Toward Releasing the Metadata Bottleneck

A Baseline Evaluation of Contributor-supplied Metadata

By Amanda J. Wilson

Metadata creation is one of the most expensive components of digital projects. Organizational expertise (the correct structure, syntax, and use of metadata elements) and subject expertise (the appropriate semantic description of a resource's content for users) are both needed to create a high-quality metadata record. Resource creators are frequently considered good metadata generators. Contributors or subject enthusiasts in a discipline are another population that may be good candidates for metadata creation. In this study, the quality of contributor-supplied metadata is evaluated. Metadata records submitted through a Web form are compared to the final published version of the record. Structural and semantic errors are noted throughout the records evaluated. Overall, semantic quality was good, reflecting subject expertise. The appearance and type of structural errors suggests that improvements in the interface can reduce contributors' need to have organizational expertise to create high-quality metadata records.

In the life cycle of digital projects such as repositories, databases, registries, and collections, one of the most expensive initial components is metadata creation for each digital resource.[1] Metadata creation requires both organizational and subject expertise to describe an object and its context for use. In this paper, organization expertise refers to the ability to apply the correct structure, syntax, and use of metadata elements, while subject expertise refers to the ability to generate appropriate semantic (or meaningful) description of a resource's content for users. High-quality metadata utilizing both expertise types is an integral part of effective searching, retrieval, use, and preservation of digital resources. Metadata professionals tend to be proficient both in organizational and subject expertise; however, they are too few to provide sufficient metadata in a timely, efficient manner for the abundance of digital resources, creating a bottleneck in a digital-project's workflow.

Rather than sacrifice the quality of metadata in digital projects, recent research to alleviate the bottleneck has explored several methods to reduce the need for either organizational or subject expertise in metadata creation, including use of creators as metadata suppliers and automatic metadata-generation processes.[2] Currently, creators seldom provide sufficient metadata for their digital resources, and scalable automatic metadata-generation techniques that produce acceptable metadata are still in development.[3] This paper explores another group outside of metadata professionals and resource creators for human metadata generation: subject enthusiasts, or those with a significant background in any discipline.[4]

Amanda Wilson (wilson.1596@gmail.com) is a doctoral student at the George Washington University, Washington, D.C.

The author would like to thank Barbara MacKenzie of RILM, Tschera Connell, Barbara Dunham, Amy McCrory, and Beth Russell for reading drafts and providing feedback in the development of this paper.

One primary benefit of enlisting contributors to provide metadata in digital projects is that the pool of potential metadata creators immediately increases. The potential for collaborative metadata generation, or "the joint production of Web resource metadata," also increases.[5] While resource authors are the primary focus of current research in human metadata generation, the purpose of this paper is to serve as a baseline study of the contributor-supplied metadata using RILM Abstracts of Music Literature (RILM), an international database of scholarly works about music.[6] RILM uses authors and subject enthusiasts (volunteers from the international music community) to create basic metadata records that are then reviewed and enhanced by metadata professionals before the final resources are published in the database. Secondarily, this paper evaluates the quality of contributor-supplied metadata when specific content guidelines for elements are available.

Literature Review

Metadata-quality evaluation is closely related to the extensive research conducted on all aspects of data quality. One of Orr's six laws of data quality states that "[the] laws of data quality apply equally to data and metadata."[7] Though the entire corpus of data-quality research will not be discussed here, some studies will be highlighted as they relate to this project. Data-quality research explores two basic tenets of data. Rothenberg identified them as correctness and appropriateness of data.[8] In metadata creation, correctness and appropriateness are comparable to organizational expertise and subject expertise, respectively. Rothenberg then elaborated on the "process of evaluating and assessing data . . . referred to as 'Verification, Validation, and Certification'" (VVC) as a method to assess data correctness.[9] In this process, data are reviewed for adherence to applicable specifications or requirements, assurance that values reflect what is represented, and authoritative endorsement at some level. Focusing on improving the appropriateness of data systems, Orr stated that to improve a system's quality (i.e., a digital-project database), only data that is relevant and used should be collected.[10] If data are not relevant, resources are wasted on its creation and maintenance. If data are not used, then they eventually become obsolete and therefore not appropriate for their original purpose.

Using financial information systems, Cappiello, Francalanci, and Pernici developed a framework to assess the appropriateness of data based on users' varying purposes and needs.[11] Their research was conducted in an effort to better define the relevance and usefulness of data, specifically taking into account "the degree to which data satisfy users' needs."[12] Research in metadata quality for digital projects and collections has examined both correctness and appropriateness of metadata using similar indicators such as timeliness and usefulness.

Recent studies of metadata quality have resulted from increased sophistication in the creation, use, and implementation of metadata. Quality assurance of metadata is essential not only for successful digital projects, but for the successful operation of metadata aggregators.[13] The following section examines the literature covering three aspects of metadata quality: its definition, indicators or evaluation metrics, and types of errors commonly found in metadata records.

Metadata Quality

What is high-quality metadata? A standard definition of metadata quality is still developing. Metadata researchers do agree, however, that metadata quality involves its fitness for a specific purpose or use.[14] In their article on improving metadata quality, Hillmann, Dushay, and Phipps of the National Science Digital Library, an aggregated database of science information, asserted that "the utility of metadata can best be evaluated in the context of services provided to end-users," focusing on appropriateness.[15] The authors proposed evaluation of metadata appropriateness at the element level, rather than record level, to best tailor information for users.

Further, Bruce and Hillman defined a framework in the form of a continuum of metadata quality for automated evaluation of metadata at an aggregator level.[16] Their framework was developed in the context of large-scale aggregated collections, which require both domain-independent indicators, such as timeliness and provenance, and three-tiered quality levels for automated quality-assurance processes. Metadata quality indicators enumerated include completeness, accuracy, provenance, conformance to expectations, logical consistency coherence, timeliness, and accessibility. These indicators address both the correctness and appropriateness of metadata. The first-tier quality level gauges the entirety of a record, a portion of organizational expertise. This tier focuses on the presence or completeness of a record. The second tier includes additional indicators, such as use of controlled vocabularies and community-defined elements, which require organizational expertise and some subject expertise. This tier begins to address appropriateness and encompasses Hillman et al.'s recommendation to evaluate quality in each element.[17] Finally, the third tier adds quality indicators such as application profiles requiring a full complement of organizational and subject-expertise skill sets.

Shreeves et al. evaluated the quality of harvested metadata at an aggregator level in order to determine how metadata quality at a local level affects searching federated collections.[18] The framework used in their study contained both intrinsic information quality (or correctness) indicators and relational information quality (or appropriateness)

indicators that mapped to each indicator in Bruce and Hillman's framework.[19] These indicators were used to evaluate records based on completeness, consistency, and ambiguity. They found that the quality of metadata varied between and among collections of metadata records.

With an aim to create a networked government information system, Moen, Stewart, and McClure completed a thorough analysis of metadata-quality literature and identified twenty-three distinct indicators to evaluate the implementation of a specific element set, the Government Information Locator Service (GILS).[20] The study focused primarily on compliance (or correctness) of metadata records created at forty-two federal agencies using GILS. The authors used the twenty-three indicators to rate compliance quality in four categories covering correctness and appropriateness: accuracy, completeness, profile, and serviceability. In both Moen, Stewart, and McClure's and Shreeves et al.'s aggregator examples, metadata records could have been created by authors, contributors, or library and information science professionals.[21]

Greenberg et al. evaluated metadata quality at the point of metadata creation.[22] Specifically, they examined the quality of eleven author-supplied metadata records using the National Institute of Environmental Health Sciences Dublin Core schema to determine baseline metadata quality. Their research argued that authors can create acceptable metadata records, as ranked by professional catalogers, using a Web template and text guidelines. Experienced professional catalogers determined the "high quality" standard for the author-generated records. They determined the acceptability of the author-supplied metadata and rated each of the records on the basis of intelligibility and correctness of the metadata.

Correctness and appropriateness are recurring themes in the metadata quality literature discussed. Specific aspects of both overarching themes appearing multiple times are completeness, accuracy, and intelligibility. These aspects apply to both the record level and element level quality. Thus metadata quality is determined by the compilation of elements used to fully describe a resource, the scheme, and the values within each element. In all but Greenberg et al., who explicitly evaluated resource creators, metadata records could have been created by authors, contributors, information professionals, or some other entity.[23] The possible effects of heterogeneity of metadata creators was not specifically identified as the cause of lower quality. Strict adherence to and availability of institutional metadata guidelines is a common recommendation to improve metadata quality.

The types of errors found in metadata quality studies are similar at both the aggregation and local levels. Within the e-print community, Barton, Currier, and Hey identified some common problems in metadata records and how they affect use and retrieval of digital resources.[24] The problem areas and elements include spelling and abbreviations, author and other contributors (e.g., editors and translators), title, subject, and date. These problems may also affect utility of metadata at an aggregated level; however, aggregators have another set of quality concerns. Dushay and Hillman also identified categories of errors found in metadata records at an aggregator level: missing data (e.g., empty elements), incorrect data (e.g., author data in language fields), confusing data (e.g., strings of names, or data culled from another source—a bibliography, for example—and not revised once in metadata record), and insufficient data.[25]

Some of these errors are obviously human errors; others, such as incorrect data, could be attributed to the usage and implementation of local metadata-element sets. For example, Hanrath noted in his usability study of a Dublin Core template that metadata creators with subject expertise desired additional specific elements to accurately describe the resources.[26] Though he was not examining the quality of records created, the absence of a logical element for a piece of information deemed relevant resulted in metadata creators supplying incorrect data in another field, thus affecting Hanrath's usability study. At both the aggregator and local level, the frequently suggested solution for many errors in metadata records is publicly available documentation developed and used at the local level that outlines use and content of metadata elements in a digital project.[27]

Research Objectives

This paper aims to evaluate the use of contributors as viable high-quality metadata creators for the RILM database. To evaluate the current baseline quality of metadata records contributed to the database, the following objectives were sought:

- gauge the completeness of each record;
- determine the types of errors (typographical/grammatical or semantic), if any, in each record; and
- identify the appearance and type of "value-added" or additional metadata in each record.

These three objectives address both the organizational and subject expertise needed for high-quality metadata by looking at quality both at the record level and within each element.

Research Methods

Rothenberg's Verification, Validation, and Certification (VVC) model for evaluating data quality guides the methodological design to evaluate each of the objectives for this study.[28] RILM's submission process produces a final

metadata record that satisfies all three criteria and ensures high-quality metadata. Currently, one way contributors can supply metadata is via online submission forms. The contributor metadata forms the basis for a short record in RILM's database. Only metadata submitted via the submission form was evaluated in this study. Based on an examination of the actual articles, trained RILM personnel, including staff, librarians, and musicologists, subsequently edit and enhance the metadata; these are the verification and validation steps.[29] RILM requires more fields for each record than appear on the submission form. These fields, including subject headings, do not appear on the submission form and are supplied by RILM personnel. The final, enhanced version of the metadata record is then published in the RILM database, replacing the initial short record; this is the certification step.

Records Evaluated

RILM provides records for various types of scholarly works related to music. Initially, the author asked RILM for a selection of raw metadata records submitted to RILM from 1998, 2000, and 2004. The submission data supplied contained records for a majority of the record types represented in RILM, including articles in proceedings, monographs, and sound recordings. Previous metadata quality studies have evaluated between 11 and 140 records.[30] In order to have enough records to enable identification of patterns in the data, the target number of records for this project is 100. The author counted the occurrence of each record type (i.e., monographs, articles appearing in periodicals) to identify one type of record for evaluation. Submissions for the articles in a periodical record type appeared most frequently; therefore, record submissions for articles appearing in a periodical were evaluated in this study. RILM is an international database containing records in many languages. Because English is the only language with which the author is familiar enough to accurately assess semantic and structural errors in metadata records, the focus of the semantic abstract evaluation in this study is English-language submissions.

Procedure

The raw submission data (the contributor-supplied metadata record) is the unit of evaluation in this study. High-quality metadata is appropriately fit for a purpose. A raw contributor-supplied metadata record's purpose for RILM is to form a temporary record in the database and the basis for a final RILM record. Thus, the high-quality metadata standard in this paper is the final RILM-enhanced record from each contributor-supplied record that appears in the RILM database as displayed via OCLC's FirstSearch interface.

Evaluating contributor-supplied metadata involved a three-step process. First, the final RILM record corresponding to each contributor-supplied record was located in the RILM database. If two records were found for the same raw submission record because the short record still remained in the database, the RILM record containing subject headings was selected for evaluation. Second, the RILM records were printed to facilitate comparative evaluation. Finally, the raw submission and final RILM-record metadata were compared element by element based on a key. Table 1 shows the metric used to evaluate each record through identification of different types of errors and value-added elements in each contributor record. For the purposes of this paper, an error occurred each time RILM made a change or addition to an element. Part A of the table contains the codes and usage notes used to mark each instance of an error type for all elements except for the "Abstract" field. Each error type in table 1 is mapped to a quality factor. Part B contains the codes and usage notes used to evaluate quality in the "Abstract" field. Analysis of the raw submission is compared to the final record based on three factors: syntax, structure, and semantic content.

In this paper, syntax refers to the overall completeness of the metadata record submitted. Each published RILM record consists of complete bibliographic information, an abstract, subject headings, and other relevant information (e.g., reference to related record in the database) if applicable. Most of these elements have corresponding fields in the submission form. Completeness is the inclusion of every required element for each type of record added to the database as defined in the RILM submitter guidelines. The appendix lists the fields required for articles appearing in a periodical. This factor is an indicator of organizational expertise.

Structure deals with the format of each element (format issues within a field). For example, values in the ISSN field should be four numbers followed by a hyphen followed by four numbers. In the submission form, text input boxes, dropdown menus, or radio buttons are provided for every field. Fields with text input boxes include "Author name," "Periodical title," "ISSN," and "Page numbers." These fields are structurally distinct from each other due to the different types of metadata that belong in each. Because the metadata may appear in different forms on the original source, variance in metadata submission may occur. Four types of structural errors were noted: non-authoritative forms, capitalization, punctuation, and other spelling and typographical errors. Structure is another indicator of organizational expertise.

Semantic content refers to the provision of content in each field. In other words, if content is supplied in a field, is that content correct, and is the content in the appropriate field? Five types of semantic errors were noted: RILM editing a portion of an element (e.g., changing "music of

Table 1. Metric used to evaluate raw submission records

Part A: Record analysis

Code	Quality factor	Notes
E	Enhanced (provided by RILM)	added to a field that is not required
Ed	Edited (edited by RILM)	portion of field is wrong
I	Incorrect data	(100% wrong data)
R	Required data missing	
U	Usage—misplaced data	can be a portion of a field
AC	Authoritative name/form needed	
G	Spelling/typographical error	
GC	Capitalization	
GP	Punctuation	
A	Added value (content)	regardless of whether used in final record
XU	Correct, unused	correct information for a field, but not used; (counts as Added value)
X	Provided correct information	

Part B: Abstract content analysis

Code	Quality factor	Score (0 or 1)
NM	First and last names	not edited (0) versus edited (1)
TI	Complete titles	not edited (0) versus edited (1)
TI(NUM)	Complete titles with index or opus numbers	not edited (0) versus edited (1)
EN	Complete entity names	not edited (0) versus edited (1)
DEF	Definition of terms not in Grove or MGG*	not edited (0) versus edited (1)
GEOG	Complete and correct place names	not edited (0) versus edited (1)
BIB	Basic bibliographic information for articles or books referred to in abstract	not edited (0) versus edited (1)
STY	Style of abstract is declarative versus descriptive	declarative (0) versus descriptive (1)

**Grove Dictionary of Music and Musicians* and *Die Musik in Geschichte und Gegenwart* (MGG) are major music encyclopedias.

the church" to "church music"), RILM adding values for a non-required element (e.g., supplying ISSN), contributors leaving out required element values (e.g., contributors not supplying an author), contributors supplying an incorrect value (e.g., wrong or incomplete ISSN, wrong date), and correct data supplied in an inappropriate field (e.g., contributors putting volume and issue information in the "Volume" field). In other words, is the correct information in the appropriate field or misplaced in a different field? For example, each "Author name" field is paired with an "Author function" drop-down menu, which contains values such as preface author, editor, and translator in addition to author. Though RILM considers persons serving as translators to be "authors," contributors may not have grasped RILM's broad concept of authorship. On the submission form, separate fields are for information that may be concatenated on the original source (e.g., volume and issue number). Other information may simply be incorrect; that is, data for a particular element in the submission record ("Year" value is 1994) is completely different from that appearing in the final published record ("Year" value is 1999).

The primary field for elaborated content is the abstract field. RILM has outlined specific guidelines for writing abstracts.[31] Essential content includes key concepts and names. Table 1, Part B: Abstract content analysis, is derived from the RILM abstracting guidelines document. Eight abstract elements are defined. In general, RILM requires complete names (persons or organizations, works, places) and titles. RILM prescribes the writing style, requesting concise, declarative abstracts. Because key concepts, themes, and names are appropriate and relevant to RILM's target audience, this factor is an indicator of subject expertise.

The appearance and type of "value-added" or additional information in the submission is also noted. A "value-added" element in contributor-supplied metadata could be keywords, notice of related RILM records, or other details the contributor may deem important, such as a Web address for an online journal. Value-added elements are also those elements for which contributors supplied values, but are not present in the final RILM record. Examples include non-English language abstracts in addition to English-language abstracts, translated titles, or notice of English-language summaries within an article. The presence of value-added information is an additional indicator of subject expertise.

Results

A total of 104 records for articles appearing in a periodical were evaluated for this study. Of those, 55 records were contributed by the person who abstracted the article. The remaining 49 records were submitted by someone other than the person who abstracted or authored the article—that is, a third party.

The contributor metadata contained 384 errors, or 3.69 errors per submission. The frequency distribution of structural and semantic errors in presented in table 2. Each submission contained between 18 and 26 elements, depending on how many authors were identified for an article. Of the total number of elements and taking all author name/function pairs as one element, 4 elements are required for each submission (see appendix for specific required elements). Table 2 shows the frequency distribution of the structural and semantic errors occurring within each element. The frequency rates show that structural errors occurred more often than semantic errors: 183 semantic errors (1.76 errors per record) and 201 structural errors (1.93 per record).

Table 3 gives the numeric frequency of errors by element for elements with the highest and lowest error rates. Elements fell into three distinct categories based on the number of errors per element: high-frequency error rate (at least 30 errors), low-frequency error rate (10-29 errors), and infrequent-error rate (less than 10 errors).

Table 4 shows the frequency of errors by specific error type. "Edited" and "Authoritative name/Form needed" (Authority Control, or AC) errors are the most frequent errors in the sample at 110 and 114 instances, respectively. Edited elements had some portion of the values changed by RILM, and the AC element values were formatted to RILM specifications. Values placed in the wrong field, "Usage—misplaced data," and values with spelling or typographical errors other than punctuation and capitalization were the least frequent errors among records evaluated. Specific results of the metadata-quality analysis based on syntax, structure, and semantic content follow.

Syntax

Syntax is the measure of completeness of the contributor-supplied metadata record. Table 5 details each element's use frequency. Of the 104 contribu-

Table 2. Frequency distribution of structural and semantic errors

Elements in order of structural error frequency rate

	Structural	Semantic	TOTAL
Periodical title*	75	2	77
Title/subtitle*	52	6	58
Serial date	35	2	37
Year of publication*	10	1	11
Volume	8	4	12
First author*	8	1	9
Issue no.	6	9	15
ISSN	6	2	8
English abstract	1	45	46
Translated title	0	53	53
Special features**	0	37	37
Second author*	0	5	5
Second-author function**	0	5	5
Page numbers	0	4	4
Name of abstractor	0	4	4
Third author*	0	1	1
Third-author function**	0	1	1
Language(s) of summaries**	0	1	1
Total:	**201**	**183**	**384**

Elements in order of semantic error frequency rate

	Structural	Semantic	TOTAL
Translated title	0	53	53
English abstract	1	45	46
Special features**	0	37	37
Issue no.	6	9	15
Title/subtitle*	52	6	58
Second author*	0	5	5
Second-author function**	0	5	5
Volume	8	4	12
Page numbers	0	4	4
Name of abstractor	0	4	4
Periodical title*	75	2	77
ISSN	6	2	8
Serial date	35	2	37
First author*	8	1	9
Third author*	0	1	1
Third-author function**	0	1	1
Year of publication*	10	1	11
Language(s) of summaries**	0	1	1
Total:	**201**	**183**	**384**

*required fields
**value-list/controlled vocabulary in radio-button or drop-down-box input type

Table 3. Elements with high- and low-frequency errors

Elements with high-frequency errors

	TOTAL	Semantic	Structural
Periodical title*	77	2	75
Title/subtitle*	58	6	52
Translated title	53	53	0
English abstract	46	45	1
Serial date	37	2	35
Special features**	37	37	0
Total:	**308**	**145**	**163**

Elements with low-frequency errors

	TOTAL	Semantic	Structural
Issue no.	15	9	6
Volume	12	4	8
Year of publication*	11	1	10
Total:	**38**	**14**	**24**

*required fields
**value-list/controlled vocabulary in radio-button or dropdown box input type

Table 4. Specific error frequencies

Frequency of errors (by type)			
Structural	AC	Authoritative name/form needed	114
	GC	Capitalization error	53
	GP	Punctuation error	32
	G	Spelling/typographical error	2
Semantic	Ed	Edited	110
	E	Enhanced	51
	R	Required data missing	10
	I	Incorrect data	7
	U	Usage—misplaced data	5
		TOTAL Errors	**384**
Added-value (by type)			
	A	Additional notes	23
	XU	Supplied, but not used	19
		TOTAL Added value	**42**

tor-supplied records, 100 contained values for all required elements. Required elements include Title/subtitle, Author name(s) and function(s), Periodical title, and Publication year. The 4 incomplete records were missing required additional author entries. These additional author entries were listed in RILM as translators and interviewees. Not only did a majority of records contain the basic required elements, but, taking all author name/function pairs as one, 16 of the 18 elements (88 percent) were used in more than half of the contributor metadata records evaluated (that is, occurred in 52 or more records).

Structure

The structure of each element is the format of the values supplied. Structural correctness for each record was determined based on the final RILM record. Table 4 shows the 4 types of structural errors noted in the evaluation. These 4 errors primarily assess spelling, punctuation, and other typographical errors, as well as AC formats for fields such as "Author name" and "Periodical title." Structural errors occurred in 6 elements: "Author name," Periodical title," ISSN," "Serial date," "Issue no.," and "Year of publication."

Table 4 gives the number of each type of structural error. AC is the most common error in the contributor metadata with 106 instances occurring throughout the records evaluated. Of the 77 structural errors in the "Periodical title" field, 45 of them were AC. RILM added and removed subtitles and other words to and from the element for corrections. Fifty-three capitalization errors occurred in two fields, Title/subtitle (35) and Periodical title (18). Within Title/subtitle, capitalization errors occurred more frequently in the subtitle (25 of 35 times or 71 percent). Thirty-two punctuation errors were documented. Twenty-seven punctuation errors occurred in both the Title and Periodical title fields.

Semantic Content

Semantic content of each metadata element refers to the correctness of values in a field. Table 4 also lists semantic errors. The most common semantic corrections performed by RILM were editing some portion of an element's value (Edited) or supplying values for non-required elements (Enhanced). Edited and Enhanced errors account for 88.48 percent of the semantic errors in contributor-supplied metadata. Edited errors occurred in the widest spread of elements—9. "English abstract" field was edited in 43 of 50 supplied values. Some of these errors were semantic, such as contributors supplying translator names in the abstract. Other errors, which will be discussed later, are the result of abstract guideline infringement. The element with the next most-frequent Edited error rate is "Special features," a controlled-value list. RILM removed terms from nine records, exchanged terms (e.g., "diagrams" to "tables," "sound recordings" to "sound files," "charts" to "illustrations") in 12 records, and added new terms to 18 records. RILM supplied values that contributors left blank for 51 non-required elements across the records evaluated.

Table 5. Contributor metadata supplied in each submission

Arranged by Element		Arranged by Frequency	
Title/subtitle*	104	Title/subtitle*	104
Translated title	68	First author*	104
First author*	104	First-author function**	104
First-author function**	104	Periodical title*	104
Second author*	6	Year of publication*	104
Second-author function**	6	Language(s) of article**	104
Third author*	1	Country of publication**	96
Third-author function**	1	Page numbers	93
Fourth author*	0	Special features**	93
Fourth-author function**	0	Name of abstractor	89
Fifth author*	0	Issue no.	88
Fifth-author function**	0	ISSN	80
Periodical title*	104	Translated title	68
ISSN	80	Serial date	65
Volume	57	Volume	57
Issue no.	88	Non-English abstract	55
Serial date	65	English abstract	50
Year of publication*	104	Language(s) of summaries published within the item**	26
Page numbers	93	Second author*	6
Country of publication**	96	Second-author function**	6
Special features**	93	Third author*	1
Language(s) of article**	104	Third-author function**	1
Language(s) of summaries published within the item**	26	Fourth author*	0
English abstract	50	Fourth-author function**	0
Name of abstractor	89	Fifth author*	0
Non-English abstract	55	Fifth-author function**	0

*required fields
**value-list/controlled vocabulary in radio-button or drop-down-box input type

Three-quarters of these values (39) were supplied in the "Translated title" field.

The remaining 3 error types were recorded 10 or fewer times. The required data missing errors (R) all surfaced in 5 author name/function pairs. RILM supplied name and function values for an interviewee, preface author, and translators. Because of the 5 missing author entries, a total of 10 R errors resulted. Incorrect errors were just that—contributor-supplied values were 100 percent wrong. In most cases the wrong year or ISSN was supplied. In 5 instances, correct values were supplied in incorrect fields (Usage—misplaced data). Four of the 5 misplaced values are related to the author elements. For example, the subject of one article was supplied as an author and translators of others were listed in abstracts. The fifth Usage—misplaced data error was unrelated to author elements—the theme for a journal issue was supplied in the "Periodical title" field.

Abstract Element Guidelines

The "English abstract" element is the only field with separate guidelines specifically outlining rules of content supplied. Guidelines for abstract elements only apply if an element is present in the summary. Analysis of "English abstract" focused solely on the explicit guidelines, which operationalize important elements of the field, and not how qualitatively good an abstract is according to RILM. Table 6 shows the number of errors in abstracts by both authors and contributors. Of the 50 supplied abstracts, 33 were contributor-authored, while 17 were author abstracts that

Table 6. Abstract content analysis data (N=50)

	First/last names	Complete titles	Titles with index/opus numbers	Entity names	Definition of terms	Complete/ correct place names	Bibliographic information	Style
Contributor abstracts								
Edited records	5	4	1	0	n/a	1	1	1
Total instances	19	14	6	3	n/a	3	3	33
Author abstracts								
Edited records	2	2	1	0	0	1	n/a	4
Total instances	11	4	1	1	1	7	n/a	17
Totals								
Edited records	**7**	**6**	**2**	**0**	**0**	**2**	**1**	**5**
Total instances	**30**	**18**	**7**	**4**	**1**	**10**	**3**	**50**

contributors provided. Overall, RILM edited 13 records for guideline errors, 39.39 percent of 33 elements, and 10 of 17 (58.82 percent) author-supplied records for a total of 23 errors for the 50 abstracts evaluated.

The most frequent abstract element supplied in the evaluated records was name, a total of 34 instances including personal and corporate names. In 7 of 30 records (23.33 percent), personal names were edited; RILM mostly supplied the fuller form of a name for the final record. All of the 4 records with corporate names remained unedited. Figure 1 shows the percentage of errors occurring in each element. No more than one third of the instances of a supplied element were edited in the records evaluated. Bibliographic information was edited most often in terms of percentage. Taking both Title elements together, Title was edited second most often at 28 percent of the time.

The last abstract element, "Style," is slightly different from the others, in that it is a subjective evaluation. RILM asks for contributors to supply abstracts in a declarative writing style. RILM edited 5 of 50 abstracts' styles to convert them from descriptive to declarative style. One example of a change in style occurs when a contributor abstract summarizes the article from a third-person point of view, while RILM's version "speaks" the summary or extract from the article. The other 3 stylistic corrections occurred in abstracts written by the paper's author. In each case, RILM transformed an author's descriptive abstract into a concise, declarative summary. For example, the following author excerpt, "Yet I have found an abundance of information about this active salon culture in diaries, letters, and journal articles. Warsaw had over forty significant salons, and direct evidence of Chopin's musical presence can be established in most of them" became "The existence of over 40 salons in Warsaw, and Chopin's participation in majority of them are documented" in RILM.

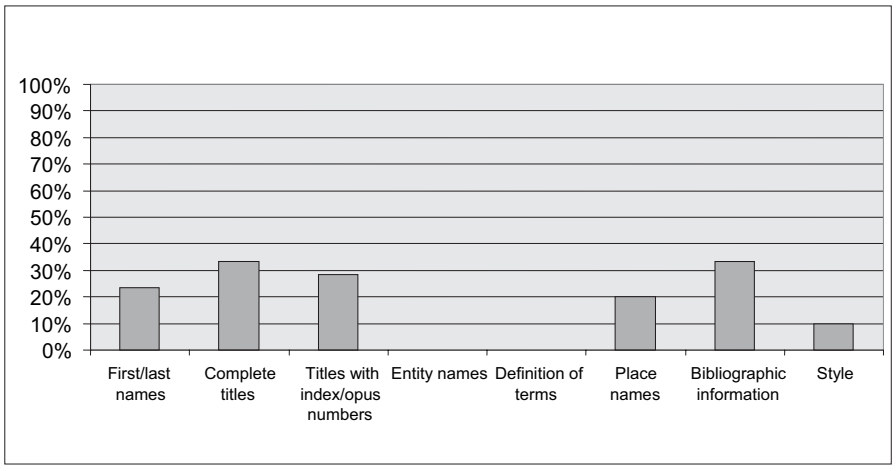

Figure 1. Percentage of supplied abstract elements edited by RILM

Value-added Elements

Value-added elements in the evaluated records are those elements containing (1) correct information for a

field but that may not be used in the final RILM record (Correct, unused) and (2) information in the "Additional information" element. RILM edited out Correct, unused values because they deemed the values unnecessary for an accurate description of the article. Several elements contained Correct, unused values, including "Translated title," "Serial date," "Special features," "Languages of summaries published within an item," and "Non-English abstract," for a total of 19 instances. "English" was provided 6 times in the "Languages of summaries" element—the most Correct, unused values provided for each element. They appeared 2, 4, 4, and 3 in the remaining elements, respectively.

Contributors supplied additional information for 23 records. The type of information fell into two categories: user value (i.e., information providing additional context for the article such as keywords and other descriptive elements) and management (i.e., information helpful for the management of records in the database). Supplied content for the user-value elements include time period, theme of volume, ISBN of volume, and keywords. The management-content values included notice of related record in RILM, notification of errata in another volume, and URL of online journal. Some information appeared in some form as a part of the final records (e.g., some keyword concepts became subject headings), while others did not.

Discussion

The objective of this study was to evaluate the baseline quality of contributor-supplied metadata to the RILM Abstracts of Music Literature database by comparing the final high-quality RILM record to the raw metadata submitted by contributors. The metadata was evaluated from several perspectives: Were all required elements supplied in each record? Were the values supplied in the correct format for its element? Were the values supplied semantically correct? Also, what value-added elements were present throughout the records evaluated?

Evaluating completeness, format, and content of each metadata record addressed the organizational and subject expertise currently reflected in the contributor-supplied metadata records. Though a few records were missing author name/function pairs, contributors completed all of the required elements. If completeness is scoped in terms of information supplied and not elements completed, then all but 2 records are complete. In other words, contributors in this study provided more than just the required elements; they also provided additional descriptive information, though the data may not appear in the most appropriate field according to RILM's standards. In a few cases, contributors supplied additional information that could assist in the management of records in the database that may have become a part of the final high-quality RILM record. Syntactically, the quality of contributor metadata records in this study was good when compared to the final RILM record.

Organizational expertise was most closely linked with the structural errors measured in this evaluation. Elements with the strictest formatting requirements were those with most of the structural errors. AC errors were by far the most common. In cases such as "Serial date" and "Periodical title," the correct format is at times different from how the information appears in the journal. Conversely, "Author name," another highly controlled field, contained very few errors. While contributors indicated awareness of the importance of full author names for searching and retrieval, the data suggest that the same may not be true of other controlled bibliographic fields. As with other elements, a dropdown menu for journal titles and author names could reduce the number of structural errors in the contributor metadata and time spent editing. Some of the capitalization and punctuation errors would also be addressed if, rather than entering values, contributors could select from a list. Structurally, the quality of contributor-supplied records was both acceptable—because contributors supplied correct values for elements with one or two standard formats such as "ISSN" and "Author name"—and not acceptable—due to the varying formats for elements such as "Periodical title" that caused more errors.

Subject expertise was related to the semantic content errors evaluated. Semantic errors occurred across a majority of fields. While many values were edited by RILM, few were completely wrong. Generally, RILM made minor edits to words and numerals (e.g., an extra integer in the "Issue no." field or replacing "music of the church" with "church music" in the "Translated title" field). Other times RILM moved information from one field to another (e.g., RILM moved translators from the "English abstract" to the author name/function fields). The "Special features" and "English abstract" fields contained the most Edited error instances, yet both are unique cases. RILM consistently exchanged terms 12 times in the "Special features" field (e.g., changing "sound recordings" to "sound files"). The new terms are not available in the contributor interface and appear to be updated terminology. In the case of "English abstract," each instance of rephrasing or substitution counted as a semantic error. However, when semantic content of abstracts was evaluated based on published guidelines, errors occurred at a rate of less than 1 in every 2 abstracts. The error rate percentage was higher in abstracts written by authors than in those written by someone other than the author of the article. The data suggest that contributors followed abstract guidelines for those abstracts they authored, but made few or no edits to ensure author abstracts conformed to guidelines, as RILM edited three-fifths of the abstracts written by the authors.

Contributors used both the abstract and additional information elements to supply other descriptive information beyond the scope of available elements. Contributors' values in the "Additional information" field support Hanrath's findings that contributors desired additional fields for their metadata records.[32] However, translators and keywords supplied in the "English abstract" field suggest that a greater understanding of certain fields in the context of the database may be needed. For instance, contributors supplied no other "author type" (e.g., interviewee, translator) than the traditional author function. The quality of metadata was slightly better than that of structural metadata. In many cases, the content was supplied in the record, but may have been in an incorrect field.

High-quality metadata is defined to contain both organizational and subject expertise. Contributors' subject expertise was reflected in the results as the content for many of the fields was supplied, though perhaps not in the appropriate RILM field. Semantically, contributor-supplied metadata evaluated in this study was high quality. To generate true high-quality metadata, however, a method to improve the structural aspect of contributor metadata is essential.

Study Limitations

Limitations to this study stem primarily from the records evaluated. First, the records are only of the "article in a periodical" type. The quality of other record types in RILM may be different and worth exploring. Second, the records evaluated do not represent a random sample of "articles in a periodical" records from the RILM database. The records are from sets of submissions from 1998, 2000, and 2004. Third, some contributors submitted multiple records at one time, resulting in many contributors having successive submissions from the same issue of a journal title. Errors made in one record tended to be repeated throughout a round of submissions. Successive submissions also limited the pool of contributors evaluated. Finally, the abstract-content analysis focused solely on English abstracts.

In the context of the metadata quality literature discussed, some aspects of the evaluation methods were not addressed or duplicated in this paper. The timeliness/currency quality factor was not examined and quality was evaluated only at one point of evaluation—a Web submission form. Metadata quality may differ with other record-creation instruments.

Implications for Further Research

In addition to ongoing development and refinement of methods to evaluate metadata quality, subsequent research should consider approaches to improving the human-generated metadata and improving the systems for creating metadata. Further exploration investigating the thought process of contributors when creating metadata may inform the design of systems. Questions such as how contributors choose the next best alternative for information and what additional information contributors need to better describe an item could help diminish error rates and increase quality. Additionally, how can instructions and further guidelines improve the quality of contributor-supplied metadata, if at all?

In terms of systems, the Web interface for metadata creation may have had an effect on the quality, particularly with regard to punctuation and capitalization. How exactly does the interface affect the quality of metadata? What aspects of the interface work well? Can new automatic elements be introduced to "pre-enhance" records? Also, when designing metadata creation systems, how can the interface best be configured to help identify corrections and streamline the review/enhancement process? Improving the system or interface should allow contributors to produce high-quality metadata records without having a great deal of organizational expertise.

Conclusion

The purpose of this paper was to serve as a baseline study of the contributor-supplied metadata using the RILM Abstracts of Music Literature. Secondarily, this paper evaluated the quality of metadata when given specific content guidelines. The baseline quality of contributor metadata from the records evaluated was semantically good, yet opportunities for structural improvement exist. The analysis of the "English-abstract" field shows that guidelines can greatly improve the semantic quality of supplied values by further defining and clarifying a field's parameters. This study demonstrates that contributors can supply high-quality basic metadata that is useful to metadata professionals in creating records of publishable quality. RILM's model of collaborative metadata generation holds promise for the greater metadata community. Because they can create good quality metadata, contributors could be a viable source for metadata generation in the effort to reduce the metadata bottleneck. The onus is on the metadata community to build systems and interfaces that harvest contributor semantic content, while leveraging a contributor's discipline knowledge. Efficient, usable systems and interfaces allowing contributors to supply subject expertise while simultaneously guiding correct structural entry of metadata should not only benefit the contributor by increasing the ease of and reducing the time needed for record creation, but also the global community as higher-quality metadata records are available earlier—a positive outcome of collaborative metadata generation.

References and Notes

1. Helen Shenton, "Real Time, Deep Time, Life Time: Spanning Digital and Traditional Collections Life Cycles," *Archiving 2005 Final Program and Proceedings* (Washington, D.C.: Society for Imaging Science and Technology, 2005), 206–14.
2. Jane Greenberg et al., "Author-generated Dublin Core Metadata for Web Resources: A Baseline Study in an Organization," *Journal of Digital Information* 2, no. 2 (2001), http://jodi.ecs.soton.ac.uk/Articles/v02/i02/Greenberg (accessed Aug. 15 2005,); Charles F. Thomas and Linda S. Griffin, "Who Will Create Metadata for the Internet?" *First Monday* 3, no. 12 (1998), www.firstmonday.dk/issues/issue3_12/ thomas/index.html (accessed Aug. 8, 2005); Jane Greenberg, Kristina Spurgin, and Abe Crystal, *Final Report for the AMeGA Project* (Submitted to the Library of Congress Feb. 17, 2005), www.loc.gov/catdir/bibcontrol/lc_amega_final_report.pdf (accessed Aug. 8, 2005).
3. Thomas and Griffin, "Who Will Create Metadata?"; Greenberg et al., "Final Report."
4. Thomas and Griffin, "Who Will Create Metadata?"; Jane Greenberg, "Metadata Generation: Processes, People and Tools," *Bulletin of the American Society for Information Science and Technology* 29, no. 2 (2003): 16–19.
5. Jane Greenberg and W. Davenport Robertson, "Semantic Web Construction: An Inquiry of Authors' Views on Collaborative Metadata Generation," in *DC-2002: Metadata for e-Communities: Supporting Diversity and Convergence. Proceedings of the International Conference on Dublin Core and Metadata for e-Communities* (Florence, Italy: Firenze Univ. Pr., 2002), 46.
6. RILM, "Welcome to the RILM Web Site," www.rilm.org (accessed Aug. 15, 2005).
7. Ken Orr, "Data Quality and Systems Theory," *Communications of the ACM* 41, no. 2 (1998): 68.
8. Jeff Rothenberg, "Metadata to Support Data Quality and Longevity," in *Proceedings of the First IEEE Metadata Conference, Silver Spring, MD, Apr. 16–18, 1996*, www.computer.org/conferences/meta96/rothenberg_Paper/ieee.data-quality.html (accessed Mar. 28, 2005).
9. Ibid., (section 3).
10. Ken Orr, "Data Quality."
11. Cinzia Cappiello, Chiara Francalanci, and Barbara Pernici, "Data Quality Assessment from the User's Perspective," in *Proceedings of the 2004 International Workshop on Information Quality in Information Systems,* ed. Felix Naumann and Monica Scannapieco (New York: ACM Pr., 2004), 68–73.
12. Ibid., 68.
13. Jane Barton, Sarah Currier, and Jessie M. N. Hey, "Building Quality Assurance into Metadata Creation: An Analysis based on the Learning Objects and e-Prints Communities of Practice," in *Proceedings of the 2003 Dublin Core Conference: Supporting Communities of Discourse and Practice—Metadata Research and Applications* (Syracuse, N.Y.: Information Inst. of Syracuse, 2003), 1–10; Marieke Guy, Andy Powell, and Michael Day, "Improving the Quality of Metadata in Eprint Archives," *Ariadne* 38 (2004), www.ariadne.ac.uk/issue38/guy (accessed Aug. 8, 2005).
14. Sarah L. Shreeves et al., "Is 'Quality' Metadata 'Shareable' Metadata? The Implications of Local Metadata Practice on Federated Collections," in *Proceedings of the Twelfth National Conference of the Association of College and Research Libraries, Apr. 7–10 2005,* ed. H. A. Thompson (Chicago: ACRL, 2005), 223–37; Guy, Powell, and Day, "Improving the Quality of Metadata."
15. Diane Hillmann, Naomi Dushay, and Jon Phipps, "Improving Metadata Quality: Augmentation and Recombination," *DC-2004: International Conference on Dublin Core and Metadata Applications Oct. 11–14, 2004* (Shanghai, China: Shanghai Scientific & Technological Literature Publishing House, 2004), 7.
16. Thomas R. Bruce and Diane I. Hillmann, "The Continuum of Metadata Quality: Defining, Expressing, Exploiting," in *Metadata in Practice,* eds. Diane I. Hillmann and Elaine L. Westbrooks (Chicago: ALA, 2004), 238–56.
17. Hillman, Dushay, and Phipps., "Improving Metadata Quality."
18. Shreeves et al., "Is 'Quality' Metadata 'Shareable' Metadata?"
19. Bruce and Hillmann, "Continuum of Metadata Quality."
20. William E. Moen, Erin L. Stewart, and Charles R. McClure, "The Role of Content Analysis in Evaluating Metadata for the U.S. Government Information Locator Service (GILS): Results from an Exploratory Study" (1997), www.unt.edu/wmoen/publications/GILSMDContentAnalysis.htm (accessed Aug. 8, 2005).
21. Moen, Stewart, and McClure, "Role of Content Analysis;" Shreeves et al., "Is 'Quality' Metadata 'Shareable' Metadata?"
22. Greenberg et al., "Author-generated Dublin Core Metadata."
23. Ibid.
24. Barton et al., "Building Quality Assurance into Metadata Creation."
25. Naomi Dushay and Diane I. Hillmann, "Analyzing Metadata for Effective Use and Re-use," in *Proceedings of the 2003 Dublin Core Conference: Supporting Communities of Discourse and Practice—Metadata Research and Applications* (Syracuse, N.Y.: Information Inst. of Syracuse, 2003), 1–10.
26. R. Scott Hanrath, "A Usability Study of a Tool for Contributor-supplied Metadata Creation: the Use of Metadata Element Definitions and Examples in Online Help" (master's thesis, University of North Carolina at Chapel Hill, 2002).
27. Holly M. Harmes, *Development of an Input Form to Capture Author-Generated Metadata for a Botanical Image Collection* (master's thesis, University of North Carolina at Chapel Hill, 2001); Moen et al., "Role of Content Analysis;" Greenberg et al., "Author-generated Dublin Core Metadata;" Guy et al., "Improving the Quality of Metadata"; Shreeves et al., "Is 'Quality' Metadata 'Shareable' Metadata?"
28. Rothenberg, "Metadata to Support Data Quality."
29. For more detailed information on RILM Web-submission process, see RILM, "Mission," www.rilm.org/about.html (accessed Aug. 15, 2005); and RILM, "The Life of a Web-submitted RILM Abstract" (Feb. 16, 2005), www.rilm.org/abstproc.html (accessed Aug. 15, 2005).

30. Greenberg et al., "Author-generated Dublin Core"; Harmes, "Development of an Input Form"; Hanrath, "A Usability Study"; Moen, Stewart, and McClure, "Role of Content Analysis"; Shreeves et al., "Is 'Quality' Metadata 'Shareable'?"
31. RILM, "How to Write a RILM Abstract," www.rilm.org/abstinfo.html (accessed Sep. 19, 2005).
32. Hanrath, "A Usability Study," 18.

Appendix. Submission Form Fields for "Articles Appearing in a Periodical"

Title/Subtitle*
Translated Title
First Author*
First Author function**
Second Author*
Second Author function**
Third Author*
Third Author function**
Fourth Author*
Fourth Author function**
Fifth Author*
Fifth Author function**
Periodical Title*
ISSN
Volume
Issue No
Serial Date
Year of Publication*
Page Numbers
Country of Publication**
Special Features**
Language(s) of article**
Language(s) of summaries published within the item**
English Abstract
Name of Abstractor
Non-English Abstract
Additional Notes
Keyer

**required fields*
***value-list/controlled vocabulary in radio button or dropdown box input type*

Statement of Ownership, Management, and Circulation

Library Resources & Technical Services, Publication No. 311-960, is published quarterly by the Association for Library Collections & Technical Services, American Library Association, 50 E. Huron St., Chicago (Cook), Illinois 60611-2795. The editor is Peggy Johnson, Associate University Librarian, University of Minnesota, 499 Wilson Library, 309 19th Ave. South, Minneapolis, MN 55455. Annual subscription price, $75.00. Printed in U.S.A. with periodicals-class postage paid at Chicago, Illinois, and at additional mailing offices. As a nonprofit organization authorized to mail at special rates (DMM Section 424.12 only), the purpose, function, and nonprofit status of this organization and the exempt status for federal income tax purposes have not changed during the preceding twelve months.

(*Average* figures denote the average number of copies printed each issue during the preceding twelve months; *actual* figures denote actual number of copies of single issue published nearest to filing date: July 2006 issue.) Total number of copies printed: *average,* 6,379; *actual,* 6,485. Sales through dealers, carriers, street vendors and counter sales: none. Mail subscription: *average,* 5,149; *actual,* 5,033. Free distribution: *average,* 50; *actual,* 62. Total distribution: *average,* 5,661; *actual,* 5,544. Office use, leftover, unaccounted, spoiled after printing: *average,* 718; *actual,* 941. Total: *average,* 6,379; *actual,* 6,485. Percentage paid: *average,* 99.12; *actual,* 98.88.

Statement of Ownership, Management and Circulation (PS Form 3526, October 1999) for 2005/2006 filed with the United States Post Office Postmaster in Chicago, October 1, 2006.