Metadata Quality in Digital Libraries

An Analysis of Survey Response Data

|  |  |  |
| --- | --- | --- |
| **Hannah Tarver** | **Meredith L. Hale** | **Rachel White** |
| *University of North Texas*  *USA*  *hannah.tarver@unt.edu*  *https://orcid.org/0000-0003-2344-9268* | *University of Tennessee, Knoxville*  *USA*  *mhale16@utk.edu*  *https://orcid.org/0000-0001-9740-7437* | *Colgate University*  *USA*  *rwhite@colgate.edu*  *https://orcid.org/0000-0001-6021-3268* |
|  |  |  |
| **Steven Gentry** | **Madison Chartier** | **Rachel J. Wittmann** |
| *University of Michigan*  *USA*  *gentrys@umich.edu https://orcid.org/0000-0003-0596-2439* | *Oklahoma State University*  *USA*  *madison.chartier@okstate.edu*  *https://orcid.org/0000-0001-9883-4932* | *University of Utah*  *USA*  *rachel.wittmann@utah.edu*  *https://orcid.org/0000-0002-7342-3907* |

**Abstract – The Metadata Quality Benchmarks group of the Digital Library Federation’s Metadata Working Group launched a survey in 2019 to gather preliminary data about metadata quality assessment and benchmarking practices used in digital libraries. Data gathered included information about the hosting organization; the size, scope, and technical aspects of the digital repositories; and quality assessment priorities and activities.  Survey analysis revealed several trends and correlations, most noticeably in the overall usage frequency of elements, assessment of required  versus optional elements, and prioritization of certain quality characteristics and evaluation methods.  The authors hope conclusions drawn from this data will spur the broader creation and implementation of metadata benchmarks, enhancing the quality and overall impact of metadata in resource access, discovery, and preservation.**

**Keywords – Digital Library Federation, Digital libraries, Digital repositories, Metadata assessment, Survey response data**

**Conference Topics – Community; Resilience**

# Introduction

Metadata quality assessment in digital libraries is a challenging and complex subject.  Successful long-term maintenance of digital resources requires high-quality metadata that documents detailed descriptions, administrative actions, preservation procedures, and other information.  However, not only do opinions differ as to what constitutes metadata assessment, but actual assessment practices may hinge upon available resources, or lack thereof, regardless of the perceived value of quality assessment.

The Digital Library Federation’s (DLF) Assessment Interest Group (AIG) was formed in 2014 to address this need for clearer assessment practices in digital librarianship.  Comprising numerous working groups that cover a broad field of topics, the AIG develops standards, tools, and practices to help institutions implement fundamental assessment practices.  The DLF AIG Metadata Working Group (MWG) specifically targets metadata and strives to “build guidelines, best practices, tools, and workflows around the evaluation and assessment of metadata used by and for digital libraries and repositories” [1].  Since its inception, the MWG has headed several projects to help establish solid metadata practices for digital collections, including a clearinghouse of metadata documentation (e.g., metadata application profiles); a publicly accessible metadata assessment Zotero group [2]; and several metadata quality analysis workshops.  More recently, the MWG has sought to gauge the efficacy of such tools and guidelines in promoting metadata quality by examining how institutions assess their metadata.

The Metadata Quality Benchmarks (MQB) group, a MWG sub-group, formed in 2018 to investigate current assessment practices and to suggest general guidelines for measuring metadata quality. This paper discusses findings from a survey released by the MQB in 2019 that gathered preliminary data about organizations’ metadata quality assessment and benchmarking practices. Institutions were asked to provide information about their organization; the size, scope, and technical aspects of their digital repositories; and their quality assessment priorities and activities.

# Literature Review

Quality assessment is integral to establishing and enforcing good metadata practices for digital library collections.  Reference [3] notes that “the quality of metadata can have significant impact in facilitating access, use, and long-term preservation to digital resources” (p. 2).  However, metadata assessment may be nebulous work because the rubric underlying assessment depends on its specific context (noted by many, e.g., [4] and [5]).  As stated in the DLF AIG Metadata Working Group’s Metadata Assessment Framework and Guidance, “[M]etadata quality is subjective.  How you define metadata quality will be unique to the core functions and mission of your institution or needs” [6] (see also [7]).

While the definition of quality may be unique to individual institutions’ needs and values, the perceived impacts of metadata quality are universally recognized.  Depending on the context, data quality may have serious consequences (e.g., medical data [8]).  Even without such stakes, studies on the efficacy and deficiencies of metadata quality assessment tools, including a 2019 study of metadata creators’ and managers’ perspectives, confirm a shared concern of compromised user accessibility due to poor quality metadata [9].  However, quality assessment may be difficult to implement due to limited staffing, time, and resources, leading to manual evaluation or sampling, even in automated processes (e.g., [10]).

To offer some guidance, the MQB created and released a metadata assessment document based on a framework established by [11], which outlined seven quality aspects: accessibility, accuracy, completeness, conformance to expectations, logical consistency and coherence, provenance, and timeliness.  This framework was updated to accommodate linked data in 2013 [12] and has been referenced by other authors exploring metadata assessment (e.g., [13], [14], [15], and [3]).  For example, reference [16] found that administrative metadata criteria (e.g., provenance) have not been studied as thoroughly as those concerning information retrieval, despite their value to institutional workflows and administrative audit trails for preservation purposes.

Metadata standards have been developed to promote consistency and shareability, including general schemas like Dublin Core (DC), as well as schemas for specific domains (e.g., Darwin Core [17]) or material types (e.g., VRA Core [18]).  Dublin Core, established in 1995 [19], has frequently been employed by digital repositories due to its early creation and wide applicability.  A 2002 study evaluated 100 Open Archives Initiative (OAI)-compliant repositories and analyzed the number of DC elements used in records and the frequency of usage [20].  A later 2014 study determined that DC was the most frequently used schema among 77 international repositories [21].

Documentation explaining how a particular institution or project implements an established schema may have varying names, such as guidelines, standards, practices, or Metadata Application Profiles (MAPs).  These resources not only provide metadata practitioners with rules and directions when creating metadata, they also govern local metadata production and influence metadata quality.  In a recent study of 24 MAPs from academic libraries in the United States, “[a] comparison of elements among the MAPs further revealed insights into the considerations and dilemmas that metadata creators face when attempting to describe disparate and unique materials” (p. 33) [22].  Although MAPs may not explicitly include metadata evaluation practices or procedures, such documentation is often closely connected to quality assessment.  Consequently, one graduate-level library science class now combines assignments for these evaluation components to help library students better understand connections between metadata evaluation and MAPs [23].

In addition to local implementations, a number of initiatives have developed guidelines for sharing metadata, including regional or cooperative projects (e.g., [24]) and aggregations.  The largest aggregation project MAPs include Europeana (for descriptive records in Europe) [25] and the Digital Public Library of America (DPLA), which uses a national network of hubs for cultural heritage materials in the United States [26].  Aggregations also provide options to test large-scale metadata evaluation.  Variations in metadata quality can become more apparent once metadata is in an aggregated environment, housed amongst collections from many institutions.  Some research has focused on specific quality aspects (e.g., a study of completeness in Europeana [27]), specific elements such as *description* [28] or *subject* [29], or general element usage [30] in DPLA.

Major aggregation entities are trying to support quality among contributing organizations, such as the Europeana Publishing Framework [31].  Similarly, DPLA has embraced community-driven approaches to improve metadata quality with the development of task forces, training, and collaborative efforts to develop and review DPLA and network hub guidelines and MAPs [32].  Analysis of metadata aggregations can also highlight quality related to shareability and users’ ability to make sense of local records outside their originating context [33], which further impacts the degree to which users may find relevant materials from different sources or understand information in simplified records [28].

Organizations may want to address quality issues in their digital repositories for a number of reasons, including findability, shareability, and long-term preservation.  Using criteria in [11] as a basis to assess metadata quality, the MQB launched a survey in the summer of 2019 and invited metadata professionals to answer questions regarding their respective institutions’ methods for measuring metadata quality.  Initial results outlining aggregated data were released online as a white paper in 2020 [34].  This paper expands on these previous findings with additional discussion of selected survey results.

# Methods

The MQB collected data through a Qualtrics survey, which the group promoted across various library-domain listservs. The survey was active from May 23-July 10, 2019.  Survey instructions asked that only one metadata professional from each institution provide responses.  Only two questions in the survey were mandatory: 1) consent to take part in the survey, and 2) how many repositories are managed by the responding institution.

Overall, 240 respondents consented to take the survey; however, 89 (37%) did not answer any subsequent questions.  Of the remaining respondents, 107 (45%) fully completed the survey, while 44 (18%) partially completed the survey, resulting in a total or partial completion rate of 63%.

Survey responses were exported as a .csv file and evaluated manually (using spreadsheets) by the researchers.  This analysis compared data across multiple responses to find correlations, which was not previously done.  It did not re-evaluate data that was fully covered in the initial findings. The survey instrument and anonymized raw data are publicly available [35].

# Results

Respondents came from a variety of backgrounds.  Academic librarians were strongly represented (55%), followed by public librarians (9%) and librarians employed by museums, consortia, and aggregation projects (7%).  Two-thirds of the responding institutions managed 1-2 digital repositories.  Respondents were asked to describe their institutions’ repositories, including whether a particular repository serves as an institutional repository (i.e., resources produced by the organization and/or constituent members), a digital collection (i.e., digitized or born-digital cultural heritage materials), or both (see Table 1).  Each repository typically contained 10,000-100,000 records, although sizes ranged from less than 100 items to 10,000,000 or more.

Additionally, the survey asked if the metadata for each repository conforms to a Metadata Application Profile (MAP) and, if so, whether that MAP is an external document (e.g., a consortial MAP to participate in an aggregation, such as DPLA) or a locally-generated MAP.  Among all repositories, local MAPs were most frequently used; however, “no MAP” was the most common response overall (see Table 1).  For 13 repositories, respondents indicated a MAP is used but did not clarify whether the repository uses a consortial or local MAP.

TABLE 1

MAP Use by Repository Type

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Institutional Repository | Digital Collection | Both | Total |
| Local MAP | 5 | 35 | 17 | 57 |
| External MAP | 1 | 5 | 6 | 12 |
| Unknown MAP | 3 | 7 | 3 | 13 |
| No Map | 13 | 28 | 19 | 60 |
| Total | 22 | 75 | 45 | 142 |

The survey asked about each repository’s schema usage as a possible factor when applying standards or sharing data between institutions.  Survey responses showed a marked preference for Dublin Core (DC)-based schemas—including simple DC, qualified DC, and locally-modified or supplemented DC—followed by Metadata Object Description Schema (MODS).  Other responses demonstrated a variety of schema usage, as well as a number of local schemas or combinations of multiple schemas.  Similarly, there were no significant patterns in controlled vocabulary usage, although very few repositories (7 of 105) reported using no controlled vocabularies at all.

## Overall Usage Frequency

The survey’s main goal was to discover whether institutions evaluate metadata and, if so, what methods and procedures they implement (e.g., whether they evaluate every element in use or only select elements).  The data was collected on a repository (rather than institutional) basis, so numbers do not correlate directly to a parent organization.

To establish a baseline, two questions provided a grid of 26 commonly-used metadata elements. The first question asked respondents to indicate for each individual repository which specific elements on the grid are required, recommended, or optional (with no answer meaning “not used”).  The second question asked respondents to indicate whether each of the same elements is evaluated or not evaluated.  Although the survey gave respondents the option of supplying local elements not represented on the grids in each question, these free-text responses were so varied that generalizations were difficult to derive from the data.  Unless otherwise noted, this paper refers only to the 26 elements listed in the grids when discussing the frequency of element availability or evaluation.

Although the initial evaluation [34] reviewed aggregated data (i.e., total responses per question), the rate of responses across the two grids was not the same.  For example, some respondents indicated whether repository elements were evaluated in one grid, but did not specify whether those elements were required, recommended, or optional in the other.  When looking at total responses in each grid question, such discrepancies made the actual number of repositories difficult to determine.  However, after adjusting the data to account for responses in either grid (assuming that “evaluated” elements are available), adjusted totals accounted for 123 individual repositories (see Table 4).  Only the *subject* element is available in all 123 repositories, followed by *creator* and *date* in 122 repositories.

Aside from general responses, the frequency of individual elements being required, recommended, or optional can also be determined based on the total availability (i.e., how often repositories prefer a particular usage for each element).  For example, *subject* is required in 36 repositories, representing 29% of total *subject* usage; *physicalLocation* is also required in 36 repositories, but is available in 23 fewer repositories than *subject*.  So required usage for *physicalLocation* is 36% of total frequency (see Table 2).  Additionally, an overall difference of only 29 repositories exists between the element used most frequently—*subject* (in 123 repositories)—and the element used least frequently—*table of contents* (in 94 repositories).

Individual responses to the grids also provided more information about total element usage and distributions across repositories.  Nearly half of the repositories (58) reported making all 26 possible elements available (see Fig. 1).  The total number of available elements in the remaining 65 repositories ranges from 6 to 25, although a majority of those repositories (57) include at least 15 elements.  Additionally, these elements can be broken down by level of usage.  For example, most repositories tend to require either relatively few elements—17 repositories require 4 elements and 11 repositories require 3 elements—or a substantive number of elements—16 repositories require 7 elements and 14 repositories require 8 elements.

Chart, histogram

Description automatically generated

Figure 1 Number of Grid Elements Reported Per Repository

Similarly, most repositories tend to recommend or make optional 5 to 10 elements.  No repositories require or recommend all elements, and only 1 repository does not require any elements and makes all 26 elements available optionally

## Metadata Evaluation

In terms of evaluation, several correlations were identified when the data was broken down by individual responses (see Table 4).  Required elements are more frequently evaluated than not evaluated.  However, when elements are only recommended, these elements are evaluated exactly half the time (13 of 26 elements), including one case—*identifier*—that is evaluated and not evaluated in an equal number of repositories (6). Optional elements are not evaluated more often than they are evaluated, except for *creator* (evaluated in 7 repositories and not evaluated in 4), as well as 2 elements that have equal numbers of repositories that do or do not evaluate them: *contributor* (23) and *title* (1).  *Creator* is also the only element that is most often evaluated regardless of usage.

Only 12 repositories lacked any kind of metadata evaluation.  One respondent clarified: all metadata records are evaluated prior to ingest; presumably, no evaluation occurs post-ingest.

TABLE 2

Total Element Frequency with Percentage of Element Use, Ordered by Frequency of Requirement

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Element Usage Frequencies | | Percentage of Repositories in Which the Element is: | | | | |
| Element Name | Required | Total | Required | Recommended | Optional | Unspecified | Not Used |
| *Title* | 115 | 121 | 0.95 | 0.02 | 0.02 | 0.00 | 0.02 |
| *Identifier* | 93 | 118 | 0.79 | 0.11 | 0.08 | 0.03 | 0.04 |
| *Rights* | 75 | 119 | 0.63 | 0.22 | 0.13 | 0.02 | 0.03 |
| *Type* | 75 | 116 | 0.65 | 0.18 | 0.13 | 0.04 | 0.06 |
| *Collection Title* | 74 | 113 | 0.65 | 0.12 | 0.19 | 0.04 | 0.09 |
| *Format* | 55 | 117 | 0.47 | 0.22 | 0.26 | 0.05 | 0.05 |
| *Date* | 54 | 122 | 0.44 | 0.44 | 0.10 | 0.02 | 0.01 |
| *Creator* | 46 | 122 | 0.38 | 0.52 | 0.09 | 0.01 | 0.01 |
| *Subject* | 36 | 123 | 0.29 | 0.46 | 0.23 | 0.02 | 0.00 |
| *physicalLocation* | 36 | 100 | 0.36 | 0.18 | 0.39 | 0.07 | 0.23 |
| *Description* | 33 | 118 | 0.28 | 0.40 | 0.30 | 0.03 | 0.04 |
| *Language* | 30 | 114 | 0.26 | 0.32 | 0.35 | 0.06 | 0.08 |
| *Extent* | 27 | 115 | 0.23 | 0.36 | 0.37 | 0.04 | 0.07 |
| *Publisher* | 26 | 113 | 0.23 | 0.29 | 0.46 | 0.02 | 0.09 |
| *Genre* | 26 | 107 | 0.24 | 0.39 | 0.29 | 0.07 | 0.15 |
| *isPartof* | 17 | 103 | 0.17 | 0.28 | 0.51 | 0.04 | 0.19 |
| *Contributor* | 16 | 121 | 0.13 | 0.43 | 0.41 | 0.02 | 0.02 |
| *Source* | 16 | 107 | 0.15 | 0.32 | 0.50 | 0.04 | 0.15 |
| *Abstract* | 15 | 110 | 0.14 | 0.35 | 0.50 | 0.02 | 0.12 |
| *Spatial* | 11 | 111 | 0.10 | 0.37 | 0.46 | 0.07 | 0.11 |
| *Coverage* | 9 | 115 | 0.08 | 0.39 | 0.49 | 0.04 | 0.07 |
| *Digitization Specs* | 8 | 99 | 0.08 | 0.19 | 0.58 | 0.15 | 0.24 |
| *Transcription* | 8 | 98 | 0.08 | 0.28 | 0.59 | 0.05 | 0.26 |
| *Relation* | 4 | 103 | 0.04 | 0.23 | 0.68 | 0.05 | 0.19 |
| *Table of Contents* | 4 | 94 | 0.04 | 0.10 | 0.76 | 0.11 | 0.31 |
| *Alternative Title* | 1 | 115 | 0.01 | 0.14 | 0.81 | 0.04 | 0.07 |

For 3 repositories, respondents did not answer either grid question, but listed local recommended or optional elements, none of which are evaluated.  This means, of 126 total repositories, 88% engage in evaluation of at least 1 element.

Although organizations are doing at least some quality control or assessment, they most commonly evaluate only a few of their total elements.  Survey respondents for 5 repositories indicated they evaluate only 1 element (2 repositories) or 2 elements (3 repositories); most respondents (representing 52% of reported repositories) are evaluating 5 to 14 elements (64 repositories).  However, these numbers are not reflective of “how much” available metadata is evaluated.  Responses for the 123 repositories using grid elements indicated that, in 66 repositories—just over 50% of total repositories—less than half of the available elements are evaluated (see Fig. 2).  Comparatively, the evaluation rate for 18 repositories is roughly half of their available elements (marked with a shaded bar in Fig. 2), and only 39 repositories reported evaluation of at least 60% of available elements.  These numbers change when broken down by usage, since required elements are more likely to be evaluated than non-required elements. For example, 8 repositories evaluate all of their elements, and 57 repositories evaluate at least half of their total elements.  However, 43 repositories evaluate all required elements, and 88 repositories (i.e., 71.5%) evaluate at least half of their required elements.  Only 3 repositories conduct evaluation but assess none of the required elements.

One interesting finding is that repositories using the fewest grid elements tend to have the lowest percentage of evaluated elements.  Original hypotheses anticipated a high assessment rate, as these repositories have less metadata to review as compared to repositories using more elements.  However, there were 6 repositories making fewer than 13 of the 26 possible elements available; among those, half indicate that no elements are evaluated (see Table 3).  Evaluation for repositories using 14-26 varied without any clear trends, including only 8 repositories that evaluate all available elements.

Chart, line chart

Description automatically generated

Figure 2 Percentage of Total (circles) and Required (diamonds) Elements Evaluated, by Repository

Significant overlap exists among the top ten respective elements that are most often available, required, or evaluated in repositories (see Fig. 3).  Elements that are in the top ten for any of these facets account for half of the 26 grid elements.  Seven DC elements—*creator*, *date*, *identifier*, *rights*, *subject*, *title*, and *type*—are in the top ten for all three areas.  This degree of overlap suggests a number of informal shared expectations regarding preferred descriptive elements.  There are also 3 elements that are frequently not evaluated, even though they are most often available or required: *description*, *format*, and *physicalLocation*.  Although *description* may be more difficult to evaluate—as it is generally a complex, free-text element—why *format* or *physicalLocation* are not evaluated is less clear. *Format* is often managed by a controlled vocabulary, and *physicalLocation* may be important for managing materials, tracking digitization projects, and scheduling preservation measures.  This overlap provides a potential starting point for identifying cross-organization expectations.

TABLE 3

Evaluation Rate for Repositories Using Fewer than 13 Grid Elements

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Number of Elements | | | | Elements Evaluated | |
| Repo-  sitory | Requ-  ired | Recom-  mended | Opt-  ional | Total | Num-  ber | Per-  cent |
| A | 1 | 1 | 4 | 6 | 1 | 0.167 |
| B | 3 | 4 | 0 | 7 | 0 | 0 |
| C | 1 | 0 | 8 | 9 | 0 | 0 |
| D | 7 | 3 | 0 | 10 | 6 | 0.600 |
| E | 1 | 4 | 6 | 11 | 0 | 0 |
| F | 7 | 4 | 1 | 12 | 5 | 0.417 |

Diagram, venn diagram

Description automatically generated

Figure 3 Venn Diagram of Overlap for the Most-Often Required, Evaluated, and Available Elements.

## Characteristics of Metadata Quality

The survey’s final questions asked about evaluation priorities and assessment methods for the seven quality aspects established by [11]: accuracy, accessibility, completeness, conformance to expectations, logical consistency and coherence, provenance, and timeliness.  One question asked respondents to express the importance of these aspects at their institutions by ranking each from most (1) to least (7) important (see Table 5).  Respondents often agreed on which rankings to give to individual aspects; however, among 87 individual responses, there were 73 unique combinations of rankings.

TABLE 4

Element Evaluation by Use (Required, Recommended, Optional, Unspecified), Ordered by Frequency of Requirement

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | Required | | | Recommended | | | Optional | | | Unspecified | |
| Element Name | Eval-  uated | Not Eval-  uated | No Answer | Eval-  uated | Not Eval-  uated | No Answer | Eval-  uated | Not Eval-  uated | No Answer | Eval-  uated | Not Eval-  uated |
| *Title* | 83 | 30 | 2 | 2 | 1 | 0 | 1 | 1 | 1 | 0 | 0 |
| *Identifier* | 57 | 30 | 6 | 6 | 6 | 1 | 4 | 5 | 0 | 1 | 2 |
| *Rights* | 58 | 13 | 4 | 16 | 9 | 1 | 3 | 10 | 3 | 0 | 2 |
| *Type* | 51 | 15 | 9 | 9 | 10 | 2 | 3 | 12 | 0 | 2 | 3 |
| *Collection Title* | 52 | 18 | 4 | 9 | 2 | 2 | 3 | 17 | 2 | 2 | 2 |
| *Format* | 35 | 14 | 6 | 7 | 14 | 5 | 8 | 22 | 0 | 2 | 4 |
| *Date* | 46 | 7 | 1 | 33 | 16 | 5 | 5 | 7 | 0 | 2 | 0 |
| *Creator* | 38 | 6 | 2 | 41 | 19 | 4 | 7 | 4 | 0 | 1 | 0 |
| *physicalLocation* | 20 | 14 | 2 | 8 | 5 | 5 | 8 | 27 | 4 | 2 | 5 |
| *Subject* | 32 | 1 | 3 | 42 | 11 | 3 | 10 | 16 | 2 | 1 | 1 |
| *Description* | 22 | 9 | 2 | 16 | 29 | 2 | 10 | 22 | 3 | 1 | 2 |
| *Language* | 19 | 10 | 1 | 14 | 18 | 5 | 8 | 25 | 7 | 2 | 5 |
| *Extent* | 13 | 12 | 2 | 15 | 20 | 6 | 6 | 31 | 5 | 3 | 2 |
| *Genre* | 19 | 3 | 4 | 27 | 11 | 4 | 7 | 20 | 4 | 4 | 4 |
| *Publisher* | 17 | 6 | 3 | 14 | 17 | 2 | 10 | 38 | 4 | 2 | 0 |
| *isPartOf* | 13 | 3 | 1 | 10 | 15 | 4 | 7 | 36 | 10 | 0 | 4 |
| *Contributor* | 15 | 0 | 1 | 32 | 17 | 3 | 23 | 23 | 4 | 2 | 1 |
| *Source* | 13 | 0 | 3 | 13 | 14 | 7 | 9 | 36 | 8 | 1 | 3 |
| *Abstract* | 8 | 7 | 0 | 19 | 18 | 1 | 12 | 30 | 13 | 1 | 1 |
| *Spatial* | 6 | 3 | 2 | 28 | 9 | 4 | 10 | 37 | 4 | 4 | 4 |
| *Coverage* | 6 | 1 | 2 | 28 | 13 | 4 | 12 | 39 | 5 | 4 | 1 |
| *Digitization Specs* | 4 | 3 | 1 | 7 | 9 | 3 | 4 | 39 | 14 | 4 | 11 |
| *Transcription* | 4 | 2 | 2 | 8 | 15 | 4 | 11 | 36 | 11 | 0 | 5 |
| *Relation* | 3 | 0 | 1 | 10 | 12 | 2 | 12 | 46 | 12 | 1 | 4 |
| *Table of Contents* | 3 | 1 | 0 | 3 | 6 | 0 | 6 | 47 | 18 | 2 | 8 |
| *Alternative Title* | 1 | 0 | 0 | 8 | 5 | 3 | 20 | 57 | 16 | 2 | 3 |

TABLE 5

Number and Percentage of Responses Ranking Quality Aspects, Grouped by High, Medium, and Low

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Ranked  1 / 2 | | Ranked  3 / 4 / 5 | | Ranked  6 / 7 | |
|  | # | % | # | % | # | % |
| Accuracy | 57 | 0.655 | 29 | 0.333 | 1 | 0.011 |
| Completeness | 52 | 0.598 | 29 | 0.333 | 6 | 0.069 |
| Consistency | 29 | 0.333 | 51 | 0.586 | 7 | 0.080 |
| Conformance to Expectations | 19 | 0.218 | 60 | 0.690 | 8 | 0.092 |
| Accessibility | 11 | 0.126 | 58 | 0.667 | 18 | 0.207 |
| Provenance | 4 | 0.046 | 21 | 0.241 | 62 | 0.713 |
| Timeliness | 2 | 0.023 | 13 | 0.149 | 72 | 0.828 |

This reflects that organizations tend to have different priorities and sometimes significant divergences in terms of which aspects are most important.

There were 3 quality aspects not ranked last by any institution: accuracy, completeness, and consistency.  Additionally, those qualities were ranked 6th by only 7 respondents (for consistency), 6 respondents (for completeness), and 1 respondent (for accuracy).  On the other end of the spectrum, neither provenance nor timeliness were ranked 1st by any organization.  As these elements are largely representative of administrative metadata, crucial only to the repository and not the end user, the revelation that these aspects are not regularly evaluated is perhaps not surprising.  It is also consistent with the findings from the literature.  Two of the aspects ranked most important (consistency and accuracy) were also listed in subsequent questions as the aspects that organizations would most like to evaluate but currently cannot.

# Discussion

This survey gathered initial information regarding institutions’ digital repositories and metadata evaluation practices.  Despite the broad spectrum of repository types and metadata implementations, review of respondent-level data revealed insightful trends and correlations.

According to respondents, around 42% of repositories do not rely on any form of Metadata Application Profile (MAP).  In the white paper documenting initial findings of the survey [34], the hypothesis was that the relatively low usage of MAPs (or the high number of repositories not using MAPs) may be due to terminology.

Some institutions may have guidance specifications or documents that effectively serve the same purpose as MAPs, even if they are not referred to as such.   This consideration is relevant since a number of respondents said they use standards and documentation as a method of quality control.   More investigation may be helpful in this area, both to determine the types of existing institutional documentation and to better understand how these are used as reference materials for validation or quality control.

Almost all repositories (93%) make at least 15 elements available in their metadata records.  Of the 123 repositories that documented element usage in the grid questions, almost half (47%) reported using all 26 listed elements.  This response rate suggests a fairly robust expectation for describing materials and content.  Most repositories also tend to require at least 3 elements, with the bulk of repositories requiring either 3-4 elements (28 repositories) or 7-8 elements (30 repositories).  If required elements can be understood to represent a “minimal-level” record, this average number of required elements is relatively sparse, but is still more than the 2 elements required by DPLA (which may serve as a baseline for institutions considering external aggregation).  There was no significant correlation between the repository type and the number of available or required elements, so variations seem to be based on local preferences rather than the expected level of description for cultural heritage versus institutional materials.

As more than half of the repositories use a DC-based schema (basic or qualified), the top ten most-frequently-available elements (required, recommended, or optional) are all part of the DC set: *subject*, *creator*, *date*, *contributor*, *title*, *rights*, *description*, *identifier*, *format*, and *type*.  The other 5 DC elements have varying levels of availability, from *coverage* (ranking 12th) to *relation*, which is 22nd out of 26 elements when ordered by frequency.  In terms of total frequency, all but one of the elements in the grid are available in at least 80% of the repositories; the least-frequently available element (*table of contents*) is still available in 76% of the repositories.  These frequencies suggest a high level of overlap regarding element usage, even among repositories of different types or those using different schemas.  This may have implications for generalizing metadata quality benchmarks or recommendations for usage and value formatting.  In fact, overlapping element usage may be more useful than relying on schemas, considering the wide array of schema applications, including a variety of combinations and qualifications of the DC elements.

The survey data also establishes that most institutions evaluate at least some of the metadata in their digital library systems, although the extent (and some respondents’ opinions on this topic) varied dramatically.  In roughly 52% of repositories, less than half of the available elements are evaluated.  Repository elements are more likely to be evaluated if they are required than if they are recommended or optional, although most repositories require relatively few elements.  In fact, 17 of the 26 elements are required in only 30% or fewer of the repositories that make them available.  These required elements may reflect a priority for findability and interface functionality (e.g., a *title* value that displays in search results for users).

Respondents also ranked completeness as one of the most important quality aspects at their institutions.  This ranking could be related to a tendency to check required elements, as the use of required elements is sometimes considered a reasonable metric for a minimally-complete record, or may simply be required by the repository system to save the metadata.  Limited personnel, unfamiliarity with evaluation tools, and other resources may also be factors, particularly given that a number of institutions reported a reliance on manual checks as their primary or only method of evaluation.

Although the data showed significant differences in quality aspect rankings in individual responses, some definite priorities can be generalized across institutions.  For example, a marked preference for accuracy, completeness, and consistency is evident, as is a relative lack of interest in timeliness or provenance.  Whether lower rankings represent a gap or need is unclear.   Perhaps a quality aspect is not prioritized for evaluation because it is too difficult to assess or define internally.  Maybe it requires less intentional review (e.g., provenance or machine-readability may be automatically generated, validated, or recorded and require no intervention).  Certain quality aspects may also have less direct effect on user needs, which often assume a high priority in element selection (e.g., completeness could affect browsing functions across a collection or system versus timeliness, which may be more important on an individual record level).  These issues surrounding priorities and resource allocation tend to be complicated but may benefit from generalized benchmarks or guidance.

# Conclusion

This survey and its analyzed results provide insights into the broad range of current metadata implementation and evaluation practices.  In the future, the authors hope to expand upon this research through efforts to identify and propose generalized metadata benchmarks, so as to provide a common standard to which all institutions may refer.  One individual shared: “[O]ur metadata quality analysis … [is] very ad hoc, irregular, and targeted to particular problems we experience… [W]e don't really make sure we're adhering to very many external guidelines” (p. 23) [34].  A common set of benchmarks may enable consistency when it comes to metadata quality, which will further enhance information institutions’ ability to preserve digital repository metadata and to make this metadata, along with its affiliated resources, more shareable and findable for a variety of users.  Based on the survey results, the MQB is currently drafting additional resources to support further benchmarking work, with plans to follow-up with respondent organizations and the digital library community.

As standards ensure successful, consistent, and shareable metadata upon creation, similar standards for quality assessment, established by a dedicated community of professionals, will prove influential in enhancing materials’ usability, accessibility, and long-term preservation on digital platforms.  Metadata evaluation practices are in need of standardization efforts similar to the community-based efforts to streamline institutional metadata creation through MAPs.  The authors hope the results presented in this survey analysis will likewise inspire fellow metadata practitioners to come together to review and develop quality assessment procedures and make them an active part of their metadata workflows.  Additionally, the authors hope efforts like the survey encourage more widespread community initiatives to identify and establish potential benchmarks that may prove useful to institutions implementing fundamental to advanced metadata.

# REFERENCES

1. DLF Metadata Assessment Working Group. (2021) About. The DLF Metadata Assessment Working Group website. [Online]. Available: https://dlfmetadataassessment.github.io/about
2. DLF Metadata Assessment Working Group. (2022) Zotero Group: Metadata Assessment. [Online]. Available: https://www.zotero.org/groups/488224/metadata\_assessment
3. O. L. Zavalina, P. Kizhakkethil, D. G. Alemneh, M. E. Phillips, and H. Tarver, “Building a Framework of Metadata Change to Support Knowledge Management,” Journal of Information and Knowledge Management, 1st ed., vol. 14, pp. 16, 2015. Available: https://digital.library.unt.edu/ark:/67531/metadc505014/
4. K. Snow, “Defining, Assessing, and Rethinking Quality Cataloging,” Cataloging & Classification Quarterly, vol. 55, no. 7-8, pp. 438-455, 2017. Available: https://doi.org/10.1080/01639374.2017.1350774
5. C. A. Reeves and D. A. Bednar, “Defining Quality: Alternatives and Implications,” The Academy of Management Review, vol. 19, no. 3, pp. 419-445, 1994. Available: https://doi.org/10.2307/258934
6. DLF Metadata Assessment Working Group. (2017) Metadata Assessment Framework and Guidance. [Online]. Available: https://dlfmetadataassessment.github.io/framework
7. G. Gueguen, C. Harper, and C. Stanton. “Perspectives on Data and Quality (Slides).” DPLAFest 2016, Washington, DC, USA, (April 15, 2016). Available: tinyurl.com/2p9v97ud
8. L. M. Schriml et al., “COVID-19 Pandemic Reveals the Peril of Ignoring Metadata Standards,” Sci Data, vol. 7, no. 188, 2020. Available: https://doi.org/10.1038/s41597-020-0524-5
9. N. T. Fox, H. Tarver, and M. E. Phillips (2019) Identifying Gaps in Tools and Interfaces for Assessing Metadata Quality [White paper]. [Online]. Available: https://digital.library.unt.edu/ark:/67531/metadc1453742/
10. M. Goovaerts and D. Leinders, “Metadata Quality Evaluation of a Repository Based on a Sample Technique,” in Metadata and Semantics Research. MTSR 2012, J. M. Dodero, M. Palomo-Duarte, and P. Karampiperis, Eds. Communications in Computer and Information Science, Vol. 343, Berlin: Springer, 2012. Available: https://doi.org/10.1007/978-3-642-35233-1\_19
11. T. Bruce and D. Hillmann, “The Continuum of Metadata Quality: Expressing, Exploiting,” in Metadata in Practice. D. Hillmann and E. Westbrook, Eds. Chicago: ALA Editions, 2004, pp. 238-256. Available: https://ecommons.cornell.edu/handle/1813/7895
12. T. Bruce and D. Hillmann (2013, Jan.) Metadata Quality in a Linked Data Context. [Online]. Available: https://blog.law.cornell.edu/voxpop/2013/01/24/metadata-quality-in-a-linked-data-context/
13. X. Ochoa and E. Duval, “Automatic Evaluation of Metadata Quality in Digital Repositories,” International Journal on Digital Libraries, Vol. 10, pp. 67-91, 2009. Available: https://link.springer.com/article/10.1007/s00799-009-0054-4
14. J-R. Park. “Metadata Quality in Digital Repositories: a Survey of the Current State of the Art,” Cataloging & Classification Quarterly, vol. 47, no. 3-4, pp. 213-228, 2009. Available: https://doi.org/10.1080/01639370902737240
15. J. Park, Y. Tosaka, S. Maszaros, and C. Lu, “From Metadata Creation to Metadata Quality Control: Continuing Education Needs Among Cataloging and Metadata Professionals,” Journal of Education for Library & Information Science, vol. 51, no. 3, pp. 158-176, 2010.
16. H. Ulrich, et al., “Understanding the Nature of Metadata: Systematic Review,” Journal of Medical Internet Research, vol. 24, no. 1, 2022. Available: https://doi.org/10.2196/25440
17. (2012) Dublin Core™ Metadata Element Set, Version 1.1: Reference Description website. [Online]. Available: https://www.dublincore.org/specifications/dublin-core/dces/
18. (2015) Darwin Core website. [Online]. Available: https://dwc.tdwg.org/
19. (2022) VRA Core Official Website. [Online]. Available: https://www.loc.gov/standards/vracore/
20. J. Ward, A Quantitative Analysis of Dublin Core Metadata Element Set (DCMES) Usage in Data Providers Registered with the Open Archives Initiative (OAI) [Master’s Thesis], Chapel Hill, NC: University of North Carolina at Chapel Hill, 2002. Available: https://cdr.lib.unc.edu/concern/masters\_papers/tq57nv59w
21. M. Curado Malta and A. A. Baptista, “A Panoramic View on Metadata Application Profiles of the Last Decade,” International Journal of Metadata, Semantics and Ontologies, vol. 9, no. 1, 2014. Available: https://doi.org/10.1504/IJMSO.2014.059124
22. A. M. Green, “Metadata Application Profiles in U. S. Academic Libraries: A Document Analysis,” Journal of Library Metadata, vol. 22, 2022. Available: https://doi.org/10.1080/19386389.2022.2030172
23. O. Zavalina, “Integrated Learning of Metadata Quality Evaluation and Metadata Application Profile Development in a Graduate Metadata Course,” Proceedings of the International Conference on Dublin Core and Metadata Applications, 2017. Available: https://dcpapers.dublincore.org/pubs/article/view/3856.html
24. C. Cronin, “Metadata Provision and Standards Development at the Collaborative Digitization Program (CDP): A History,” First Monday, Vol. 13, No. 5, 2008. Available: https://firstmonday.org/ojs/index.php/fm/article/view/2085/1957
25. (2017) Europeana Data Model - Mapping Guidelines v2.4. [Online]. Available: https://pro.europeana.eu/page/edm-documentation
26. (2017) Metadata Application Profile. Digital Public Library of America website. [Online]. Available: https://pro.dp.la/hubs/metadata-application-profile
27. P. Király and M. Büchler, “Measuring Completeness as Metadata Quality Metric in Europeana,” 2018 IEEE International Conference on Big Data (Big Data), 2018, pp. 2711-2720. Available: https://doi.org/10.1109/BigData.2018.8622487
28. H. Tarver, O. L. Zavalina, and M. E. Phillips, “An Exploratory Study of the Description Field in the Digital Public Library of America,” Proceedings of the International Conference on Dublin Core and Metadata Applications, 2016, pp. 34-44. Available: https://digital.library.unt.edu/ark:/67531/metadc910346/
29. M. E. Phillips and H. Tarver, “Investigating the Use of Metadata Record Graphs to Analyze Subject Headings in the Digital Public Library of America,” The Electronic Library, vol. 39, no. 3, 2021. Available: https://doi.org/10.1108/EL-11-2020-0317
30. C. A. Harper, “Metadata Analytics, Visualization, and Optimization: Experiments in statistical analysis of the Digital Public Library of America (DPLA),” Code4Lib, no. 33, 2016. Available: https://journal.code4lib.org/articles/11752
31. B. Daley, H. Scholz, and V. Charles, “Developing a Metadata Standard for Digital Culture: the Story of the Europeana Publishing Framework.” [Online]. Available: https://pro.europeana.eu/post/developing-a-metadata-standard-for-digital-culture-the-story-of-the-europeana-publishing-framework
32. G. Guegen, “Metadata Quality at Scale: Metadata Quality Control at the Digital Public Library of America,” Journal of Digital Media Management, vol. 7, no. 2, 2019, pp. 115-126. Available: https://www.ingentaconnect.com/content/hsp/jdmm/2019/00000007/00000002/art00003
33. S. L. Shreeves, et al., “Is Quality Metadata Shareable Metadata? The Implications of Local Metadata Practices for Federated Collections,” in Proceedings of the Twelfth National Conference of the Association of College and Research Libraries, H. A. Thompson, Ed. Chicago: Association of College and Research Libraries, 2005, pp. 223-237. Available: https://www.ideals.illinois.edu/handle/2142/145
34. S. Gentry, et al. (2020, May) Survey of Benchmarks in Metadata Quality: Initial Findings [White Paper]. [Online]. Available: http://dlfmetadataassessment.github.io/assets/2020-dlf-mawg-mqb-white-paper.pdf
35. DLF Metadata Assessment Working Group. (2020) Projects: Metadata Quality Benchmarks. The DLF Metadata Assessment Working Group website. [Online]. Available: https://dlfmetadataassessment.github.io/benchmarks