Facet Analysis of Archival Metadata Standards to Support Appropriate Selection, Combination and Use of Metadata Schemas

Jae-Eun Baek
University of Tsukuba, Japan
jaeeunb@slis.tsukuba.ac.jp

Shigeo Sugimoto
University of Tsukuba, Japan
sugimoto@slis.tsukuba.ac.jp

Abstract

Metadata is one of the keys for digital archiving and preservation. This is well recognized as an important issue in our networked information society. There are several standards for archival and preservation metadata, e.g. ISAD(G), EAD, AGRkMS, PREMIS, and OAIS. This leads to selection and interoperability issues for metadata standards in the design of metadata schemas for particular archival systems. This paper identifies features of metadata standards to appropriately select, combine and use them in the resource lifecycle.

We present a feature analysis of the metadata standards by identifying the primary resource lifecycle stage(s) where a standard would be applied. Based on this feature analysis, this paper proposes a framework to help selection, combination and use of metadata schemas for digital archiving and preservation. Then, we propose to categorize metadata elements using 5W1H attributes – What, Why, Where, Who, When and How – coupled with a task model derived from the resource lifecycle. In this study, metadata elements of the chosen standards are categorized using the 5W1H attributes and mapped to each other. The mappings are grouped and sorted in accordance with the task model. The 5W1H and task models are applied to six element sets chosen from major metadata standards. Thus, the proposed models help us identify contexts of descriptive elements and define crosswalks among standards.

Keywords: Archives, Archival Metadata, Preservation Metadata, Records Lifecycle, Metadata Schemas, Metadata Standards, Task Model, 5W1H Model, Metadata Interoperability

1. Introduction

Archiving and preservation of digital resources is widely recognized as a crucial issue in the networked information society, especially for memory organizations such as libraries and archives. Those memory organizations select, collect, organize, preserve and provide access to resources. As metadata is essential for all organizations to properly perform these tasks, it is important for them to use appropriate metadata schemas.

There are several major metadata standards used for recordkeeping, archiving and preservation of digital resources – AGRkMS, MoReq2, EAD, OAIS, PREMIS, and so forth. In addition, metadata schemas for other purposes such as finding aids, rights management and accessibility description are used in accordance with requirements given in a particular lifecycle stage. Thus, many metadata schemas are related to archiving and preservation tasks throughout the whole resource lifecycle. This means that, on one hand, we need to appropriately choose metadata standard(s) to define a metadata schema for a particular application system, and, on the other hand, we may need to combine different metadata standards to define an application profile in accordance with the requirements given to the application system. In addition, we may need to define crosswalks between metadata schemas for data exchange.

In general, every metadata schema has its base data model. Every metadata element of a metadata schema is defined as a property (or an attribute) of an entity included in the model. For example, PREMIS has five types of entities in its data model - intellectual entity, digital object, event, right, and agent, and elements, which are called semantic units in PREMIS, are defined for these entities in accordance with the purpose of PREMIS. On the other hand, generally speaking, the resource lifecycle is generally not explicitly referred in the metadata schema. In other words, it is not explicitly defined when a descriptive element should be assigned or have its value revised in the lifecycle. For example, some elements of an intellectual entity of PREMIS such as *title* and *creator* are assigned when the entity is created, which is in the very early stages of the lifecycle, whereas PREMIS is primarily for preservation. Thus, the data model of a metadata standard does not explicitly reflect lifecycle stage(s) for which the standard is primarily designed.

Based on this observation of metadata schemas for archiving and preservation of digital resources, this paper proposes a methodology to analyze metadata schemas in order to help selection, combination and use of metadata schemas used throughout the whole lifecycle of resources, i.e. from creation to preservation and re-use. Prior to this paper, we proposed a feature analysis of metadata schemas for archiving and preservation of digital resources (Baek, 2010), where we used attribute sets from AGLS, EAD, ISAD(G), OAIS, PREMIS and a set of attributes extracted from the decision tree for a preservation process defined by the Digital Preservation Coalition (DPC). Although AGLS and the DPC attribute set are not designed as a metadata schema for archiving or preservation, we have included them in this study to show the characteristics of metadata standards for archiving and preservation.

The result of the feature analysis of those metadata element sets shows relationships between the lifecycle stages and the element sets of the metadata standards. This result made us to further analyze the element sets from the viewpoint of tasks performed in the resource lifecycle. This paper proposes a task-oriented model of the resource lifecycle for more detailed analysis of the element sets. We used 5W1H attributes –what, why, where, who, when and how – to characterize metadata elements in the task oriented model. The 5W1H attributes are widely known as basic attributes to explain an event, e.g. in news articles. We adopted the 5W1H attributes in order to characterize a metadata element in the context of the task where the element is used. The 5W1H attributes help to uniformly categorize the metadata elements of each metadata schema standard. This paper shows mapping examples of the descriptive elements of the standards – AGLS, AGRkMS, EAD, ISAD(G), OAIS, PREMIS and the DPC attribute set.

The metadata standards referred in the paragraphs above are different - OAIS defines a reference model for archival systems in which metadata for preservation is included; PREMIS defines a set of elements based, i.e., dictionary of semantic units; AGRkMS and MoReq2 are a metadata schema standard defined based on ISO 23081; and EAD defines encoding scheme based on ISAD(G). In the rest of this paper, we use the term *metadata standard* as a comprehensive term which implies standards of metadata element set, metadata schema, and reference model.

The rest of this paper is organized as follows. Section 2 describes basic issues of archival and preservation metadata standards. Section 3 shows our previous work detailing features of archival and preservation metadata standards based on the lifecycle model and across the standards. Section 4 proposes the basic models used in this study. Section 5 shows several example mappings among the standards, as the limits of this paper do not allow us to include the entire set of mappings. In section 6 and 7, we present related works, discussions and conclusion.

2. Basic Concepts for Feature Analysis of Archival Metadata Standards

2.1 Resource Lifecycle and Metadata for Archival and Preservation Tasks

There are several categories of metadata schemas, e.g. resource discovery, archival, preservation, resource management, and so forth. Archival and preservation metadata schemas are used primarily to manage resources in accordance with information resource management policy and the lifecycle of the resources. The information resource lifecycle includes several stages such as creation, publication, use, and archiving. Tasks performed on each information resource depend on the lifecycle stages. For example, resource search is performed at all stages but revision is primarily done only in the creation stage in the case of records management.

A lifecycle model is determined based on the type of information resource, so that a lifecycle model for a particular type of resource need not be the same as for another resource type. One or more metadata schemas are used throughout the whole resource lifecycle in accordance with the functional requirements – search, access control, accessibility management, preservation, and so forth. It is crucial for us to be able to define a metadata schema or a scheme to use in accordance with the functional requirements which may be changed over time.

A metadata schema for a domain should be designed based on a standard but it has to satisfy the requirements of the domain. The application profile concept enables us to choose appropriate metadata description elements from one or more base metadata vocabularies in order to better meet such requirements. Selection of appropriate description elements is a key for designing metadata schemas for the application and for enhancing metadata interoperability. It is crucial to be able to systematically map metadata vocabularies to each other. This paper proposes a framework to characterize descriptive elements of metadata vocabularies and improve mapping among them.

2.2 Resource Lifecycle

All resources are created, used, preserved or discarded in accordance with the resource management policy of an organization or institution. This process is called a resource lifecycle or a records lifecycle. In this paper, we use the records lifecycle model of the US National Archives and Records Administration (NARA) as a base lifecycle model. The NARA lifecycle is comprised of seven stages regardless of resource types – (1) Creation, (2) Maintenance and use, (3) Disposition, (4) Arrangement and description, (5) Preservation, (6) Reference and (7) Continuing use. The last two stages are gathered in this study because there is no significant difference between those stages from the viewpoint of metadata. Thus, we have slightly revised the NARA model as shown in Figure 1.

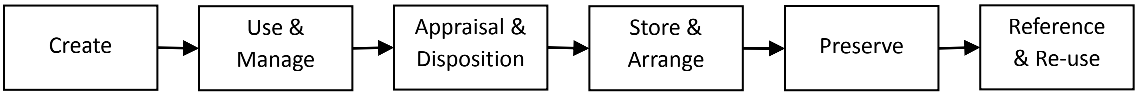


FIG. 1. Lifecycle of This Research

3 Feature Analysis of Metadata Standards based on Resource Lifecycle

3.1 Metadata Schema for Record Management and Archives

This section briefly introduces the metadata schema standards referred in this study. The standards are EAD, ISAD(G), AGRkMS, OAIS and PREMIS, which are designed for archiving. In addition, this study includes AGLS which is primarily designed as a finding aid for Web resources, and an attribute set which is extracted by the authors from a decision procedure for preservation designed by the Digital Preservation Coalition, which is called the DPC Decision Tree. This attribute set, which is not designed as a metadata schema originally, is included in this study because the DPC decision process includes crucial attributes for preservation.

(1) AGLS Metadata (Australian Government Locator Service Metadata)

AGLS Metadata is made to improve the search of both digital and non-digital resources supplied by the Australian Government. It is designed to facilitate, discover, and search for resources online (National Archives of Australian, 2006).

(2) AGRkMS (Australian Government Recordkeeping Metadata Standard)

AGRkMS is issued for national archives and is based on the AGLS standard (National Archives of Australian, 2010). This standard describes the metadata properties that Australian Government agencies should adopt to describe the different entities involved in their business and records management processes (National Archives of Australian, 2008). The data model has five entities - Record, Agent, Business, Mandate and Relationship – following the metadata standard for record keeping (ISO 23081).

(3) EAD (Encoded Archival Description) and ISAD(G) (General International Standard Archival Description)

EAD is a metadata schema for archiving digital resources, keeping compatibility with ISAD(G). In addition to the content description of digital resources, EAD has elements for structural description (Library of Congress, 2002). ISAD(G) is designed for traditional archives and is not specific to digital resources. (International Council on Archives, 2000)

(4) OAIS (Open Archival Information System)

OAIS is an international standard for preservation of digital resources. OAIS is a reference model for archive systems to guarantee access (Harvard University Library, 2008). The OAIS reference model outlines the functions required to access information objects and guarantee efficient long-term preservation (CCSDS, 2002).

(5) PREMIS (Preservation Metadata and Implementation Standard)

PREMIS is a metadata standard primarily defined for the preservation of digital resources. It supplies the data model for preservation and the data dictionary. The PREMIS data model consists of five entities – intellectual entity, digital object, agent, rights and event (Online Computer Library Center, 2008).

(6) Attribute Set of DPC Decision Tree (DPC set)

The decision process for preservation defined by DPC gives guidelines for an evaluation process for preservation of digital resources. It is aligned in a tree structure composed of sequences of Questions-and-Choices (Digital Preservation Coalition, 2006). The decision process is divided into four sections – Selection, Rights & Responsibility, Technology & Metadata, Documents & Costs. In this study, every question is transformed into a metadata attribute or attributes. The attributes transformed are useful in identifying resource attributes used for preservation. Comparison of this attribute set with other metadata standard element sets helps us understand what attributes are examined or assigned during the preservation process. Figure 2 shows a few examples of the transformation of a decision question into a metadata attribute. In this study, we have defined 27 attributes extracted from the questions in the four sections of the decision tree.

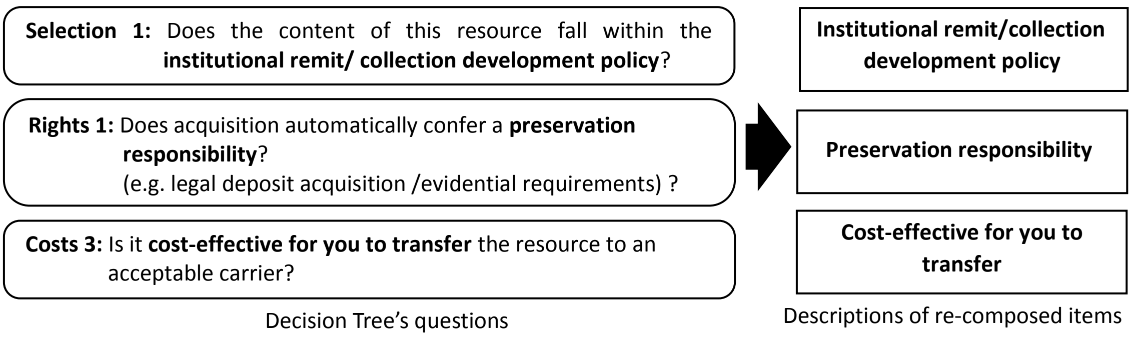


FIG. 2.Transformation of Decision Questions to Metadata Attributes

3.2 Mapping Metadata Elements to Resource Lifecycle

In our previous study, we showed a simple feature analysis based on the relationships among the elements and the lifecycle stages (Baek, 2010). Metadata is created at some stage in the resource lifecycle. The metadata creation schedule in the lifecycle, in other words the timing with which a metadata instance should be created, is determined based on the metadata management policy.

TABLE 1. Metadata standard shown by figures (%)

Lifecycle \ Metadata	AGLS	DPC decision tree	EAD	ISAD(G)	OAIS	PREMIS
Create	16		11	11	1	5
Use & Manage	**28**		13	6	2	22
Appraisal & Disposition	5	**61**	14	15	13	
Store & Arrange	18	39	**33**	**43**	30	21
Preserve	15		20	19	**39**	**45**
Reference & Re-use	18		9	6	15	7

A value for each metadata element can be determined in an earlier stage in the lifecycle. For example, elements such as title and creator can be determined at the stage of resource creation but the elements such as provenance and fixity are determined after the appraisal stage.

In this study, we mapped the lifecycle stages to metadata elements extracted from the metadata standards. For this mapping table, we determined the primary stages where the element value is initially given or revised for every element extracted from metadata standards. Table 1 outlines the mapping of metadata elements to lifecycle stages. The dark portions show the primary stages where a significantly large number of elements from each standard are assigned or reassigned their values.

4. Basic Models for Facet Analysis—Task and 5W1H Models

In this study, we propose to use 5W1H attributes—What, Why, Where, Who, When and How—to categorize descriptive elements of metadata standards for archival and preservation tasks. These attributes are well known for their use in describing news articles. In this chapter, we first propose a task model for the resource lifecycle in section 4.1. Section 4.2 describes the underlying concepts to define the 5W1H model described in section 4.3.

4.1 Task oriented View of Resource Lifecycle – Task Model

The records lifecycle defines stages of records – from creation at offices to preservation in archives. The task model is defined in parallel to the records lifecycle in a previous section. Figure 3 shows the task model is composed of six task groups (T1-T6), defined as follows:

T1: Creation tasks: Tasks used for initial creation including those for the approval process,
T2: Primary Usage tasks: Tasks for primary users to find and browse resources,
T3: Appraisal and Retention tasks: Tasks to select and discard resources,
T4: Archival Transformation tasks: Conversion and transformation tasks for archival storage,
T5: Preservation tasks: Maintenance tasks for archival storage, and
T6: Archival Usage tasks: Tasks to find and use archived resources.

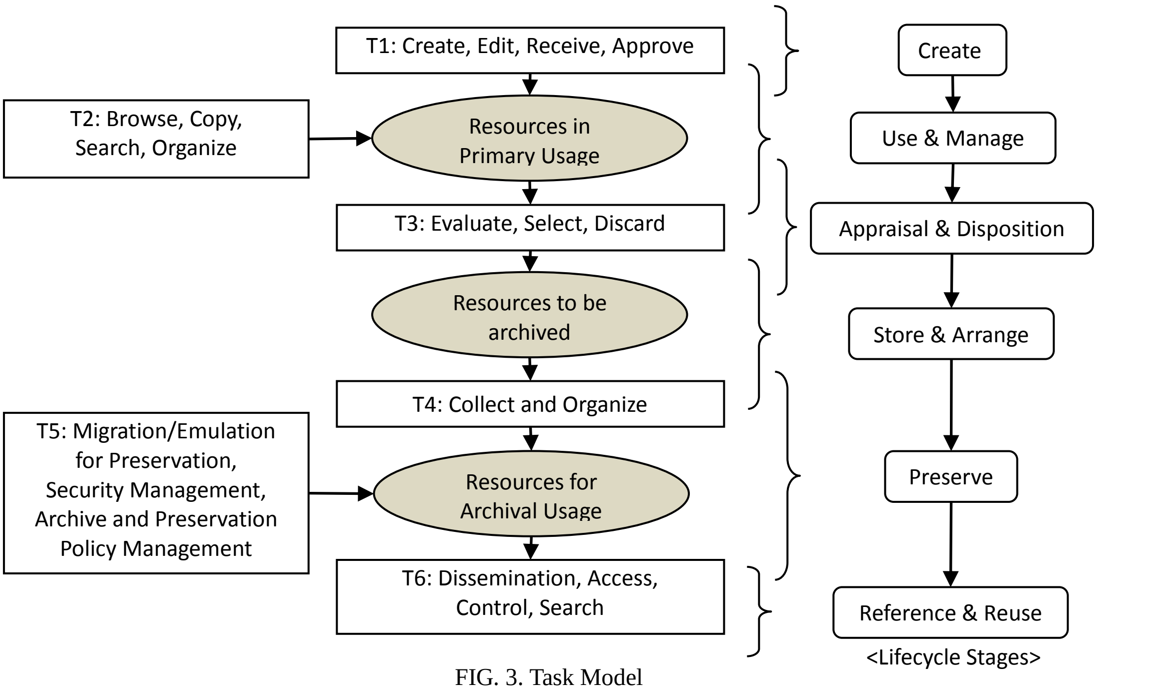
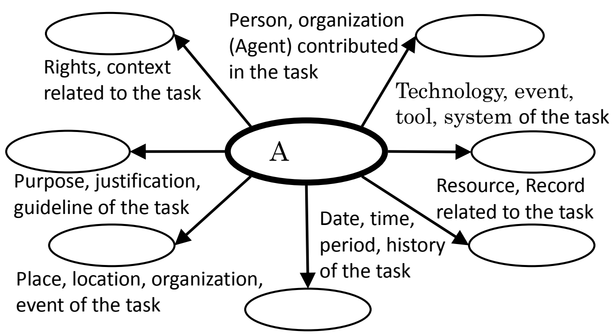


FIG. 3. Task Model

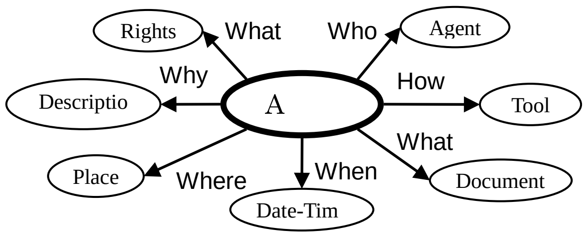
The lifecycle stages are shown to the right of the task model in Figure 3. The task model complements the lifecycle model in the aspects of tasks performed at each stage of the lifecycle and explicitly shows the transition in status of the resources.

4.2 Combining Task oriented Model and Metadata Elements

Several tasks are carried out on a resource during the lifecycle, e.g., creation, edition, search, revision, appraisal, disposal, conversion, and so forth. The metadata of a resource is used to carry out each task. As shown in Figure 4a, every single task is associated with entities shown as a circle. These entities are agents play some role in the task, locations or institution where the task is performed, reasons or guidelines to perform the task, and so forth. Generally, the relationships between a task and its associated entities are determined task-by-task, but we need an appropriate categorization of these tasks. In this study, we propose to use 5W1H attributes as generalized categories to express those relationships as shown in Figure 4b.



a. A Task-centric View



b. A Generalized Task-centric View – 5W1H

FIG.4. Task-centric view of metadata and 5W1H

Many, but not all, of the entities associated with a task are recorded as a metadata value in accordance with the schema used in a particular system. However, in general, data models of metadata standards are defined based on data entities but not tasks. This means that the metadata elements are not explicitly related to the tasks, in spite of the correspondence between lifecycle stages and metadata elements, which we found in our previous study. In addition, the difference of data models of metadata standards have to be taken into account to map their metadata elements. The underlying idea of this study is to use the generalized task-centric view of metadata to map metadata schemas instead of the data entity-centric view in conventional mapping.

In this study, we categorize every metadata element using the 5W1H attributes – who, where, when, what, why and how. This is a reversed view of Figure 4b which shows links labeled with the 5W1H attributes, i.e., an input link to an entity is reversed as a metadata element of the entity.

4.3 5W1H Model

As described in the previous section, 5W1H attributes are used to identify categories of metadata elements. A metadata element category represented by a 5W1H attribute is called a 5W1H category in the rest of this paper. The paragraphs below show definitions of the 5W1H categories:

1) What: Information about preservation processes and tasks such as resources used for reservation, rights and rules for preservation.
2) Why: Reason for an operation on a resource, e.g., purpose of creation, criteria for preservation.
3) When: Time, date, period and era when the task was performed, e.g. date of creation or expiration.
4) Where: Place, location, organization, or institution where the task was performed.
5) Who: Agent related to a resource, e.g., a person or an organization that has made a contribution to the task.
6) How: Operations performed on a resource and related information, e.g., file formats, software tools, rights management, and so forth.

4.4 Categorizing Metadata Elements using 5W1H

In this study, we mapped metadata elements to the 5W1H categories in order to identify semantic relationships between elements of the metadata standards, e.g., same meaning, broader/narrower meaning. The mapping was done manually but we used two sets of keywords to help categorization tasks shown in Table 2 and 3. The keywords included in these tables are words typical in each category or task.

TABLE 2. Classification`s vocabulary with 5W1H model (Example)

5W1H model	Keywords
Who	Agent, Institution, Name, Organization, People, Person etc
When	Date(s), Period, Time etc
Where	Agent, Country, Institution, Location, Organization, Place etc
What	Bibliography, Description, History, Relationship, Right etc
How	Action, Event, Technique, Tool, Transference etc
Why	Purpose, Reason etc

TABLE 3. Classification`s vocabulary with Task model (Example)

Task model	Keywords
T1: Create, Receive, Approve	Create, Make, Produce etc
T2: Browse, Copy, Search, Organize	Access, Manage, Use etc
T3: Evaluate, Select, Discard	Accept, Appraise, Destruct, Select etc
T4: Collect and Organize	Archive, Collect, Manage, Store etc
T5: Migration/Emulation for Preservation, Archive/ Preservation Policy Management	Archive, Manage, Store, Preserve etc
T6: Dissemination, Access, Control, Search	Access, Search, Use etc

5. Mapping Metadata Standards Using the 5W1H model

This section shows 5W1H model by example mappings among the elements of metadata standards chosen for the comparison – AGLS, EAD, AGRkMS, OAIS, PREMIS and the attribute sets of DPC. The full mapping tables are uploaded on the authors' Web site because they are too large to include in this paper. (See http://www.slis.tsukuba.ac.jp/digitalarchive/DC2011/ for the mapping tables.)

5.1 Classifications of Descriptive Element

The paragraphs and tables below show the classification and mapping examples of elements chosen from the metadata standards.

(1) *Publisher* of AGLS Metadata

Publisher element of AGLS means an entity responsible to make a resource available. According to *Describing services* of AGLS, this element may be used to provide details of the organization that provides access to the service. As shown in Table 2, agents such as organizations and institutions are often used as a location. Therefore, Table 4 includes both Who and Where. Corresponding elements of EAD and OAIS in these categories are shown in the table. Other standards have no corresponding element, which means that there is no corresponding element in their own metadata elements vocabularies. (note: An empty field in the table means no corresponding element.)

TABLE 4. AGLS: *Publisher*

5W1H model	Metadata Standards					
	AGLS	AGRkMS	DPC	EAD	OAIS	PREMIS
Who	**Publisher**			Publication Statement	Name of publisher	
				Publisher		
Where	**Publisher**			Publication Statement	Place of Publication	
				Publisher	Name of publisher	

(2) *Date Range* of Australian Government Recordkeeping Metadata Standard (AGRkMS)

Date Range element of AGRkMS means Date and Time associated with an entity. It has *Start Date* and *End Date* as its sub-elements. In the definition of *Date Range*, it includes "Dates", "Times" of keyword of 5W1H model. The category of these elements is obviously *When*. Table 5 shows corresponding elements of AGLS, EAD, OAIS and PREMIS.

TABLE 5. AGRkMS: *Date Range*, *Start Date* and *End Date*

5W1H model	Metadata Standards					
	AGLS	**AGRkMS**	DPC	EAD	OAIS	PREMIS
When	Date	**Date Range**		Date	Date of Publication	dateCreatedByApplication
		Start Date			Change History Before Archiving	
						PreservationLevelDateAssigned
		End Date				

(3) *Multiple media formats* of DPC Decision Tree Attributes

Multiple media formats element of the DPC attribute set, which is created in this research based on the rules shown in section 3.1, means that a resource could have more than one media format, which may be digital or non-digital. Format means a type of media of a resource and also a technology required to render a resource. The former is categorized in What and the latter in How in 5W1H categories.

TABLE 6. DPC Attribute Set: *Multiple media formats*

5W1H model	Metadata Standards					
	AGLS	AGRkMS	**DPC**	EAD	OAIS	PREMIS
What	Format	Format	**Multiple media formats**			
How	Format	Format	**Multiple media formats**			Format

(4) Title of the Unit of EAD

Title of the Unit element of EAD means the name of the described materials. As *Title of the Unit* expresses a name of a resource handled in a task, it is categorized in *What*.

TABLE 7. EAD: *Title of the Unit*

5W1H model	Metadata Schemas for Archive					
	AGLS	AGRkMS	DPC	**EAD**	OAIS	PREMIS
What	Title	Name		**Title of the Unit**	Resource description	

(5) Reason for Creation of OAIS

Reason for Creation element of OAIS is used to specify a reason(s) of creation of a resource. This element is categorized in *Why*. *Description* of AGLS is included here as an element of broader meaning.

TABLE 8. OAIS: *Reason for Creation*

5W1H model	Metadata Schemas for Archive					
	AGLS	AGRkMS	DPC	EAD	**OAIS**	PREMIS
Why	Description				**Reason for Creation**	

(6) *Size* of PREMIS

Size element of PREMIS expresses a technical value such as file size. Elements to express technical values are primarily categorized in *How* in this study. It is mapped to *Description* of AGLS which has a broader meaning and to *Format* of AGRkMS and *Extent* of EAD.

TABLE 9. PREMIS: *Size*

5W1H model	Metadata Schemas for Archive					
	AGLS	AGRkMS	DPC	EAD	OAIS	**PREMIS**
How	Format	Format		Extent		**Size**

5.2 Mapping of Elements in a Task Group

In this study, we organized the mapping table in accordance with the task groups shown in section 4.1. The task groups are used to classify descriptive elements of metadata standards in combination with the 5W1H model. Tasks and 5W1H attributes give contexts for the mappings in this study. The mapping table is not included in this paper because of its size and included in the mapping table online.

6. Related Works

The survey given by Haslhofer and Klas (2010) first describes the metadata used in current information systems and its concepts, and then, metadata interoperability and its problems are explained. Day (2001) described recent developments relating to digital preservation metadata, and introduces digital preservation problems, and the importance of metadata for preservation strategies.

Evans et al. (2005) analyzed and explored development of metadata for multiple archival purposes and relevance to future archival systems in the Clever Recordkeeping Metadata Project. Mckemish et al. (2009) describes a conceptual framework for recordkeeping metadata based on the records continuum model.

Chan and Zeng (2006) studied interoperability problems with multiple metadata schemas, such as having the same subject domain and resources of same type. It then explains three levels - Schema level, Record level, Repository level - from the same interoperability viewpoint. Baca (2003) focuses on the selection of appropriate metadata schemas for Cultural Heritage Information, and describes the metadata mapping and crosswalks among various element sets such as CDMA, EAD, MARC, and VRA Core.

Zeng (1999) discussed application of existing metadata formats to a historical fashion collection and developed a catalog for digitized historical fashion collection objects. Three schemes – AACR, Dublin Core, and Visual Resources Associations (VRA) core – were used in this study. Shimazu et al. (2006) studied metadata exchange interfaces for interdisciplinary contents-sharing. The paper shows the interface module which converts tag-labels using 5W1H attributes.

Our study uses the task model and 5W1H model to identify contexts given to resources which are objectives of metadata description. This is a unique feature of this study. In comparison with those works surveyed in survey papers and those listed in the paragraphs above.

7. Discussions and Conclusion

In this research, we proposed the 5W1H and task models to analyze the features of descriptive elements of archival and preservation metadata standards, and also to create mappings among the standards. This study has identified features of the standards in accordance with the lifecycle stages and the mappings as well. We learned that it is crucial to combine metadata standards for archiving and preservation of digital resources. As is well recognized in the metadata community, the concept of application profiles is crucial for metadata interoperability. The fundamental point of this study is to see metadata standards from a task-oriented view derived from the resource lifecycle. Semantics of metadata elements is primarily given by their underlying data model. The data model is defined based both on analysis of entities included in the domain and tasks on the entities. However, resource lifecycle has to be taken into account in addition to the data models in the case of archival and preservation to combine more than one metadata standard.

The contribution of this study is the use of contextual information extracted from the records lifecycle model. We consider that the two models –Task and 5W1Hmodels – are useful because they provide simple semantics which help to identify meanings of descriptive elements from the viewpoint of tasks in the lifecycle and aspects required to identify the tasks, respectively.

We consider that we need to evaluate the full mapping table given online. This table was made manually, which was a very labor intensive task. We consider that tools to support use of the mappings are crucial issues. These issues are left as our next step. We consider that the contextual information such as tasks and 5W1H attributes is crucial to semantically link metadata elements across the standards.

Acknowledgements

The authors are grateful to Jan Askhoj, a PhD student at the authors' laboratory for his help and contribution in the discussion of this study. The authors thank Prof. Mitsuharu Nagamori and all members of our laboratory for their useful comments and supports. This study was carried our as a part of our research supported by Grant-in-Aid for Scientific Research (A) (#22240023).

Reference

Baca, Murtha. (2003). Practical Issues in Applying Metadata Schemas and Controlled Vocabularies to Cultural Heritage Information. Retrieved April 29, 2011, from http://polaris.gseis.ucla.edu/gleazer/260_readings/Baca.pdf.

Baek, Jaeeun, and Shigeo Sugimoto. (2010). Feature Analysis of Metadata Schemas for Records Management and Archives from the Viewpoint of Records Lifecycle. Journal of Korean Society of Archives and Records Management, 10(2), 75-100.

Chan, Lois Mai, and Marcia Lei Zeng. (2006). Metadata Interoperability and Standardization – A Study of Methodology Part 1. D-Lib Magazine, 12(6). Retrieved April 29, 2011, from http://www.dlib.org/dlib/june06/chan/06chan.html.

Consultative Committee for Space Data Systems (CCSDS). (2002). Reference Model for an Open Archival Information System (OAIS). Retrieved April 29, 2011, from http://public.ccsds.org/publications/archive/650x0b1.pdf.

Day, Michael. (2001). Metadata for Digital Preservation: A Review of Recent Developments. ECDL 2001, LNCS 2163, 161-172.

Digital Preservation Coalition. (2006). Decision Tree for Selection of Digital Materials for Long-term Retention. Retrieved April 29, 2011, from http://www.dpconline.org/advice/preservationhandbook/decision-tree.

Evans, Joanne, Sue Mckemmish and KarunaBhoday. (2005). Create Once, Use Many Times: The Clever Use of Recordkeeping Metadata for Multiple Archival Purposes. Archival Science, 2005(5), 17-42.

Harvard University Library. (2008). Metadata Standards. Retrieved April 29, 2011, from http://hul.harvard.edu/ois/digproj/metadata-standards.html.

Haslhofer, Bernhard, and Wolfgang Klas. (2010). A Survey of Techniques for Achieving Metadata Interoperability. ACM Computing Surveys, 42(2), Article 7, 1-37.

International Council on Archives. (2000). ISAD(G): General International Standard Archival Description, Second edition. Retrieved April 29, 2011, from http://www.ica.org/sites/default/files/isad_g_2e.pdf.

International Organization for Standardization. (2006). ISO 23081-1:2006: Information and documentation - Records management processes - Metadata for records - Part 1: Principles.

International Organization for Standardization. (2009). ISO 23081-2:2009: Information and documentation - managing metadata for records - Part 2: Conceptual and implementation issues.

Library of Congress. (2002). Encoded Archival Description. Retrieved April 29, 2011, from http://www.loc.gov/ead.

McKemmish, Sue, Glenda Acland, Nigel Ward and Barbara Reed. (2009). Describing Records in Context in the Continuum: the Australian Recordkeeping Metadata Schema. Retrieved April 29, 2011, from http://infotech.monash.edu/research/groups/rcrg/publications/archiv01.html.

National Archives and Records Administration. Records have Lifecycle. Retrieved April 29, 2011, from http://www.archives.gov/about/info/whats-a-record.html.

National Archives of Australia. (2006). Australian Government Implementation Manual : AGLS Metadata. Retrieved April 29, 2011, from http://www.naa.gov.au/Images/AGLS_Manual_v2_tcm2-914.pdf.

National Archives of Australia. (2008). Australian Government Recordkeeping Metadata Standard Version 2.0. Retrieved April 29, 2011, from http://www.naa.gov.au/Images/AGRkMS_Final%20Edit_16%2007%2008_Revised_tcm2-12630.pdf.

National Archives of Australia. (2010). AGLS Metadata Australian Government Implementation Manual. Retrieved April 29, 2011, from http://www.naa.gov.au/Images/AGLS%20Australian%20Government%20Implementation%20Manual%20v3_0_tcm2-913.doc.

Online Computer Library Center. (2008). Data Dictionary for Preservation Metadata: Final Report of the PREMIS Working Group. Retrieved April 29, 2011, from http://www.loc.gov/standards/premis/v2/premis-2-0.pdf.

Shimazu, Keiko, atsya Arisawa, Isao Saito. (2006). Interdisciplinary contents management using 5W1H interface for metadata. IEEE/WIC/ACM International Conference on Web Intelligence (WI'06), 909 – 912.

Zeng, Marcia Lei. (1999). Matadata Elements for Object Description and Representation: A Case Report form a Digitized Historical Fashion Collection Project. Journal of the American Society for Information Science. 50(13), 1193-1208.