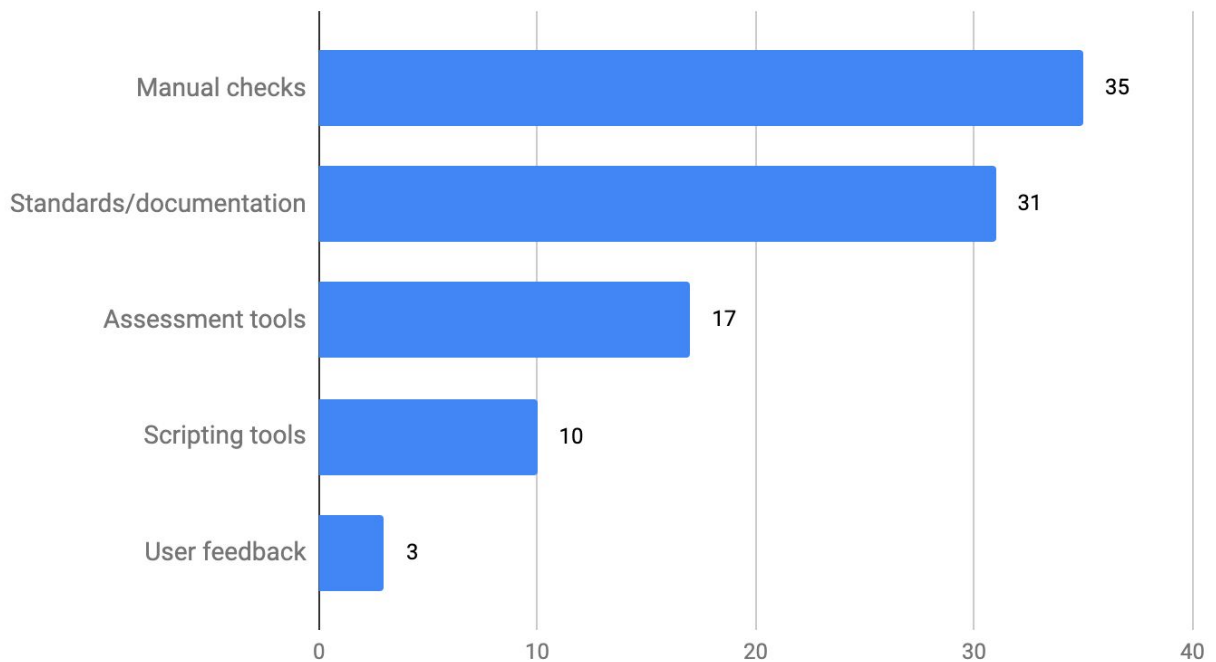
The leading evaluation method of manual review may be time consuming, but the reasons for this particular method include the necessity for human review of metadata accuracy (e.g., to ensure that information matches the item described) and technological challenges. Highlighting the need for multi-human review, one respondent indicated the manual checks are done by a person who did not create the metadata for an added level of evaluation. Another respondent indicated limited validation options due to a non-XML metadata schema which resulted in manual checks. Inversely, another respondent stressed the exclusive responsibility of evaluating metadata:

> "[...] as the sole metadata archivist on staff, I review every record and edit for
> clarity, semantically and syntactically."

---

[28] Please note, respondents may have indicated multiple evaluation methods per response.

In addition to manual review, 65% of the respondents use established documentation and standards to evaluate metadata quality; some cite the inclusion of required and recommended metadata elements as a baseline, e.g., "For completeness we have a metric calculated when items are indexed that checks for the presence of fields that are required for a minimally viable record which we refer to as 'complete' or more reasonably 'minimally complete.'" Quality assessment tools mentioned in response data include OpenRefine (6 references), DPLA OAI aggregation tools, functionality in local systems (including ContentDM, DSpace and ArchivesSpace), Combine, and--according to one respondent--a Ruby-based tool available on GitHub.

The survey then asked respondents to describe characteristics that they would like to measure, but are unable to evaluate in their repositories.[29] Of the 35 free-text responses to this question, the characteristics that respondents would like to measure include:

- Consistency ---- 10 institutions (29%)
- Accuracy --------- 9 institutions (26%)
- Timeliness -------- 7 institutions (20%)
- Accessibility ------ 4 institutions (11%)
- Provenance ------ 3 institutions (9%)
- Completeness --- 2 institutions (6%)

In some cases, these aspirational responses aligned somewhat with quality aspects that most respondents ranked highly at their organizations (e.g., consistency was ranked first by about 17% or second by another 16% of respondents and is also at the top of this list).[30] However, around a third of respondents said that their organization ranks completeness as most important (i.e., ranked #1), but only two respondents mentioned completeness in this question. This may also reflect the fact that the quality aspects that respondents most often want to measure *but are unable to* tend to be aspects that may require more human intervention--e.g., accuracy and timeliness--rather than aspects that may be easily evaluated in systematic ways, such as the completeness calculation mentioned by a respondent in the previous question.[31]

Evaluating accuracy again calls out the need for multiple sources of human intelligence, e.g., "more thorough evaluation for content accuracy - specialized knowledge held by the contributors, but not necessarily in the metadata unit where the review happens." Respondents acknowledged the difficulty of evaluating certain metadata elements over others. They particularly mentioned measuring consistency among free-text elements, which tend to be more difficult to evaluate in systematic ways compared to elements that have more standardized formatting (such as names or dates).

---

[29] §4.25
[30] §4.23
[31] §4.24

Respondents' emphasis on evaluating metadata production rates and provenance could suggest a desire for more sustainable/efficient work models. Some respondents mentioned various ways of keeping track of metadata work, for example, tracking changes to a metadata record from an administrative perspective--e.g., "provenance over time, who edited what, which collections have been edited recently and which have not received a recent review"--as well as time spent on metadata work per item by metadata specialists reflecting a productivity standpoint: "Tracking time per item is a challenge; I have specialists track overall time and counts to get an average." In addition to comments regarding the seven quality aspects, some respondents' answers to this question also touched on discoverability/searchability:

> "I would most like to know which metadata elements are most useful in deriving search results"

The final question[32] solicited any thoughts respondents would like to share, which collected 18 respondents' closing feedback. Although nearly half of these related to the survey itself--for example, further explanation of earlier responses--a number of themes emerged:

- Comments/clarification regarding survey and responses ------ 7 respondents (39%)
- Issues with workflow, documentation, and/or standards ------- 7 respondents (39%)
- Lack of staff, expertise, and/or resources ------------------------- 5 respondents (28%)
- Unsure of need for metadata quality assessment --------------- 3 respondents (28%)

Some respondents addressed workflow, documentation, or standards issues. Not all respondents currently have established standards, or metadata application profiles, or have not yet formulated expectations for evaluation. Other organizations have evaluation practices that

---

[32] §4.26

may not be consistent, making it difficult to agree on standards, or have a systematic approach. Additionally, several responses reflected a lack of available staff, expertise, and/or resources, further complicating metadata quality assessment activities.

Statements regarding challenges in metadata quality analysis:

"We are still developing the metadata profile of our digital collections and thus have no assessment standards established"

"our metadata quality analysis are very ad hoc, irregular, and targeted to particular problems we experience...we don't really make sure we're adhering to very many external guidelines"

"I am the only one who does anything with non-MARC metadata, and my experience is still quite limited"

Finally, three responses questioned the necessity or importance of metadata evaluation. For example, respondents questioned various kinds of metadata evaluation directly, as exemplified by one respondent's comment that "our issues are finding anomalies and efficient correction, *not measurement* [our emphasis]," while another stated "we are creating our own metadata, so I'm not sure how this 'evaluation' process is relevant…we do not evaluate metadata contributed by [other organizations]." One respondent noted that,

"as a practitioner of digital libraries, I would say I haven't thought about metadata in the ways in which you presented it in this survey [...] I most care about 'Is this metadata working for our users?'"

Discussion

Based on the respondent data in section 1, a predominant profile emerges. Respondents were mostly working in academic libraries with over 100 employees that have 5 or fewer non-MARC metadata FTEs, and they typically have 5-10 years experience working with non-MARC metadata (and some experience with MARC metadata). This indicates mostly mid-career metadata professionals. While this profile highlights the majority of respondents, it does not diminish the diversity of roles from differing types of organizations that completed the survey. Respondents' answers ran the gamut, whether a question addressed metadata responsibilities,[33] metadata schema or controlled vocabulary employed,[34] metadata elements used in repositories,[35] or metadata assessment tools.[36] Even in cases where responses to a question showed some uniformity, there still existed significant diversity among a portion of the responses. While this reality is unsurprising as standards, tools, elements, and tactics are selected to fulfill specific needs--and modified as needed to best suit specific projects--it also suggests that future metadata assessment evaluation options must reflect a similar flexibility to remain accessible and usable regardless of project type, available resources, or user expertise.

In terms of specific response data from section 3 there is an obvious trend concerning the metadata elements typically required by repositories. For example, the two elements that respondents most often selected as being required in a repository were *title* and *identifier*. Although both are fairly common elements that ensure staff and users can access digital records, commonalities in element usage could indicate that many organizations share best practices or guidelines, or that MAPs are influencing each other (e.g., MAP creators could be drawing upon other MAPs to customize their own schemas). Another possibility is the influence of the Digital Public Library of America's (DPLA) metadata requirements. For example, the DPLA's Metadata Application Profile version 5.0 has minimal required metadata elements that include *title*, which was also overwhelmingly the most required metadata element (95% required) in the survey.

Despite our efforts at clarity, some of the data reflects that respondents did not all interpret questions the same way. For example, when discussing controlled vocabularies, there were distinct differences in respondents' answers, which may affect overall data. Similarly, the types of internal documentation used for metadata creation may have resulted in unexpected response data. Slightly more than half of respondents indicate that they are implementing a MAP, although a higher response was anticipated. While "Metadata Application Profile" was defined in the beginning of the survey, the question "Are you implementing a Metadata Application Profile"[37] could have been followed up with additional questions to clarify if some organizations might be implementing similar guidelines or best practices without using the term

---

[33] §1.8

[34] §2.15 & §2.16

[35] §3

[36] §4.22

[37] §2.13

'metadata application profile' for local documentation. While these differences were noticeable, further research may uncover minor anomalies for some of the other data that were potentially influenced by varying local practices. This also underscores the need for further information and flexibility for benchmarking that would fit a variety of needs.

Finally, respondents revealed that some metadata characteristics remain difficult to computationally measure. This is exemplified in section 4, in which respondents emphasized both the primacy of certain metadata characteristics over others (e.g., accuracy) as well as a wide range of metadata evaluation techniques--such as manual checking and relying upon standards and documentation. These answers reinforced that while some tools--such as applications and scripting--can help successfully evaluate some metadata qualities, a human element is still required during review to ensure that metadata records appropriately describe their resources (at least for the time being). The current necessity of manual metadata assessment reinforces the need for streamlined tactics that partner efficient resources with human effectiveness to maximize metadata evaluation impact.

Conclusion

This paper provides a breakdown of data, general analysis of the survey results, and synthesizes information about how organizations already do metadata quality assessment and benchmarking in line with the initial research questions:

What metadata requirements and standards are commonly implemented in libraries, archives, museums, and other cultural heritage organizations. Aside from documenting the various responses regarding required elements and standards,[38] this survey discovered that, while there is often disparity among organizations, there is also significant overlap that could serve as a foundation for benchmarking. For example, more than half of the repositories share one of the two most-common schemas (Dublin Core and MODS) though there was little to no agreement among the other half of responses. Alternately, individual elements are widely shared among digital libraries,[39] which may suggest that element-based benchmarks or quality measurements not tied to particular schemas could be useful to the community.

The methods and criteria used to evaluate metadata quality in these institutions. Despite technological advances, evaluating metadata accuracy seems to still rely heavily on manual assessment, though there are a number of tools[40] in use by some organizations. One issue with tools is that they may be reliant on a particular DAMS or schema, and those were areas where organizations often use local implementations,[41] making it difficult to share. In terms of criteria, the survey primarily relied on free-response answers, and most respondents were not explicit about evaluation criteria (although some answers mentioned specific aspects of quality, e.g.,

---

[38] §2 & §3
[39] §3.19
[40] §4.22
[41] §3.12 & §3.15

using required elements as a concrete measure, or at least a component, of completeness). It may be useful to determine if these kinds of methods are generalizable--or, if not, where they become problematic--as a starting point for certain areas of quality.

Gaps in knowledge and practice related to metadata quality. Information related to this research question was addressed primarily through free-text responses and few respondents directly discussed this topic. However, evaluation strategies referenced by respondents largely rely on manual assessment; this poses potential barriers due to the necessity of staff and time for this task, especially given that 49% of respondents report having 3 or fewer FTEs dedicated to non-MARC metadata. Because the time needed to manually evaluate metadata could quickly exceed departmental bandwidth, benchmarks may need to take into account limits on staffing as well as ways that organizations may be able to maximize metadata effectiveness and efficiency. As one respondent poignantly commented:

> "It's not magic. It's manual. Because even automated work needs time and attention to set up, run, evaluate, etc."

All of the data collected in this survey is useful to the digital library community to understand the ways that institutions implement metadata. It also assists in directing possible areas and most useful methods or frameworks for approaching quality assessment. However, more work will need to be done to start establishing generalized benchmarks for evaluating metadata quality.

Further Research

Since this survey is an initial step, the Benchmarks sub-group intends to follow up on these results using more targeted interviews with specific organizations and possibly other methods of data collection to clarify survey responses. Some areas of further study include:

- **How organizations are defining evaluation.** We purposely did not try to define or limit types of evaluation in this context (e.g., existence of a value, validation of values, review for formatting or spelling, etc.) or methods used (e.g., automated validation, manual evaluation, both/neither, etc.). However, it may be useful to gather more specific data.

- **More specifics about the mechanics of evaluation.** We may also want to explore how elements are chosen for evaluation; what methods are used for evaluating different types of elements, especially free-text type fields, and whether there are commonalities among organizations; and the importance or role of documentation (e.g., MAPs or guidelines) as a tool for evaluation.

- **Quality components and metrics.** Additional clarification on the particular aspects of quality that organizations are currently evaluating (e.g., completeness, accuracy, provenance, etc.) and what metrics, if any, are used to determine the level of quality.

- **Evaluation tools**. To address particular quality aspects, we may want to look into tools that are currently being used by individual organizations, or where tools fall short--e.g., specific desired functionality that is not yet available, or usage difficulties--to determine how tools might affect benchmarking and quality assessment. As part of the wider goals of the Metadata Working Group, it might also be useful to add any new tools to the group's assessment toolkit and to talk with respondents about resources compiled by the group to determine where there are gaps in familiarity with tools that could assist current workflows.

Working toward common, recognizable benchmarks would be beneficial to the wider digital library community to provide common points of reference, as well as to potentially improve metadata shareability. Additionally, benchmarks would serve as a useful tool for organizations that want to improve metadata quality for their users but do not currently have a framework for evaluation or remediation/enhancement of existing values.

Appendix A: Survey Questions

This table lists the questions from the survey with the possible options that respondents could choose (when applicable). There is also a PDF version which preserves additional formatting of the original survey.[42]

Consent		
Q1*	By selecting "Yes, I agree", I confirm that I am 18 years or older, have read the information in this consent form, and have had the opportunity [to ask] questions. I voluntarily agree to take part in this study.	Yes, I agree No, I do not agree
Part 1: Respondent Profile		
Q2	What kind of organization do you work for?	Libraries - Academic Libraries - Public Libraries - Special Archives Museum Consortium Aggregation Project Other <<free text field>>
Q3	If you are willing, please fill in the name of the organization, consortium, or aggregation project you are representing	<<free text field>>
Q4	How many total employees work for your organization?	1-10 11-50 51-100 101+
Q5	How many full time employees in your organization work with non-MARC metadata? If half-time, indicate .5, etc.	<<free text field>>
Q6	How long have you been working with non-MARC metadata?	Never 0 to 4 years 5 to 9 years 10+ years
Q7	How long have you worked with MARC metadata?	Never 0 to 4 years 5 to 9 years 10+ years

---

[42] http://dlfmetadataassessment.github.io/assets/Survey_of_Metadata_Quality_Benchmarks.pdf

Q8	What tasks are your responsibility when working with metadata? Select all that apply.	Creating descriptive metadata Setting guidelines and best practices Supervising metadata creators Quality control checks Managing existing metadata (migration, remediation, enhancements) Other <<free text field>>
Part 2: Metadata Basics		
Q9	How many repositories does your organization manage?	1-2 3-4 5-6 Other <<free text field>>
Q10*	For how many repositories would you like to fill out the following section?	<<free text field>>
Q11	Does this repository serve as an institutional repository, a platform for digital collections, or both?	Digital Collections Institutional Repositories Both
	(Optional) Name of Repository	<<free text field>>
Q12	What type of system is being used?	Bepress CollectiveAccess CONTENTdm DSpace Eprings Islandora Omeka Samvera Other <<free text field>>
Q13	Are you implementing a Metadata Application Profile (MAP)?	Yes No
Q14	If you answered yes to implementing a Metadata Application Profile, is it from a governing body (digital library/consortia) or created specifically for your digital library?	Using a MAP created by an external consortia (example, DPLA hub; Mountain West Digital Library) Using a MAP created specifically for local repository
Q15	What metadata schema is being used?	Dublin Core EDM MODs PBCore Premis Qualified Dublin Core VRA Core Other <<free text field>>

Q16	Which controlled vocabularies are being used? Select all that apply.	FAST Subject Headings GeoNames.org Getty Art & Architecture Thesaurus Getty Thesaurus of Geographic Names Getty Union List of Artist Names Library of Congress Genre/Forms Library of Congress Subject Headings Library of Congress Thesaurus of Graphic Materials Medical Subject Headings MeSh Other <<free text field>>
Q17	Approximately how many descriptive metadata records are in this repository?	<<free text field>>
Q18	Do you use local or regional controlled vocabularies?	Yes (please specify) <<free text field>> No
Part 3: Metadata Elements Grids		
Q19	Please indicate if an element is required, optional, or recommended for your repository or project. Select "Required" for elements that must be present in a metadata record in order for a resource to be published online and/or harvested. Select "Recommended" for elements that are strongly encouraged. Select "Optional" for elements that are only included when applicable. **If an element in the list is not relevant to your repository, please do not select any options for that element.** If there are elements missing from this grid that are required, recommended, or optional for your project, please add these in the free text field below.	Abstract Alternative Title Collection Title Contributor Coverage Creator Date Description Digitization Specifications Extent Format Genre Identifier isPartOf Language physicalLocation Publisher Relation Rights Source Spatial Subject Table of Contents Title Transcription Type [Note: The grid included definitions for each element; see below (♦) for more information]
	Please name and define any other **required** metadata elements not listed above:	<<free text field>>

	Please name and define any other **recommended** metadata elements not listed above:	<<free text field>>
	Please name and define any other **optional** metadata elements not listed above:	<<free text field>>
Q20	Which metadata elements do you evaluatie for quality? Select "Evaluated" for elements that are in your system and evaluated for any measure of quality. Select "Not Evaluated" for elements that are in your system but not measured for quality. **If an element in the list is not relevant to your repository, please do not select any options for that element.** If there are elements missing from this grid that are evaluated for quality, please add these individually in the free text field.	Abstract Alternative Title Collection Title Contributor Coverage Creator Date Description Digitization Specifications Extent Format Genre Identifier isPartOf Language physicalLocation Publisher Relation Rights Source Spatial Subject Table of Contents Title Transcription Type [Note: The grid included definitions for each element; see below (♦) for more information]
	Please name and define any other **evaluated** metadata elements not listed above:	<<free text field>>
	Please name and define any other **not evaluated** metadata elements not listed above:	<<free text field>>
Q21	This section will repeat based on the number of repositories you indicated earlier. If you do not want to fill this section out again, please indicate below.	I do not want to repeat this section

Part 4: Metadata Quality Assessment		
Q22	Does your organization use any tools for metadata quality assessment? Select all that apply.	DPLA OAI Aggregator Tools Gadget LibreCat/Catmandu LODrefine MARCEdit Metadata Quality Control (MDQC) by AVP Metadata Breakers OpenRefine Python pandas Spreadsheet based software (Microsoft Excel, LibreOffice Calc, Google Sheets) Other <<free text field>>
Q23	When judging the quality of a metadata record, what aspects are most important to your organization? By dragging and dropping, please rank (1 being most important) the characteristics of quality per the DLF AIG Metadata Assessment Working Group Toolkit[43]).	Completeness Accuracy Accessibility Conformance to expectations Consistency Timeliness Provenance [Note: This question included definitions for each aspect; see below (✧) for more information]
Q24	How do you measure for the characteristics described in the previous question?	<<free text field>>
Q25	What characteristics would you like to measure but are unable to?	<<free text field>>
Q26	(Optional) Any thoughts you would like to share?	<<free text field>>
	May we follow up with you if we have further questions? Would you be open to an informational interview? If yes, please leave your name and email address below:	<<free text field>>

* Required question

---

[43] Original text linked to metadata quality assessment information on the Metadata Working Group github site; see also: https://dlfmetadataassessment.github.io/Framework

Definitions

✦ Element definitions from Part 3.

Abstract: A summary of the resource.
Alternative Title: An alternative name for the resource.
Collection Title: Name of a group of related resources which the described resource belongs to.
Contributor: An entity responsible for making contributions to the resource.
Coverage: Describes the spatial and temporal characteristics of the resource.
Creator: An entity responsible for making the resource.
Date: The date of the creation of the original resource.
Description: An account of the resource, including item's history, appearance, contents, etc.
Digitization Specifications: Description of process, equipment, and specifications used to convert resource to digital format.
Extent: The size or duration of the resource.
Format: File format of the digital resource.
Genre: Nature of original resource.
Identifier: Unambiguous reference to the resource.
isPartOf: A related resource(s) in which the described resource is physically or logically included.
Language: Language of the resource.
physicalLocation: The institution or repository that holds the resource or where it is available.
Publisher: An entity responsible for making the resource available.
Relation: A related resource(s).
Rights: Information about rights held in and over the resource.
Source: A related resource from which the resource is derived.
Spatial: The geographic topic or applicability of the resource.
Subject: Topic that describes what the resource is about.
Table of Contents: A list of subunits of the resource.
Title: A name given to the resource.
Transcription: Transcription or full text of resource.
Type: The nature of the resource (StillImage, MovingImage, Sound, or Text).

⬦Metadata quality aspect definitions[44] from Part 4.

Completeness: The element, property, and/or attribute is present.
Accuracy: Information is correct both semantically and syntactically.
Accessibility: Metadata can be read by both humans and machines.
Conformance to expectations: Values adhere to the expectations of your defined user communities (both internal and external)
Consistency: Semantic and structural values and elements are represented in a consistent manner across records. Values are consistent within your domain.
Timeliness: When the resource changes, the metadata is updated accordingly. When additional metadata becomes available or when metadata standards change, the metadata associated with the resource changes.
Provenance: You have information about the source of the metadata, and you can track metadata transformations back to the original form of the metadata record.

Appendix B: Response Data

A complete set of anonymized data collected in this survey is available for download.[45] Date, time, IP addresses, and geographic data has been omitted. Responses that included project, organization, and/or repository names were removed from this data. Any potentially identifying names, acronyms, and/or links were removed from free text responses as well.

---

[44] Also available at: http://dlfmetadataassessment.github.io/Framework
[45] Anonymized response data available at:
http://dlfmetadataassessment.github.io/assets/DLFMetadataQualityBenchmarksSurveyTextResponse.xlsx

Appendix C: Distribution List

The survey invitation was promoted via e-mail to professionals in the digital library community, including the following listservs:

- Association for Library Collections and Technical Services (ALCTS) Central, alctscentral@lists.ala.org

- Association of Moving Image Archivists Listserv, AMIA-L@LSV.UKY.EDU

- Art Libraries Society of North America (ARLISNA), ARLIS-L@lsv.arlisna.org

- Autocat, autocat@listserv.syr.edu

- Bibliographic Framework (BIBFRAME), BIBFRAME@LISTSERV.LOC.GOV

- Code4Lib, CODE4LIB@LISTS.CLIR.ORG

- Digital Scholarship Section (American Library Association), dss-l@lists.ala.org

- Digital Library Federation (DLF)-Announce, DLF-ANNOUNCE@lists.clir.org

- Digital Library Federation Assessment Interest Group (DLF-AIG) Google Group, digital-library-assessment@googlegroups.com

- Digital Public Library of America (DPLA) Hubs, allhubs@dp.la

- DSpace, dspace-community@googlegroups.com

- Metadata Interest Group (American Library Association), alcts-nrmig@lists.ala.org

- Metadata Librarians, metadatalibrarians@lists.monarchos.com

- Program for Cooperative Cataloging, PCCLIST@LISTSERV.LOC.GOV

- Society of American Archivists (SAA) metadata and digital objects, ARCHIVISTS-metadata@ConnectedCommunity.org

- Troublesome Catalogers Facebook page, https://www.facebook.com/groups/161813927168408/

- Visual Resource Association (VRA) Listserv, VRA-L@LISTSERV.UARK.EDU